



2008

Proceedings of

11th International Conference on Computer and Information Technology (ICCIT 2008) & Workshops

24 - 27 December, 2008

ICCIT 2008

IDCS
2008

DMAI
2008

IEEE Catalog Number: CFP0817D
ISBN: 978-1-4244-2136-7
Library of Congress: 2008900703

©2008 IEEE

International Workshop on
Data Mining and Artificial Intelligence
(DMAI'08)

International Workshop on
Internet and Distributed Computing Systems
(IDCS'08)



Organized by

Faculty of Electrical & Electronic Engineering

Khulna University of Engineering & Technology
Khulna-9203, Bangladesh

Technical Co-sponsor



IEEE Bangladesh Section
Dhaka, Bangladesh

ICCIT 2008

The International Conference On Computer and Information Technology and Workshops

24-27 December, 2008

Khulna University of Engineering and Technology (KUET), Bangladesh



CONFERENCE PROCEEDINGS



Organized by

Faculty of Electrical and Electronic Engineering (EEE)

Khulna University of Engineering and Technology (KUET)

Khulna-9203, Bangladesh

<http://www.kuet.ac.bd/>

ICCIT 2008

The International Conference On Computer and Information Technology and Workshops

Editors

Prof. Dr. M.M.A Hashem
Dr.M.M. Hafizur Rahman
Pintu Chandra Shill
Rushdi Shams
Md. Jahidul Islam

Published by

Faculty of Electrical and Electronic Engineering (EEE)
Khulna University of Engineering and Technology (KUET)
Khulna-9203, Bangladesh

Additional Copies of the Proceedings can be obtained from

ICCIT 2008 Secretariate
Department of Computer Science & Engineering (CSE)
Khulna University of Engineering and Technology (KUET)
Khulna-9203, Bangladesh

IEEE Catalog Number: CFP0817D-PRT

ISBN: 978-1-4244-2135-0

Library of Congress: 2008900703

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Operations Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved. Copyright ©2008 by the Institute of Electrical and Electronics Engineers.

ICCIT 2008 Advisory Committee



Chair

Prof. Dr. Md. Nowsher Ali

Vice-Chancellor

Khulna University of Engineering & Technology

Co-Chair

Prof. Dr. Bashudeb Chandra Ghosh

Dean, Faculty of Electrical & Electronic Engineering, KUET

Members

Prof. Dr. Muhammed Alamgir

Dean, Faculty of Civil Engineering, KUET

Prof. Dr. Naseem Ahmed

Dean, Faculty of Mechanical Engineering, KUET

Prof. Dr. Md. Mahbub Alam

Director, Students' Welfare, KUET

Prof. Dr. Md. Monjur Hossain

Director, Planning & Development, KUET

Prof. Dr. A. N. M. Mizanur Rahman

Director, Research & Extension, KUET

Prof. Dr. Md. Abul Bashar

Head, Department of Civil Engineering, KUET

Prof. Dr. Tarapada Bhowmick

Head, Department of Mechanical Engineering, KUET

Prof. Dr. Md. Kutub Uddin

Head, Department of Industrial Engineering & Management, KUET

Prof. Dr. Md. Nurunnabi Mollah

Head, Department of Biomedical Engineering, KUET

Prof. Dr. Sayed Ali Molla

Head, Department of Energy Technology, KUET



KUET



ICCIT 2008 International Program Committee

Chair

Mohammad A. Karim

Vice-President for Research
Old Dominion University
Norfolk, VA 23508, USA

Jemal Abawajy

Deakin University

Mohammad Abdullah

IBM Global Services

Sheikh I. Ahamed

Marquette University

Aftab Ahmad

Norfolk State University

Iftekhar Ahmad

Edith Cowan University

Farid Ahmed

The Catholic University of America

Rakib Ahmed

Australia National University

M. Ferdous Alam

LSI Corp

Mohammad S. Alam

University of South Alabama

Shawkat Ali

Central Queensland University

Abdul Rahman Alsamman

University of New Orleans

Vijayan Asari

Old Dominion University

Mohammed Atiquzzaman

University of Oklahoma

Abdul S. Awwal

Lawrence Livermore National Laboratory

S. Mahfuz Aziz

University of South Australia

Peter Brass

The City College of the City University of New York

Arshad Chowdhury

Georgia Inst. of Technology

Rezaul A. Chowdhury

University of Texas Austin

Tom Dhaene

University of Antwerp

Manzur Murshed

Monash University, Gippsland

Akiko Nakaniwa

Osaka Sangyo University

Suman K. Nath

Microsoft Research

Stephan Olariu

Old Dominion University

Janos Pach

City College of CUNY

Yi Pan

Georgia State University

Manoranjan Paul

Monash University

Members

Hairong Qi

University of Tennessee

Abdul Wahab A. Rahman

Nanyang Technological University

M. A. Rashid

Massey University at Albany, New Zealand

M. Mamun-or-Rashid

Kyung Hee University

Mohammad A. Salam

Southern University and A&M College

Salil K. Sanyal

Jadavpur University

Ruhul A. Sarker

University of New South Wales

Badrul M. Sarwar

eBay Inc.

M. Hasan Shaheed

Queen Mary, University of London

Serhan Dagtas

University of Arkansas at Little Rock

Sekharjit Datta

Loughborough University

Phalguni Gupta

Indian Institute of Technology, Kanpur

Ekram Hossain

University of Manitoba

M. Alamgir Hossain

University of Bradford

Mostofa K. Howlader

Oakridge National Lab

M. Tamjidul Hoque

Griffith University

M. Julius Hossain

Dublin City University

Mohammed N. Islam

ODU Research Foundation

Abusaleh M. Jabir

Oxford Brookes University

Hai Jin

Huazhong University of Science & Technology

Mustafa Kamal

University of Central Missouri

Joarder Kamruzzaman

Monash University

Gour Karmakar

Monash University

Iffat H. Kazi

Sun Microsystems

Jamil Y. Khan

University of Newcastle

Javed I. Khan

Kent State University

Latifur Khan

University of Texas at Dallas

Mariofanna Milanova

University of Arkansas at Little Rock

Ferdous Sohel

University of Western Australia

Min Song

Old Dominion University

Golam Sorwar

Southern Cross University

Mahbubur R. Syed

Minnesota State

University, Mankato

Tarek M. Taha

Clemson University

Koichi Wada

Nagoya Institute of Technology

Kuang-Ching Wang

Clemson University

Yong Xue

London Metro University

Mohammed Yeasin

University of Memphis

Salahuddin Zabir

Kyung Hee University

Panasonic R&D

Sajjad Zahir

University of Lethbridge

Shuqun Zhang

College of Staten Island of CUNY

Xiao Zhou

Tohoku University

Zhigang Zhu

City College of CUNY

Elena Zudilova-Seinstra

University of Amsterdam

Muhammad Zulkernine

Queen's University





Organizing Committee

Chair

Dr. M.M.A. Hashem

Professor & Head of the Dept.,
Dept. of Computer Science & Engineering (CSE)
Khulna University of Engineering & Technology (KUET)
Khulna - 9203, Bangladesh

Secretary

Dr. Md. Rafiqul Islam

Professor & Head of the Dept.,
Dept. of Electrical & Electronic Engineering (EEE)
Khulna University of Engineering & Technology (KUET)
Khulna - 9203, Bangladesh

Asstt. Secretary

Dr. Md. Osman Goni

Associate Professor & Head of the Dept.,
Dept. Electronics & Communication Engineering (ECE)
Khulna University of Engineering & Technology (KUET)
Khulna - 9203, Bangladesh

Members

Professor Dr Farruk Ahmed
Professor Dr Haider Ali
Professor Dr Muhammad Masroor Ali
Professor Dr Mohammed Anwer
Professor Dr. Md. Aynal Haque
Professor Dr Abul L Haque
Dr. K.M. Azhaul Hasan
Professor Dr Syed Akhter Hossain
Professor Dr M Zafar Iqbal
Professor Dr Md Rafiqul Islam
Brigadier General Md. Tajul Islam
Prof. Dr. S. M. Lutful Kabir
Professor Dr Abul Kashem
Dr. Mohammod Abul Kashem
Professor Dr M Kaykobad
Professor Dr Mozammel Haque Azad Khan
Professor Dr Abdullah-Al Mamun
Professor Dr M A Mottalib
Dr. Ashfaque Rahman
Dr. M.M. Hafizur Rahman
Professor Dr M Lutfar Rahman
Dr. Miftahur Rahman
Professor Dr Chowdhury Mofizur Rahman
Professor Dr. Md. Saidur Rahman
Professor Dr Chowdhury Mofizur Rahman
Professor Dr Zahidur Rahman
Professor Dr Sayeed Salam
Professor Dr M A Sobhan
Dr. Shahid-Uz-Zaman

North South University
University of Dhaka
Bangladesh University of Engineering & Technology
Independent University Bangladesh
Bangladesh University of Engineering & Technology
North South University
Khulna University of Engineering & Technology
University of Liberal Arts Bangladesh
Shahjalal University of Science and Technology
Khulna University
Military Institute of Science & Technology, MIST
Bangladesh University of Engineering & Technology
Bangladesh University of Engineering & Technology
Dhaka University of Engineering & Technology
Bangladesh University of Engineering & Technology
East West University
Ahsanullah University of Science & Technology
Islamic University of Technology
American International University Bangladesh
Khulna University of Engineering & Technology
University of Dhaka
North South University
United International University
Bangladesh University of Engineering & Technology
United International University
Jahangirnagar University
BRAC University
Independent University Bangladesh
Rajshahi University of Engineering & Technology





ICCIT 2008 REVIEWERS

<i>Name</i>	<i>Organization</i>	<i>Country</i>
Jemal Abawajy	Deakin University	Australia
Mohammad Abdullah	IBM Global Services	USA
Sheikh I. Ahamed	Marquette University	USA
Aftab Ahmad	Norfolk State University	USA
Iftekhar Ahmad	Edith Cowan University	Australia
Farid Ahmed	Johns Hopkins University Applied Physics Laboratory	USA
Rakib Ahmed	Australian National University	Australia
Sabbir Ahmed	Monash University	Australia
M. Ferdous Alam	LSI Corp	USA
Mohammad S. Alam	University of South Alabama	USA
Miraz Al-Mamun	Monash University	Australia
Mortuza Ali	Monash University	Australia
Shawkat Ali	Central Queensland University	Australia
Abdul Rahman Alsamman	University of New Orleans	USA
Vijayan Asari	Old Dominion University	USA
Mohammed Atiquzzaman	University of Oklahoma	USA
Mohammad Awrangjeb	University of Melbourne	Australia
Abdul S. Awwal	Lawrence Livermore National Laboratory	USA
S. Mahfuz Aziz	University of South Australia	Australia
Peter Brass	The City College of the City University of New York	USA
Nathan Brewer	Australian National University	Australia
Arshad Chowdhury	Georgia Institute of Technology	USA
Rezaul A. Chowdhury	University of Texas Austin	USA
Tom Dhaene	University of Antwerp	Belgium
Serhan Dagtas	University of Arkansas at Little Rock	USA
Pham Minh Duc	University of South Australia	Australia
Omar Farooq	Aligarh Muslim University	India
Phalguni Gupta	Indian Institute of Technology, Kanpur	India
Nilimesh Halder	University of Western Australia	Australia
Ahsanul Haque	Monash University	Australia
Ekram Hossain	University of Manitoba	Canada
M. Alamgir Hossain	University of Bradford	United Kingdom
Mostofa K. Howlader	Oakridge National Laboratory	USA
M. Tamjidul Hoque	Griffith University	Australia
M. Julius Hossain	Dublin City University	Ireland
Tasadduq Imam	Monash University	Australia
Md. Kamrul Islam	Monash University	Australia
Mohammed N. Islam	ODU Research Foundation	USA
Muhammad Sirajul Islam	The University of Melbourne	Australia
Abusaleh M. Jabir	Oxford Brookes University	United Kingdom
Hai Jin	Huazhong University of Science & Technology	China
Mustafa Kamal	University of Central Missouri	USA
Joarder Kamruzzaman	Monash University	Australia
Mohammad A. Karim	Old Dominion University	USA
Gour Karmakar	Monash University	Australia
Iffat H. Kazi	Sun Microsystems	USA
Jamil Y. Khan	University of Newcastle	Australia





ICCIT 2008 REVIEWERS

<i>Name</i>	<i>Organization</i>	<i>Country</i>
Latifur Khan	University of Texas at Dallas	USA
Quazi Ehsanul Kabir Mamun	Monash University	Australia
Mariofanna Milanova	University of Arkansas at Little Rock	USA
Manzur Murshed	Monash University, Gippsland	Australia
Suman K. Nath	Microsoft Research	USA
Janos Pach	City College of CUNY	USA
Yi Pan	Georgia State University	USA
Manoranjnan Paul	Monash University	Australia
Hairong Qi	University of Tennessee	USA
Abdul Wahab A. Rahman	Nanyang Technological University	Singapore
M. Ziaur Rahman	Edith Cowan University	Australia
Masud Rahman	Department of ITRDLG	Australia
M. A. Rashid	Massey University, Albany	New Zealand
M. Mamun-or-Rashid	Kyung Hee University	South Korea
Anwar Sadat	Monash University	Australia
Mohammad A. Salam	Southern University and A&M College	USA
Salil K. Sanyal	Jadavpur University	India
Ruhul A. Sarker	University of New South Wales	Australia
Badrul M. Sarwar	eBay Inc.	USA
M. Hasan Shaheed	Queen Mary, University of London	United Kingdom
Md. Atiur Rahman Siddique	Monash University	Australia
Ferdous A. Sohel	University of Western Australia	Australia
Min Song	Old Dominion University	USA
Golam Sorwar	Southern Cross University	Australia
Mahbubur R. Syed	Minnesota State University, Mankato	USA
Tarek M. Taha	Clemson University	USA
Kuang-Ching Wang	Clemson University	USA
Yong Xue	London Metro University	United Kingdom
Mohammed Yeasin	University of Memphis	USA
Salahuddin Zabir	Kyung Hee University	South Korea
Sajjad Zahir	University of Lethbridge	Canada
Shuqun Zhang	College of Staten Island of CUNY	USA
Xiao Zhou	Tohoku University	Japan
Zhigang Zhu	City College of CUNY	USA
Elena Zudilova-Seinstra	University of Amsterdam	The Netherlands
Muhammad Zulkernine	Queen's University	Canada





ICCIT 2008 Local Management Committee

Information and Banner

Muhammad Arif Hossain (MATH)
Md. Abdus Salim Mollah (CSE)
Md. Mohsin Ali (CSE)
Md. Nazmul Haque (EEE)
Md. Abdur Rahman (PRO)

Session Management

Ashraful Ghani Bhuiyan (EEE)
Mohiuddin Ahmad (EEE)
Mohammad Mashud (ME)
Md. Zulfikar Hossain (ECE)
Rushdi Shams (CSE)
Md. Jahidul Islam (CCC)
Mehirlal Sarkar (CSE)

Food and Refreshment

Md. Toyebur Rahman (ADMIN)
A.N.M. Enamul Kabir (EEE)
Md. Shafiul Islam (CE)
Md. Nazrul Islam (EEE)
Sheikh Md. Masudul Ahsan (CSE)
Md. Zaidur Rahman (MATH)
Md. Atikul Islam (CE)
Md. Hasanuzzaman (MATH)
Kingshuk Jubaer Islam (IEM)
Md. Elias Enam (ME)
Md. Nuruzzaman (ADMIN)
Md. Enamul Haque (ADMIN)

Souvenir and Certificates

Mihir Ranjan Halder (ME)
Md. Abdur Rafiq (EEE)
Shaheena Sultana (CSE)
Naruttam Kumar Roy (EEE)
Hussain Md. Abu Nyeem (ECE)
Khaled Mahbub Morshed (ECE)
Kheya Banerjee (ECE)

Publication

Mohammad Mashud (ME)
Md. Zulfikar Hossain (ECE)
Md. Rezaur Raihan (EEE)
Md. Mahfuz Sarwar (ME)
Rushdi Shams (CSE)
Md. Jahidul Islam (CCC)

Finance Management

Pintu Chandra Shill (CSE)
Debabrata Kumar Karmokar (EEE)
Hussain Md. Abu Nyeem (ECE)

Registration and Kit Distribution

Ashraful Ghani Bhuiyan (EEE)
Mohammad Abu Yousuf
Md. Abdur Rafiq (EEE)
G.G. Md. Nawaz Ali (CSE)
Md. Alamgir Hossain (EEE)
Kalyan Kumar Halder (EEE)
Md. Selim Hossain (EEE)

Hotel & Transportation

A.N.M. Mizanur Rahman (ME)
K.M. Azharul Hasan (CSE)
Rushdi Shams (CSE)
Md. Kowsar Hossain (CSE)
Md. Habibullah (EEE)

Invitation

Md. Mahbub Alam (PHY)
Mohammad Arif Hossain (MATH)
Md. Shahjahan (EEE)
Md. Sherajul Islam (EEE)
Azizur Rahman (IEM)
Md. Amjad Hossain (CSE)

Foreign Guest Handling

Ashraful Ghani Bhuiyan (EEE)
M.M. Hafizur Rahman (CSE)
K.M. Azharul Hasan (CSE)
Mohiuddin Ahmad (EEE)
Tanay Dey (CSE)
Md. Mahbub Hasan (EEE)
Md. Sherajul Islam (EEE)
Tanvir Hassan (Daffodil Univ)

Volunteer Coordination

Md. Nazrul Islam (CSE)
Pintu Chandra Shill (CSE)
Khaled Mahabub Morshed (ECE)
Debabrata Kumar Karmokar (EEE)
Md. Amjad Hossain (CSE)

Inauguration Session Management

Ashraful Ghani Bhuiyan (EEE)
Tanay Dey (CSE)
Rushdi Shams (CSE)
Kheya Banerjee (ECE)



ACKNOWLEDGEMENTS

The organizing committee acknowledges the efforts of all those who have contributed their valuable time and efforts as reviewers in ensuring high quality technical papers for the 11th International Conference on Computer and Information Technology (ICCIT 2008).

Deepest appreciation to all faculty members of the Faculty of Faculty of Electrical and Electronic Engineering, Khulna University of Engineering and Technology (KUET) for their sincere cooperation in making the conference successful. Appreciation also goes to all parties who have contributed to the success of 11th ICCIT 2008 in KUET.

Finally, the organizing committee would like to express their thanks to the followings for sponsoring this conference:

University Grants Commssion of Bangladesh
United International University (UIU), Bangladesh
Eastern University, Bangladesh
University of Liberal Arts (ULAB), Bangladesh
East West University (EWU), Bangladesh
North South University (NSU), Bangladesh
BRAC University, Bangladesh
Visual Magic Corporation(VMC) Limited, Bangladesh
Apartment design & development ltd.(ADDL), Bangladesh
Link3 Technologies Ltd.
Islamic University of Technology(IUT), Bangladesh
Meghna Cement Mills Ltd., Bangladesh
Energypac Power Generation Ltd., Bangladesh
Independent University, Bangladesh
American International University-Bangladesh (AIUB)
Military Institute of Science & Technology, Bangladesh



MESSAGES

Chairman

University Grants Commission
of Bangladesh



MESSAGE

It is heartening to note that the Faculty of Electrical & Electronic Engineering, Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh is organizing the 11th International Conference on Computer and Information Technology (ICCIT 2008) during December 25-27, 2008.

Knowledge acquisition is a continuous process that involves generation of new ideas, thoughtful outlining of possibilities and consummate implementation. The attainment is fulfilled when the knowledge is shared with others. The recent rapid development in Information and Communication Technology has urged scholars, scientists, researchers, academicians, theorists, professionals, technologists, and entrepreneurs to share the improvement they are inducing. Conference like ICCIT is one of the strong and prestigious platforms for them to sustain the newer technologies they are inventing. One of the aspirations of the ICCIT 2008 is to foster co-operations among the researchers of the same and interdisciplinary fields from Bangladesh and abroad. The knowledge intakes from the conference will have a meticulous effect to broaden our views and to contemplate for inventing more relevant ideas in near future.

I hope that the research papers in different fields of ICT to be presented in the conference, discussions in the technical and Keynote sessions will benefit the scientific and engineering community of Bangladesh and around the globe. I appreciate the sincere efforts of the organizers in holding the conference on a regular basis. I wish ICCIT 2008 all success.

Nazrul Islam

(Professor Nazrul Islam)

Chief Guest

ICCIT 2008



KUET

ICCIT



2008



Vice Chancellor

Khulna University of Engineering & Technology,
Khulna 9203, Bangladesh



MESSAGE

It is an honor and privilege for Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh to have the opportunity to host the 11th International Conference on Computer and Information Technology (ICCIT 2008) and Workshops to be held on 24-27 December, 2008. It is my utmost pleasure to welcome participants to a congregation of researchers and scholars to share knowledge and ideas for promoting knowledge-based society.

The essence of Information and Communication Technology has been pervaded greatly over the years. ICT is now penetrating every walk of our lives- health, education, national security, economy, business, e-governance. The geographically sparse world has been commuted into a single logical demographic area amid ICT and ICCIT 2008 is the one of such exercises that made it possible. Since its commencement in 1998, ICCIT became the platform for the researchers to stand upon and to share the inwardness and pith of ICT which boosts up the spirit of intellectual growth.

KUET is committed to uphold the heritage of the conference and has indulged a new dimension for researchers, professionals, academicians, entrepreneurs, and students from diversified areas. I heartily welcome all the participants in ICCIT 2008 with ardent hope that the bonding among the researchers, faculties, and students will be strengthened through the congress.

I wish participants from home and abroad a pleasant time during the conference and workshops as Bangladesh is a land of natural beauty and heritage. I wish every success of the conference and workshops.

(Prof. Dr. Md. Nawsher Ali)
Chair, Advisory Committee
ICCIT 2008





Dean

Faculty of Electrical & Electronic Engineering
Khulna University of Engineering & Technology,
Khulna 9203, Bangladesh



MESSAGE

I am highly pleased to write few words for 11th International Conference on Computer and Information Technology (ICCIT 2008) to be held in KUET from December 24-27, 2008. For the first time the conference is being organized at Khulna, a place of about 300 kilometers from Dhaka organized by the Faculty of Electrical & Electronic Engineering, KUET. Khulna, an important divisional city of the south west region of Bangladesh, has its heritage for the Sunderbans, the Mongla port and conventional industries.

The Information and Communication Technology (ICT) aided by computers has emerged with its intensive applications for the global human life. A wide window has been opened by this technology to contribute to the society. New fields of research and areas of applications for the improvement of life style of human race and to alleviate unemployment problem are the benefits of this technology. In Bangladesh, we have a huge population and unemployment. I feel the conference will focus on this problem and deduce suggestions to overcome the problem and help advancement of the technology.

I come to know that the conference is enriched by 138 technical papers and 6 invited lectures. About 200 participants from home and abroad are going to participate in the conference. I think it will be a huge gathering of experts and fresh researchers. I hope the concluding round table discussion will find out some causes and probable solutions that will help to develop our ICT.

I feel it as a huge involvement to organize an international conference like ICCIT at Khulna. A group of teachers of the faculty have been actively working for a long period to make the conference successful. In spite there may be some unseen inconveniences and I request all the participants to overlook and forgive if any such by your generous heart.

At last, I wish grand success of the conference and pleasure stay of all participants at KUET.

(Prof. Bashudeb Chandra Ghosh)
Co-chair, Advisory Committee
ICCIT 2008



KUET

2008 11th International Conference on Computer and
Information Technology (ICCIT 2008) & Workshops



IEEE Bangladesh Section



MESSAGE

I am delighted to know that the Faculty of Electrical and Electronic Engineering, Khulna University of Engineering and Technology (KUET) is organizing 11th International Conference on Computer and Information Technology (ICCIT 2008) and Workshops. It is also pleasurable to note that ICCIT is going to be held outside Dhaka for the first time. I congratulate KUET authority and organizers for this historic event. ICCIT has prodigious reputation for being the platform of researchers and scholars from home and abroad. I hope this congregation of ICCIT will offer tangible mantle towards a fertile exchange of knowledge on the broad spectrum of ICT and will help participants in their endeavor to develop a knowledge-based society.

IEEE, the largest professional association in the world, always fosters technological innovation and excellence for the benefit of humanity. Besides supporting its members, IEEE is supporting education, training, research and socio-culture activities all over the world. KUET has a strong and active member group of IEEE and it has established student branch long back. In fact, out of 7 IEEE student branches in Bangladesh, only KUET branch is outside of Dhaka. With the technical co-sponsorship of IEEE Bangladesh Section, ICCIT 2008 is going to be the largest congress of scholars of ICT and related fields. IEEE is committed to local technical bodies to provide support to replete the global technical needs.

On behalf of IEEE Bangladesh Section, I like to welcome all the authors and participants. I would like to thank the organizers, participants, sponsors and others whose sincere efforts will constitute a fruitful communication between national and international experts in computer and IT.

I wish a grand success of ICCIT 2008.


(M. Aynal Haque)
Chair, IEEE Bangladesh Section
Professor, Dept. of EEE, BUET



Office of the Vice President for Research

2035 Hughes Hall
Norfolk, Virginia 23529-0226
Phone (757) 683-3460, Fax (757) 683-5902



MESSAGE

In spite of the economic signs of doom and gloom, we are faced with dramatic changes as far as computing and information science are concerned. These two fields are intertwined and are going to be impacted by the technology that is still evolving, by the science that needs to be re-interpreted, simulations that need to be designed, visualization that needs further exploration and systems that are still being re-engineered. It is in this backdrop, the 11th International Conference on Computer and Information Technology is taking place, this time being hosted by Khulna University of Engineering & Technology.

This conference is international in its scope and reach; yet of serious technical significance to Bangladesh and Bangladeshi institutions of higher learning. It continues to serve as an engine that is propelling forward a large number of energized students, researchers, and faculty members of Bangladeshi universities. It is heartening to see that each year, more professionals and institutions are joining this bandwagon. Consequently, this conference is getting continually re-energized, refocused and retooled. This year's conference received 538 submissions of which 158 were finally accepted. This amounts to an acceptance rate of 29.4%. Of the accepted papers, 42.4% came from authors from outside of Bangladesh. A team of 83 reviewers – all from outside of Bangladesh evaluated these manuscripts; 89.03% of the assignments made to the reviewers were completed. The true measure of the success of this international conference, however, will be judged by the newer ideas and techniques so disseminated through it, from the impact it will have by nurturing a technological climate in Bangladesh. Starting this year, ICCIT papers will get included in the IEEE Xplore Digital Library and of those which will be enhanced further, upon additional review and revision, may be eligible to be included in one of the five Academy Publisher journal special issues in 2009.

On behalf of the international program committee, I would like to thank each one of the participants, authors, reviewers, speakers, members of the international program and organizing committees, advisors, sponsors, session chairs, volunteers, and supporting organizations and universities for their respective contributions in making this collective effort a genuine success. Last but not the least, ICCIT 2008 owes a particular "thank you" to KUET and its leadership for hosting this international event.



(Mohammad A. Karim)

Chair, International Program Committee
ICCIT 2008

2008 11th International Conference on Computer and
Information Technology (ICCIT 2008) & Workshops



Head of the Department

Dept. of Computer Science and Engineering
Khulna University of Engineering & Technology,
Khulna 9203, Bangladesh



MESSAGE

With delicate pleasance, exquisite privilege and honor, on behalf of the Organizing Committee, I am welcoming you to the 11th International Conference on Computer and Information Technology (ICCIT 2008) and Workshops. The conference is organized by the Faculty of Electrical & Electronic Engineering, Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh and will be held on December 24-27, 2008.

The continuum of ICCIT since 1998 has proved the conference to be the largest and most prestigious international ICT conference in Bangladesh. ICCIT 2008 will surely indulge the intellectual community of home and abroad by presenting the newest ideas, technological advancements, thoughtful researches, and generative implementations.

We leaned on 83 prominent reviewers from 13 different countries to select the papers. The selected papers are categorized into 26 different areas. As we have pervaded the call for paper around the globe efficiently, researchers from 24 different countries have submitted their papers initially and papers from 17 different countries will be presented in the conference which adds an extra relish of diversity to the conference. The papers presented in the conference will be published in IEEEExplore of which some of them, if selected, will get the opportunity to be placed in special issue of five distinguished and well recognized journals published by Academy Publisher. Moreover, for the first time in the history of ICCIT, two workshops- Internet and Distributed Computing Systems (IDCS 2008) and Data Mining and Artificial Intelligence (DMAI 2008) organized by University of Melbourne and University of Central Queensland, respectively will be hosted by KUET in conjunction with ICCIT 2008. 14 papers will be presented in IDCS 2008 and 16 papers will be presented in DMAI 2008.

I would like to express my gratefulness to the IPC Chairman Professor Dr. Mohammad A. Karim and his team for successfully reviewing large number of papers in a short span of time. My heartfelt gratitude goes to the sponsors as without their assistance to this noble endeavor, it would be difficult for KUET to deplete the conference and the workshops successfully. Lastly, I wish my solemnest thanks to the Organizing Committee Members for their affable exertions to support and mentor the ICCIT 2008.

(Prof. Dr. M.M.A. Hashem)
Chair, Organizing Committee
ICCIT 2008



Table of Contents

Title and Authors Name	Page No.
Algorithms	
Cutting a Cornered Convex Polygon Out of a Circle <i>Syed Ishtiaque Ahmed, Md. Ariful Islam, Masud Hasan</i>	1
On the Minimization of Complete Test Set of Reversible k -CNOT Circuits for Stuck-at Fault Model <i>Muhammad Ibrahim, Ahsan Raja Chowdhury and Hafiz Md. Hasan Babu</i>	7
Two-Level Dictionary-Based Text Compression Scheme <i>Md. Ziaul Karim Zia¹, Dewan Md. Fayzur Rahman², and Chowdhury Mofizur Rahman</i>	13
Generation of Random Channel Specifications for Channel Routing Problem <i>Achira Pal , Tarak N. Mandal , Alak K. Datta &, Debojit Kundu and Rajat K.Pal</i>	19
Efficient Generation of Combinatorial Families <i>Md. Shariful Islam Bhuyan and Md. Saidur Rahman</i>	25
Information Systems and Software engineering	
City Planning and Development Using Geographic Information Systems <i>Mustafa A. Kamal</i>	31
An Application Program Interface for vBulletin <i>Saumen Seal Sami, Tazilur Rahman, Khondker Shajadul Hasan, Javed Siddique</i>	36
Structural Operational Semantics of Packages in Java <i>Abdullah Al Farook, Mohammad Shamsul Arefin and Md. Moshikul Hoque</i>	40
Automated Digital Archive for Land Registration and RecordSystem <i>Mabruka Toaha, Shahriar Khan</i>	46
Development of a Highly Optimized Preemptive Real Time Operating System <i>Md. Al Mehedi Hasan, Shamim Ahmad</i>	52
Bio-Informatics	
Resource Capability Discovery and Description Management System for Bioinformatics Data and Service Integration – An Experiment with Gene Regulatory Networks <i>Emdad Ahmed</i>	56
A Data Hiding Capacity Estimation Technique for Biometric Images <i>Ghazaleh Sarbisheim, Saied Hossaini Khayat</i>	62
Regression Diagnostics in Large and High Dimensional Data <i>A. M. Nurunnabi, and Mohammed Nasser</i>	67
Comparisons of Maximum System Lifetime in Diverse Scenarios for Body Sensor Networks <i>Md. Nazrul Islam Mondal Md. Shahid Uz Zaman, Abdullah Al Masud, Md. Jahangir Alam</i>	73
Cluster Analysis of Microarray data of <i>Shewanella oneidensis</i> <i>Sharmin Nilufar and Jannatul Ferdousi Ara</i>	79

Wireless Communication

Performance Evaluation of DODE of a Voice/Data Integrated Wireless Mobile Network <i>H. M. Rahat , Runa Laila , Sabiha Rahman Juthy , Md. Imdadul Islam , and M. R. Amin</i>	85
A New Weighted Centroid Localization Algorithm in Wireless Sensor Networks <i>Yahya Forghani</i>	89
Performance evaluation of a wireless Orthogonal Frequency Division Multiplexing system under various concatenated FEC channel-coding schemes <i>Md. Dulal Haque, Shaikh Enayet Ullah, and Md.Razu Ahmed</i>	94
Cost Effective Multi-granular Location Information Strategy for MANET <i>Koushik Majumder, S. R. Biradar, Subir Kumar Sarkar</i>	98
QoS based Fair Load-Balancing: Paradigm to IANRA Routing Algorithm for Wireless Networks (WNs) <i>Ayyoub Akbari Moghanjoughi , Sabira Khatun, Borhanuddin M. Ali, Raja S. A. R. Abdullah</i>	104

Artificial Intelligence

An Artificial Ant Based Novel and Efficient Approach of Regular Geometric Shape Detection from Digital Image <i>Ayan Acharya, Kaushik Chattopadhyay, Deepyaman Maiti, and Amit Konar</i>	110
Codebook Design Method for Speaker Identification based on Genetic Algorithm <i>Md. Rabiul Islam, Md. Fayzur Rahman, Md. Ruhul Amin and Md. Shahid Uz Zaman</i>	116
Emotion-Driven Learning Agent for Setting Rich Presence in Mobile Telephony <i>Sumanta Saha and Rumana Quazi</i>	121
An HPSG Analysis of Arabic Passive <i>Md. Shariful Islam Bhuyan and Reaz Ahmed</i>	127
WalkSAT Approach in Solving the Staff Transfer Problem <i>Sriyankar Acharyya</i>	132

Computer Networks and Data communication

Simulation and Bit Error Rate Performance Analysis of 4G OFDM Systems <i>Md. Sadek Ali, Md. Mahbubur Rahman, Md. Delowar Hossain, Md. Shariful Islam, Md. Moinul Islam</i>	138
Lightpath Reconfigurations in IP-over-CWDM Networks with Stackable ROADMs <i>Md. Nooruzzaman, Yuichi Harada, Michio Hashimoto, Osanori Koyama, and Yutaka Katsuyama</i>	144
Evaluation of TCP Performance over Mobile IP Wired-cum-Wireless Networks <i>Md. Asif Nashiry, Md. Abdullah-Al-Mahmud, Md. Mahbubur Rahman, Md. Nashid Anjum</i>	150
A System for Locating Users of WLAN using Dynamic Mapping in Indoor and Outdoor Environment-LOIDS <i>K. Krishna naik and dr. M.N. Giri prasad</i>	156
Evaluation of Traffic Parameters of Multidimensional Traffic of a Combined Link Using a Tabular Method <i>Md. Imdadul Islam, Md. Shahriar Karim, J. K. Das and M. R. Amin</i>	161

Bangla Language Processing

Analysis and Synthesis of Bangla Dental Consonants	166
--	-----

<i>Syed Akhter Hossain, M Lutfar Rahman and Farruk Ahmed</i>	
An approach to Recognize Hand written Bengali Numerals for Postal Automation.	171
<i>Md. Saidur Rahman, G. M. Atiqur Rahaman, Asif Ahmed and G. M. Salahuddin</i>	
An Algorithm for Segmenting Modifiers from Bangla Text	177
<i>Nasreen Akter, Saima Hossain, Md. Tajul Islam, and Hasan Sarwar</i>	
An Empirical Framework for Translating of Phrasal Verbs of English Sentence into Bangla	183
<i>Mohammed Moshiul Hoque, Md Mokaddem Ahamad, Mohammad Shamsul Arefin, Md. Monjur-Ul-Hasan and Md. Gahangir Hossain</i>	
Design of an Interactive Bangla Mobile Keypad Based on Ergonomics	189
<i>Muhammad Raisul Alam, Md. Masum, Shafiul Hasan Md. Tareq and Mostafa Wasiuddin Numan</i>	

Computer Vision

Crowd Behaviour Monitoring on the Escalator Exits	194
<i>Md. Haidar Sharif, Nacim Ihaddadene, and Chabane Djeraba</i>	
Solutions to Motion Self-Occlusion Problem in Human Activity Analysis	201
<i>Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim, and S. Ishikawa</i>	
Performance Analysis on an Efficient Human Motion Database with Various Motion Representations	207
<i>S. M. Ashik Eftakhar, Joo Kooi Tan, Hyoungseop Kim, and Seiji Ishikawa</i>	
Action Recognition with Various Speeds and timed-DMHI Feature Vectors	213
<i>Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim, and S. Ishikawa</i>	
Intelligent Video Surveillance for Monitoring Fall Detection of Elderly in Home Environments	219
<i>Homa Foroughi, Baharak Shakeri Aski, and Hamidreza Pourreza</i>	

Knowledge and Data Engineering

An Improved Keyword Extraction Method Using Graph Based Random Walk Model	225
<i>Md. Rafiqul Islam and Md. Rakibul Islam</i>	
Performance Maximization for Question Classification by SubSet Tree Kernel using Support Vector Machines	230
<i>Muhammad Arifur Rahman, Vitalie Scurtu</i>	
On the Design of an Effective Corpus for Evaluation of Bengali Text Compression Schemes	236
<i>Md. Rafiqul Islam, and S. A. Ahsan Rajon</i>	
A Corpus-based Evaluation of Lexical Components of a Domainspecific Text to Knowledge Mapping Prototype	242
<i>Rushdi Shams and Adel Elsayed</i>	
Density Based Clustering Technique For Efficient Data Mining	248
<i>Md. Asikur Rahman, A.K.M Rasheduzzaman Chowdhury, Daud Md. Jamilur Rahman, Abu Raihan Mostofa Kamal</i>	

Digital Logic and System Design

An Intelligent Aviation Obstruction Warning light	253
<i>Muhammad Mehedi Al Emran Hasan, and Badal Chandra Sarker</i>	
Low-Cost Realization of Toffoli Gate for the Low-Cost Synthesis of Quantum Ternary Logic Functions	259
<i>Md. Mehedi Hasan</i>	

Design and Implementation of Microprocessor Based Electronic Voting System <i>Muhammad Raisul Alam, Md. Masum, Md. Mostafizur Rahman, Md. Arifur Rahman</i>	264
Genetic Algorithm Based Synthesis of Ternary Reversible/Quantum Circuit <i>Rashida Khanom Tahseen Kamal ,and Mozammel H. A. Khan</i>	270
New Methodologies for High Level Modeling and Synthesis of Low Density Parity Check Decoders <i>Syed Mahfuzul Aziz and Sunil Sharma</i>	276

Fuzzy Logic Systems

The Study of Tracking Control of a Robotic Manipulator Actuated by Shape Memory Alloy <i>S. M. Mahfuzur Rahman and Kyoung Kwan Ahn</i>	282
A Fuzzy Logic-Based Adaptive Handoff Management Protocol for Next-Generation Wireless Systems <i>Presila Israt, Namvi Chakma, and M. M. A. Hashem</i>	288
Automatic Life-Logging: A novel approach to sense real-world activities by environmental sound cues and common sense <i>Mostafa Al Masum Shaikh, Md. Khademul Islam Molla, and Keikichi Hirose</i>	294
A Novel Fuzzy Method to Traffic Light Control Based on Unidirectional Selective Cellular Automata for Urban Traffic <i>Moein Shakeri, Hossein Deldari, Alireza Rezvanian, and Homa Foroughi</i>	300
Bangla Vowel Sign Recognition by Extracting the Fuzzy Features <i>Mohammad Sarwar Kamal, Mohammed Moshikul Hoque, Md. Monjur Ul Hasan and Mohammad Shamsul Arefin</i>	306

Digital Signal and Image Processing

Active Noise Control for Industrial Applications <i>Syed Foysol Islam, A.H.M. Saiful Islam, M. M. Rajib Billah</i>	312
A Lossless Image Compression Technique Using Generic Peano Pattern Mask Tree <i>Mohammad Kabir Hossain, Shams M Imam , Khondker Shajadul Hasan , William Perrizo</i>	317
A Combined Local-Global Optical Flow Approach for Cranial Ultrasonogram Image Sequence Analysis <i>Mohiuddin Ahmad, Taohida Taslima, Lutfunnesa Lata, and Masayoshi Yamada</i>	323
Distortion Measurement Using Arc-Length-Parameterisation within a Vertex-Based Shape Coding Framework <i>Ferdous Ahmed Sohel and Mohammed Bennamoun</i>	329
Time-Frequency Representation of Audio Signals Using Hilbert Spectrum with Effective Frequency Scaling <i>Md. Khademul Islam Molla, Mostafa Al Masum Shaikh and Keikichi Hirose</i>	335

Distributed and Parallel Processing

Server Consolidation Using OpenVZ: Performance Evaluation <i>Mohiuddin Ahmed, Showayb Zahda, Majed Abbas</i>	341
A New Distributed Evolutionary Computation Technique for Solving Large Number of Equations <i>Moslema Jahan, M. M. A Hashem, and Gazi Abdullah Shahriar</i>	347
Achieving Guaranteed Performance on Computational Grid <i>Ajanta De Sarkar, Dibyajyoti Ghosh, Rupam Mukhopadhyay and Nandini Mukherjee</i>	353

On Hot-Spot Traffic Pattern of TESH Network	359
<i>M.M. Hafizur Rahman, Yasushi Inoguchi, Yukinori Sato, Yasuyuki Miura, Susumu Horiguchi</i>	
Dynamic Communication Performance of a TESH Network under the Nonuniform Traffic Patterns	365
<i>M.M. Hafizur Rahman, Yasushi Inoguchi, Yukinori Sato, Yasuyuki Miura, Susumu Horiguchi</i>	

System Security and Control

SAGLET- Secure Agent Communication Model	371
<i>Elhum Nusrat, Abu Shohel Ahmed, Gazi Mushfiqur Rahman, Lafifa Jamal</i>	
Side Channel Attack Prevention for AES Smart Card	376
<i>Mohammad Zahidur Rahaman and Mohammad Akram Hossain</i>	
An Energy-Efficient Data Security System for Wireless Sensor Network	381
<i>Md. Anisur Rahman and Mitu Kumar Debnath</i>	
MIMOME channel secrecy capacity	387
<i>Mohammad Rakibul Islam, Jinsang Kim, Md. Shamsul Arefin</i>	
A Practical Analysis of the Robustness and Stability of the Network Stack in Smartphones	393
<i>Sheikh Mahbub Habib, Cyril Jacob and Tomas Olovsson</i>	

Pattern Recognition

Sensitivity Analysis in Robust and Kernel Canonical Correlation Analysis	399
<i>Md. Ashad Alam, Mohammed Nasser and Kenji Fukumizu</i>	
An Efficient System for Recognition of Human Face in Different Expressions by Some Measured Features of the Face Using Laplacian Operator	405
<i>Abu Shamim Mohammad Arif, G. M. Atiqur Rahman, Gautam Kumar Biswas, Syed Nurul Islam</i>	
An Effective Term Weighting Method Using Random Walk Model for Text Classification	411
<i>Md. Rafiqul Islam and Md. Rakibul Islam</i>	
Vision System for Human-Robot Interface	415
<i>Md. Ezharul Islam, Nasima Begum and Md. Al-Amin Bhuiyan</i>	
Rice Disease Identification using Pattern Recognition Techniques	420
<i>Santanu Phadikar and Jaya Sil</i>	

Software and Web Engineering

Geospatial Web Services in Environmental Planning	424
<i>Iftikhar U. Sikder</i>	
Implementation of E-Governance: Only Way to Build a Corruption-Free Bangladesh	430
<i>S. A. AHSAN RAJON and SK. Ali Zaman</i>	
File Sharing in Advanced Collaborating Environment: Node Registration & Cache Optimization Techniques	436
<i>Md. Rezwanur Rahman, Md. Ali Ashik Khan, Tamkin Khan Avi, Mohammad Rezwanul Huq</i>	
A New Approach to Wrap Legacy Programs into Web Services	442
<i>Saeed.Parsa, Leila.Ghods</i>	
Speech Enabled Operating System Control	448
<i>Md. Abdul Kader, Biswajit Singha, and Md. Nazrul Islam</i>	

Algorithms

Short Text Compression for Smart Devices	453
<i>Md. Rafiqul Islam, S. A. Ahsan Rajon and Anonda Podder</i>	
Smallest and Some New Equiprojective Polyhedra	459
<i>Masud Hasan, Mohammad Monoar Hossain, Sabrina Nusrat and Alejandro L´opez-Ortiz</i>	
A comprehensive Analysis of Degree Based Condition for Hamiltonian Cycles	465
<i>Md. Kamrul Hasan, Mohammad Kaykobad, Young-Koo Lee¹, Sungyoung Lee</i>	
Design of a Fractional-order Self-tuning Regulator using Optimization Algorithms	470
<i>Deepyaman Maiti, Mithun Chakraborty, Ayan Acharya, and Amit Konar</i>	
A Concurrent Approach to Clustering Algorithm with Applications to VLSI Domain	476
<i>Md. Asif Iqbal, Mohammad Iftekhar Monir, Taufique Sayeed, Abu Hamed Mohammad Misbah-Uddin</i>	

Computer Networks and Data Communication

Proposal for an Efficient Quantum Key Distribution System Using Entanglement	481
<i>M.A. Bashar, M. A. Chowdhury, R. Islam, M. S. Rahman, S. K. Das</i>	
Symbol Timing Estimation using Near ML Techniques and Statistical Performance Evaluation for Binary Communications	486
<i>Md. Tofazzal Hossain and Sithamparanathan Kandeepan</i>	
Layer-2 Protocol Adaption Method to Improve Fast Handoff for Mobile IPv6 for Vertical Handoffs	492
<i>Vahid Solouk¹, Borhanuddin M. Ali, Sabira Khatun, Daniel Wong, and M. Adzir Mahdi</i>	
BER Performance of Coded MSK/Multichannel CDMA Fourth Generation Synchronous Downlink Cellular Systems in the Presence of MUI and FO	498
<i>M. Abou El-Nasr</i>	
A New SS7-SIGTRAN protocol Interchanger Software and Hardware to Implement an Improved Distributed Database Based ATM Network Using Existing IP Network	504
<i>M. S. Munir, K. Ahmed, A.S.M. Shihavuddin, S. M. Nizamuddin and M. J. Sarker.</i>	

Computer Vision

Face Recognition using Gabor Filters	510
<i>Md. Tajmilur Rahman and Md. Alamin Bhuiyan</i>	
High-speed distance measurement between moving vehicles with NIR-LED markers	516
<i>M. Fukuzawa, H. Hama, N. Nakamori, and M. Yamada</i>	
SEI and SHI Representations for Human Movement Recognition	521
<i>Mohiuddin Ahmad and Md. Zulfikar Hossain</i>	
Generation of Mountain Ranges by Modifying a Controlled Terrain Generation Approach	527
<i>K. Raiyan Kamal and M. Kaykobad.</i>	
A Machine Vision Based Automatic System for Real Time Recognition and Sorting of Bangladeshi Bank Notes.	533
<i>Raihan Ferdous Sajal, Mohammed Kamruzzaman, Faruq Ahmed Jewel.</i>	

Knowledge and Data Engineering

An Efficient Implementation Scheme for Multidimensional Online Analytical Processing	536
<i>Nilufar Easmin, K. M. Azharul Hasan, Nilufa Parvin</i>	
Decision Tree Based Routine Generation (DRG) Algorithm: A Data Mining Advancement to Generate Academic Routine for Open Credit System	542
<i>Md. Ashiqur Rahman, Shafaat Sheik Giasuddin, Rashedur Rahman</i>	

Extended Algorithm for Spatial Characterization and Discrimination Rules	548
<i>Md. Rashidul Hasan, Md. Zakir Hossain, Fahim Md. Chowdhury, Mohammad Khalad Hasan</i>	
Isolating Significant Phrases in Common Natural Language Queries to Databases	554
<i>Mohammad Moinul Hoque, Md. Shahriar Mahbub, and S.M. Abdullah Al-Mamun</i>	
Meliorated Approach for Extracting Bilingual Terminology from Wikipedia	560
<i>Anand Gupta, Akhil Goyal, Aman Bindal and Ankuj Gupta</i>	

Wireless Communication

Comparing FLV and MPEG4 (H.264/AVC) Multimedia File Format with Wireless Network Parameters	566
<i>Mehedi Mahmud, 1Sadia Monir, Tanjila Khan, Mohammad Abdur Rahim Mustafa and Hasan Sarwar</i>	
Direction of Arrival Algorithms for Adaptive Beamforming in Next Generation Wireless Systems	571
<i>Mohammad Ismat Kadir, Md. Shadiul Hoque, Saiful Islam</i>	
The Channel Capacity of a GSM Frequency Hopping Network with Intelligent Underlay-Overlay	576
<i>Md. Majharul Islam Rajib, Wahida Nasrin</i>	
Cell Database Based Time Specific Resource Reservation Scheme	581
<i>S. M Didar-Al- Alam, Md. Nazmul Hasan, Md. Mahfuzul Islam, Akramul Azim</i>	

E-Commerce and E-Governance

Developing a Model of E-governance for Urban and Rural Areas of Bangladesh	587
<i>Gazi Zahirul Islam, Md. Ibrahim Khan, Md. Anas Boksh Mazady</i>	
Secure Online Sealed Bid Auction	593
<i>Md. Nazrul Islam, Mohammad Zahidur Rahman</i>	
Short Messaging Service (SMS) Based m-Banking System in Context of Bangladesh	599
<i>Md. Subrun Jamil, Fouzia Ashraf Mousumi</i>	
A study of major Mobile payment systems' functionality in Europe	605
<i>S. Mohammadi, H. Jahanshahi</i>	
The opportunities and challenges of using ICT by State Owned Enterprises in Bangladesh: Case of SOEs under Privatization Commission	611
<i>Ahsanullah M Dewan, Saifullah M Dewan, Shams Ara Nazmin</i>	

Distributed and Parallel Processing

Parallel Performance of Domain Decomposition Method on Distributed Computing Environment	617
<i>A.M.M. Mukaddes and Atsushi Uragami</i>	
Routing Algorithm for Vertically Stacked Optical Banyan Networks with the Time Complexity $O(N)$ using Pipelined Processors	623
<i>Md. Mamun-ur-Rashid Khandker</i>	
A Crosstalk Free Routing Algorithm of Generalized Recursive Non-blocking Network	629
<i>Most. Arjuman Sultana, Gita Chowdhury, M. M. Hafizur Rahman</i>	
Lower Bound on Blocking Probability for Vertically Stacked Optical Banyan Networks with Given Crosstalk Constraint	635
<i>Basra Sultana and M. R. Khandker</i>	
Performance of Fast Routing Algorithm for Vertically Stacked Optical Banyan Networks with	640

Link Failures and Given Crosstalk Constraint
Basra Sultana and M. R. Khandker

Computer Networks and Data Communication

Using Log Likelihood Relation for BER Analysis of QAM in Space Diversity <i>Md. Zahangir Alam, Chowdhury Sajadul Islam, Md. Zahangir Alom, M Abdus Sobhan</i>	646
Learning Intrusion Detection Based on Adaptive Bayesian Algorithm <i>Dewan Md. Farid, Mohammad Zahidur Rahman</i>	652
Analytical Performance Evaluation of Space Time Coded MIMO OFDM Systems Impaired by Timing Jitter <i>M. R. H. Mondal, and S. P. Majumder</i>	657
Performance Comparison of IPv4 and IPv6 on Various Windows Operating Systems <i>Shaneel Narayan, Samad S. Kolahi, Yonathan Sunarto, Du D. T. Nguyen, Paul Mani</i>	663
A New Proposal for Choosing a Deflection Link in a Deflection Network <i>M. Atiqur Rahman, M. Tanvir Hossain, and Shahed Anwar</i>	669

Wireless Communication

Design and Construction of Direct Sequence Spread Spectrum CDMA Transmitter and Receiver <i>M. Habib Ullah, Akhmad Unggul Priantoro, M. Jasim Uddin,</i>	675
Comparative performance analysis of DSR and AODV protocol based on Mobility factor <i>A.S. M. Ashique Mahmood, Mohammad Shahid Hossain, Faisal Aman Aziz</i>	681
A Least Square Approach for TDOA/AOA Wireless Location in WCDMA System <i>Mohammad Nur-A-Alam, Mohammed Mezbaul Haque</i>	686
Prediction of the Density of Active Wireless Device Using Markov Model <i>Md. Osman Gani, Iftekhar Jahan Mehedi, Mehnaz Seraj, Hasan Sarwar and Chowdhury Mofizur Rahman</i>	691

Computer Networks and Data Communication

End-to-End Mobility Management Solutions for TCP: An Analysis <i>Peer Azmat Shah, Muhammad Yousaf</i>	696
Peer-to-Peer Mobile Applications Using JXTA/JXME <i>Tahmina Tahsin, Lazeeb Faiz Choudhury, Md. Lutfur Rahman</i>	702
Connecting Telephone Exchange Centers over Metro Ethernet Networks by using Diffserv-MPLS <i>Ghasem Mirjalily, Parisa Bastani, and S.M.T. Almodarresi</i>	708
Security Enhancement of MD5 Hashed Passwords by Using the Unused Bits of TCP Header <i>Md. Didarul Alam Chawdhury, and A.H.M. Ashfak Habib</i>	714
Crosstalk and Bit Error Rate Performance of a Proposed Optical Cross-Connect Topology <i>Md. Prince Hossen, Md. Abubakar Siddique and Md. Rafiqul Islam</i>	718

Neural Networks and Others

Solar Cell Efficiency Improvement Using Compound Parabolic Concentrator and an Implementation of Sun Tracking System <i>Eklas Hossain, Riza Muhida, Ahmad Faris Dzulkipli and Khairul Azizi Abdul Rahman</i>	723
Comparison of GARCH and Neural Network Methods in Financial Time Series Prediction	729

<i>Altaf Hossain and M. Nasser</i>	
Neural Network and Genetic Algorithm Approaches for Forecasting Bangladeshi Monsoon Rainfall	735
<i>Shipra Banik, Farah Habib Chanchary, Khodadad Khan, Rifat Ara Rouf and Mohammed Anwer</i>	
On the Multi-Channel Noise Cancellation Using Beamforming Algorithm	741
<i>Md. Zahangir Alom, Md. Zahangir Alam, M Abdus Sobhan</i>	
Recurrent Neural Network Classifier for Three Layer Conceptual Network and Performance Evaluation	747
<i>Md. Khalilur Rhaman and Tsutomu Endo</i>	

Antenna and Propagation

An Energy Saving Program for Bangladesh, for Reducing Load Shedding, and for Continuity of Power for IT Sector	753
<i>Shahriar Khan.</i>	
Analysis of Broadband Slotted Microstrip Patch Antenna	758
<i>Mohammad Tariqul Islam, Mohammed Nazmus Shakib, Norbahiah Misran, and Baharudin Yatim</i>	
Analysis of Antennas for the Wireless Communication and Radar Applications	762
<i>Md. Osman Goni, Abu Md. Numan-Al-Mobin, Md. Hasanur Rahman</i>	
Amplitude Banded Technique and Parallel Structure for Godard and Sato Algorithms of Blind Channel Equalization	768
<i>Muhammad Lutfor Rahman Khan, Mohammed H. Wondimagegnehu and Tetsuya Shimamura</i>	
Performance Investigation of Earth-to-Satellite Microwave Link Due to Rain Fade in Bangladesh	773
<i>Md. Rafiqul Islam, Md. Arafatur Rahman, Farhat Anwar and Md. Mahfujur Rashid</i>	

Cutting a Cornered Convex Polygon Out of a Circle

Syed Ishtiaque Ahmed, Md. Ariful Islam, Masud Hasan

Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology, Dhaka, 1000, Bangladesh
Email: ishtiaque@csebuat.org, arif.liton@gmail.com, masudhasan@cse.buet.ac.bd

Abstract—The problem of cutting a convex polygon P out of a piece of paper Q with minimum total cutting length is a well studied problem in computational geometry. Researchers studied several variations of the problem, such as P and Q are convex or non-convex polygons and the cuts are line cuts or rays cuts. In this paper we consider yet another variation of the problem where Q is a circle and P is convex polygon such that P is bounded by a half circle of Q and all the cuts are line cuts. We give a simple linear time $O(\log n)$ -approximation algorithm for this problem where n is the number of vertices of P .

Index Terms—Algorithm, circle, computational geometry, line cut, polygon cutting.

I. INTRODUCTION

A *cornered convex polygon* P inside a circle Q is a convex polygon which is positioned completely on one side of a diameter of Q . See Figure 1(a).

A *line cut* is a line that does not run through the polygon P and divides Q into two pieces. After a cut is made, Q is updated to the piece containing P . The *cost* of a cut is the length of the intersection of the cut with Q . A *cutting sequence* is a sequence of cuts such that after the last cut in the sequence we have $P = Q$. In this paper we consider the problem of cutting P out of Q by line cuts with total cutting cost as small as possible. See Figure 1(b).

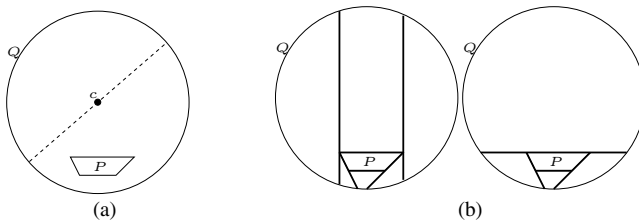


Fig. 1. (a) A cornered convex polygon. (b) Two different cutting sequences to cut P out of Q . The cutting lengths are shown in bold. The cost of the sequence in left figure is more than that in the right figure.

A. Known results

If Q is another convex polygon with m edges, this problem has been approached in various ways by many researchers in computational geometry community [1], [2], [4], [5], [6], [7], [8], [9]. Overmars and Welzl first introduced this problem of cutting a convex polygon out of

another convex polygon optimally with line cuts [8]. If the cuts are allowed only along the edges of P , they proposed an $O(m+n^3)$ -time algorithm for this problem with optimal cutting length, where n is the number of edges of P .

The problem seems to be more difficult if the cuts are more general, i.e., they are not restricted to touch only the edges of P . In that case Bhaduri and Chandrasekaran showed that the problem has optimal solutions that lie in the algebraic extension of the input data field [1] and due to this algebraic nature of this problem, an approximation scheme is the best that one can achieve [1]. They also gave an approximation scheme [1] with pseudo-polynomial running time.

After the indication of Bhaduri and Chandrasekaran [1] on the hardness of the problem several researchers have given polynomial time approximation algorithms for this problem. Dumitrescu proposed an $O(\log n)$ -approximation algorithm with $O(mn + n \log n)$ running time [4], [5]. Then Daescu and Luo [7] gave the first constant factor approximation algorithm for this problem. Their algorithm has a running time of $O(n^3 + (n + m) \log(n + m))$. The best known constant factor approximation algorithm is due to Tan with an approximation ratio of 7.9 and running time of $O(n^3 + m)$ [9]. In the same paper [9], he also proposed an $O(\log n)$ -approximation algorithm with improved running time of $O(n + m)$. As the best known result so far, very recently, Bereg, Daescu and Jiang gave a simple $O(m + \frac{n^6}{\epsilon^{12}})$ -time $(1 + \epsilon)$ -approximation algorithm for this problem [2].

This problem has also been studied when Q is a non-convex polygon [8], P is a non-convex polygon [3], [7], [9], and/or the cuts are ray cuts [3], [7], [9].

B. Our Result

All the previous results consider Q as a polygon (either convex or non-convex). However, to our knowledge, no algorithm is known when Q is a circle. In this paper, we consider the problem where Q is a circle and P is a convex polygon such that P can be separated from the center of Q by a chord. Our algorithm has an approximation ratio of $O(\log n)$ and runs in $O(n)$ time.

When both P and Q are convex polygons, almost all the existing algorithms on line cuts use two major steps: cutting a minimal rectangle or triangle from Q that bounds

P and then cutting P out of that bounding box. Our algorithm also follows similar approach. However, we observe that the existing algorithms cannot be applied directly to solve our problem. Moreover, the running times of those algorithms are too high compared to our algorithm. For example, Tan's constant factor approximation algorithm takes $O(n^3 + m)$ time and the PTAS of Bereg et. al. takes $O(m + \frac{n^6}{\epsilon^{12}})$ time. We also observe that in those algorithms, increasing the value of m for approximating Q to a circle makes them inefficient.

C. Outline

The rest of the paper is organized as follows. We give some definitions and preliminaries in Section II. Then we present our algorithm in Section III. Finally, Section IV concludes the paper with some future works.

II. PRELIMINARIES

A line cut is a *vertex cut* through a vertex v of P if it is tangent to P at v . Similarly, a line cut is an *edge cut* through an edge e of P if it contains e .

At any time the edges of P through which an edge cut has passed are called to be *cut edges* of P and other edges of P are called to be *uncut edges* of P . To cut P out of Q all n edges of P must be cut and for that we require exactly n edge cuts. However, applying only edge cuts may not give an optimal solution and that is why we need vertex cuts as well.

In the rest of this section we give some elementary geometry that plays important role in our paper. Let c be the center of Q . An edge e of P is *visible* from c if for every point p of e the line segment cp does not intersect P in any other point. So, if e is collinear with c , then we consider e as invisible. Similarly, a vertex v of P is *visible* from c if the line segment cv does not intersect P in any other point.

In this paper we do not consider the diameter of Q as a *chord*, i.e., a chord is always smaller than a diameter. Let ll' be a chord of Q . ll' divides Q into two circular segments, one is bigger than a half circle and is called the *bigger segment* created by ll' , and the other one is called the *smaller segment* created by ll' . Let tt' be another chord intersecting ll' at x such that tx is in the smaller segment of ll' . See Figure 2(a). Then the following lemma holds.

Lemma 1: Length of xt is no bigger than length of ll' .

Following lemma is also obvious, whose illustration can be found in Figure 2(b).

Lemma 2: Let $\triangle abc$ be an obtuse triangle with $\angle bac$ being the obtuse angle. Consider any line segment connecting two points b' and c' on ab and ac , respectively, possibly one of b' and c' coinciding with b or c respectively. Then the angle $\angle bb'c'$ and $\angle cc'b'$ are obtuse.

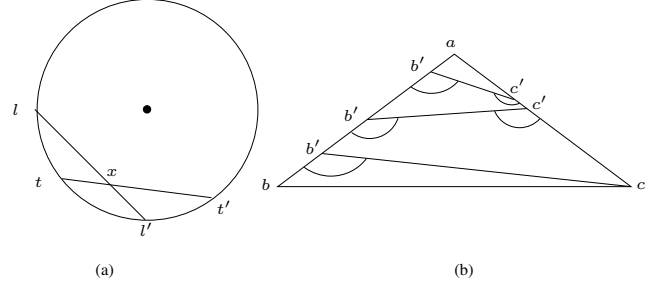


Fig. 2. Some preliminaries.

III. ALGORITHM

Our algorithm has four phases : (1) D-separation, (2) triangle separation, (3) obtuse phase and (4) curving phase. In *D-separation phase*, we cut out a small portion of Q (less than the half of Q in size) which contains P (and looks like a “D”). Then in *triangle separation phase* we reduce the size of Q even more by two additional cuts and bound P by almost a triangle. In *obtuse phase* we assure that all the portions of Q that are not inside P are inside obtuse triangles. Finally, in *curving phase* we cut P out of Q by cutting in rounds.

Let C^* be the optimal cutting length of P from Q . Clearly $C^* \geq |P|$, where $|P|$ is the length of the perimeter of P .

A. D-Separation

A *D* of the circle Q is a circular segment of Q which is smaller than an half circle of Q . By a *D-separation* of P from Q we mean a line cut along a chord of Q that creates a *D* containing P . In general for a circle Q and a convex polygon P there may not exist any *D-separation* of P . But in our case since P is cornered, there always exists a *D-separation* of P .

We will first observe that there exists only one minimum-cost *D-separation* C_1 of P and then we will find that C_1 . (For our algorithm, however, showing the *uniqueness* of the minimum *D-separation* is not required).

Lemma 3: The minimum-cost *D-separation* C_1 for P can be found in $O(n)$ time.

Proof: Clearly C_1 must touch P . So, C_1 must be a vertex cut or an edge cut of P . Observe that any vertex cut or edge cut that is a *D-separation* must be through a visible vertex or a visible edge. Let e be a visible edge of P . Let ll' be a line cut through e . Let cp be the line segment perpendicular to ll' at p . If p is a point on e , then we call e a *critical edge* of P and ll' a *critical edge cut* of P . Observe that since P is convex, it can have at most one critical edge. Similarly, let v be a vertex of P . Let tt' be a vertex cut through v . Let cp be the line segment perpendicular to tt' at p . If tt' is such that $p = v$, then we

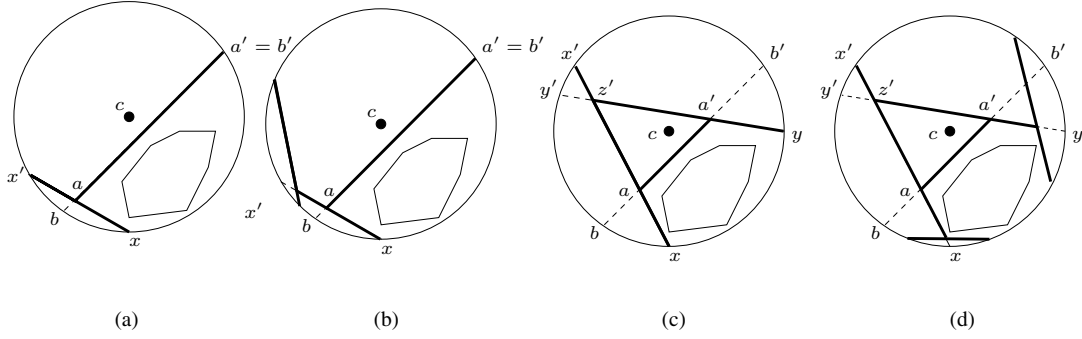


Fig. 3. Separating P from c by using more than one cut

call v a *critical vertex* of P and tt' a *critical vertex cut* of P . Again observe that P can have at most one critical vertex. Moreover, P has a critical edge if and only if it does not have a critical vertex. See Figure 4. Now, if P has the critical edge e , then C_1 is the corresponding critical edge cut ll' . C_1 is minimum, because any other vertex cut or edge cut of P is either closer to c (and thus bigger) or does not separate c from P . On the other hand, if P has the critical vertex v , then C_1 is the corresponding critical vertex cut tt' of P . Again, C_1 is minimum by the same reason.

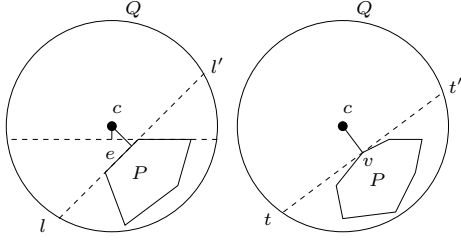


Fig. 4. Critical edge and critical vertex.

For the running time note that all visible vertices and visible edges of P can be found in linear time. Then finding whether an edge of P is critical takes constant time. Over all edges finding the critical edge, if it exists, takes $O(n)$ time. Similarly, for each visible vertex v we can check in constant time whether v is critical or not by comparing the angles of cv with two adjacent edges of v , which takes constant time. Over all visible vertices, it takes linear time. ■

Lemma 4: Cost of C_1 is at most C^* .

Proof: Consider any optimal cutting sequence $\mathcal{C}$. $\mathcal{C}$ must separate P from c . However, it may do that by using a single cut or by using more than one cut. If it uses a single cut, then it is in fact doing a D-separation. By Lemma 3, it can not do better than C_1 , and therefore, cost of C_1 is at most C^* .

Now we assume that $\mathcal{C}$ separates P from c by more than one cut. There are several cases here. We will prove that

in any case the cutting cost of separating P from c is even higher than that for a single cut. Let C be the first cut in $\mathcal{C}$ that separates c from P . C can not be the very first cut of $\mathcal{C}$, otherwise, we are in the previous case considered above. It implies that C smaller than a chord of Q . See Figure 3. Let a and a' be the two end points of C . Let bb' be the chord of Q that contains aa' where b is closer to a than to a' . Now, at least one of a and a' is not incident on the boundary of Q . We first assume that a is not incident to the boundary of Q but a' is. Let $C_x = xx'$ be the cut that was applied before C and intersects C at a . Since C_x does not separate c from P , the bigger segment of C_x contains P , c and a . Now if x' is an end point of C_x , then by Lemma 1 ab is smaller than xx' . It implies that having C_x in addition to C_1 increases the cost of separating P from c (see Figure 3(a)). Similarly, if x' is not an end point of C_x , and thus another cut is involved, then by Lemma 1 the cost will be even more (see Figure 3(b)).

Now consider the case when both a and a' are not incident on Q . So at least two cuts are required to separate ab and $a'b'$ from bb' . Let those two cuts be C_x and C_y respectively. Let xx' and yy' be the two chords of Q along which C_x and C_y are applied. Now if C_x and C_y do not intersect, then by applying the above argument for each of them we can say that the cutting cost is no better than that for a single cut. But handling the case when C_x and C_y intersect is not obvious. (See Figure 3(c,d)). Let z' be their intersection point. Again there may be several subcases: x and y may or may not be end points of C_x and C_y . Assume that both x and y are end points of C_x and C_y . Remember that none of C_x and C_y separates P from c . So region bounded by $xz'y$ must contain P and c inside of it and in that case the total length of xz' and yz' is at least the diameter of Q , which is bigger than bb' (see Figure 3(c)). For the other cases, where at least one of x and y is not an end point of C_x and C_y , it can be recursively proven that the cost is even more (For example, see Figure 3(d)). ■

B. Triangle Separation

In this section we apply two more cuts C_2 and C_3 and “bound” P inside a “triangle”. From there we achieve three triangles inside which we bound the remaining uncut edges of P (in the D-separation phase, at most one edge of P becomes cut).

Let $C_1 = aa'$ be the cut applied during D-separation. We apply $C_2 = at$ such that at is tangent to P . Then we apply the third cut $C_3 = a't'$ such that $a't'$ is also tangent to P . If $a't'$ intersects at inside or on the boundary of Q , then let z be the intersection point. (See Figure 5(a)). If $a't'$ and at do not intersect, then let z be the intersection point of the lines containing at and $a't'$. (See Figure 5(b)).

We get three resulting triangles $T_a, T_{a'}, T_z$ having a, a' and z , respectively, as a vertex as follows. We describe only T_a , description of $T_{a'}, T_z$ are analogous. If C_1 is an edge cut, then let rr' be the corresponding edge such that a is closer to r than to r' . (See Figure 5(a)). If C_1 is a vertex cut, then let r be the corresponding vertex of P . Let s be the similar vertex due to C_2 . (See Figure 5(b)). Then $T_a = \triangle ars$. The polygonal chain of P bounded by this triangle is the edges from r to s that reside inside T_a .

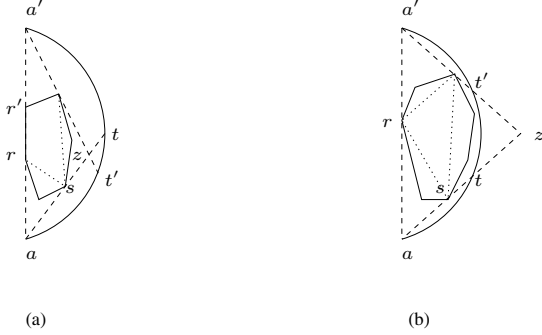


Fig. 5. Triangle separation.

Lemma 5: Total cost of C_2 and C_3 to achieve $T_a, T_{a'}$ and T_z is at most $2C^*$. Moreover, C_2 and C_3 can be found in linear time.

Proof: Whether z is within Q or outside Q , the length of at and $a't'$ can not be more than twice the length of aa' , which, by Lemma 3, is no more than C^* . So the total cost of C_2 and C_3 is at most $2C^*$.

To find at linearly, we can simply scan the boundary of P starting from the vertex or edge where aa' touches P and check in constant time whether a tangent of P is possible through that vertex or edge. Similarly we can find $a't'$ within the same time. ■

C. Obtuse Phase

Consider the triangle T_a obtained in the previous section. We call the vertex a of T_a the *peak* of T_a and the edge rs the *base* of T_a . Let the polygonal chain of P that is bounded by T_a be H_{rs} . Observe that the angle of T_a

at a is acute. Similarly, the angle of $T_{a'}$ at its peak a' is also acute. However, the angle of T_z at its peak z may be acute or obtuse.

We obtain, by applying a cut, for each of $T_a, T_{a'}$ and T_z (if T_z is acute) one or two triangles such that the angle at their peaks are obtuse and they combinely bound the polygonal chain of the corresponding triangle. We describe the construction of the triangle(s) from T_a only, description for the triangles obtained from $T_{a'}$ and T_z are analogous.

Let P_a be the polygonal chain bounded by T_a . Length of ar and as may not be equal. W.l.o.g. assume that as is not larger than ar . Let s' be the point on ar such that length of as' equals to the length of as . Connect ss' if s' is different from r . We will find a line segment inside T_a such that it is tangent to P_a and is parallel to ss' . We have two cases: (i) ss' itself is a tangent to P_a , and (ii) ss' itself is a tangent to P_a . Case (i): If ss' itself is tangent to P_a (at s), we apply our cut along ss' . This may be a vertex cut through s . In that case we get the resulting obtuse triangle $\triangle rss'$ and the polygonal chain bounded by this triangle is same as P_a . (See Figure 6(a)). On the other hand, this may be an edge cut through the edge of P_a that is incident to s . Let u be the vertex other than s of the edge of the edge cut. Then our resulting obtuse triangle is $\triangle rus'$ and the polygonal chain bounded by this triangle is the edges from r to u . (See Figure 6(b)). Case(ii): For this case, let u, u' be two points on as and ar , respectively, such that uu' is tangent of P_a and is parallel to ss' . We apply the cut along uu' . Again this cut may be a vertex cut or an edge cut. If it is a vertex cut, let g be the vertex of the cut. Then we get two obtuse triangles $\triangle ru'g$ and $\triangle sug$ and the polygonal chains bounded by them are the sets of edges from r to g and from g to s respectively. (See Figure 6(c)). If it is an edge cut, then let gg' be the edge of the edge cut with u being closer to g than to g' . Then we get two obtuse $\triangle ru'g'$ and $\triangle sug$ and the polygonal chains bounded by them are the sets of edges from r to g' and from g to s respectively. (See Figure 6(d)).

Lemma 6: Total cost of obtaining obtuse triangles from $T_a, T_{a'}$, and T_z is at most C^* . Moreover, they can be found in $O(\log n)$ time.

Proof: Consider the construction of the obtuse triangle(s) from T_a . Length of the cut ss' or uu' is at most the length of their base ss' , which is bounded by the length of P_a . Over all three triangles $T_a, T_{a'}$ and T_z , the total cutting length is bounded by the length of the perimeter of P . Since C^* is at least the length of the perimeter of P , the first part of the lemma holds.

For running time, in T_a , we can find the tangent of P_a in $O(\log |P_a|)$ time by using a binary search. Over all three triangles $T_a, T_{a'}$ and T_z , we need a total of $O(\log n)$ time. ■

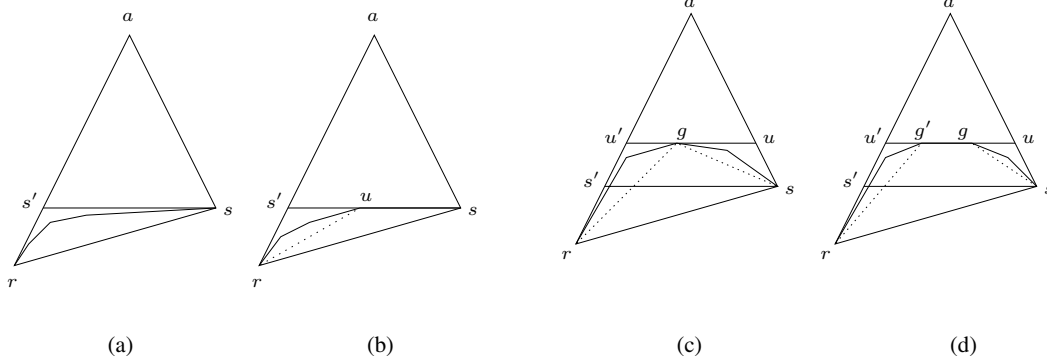


Fig. 6. Obtaining obtuse triangle(s) from T_a .

D. Curving Phase

After the obtuse phase the edges of P that are not yet cut are partitioned and bounded into polygonal chains with at most six obtuse triangles. In this section we apply the cuts in rounds until all edges of P are cut. Our cutting procedure is same for all obtuse triangles and we describe for only one.

Let $T_u = \triangle gus$ with peak u and base gs be an obtuse triangle. See Figure 7. Let the polygonal chain bounded by T_u be P_u . Let the edges of P_u be $e_1, \dots, e_k$ with $k \geq 2$. We will apply cuts in rounds and all cuts are edge cuts. In the first round we apply an edge cut C' along the edge $e_{k/2}$. Then we connect g and s with the two end points of $e_{k/2}$ to get two disjoint triangles. Since T_u is obtuse, by Lemma 2 these two new triangles are also obtuse. So, as a result we cut out one polygonal edge and get two new obtuse triangles. In the next round we work on each of these two triangles recursively, then in the next next round we work on four triangles and so on. We continue in this way until all edges of P_u are cut.

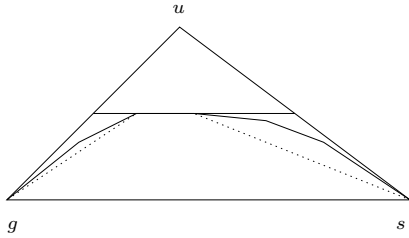


Fig. 7. Curving phase

Lemma 7: There are at most $O(\log k)$ rounds of cuts to cut all edges of P_u . Moreover, the total cost of these cuts is $|P_u| \log k$, where $|P_u|$ is the length of P_u , and the total running time is $O(|P_u|)$.

Proof: At each round the number of triangles get doubled and the number of edges that become cut also get doubled. So after $\log k$ rounds all k edges are cut.

At each round the total length of the bases plus the length of the edges that are being cut is no more than $|P_u|$. So, the total cutting cost is $|P_u| \log k$ over all rounds.

For running time, for applying an edge cut we always move to the middle edge of a polygonal chain. So, finding an edge cut takes constant time. Moreover, after each cut the corresponding edge is never considered again. So, over all rounds we can have at most $|P_u|$ cuts, which gives a running time of $O(|P_u|)$. ■

Corollary 1: The total cost of curving phase is $C^* \log n$ and the running time is $O(n)$.

Combining the results of all four phases, we get the following theorem.

Theorem 1: Given a circle Q and a cornered convex polygon P of n edges within Q , P can be cut out of Q by using line cuts in $O(n)$ time with a cutting cost of $O(\log n)$ times the optimal cutting cost.

IV. CONCLUSION

In this paper we have given a linear time $O(\log n)$ -approximation algorithm to cut out a convex polygon P out of a circle Q where P resides in one side of a diameter of Q . The general case of this problem, where P is not necessarily in one side of a diameter of Q , is still to be solved. We also think it would be interesting to see constant factor approximation algorithms or approximation schemes for these problems. One approach to achieve any improvement is to use rotating calipers around the boundary of P to cut out a bounding rectangle from Q instead of separating a D and then apply curving phase to that rectangle.

Finally, we observe that our algorithm also works if P touches the center of Q , with small modification.

REFERENCES

- [1] J. Bhadury and R. Chandrasekaran, "Stock cutting to minimize cutting length," *European Journal of Operations Research*, 88:69-87, 1996.
- [2] Sergey Bereg, Ovidiu Daescu, and Minghui Jiang, "A PTAS for Cutting Out Polygons with Lines," *COCOON 2006*, LNCS 4112, pp. 176-185, 2006.

- [3] Ramaswamy Chandrasekaran, Ovidiu Daescu, and Jun Luo, "Cutting Out Polygons," *In Proceedings of the 17th Canadian Conference on Computational Geometry*, pp. 183-186, 2005.
- [4] Adrian Dumitrescu, "An approximation algorithm for cutting out convex polygons," *Computational Geometry: Theory and Applications*, 29:223-231, 2004. (Preliminary version in *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'03)*, pp. 823-827, 2003.)
- [5] Adrian Dumitrescu, "The cost of cutting out convex n-gons," *Discrete Applied Mathematics*, 143(1-3):353-358, 2004.
- [6] Erik D. Demaine, Martin L. Demaine, and Craig S. Kaplan, "Polygons cuttable by a circular saw," *Computational Geometry: Theory and Applications*, 20:69-84, 2001.
- [7] Ovidiu Daescu and Jun Luo, "Cutting out polygons with lines and rays," *International Journal of Computational Geometry and Applications*, 16(2-3):227-248, 2006. (Preliminary version in *Proceedings of the 15th Annual International Symposium on Algorithms and Computation (ISAAC'04)*, LNCS 3341, pp. 669-680, 2004.)
- [8] Mark H. Overmars and Emo Welzl, "The complexity of cutting paper," *Proceedings of the 1st Annual ACM Symposium on Computational Geometry (SoCG'85)*, pp. 316-321, 1985.
- [9] Xuehou Tan, "Approximation algorithms for cutting out polygons with lines and rays," *Proceedings of the 11th International Computing and Combinatorics Conference (COCOON'05)*, LNCS 3595, pp. 534-543, 2005.

On the Minimization of Complete Test Set of Reversible k -CNOT Circuits for Stuck-at Fault Model

Muhammad Ibrahim, Ahsan Raja Chowdhury and Hafiz Md. Hasan Babu

Department of Computer Science and Engineering

University of Dhaka, Dhaka-1000, Bangladesh

E-mail: ibrahimmuhammad313@yahoo.com, farhan717@univdhaka.edu, hafizbabu@hotmail.com

ABSTRACT — In this paper, we propose an algorithm that produces the Complete Test Set (CTS) of a reversible circuit for Single Stuck-at Fault (SSF) and Multiple-Stuck-at Fault (MSF) models. Our algorithm works only for an important subclass of reversible circuits – the circuits consisting of k -CNOT gates ($k \geq 2$) though, any n -wire circuit having 0-CNOT or 1-CNOT gates can be converted to a $(n + 2)$ wire circuit having only k -CNOT gates with $k \geq 2$ with some additional hardware cost. Generated complete test set is not necessarily optimal, but minimizing the size of the complete test set is our key concern. Finally we provide some experimental results for the proposed method and compare it with existing methods to show how it outperforms almost all of the existing algorithms in terms of number of elements of CTS but is outperformed by some of the existing ones in terms of hardware cost.

Index Terms — Automatic Test Pattern Generator (Atpg), Complete Test Set (CTS), Design for Testability (DFT), k -CNOT Circuit, Multiple Stuck at Fault (MSF), Single Stuck at Fault (SSF), Test Vector (TV).

I. INTRODUCTION

The more logic elements are packed in conventional computer, the more heat is dissipated. When a computational system erases a bit of information, it must dissipate $\ln 2 \times kT$ energy, where k is Boltzmann's constant and T is the temperature. In room temperature, this is about 2.9×10^{-21} joules [1, 2]. Reversible circuits dissipate very low power [3].

A test set is a set of test vectors that are applied to the circuit in question to detect faults, if any. Generating efficient test sets is hard for conventional irreversible circuits though, it is significantly simpler for the case of reversible circuits.

Research reported so far in the field of reversible circuit testing regarding the topic in question is not bountiful. Patel et al. [4] formulated the problem of constructing minimal test set as an Integer Linear Program (ILP) with binary variables. They [4] also used an algorithm that simplifies the aforementioned method for a large circuit by breaking it into smaller sub-circuits [4]. Chakraborty [5] showed that all stuck-at faults in an n -wire k -CNOT circuit $C(n)$ with $k \geq 2$ can be detected using at most n test vectors (TV), namely, n weight-1 vectors. $C(n)$ has a complete test set (CTS) of size n for multiple stuck-at faults (MSF). He [5] proved this using *fixing* property. He has also shown that if an n -wire k -CNOT circuit ($k \geq 0$) is

modified by introducing an extra wire that feeds a new control input added to every k -CNOT gate, the resulting circuit has CTS of size 3 for MSFs. Allen et al. [6] proposed an algorithm that detect and localize all Missing Gate Fault (MGF), Lock Gate Fault (LGF) and Single Stuck-at Fault (SSF). Ito et al. [7] showed that it is NP-hard to generate a minimum complete test set for stuck-at faults on a set of wires of a reversible circuit.

We propose a method which exploits *fixing property* of digital circuit, like [5], to produce CTS for a given circuit. We apply a test vector at input of the circuit and observe if the output is same as the input. If it does, then we proceed to the next test vector. If it does not, then some wire(s) of the circuit must be faulty. Our target here is to reduce the size of the complete test set as less as possible without modifying the original k -CNOT circuit with $k \geq 2$. We do not consider fault diagnosis in this paper.

II. ASSUMPTION

We follow the following assumptions throughout the whole paper if otherwise not stated:

- An n -wire reversible circuit consisting of m number of k -CNOT gates is given ($k \geq 2$ and $m \geq 0$). If we have 0-CNOT and/or 1-CNOT gates in the original circuit, then a design-for-test (DFT) of the circuit can be constructed easily by adding at most 2 extra wires (so that 1-CNOT and 0-CNOT gates becomes 2-CNOT gates having necessary extra control input(s) from the extra wire(s)), so number of wires becomes $n + 2$ and all the methods are applicable on the DFT. Thus, without losing any generality, we assume that we have only k -CNOT gates with $k \geq 2$. When we used any k -CNOT gates with $k = 0$ or 1, we specify explicitly.
- Let any test vector (TV) be $t_1 t_2 \dots t_n$ where t_i is i^{th} bit.
- In spite of altering any bit in any level of the circuit except the last, if the final output of the circuit is equal to the input, then the input vector is *fixed* by definition. But this concept leads to some extra complexity. So throughout the whole paper, when we say a vector is *fixed* by a circuit, we mean that no bit of the input vector is altered by any gate across the whole circuit.
- The whole paper is *fixing property* oriented unless otherwise stated. For example, if we say a test vector is

invalid for a circuit, it tacitly implies that this test vector is not *fixed* by that circuit and vice versa.

III. PROPOSED METHOD

A. Description of a Gate and Description Table of a k -CNOT Circuit

Let a gate g has wire indices $g_1, g_2, \dots, g_p$ as control inputs. Then we call the sequence $g_1, g_2, \dots, g_p$ the description of the gate g . If g is 0-CNOT gate, then it has no description as we are concerned here with only the control inputs. If we construct a table where the rows are descriptions of the gates of a circuit, then we call this data structure “**Description Table**” of the circuit.

Example 3.1. Fig. 3.1(b) shows the description table of the circuit shown in Fig. 3.1(a).

B. Algorithm

We will use Description table, Status table and $TV_{(temp)}$ (Temporary Test Vector). Description table is already defined. The Status table is an $n \times 2$ array. Each row holds the current test status for both SA-0 and SA-1 faults of i^{th} wire. (For example, 0 if not tested, 1 if tested) $TV_{(temp)}$ is an n element array. This is the temporary TV of which our algorithm converts some of its ‘1’ bits to 0.

Procedure 1: Process_description_table

Input: Description table of a circuit.

Output: Description table of the circuit with some sorting applied (the type of sorting is discussed later).

Step 1: Sort each row of the description table in increasing order of their elements.

Step 2: Treat a row of a description table as a single element and then sort all rows in decreasing order of their first element.

return Description table.

Procedure 2: Adjust_status_table

Input: Status table and a TV

Output: Adjusted status table.

Step 1: Adjust the status table as follows:

for all i in $\{1, 2, \dots, n\}$, **do**
 if i^{th} bit of TV is 0, then set 1st element of i^{th} row of the status table to 1.
 else set 1st element of i^{th} row of the status table to 0.

return Status table.

Procedure 3: Get_New_TV

Input: Status table.

Output: A TV that acts as the $TV_{(temp)}$ for the next iteration.

Step 1: Calculate the TV as follows:

for all i in $\{1, 2, \dots, n\}$, **do**
 if i^{th} row of status table has 0 in the second element, set t_i to 1.

else set t_i to 0.

return TV.

Algorithm 1: Efficient_Complete_Test_Set_Generator (ECTSG)

Input: The configuration of a k -CNOT ($k \geq 2$) circuit is in description table as input. Input pattern is as follows: each gate occupies 1 row in the description table. The elements of each row are the wire indices of a particular gate where this gate has control inputs. Inversion input wire index is not entered in the row. Each row is terminated with a -1 value which works as sentinel.

Output: A set CTS of TVs that constitute the CTS for the given circuit. CTS is initially empty.

Step 1: (a) Initialize $TV_{(temp)}$ with all 1 i.e. 1^n .

(b) Initialize status table by filling all n entries of its both columns with 0. First and second columns represent SA-1 and SA-0 fault respectively.

Step 2: call Process_description_table.

Step3: **for** each gate g in description table **do**

if g does not fix $TV_{(temp)}$, **then do**
 let w be the first wire index of this gate in description table i.e. first element of the row meant for this gate. Set t_w in $TV_{(temp)}$ to 0.

Step 4: Now after possible modification, $TV_{(temp)}$ is a valid TV of CTS. So include it in the set CTS.

Step 5: call Adjust_status_table with $TV_{(temp)}$ as its argument.

Step 6: Check if status table is full or not i.e. if there is any element in any row is 0 still, indicating that this wire needs further testing. If no such element found, then the algorithm halts at this point.

Step 7: If the algorithm does not halt in step 6, then it indicates that there exist some i such that i^{th} wire is not tested for SA-0 and/or SA-1 fault. So we need at least one more TV.

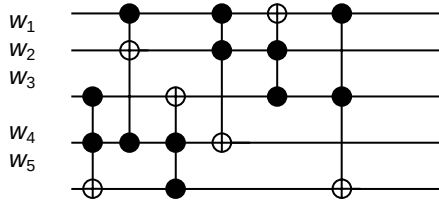
(a) Call Get_New_TV $_{(temp)}$. The returned TV by Get_New_TV $_{(temp)}$ is the TV $_{(temp)}$ for the next iteration.

(b) Go to step 3.

return CTS.

The following example demonstrates how ECTSG works.

Example 3.2: Consider a circuit whose description table contains the rows (4, 5), (3, 4), (2, 3), (1, 4), (1, 3) and (1,2)



Gate	Indices of Control Input	
1	3	4
2	1	4
3	4	5
4	1	2
5	2	3
6	1	3

(a) (b)
Fig. 3.1 (a) Original circuit (b) Corresponding description table

respectively, after sorting i.e. after calling the *process_description_table* procedure.

We have shown some initial tables maintained by ECTSG in Fig. 3.2 and 3.3. The reader can verify that ECTSG generates 3 TVs as CTS for this circuit: $TV_1 = 0\ 0\ 1\ 0\ 1$, $TV_2 = 0\ 1\ 0\ 1\ 0$, $TV_3 = 1\ 0\ 0\ 0\ 0$; hence $|CTS| = 3$. $\square$

Wire index	Tested for SA-1 fault (0 if not tested, 1 otherwise)	Tested for SA-0 fault (0 if not tested, 1 otherwise)
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0

Fig. 3.2. Initial Status Table

$TV_{1(temp)}$	TV_1
1	0
1	0
1	1
1	0
1	1

(a)

Wire index	Tested for SA-1 fault	Tested for SA-0 fault
1	1	0
2	1	0
3	0	1
4	1	0
5	0	1

(b)

Fig. 3.3 (a) Test Vector (TV_1) after **Pass 1** (b) Adjusted status table for TV_1 's effect

If no sorting is performed on the *description_table* i.e. if *Process_description_table* is not called in step 2 of ECTSG, the correctness of the proposed method will not be hampered but possibly different CTS will be yielded. What is the rationale for sorting then? To be convinced, consider the following arguments.

In the first place, we sort each row of *description_table* in increasing order of its elements. Now we will investigate the effect if the rows of *description_table* are sorted by *description_table[i][1]*. Consider the *description_table* of Fig. 3.4(b) and Fig. 3.4(c) where the rows of *description_table* are sorted by increasing and decreasing order of the *description_table[i][1]* respectively ($1 \leq i \leq m$). Fig. 3.4(a) shows the original *description_table*.

Now we assume that ECTSG does not call *Process_description_table* in Step 2 rather works directly on the input *description_table*. Table I shows that the sorting style adopted in Fig. 3.4 (c) yields most efficient result i.e. $|CTS| = 3$ that is less than the other two.

The reason is: in both cases we choose the first control input k of a gate (i.e. first element of a row in *description_table* i.e. *description_table[i][1]*) to set t_k in $TV_{(temp)}$ to 0 if it is not already 0. In the first case of our example given above, the abovementioned first control input k may appear only in the rows where *description_table[i][1]* is the first element – possibly few because it is not natural to have all the gates connected with 1st wire.

But in the second case, the first control input k may appear throughout all subsequent rows in the *description_table* because the elements of each row are sorted in increasing order, therefore possibly reducing the total number of changes in $TV_{(temp)}$. You can recall here that ECTSG algorithm's basic target is to retain $TV_{(temp)}$ as much as possible in a pass to generate the corresponding actual TV.

We can conclude from the above discussion that ECTSG with sorting according to the discussed style *may* outperform ECTSG without sorting. We say 'may' because we are not providing any rigorous proof of our conjecture. So we are not ruling out the possibility of getting worse result after sorting according to the abovementioned style for a particular circuit. What we have shown using the above examples is based on empirical observations.

C. Evaluation of Proposed Method

i) Optimality of ECTSG

ECTSG algorithm does not necessarily return optimal result but it does in many of the cases. We mentioned earlier that it is NP-hard to generate a minimum complete test set for stuck-at faults on a set of wires of a reversible circuit [7].

Gate	Indices of Control Input Gate	
1	3	4
2	1	4
3	4	5
4	1	2
5	2	3
6	1	3

(a)

Gate	Indices of Control Input Gate	
1	1	2
2	1	3
3	1	4
4	2	3
5	3	4
6	4	5

(b)

Gate	Indices of Control Input Gate	
1	4	5
2	3	4
3	2	3
4	1	2
5	1	4
6	1	3

(c)

Fig. 3.4 (a) Original *description_table* for the sample circuit Fig. 3.1(a), (b) Rows sorted in increasing order of the first elements and each row is also sorted in increasing order (c) Rows sorted in decreasing order of the first elements and each row is also sorted in increasing order.

TABLE I
COMPARISON OF DIFFERENT SORTING STYLES ON *DESCRIPTION TABLES* SHOWN IN FIG. 3.4

Description Table	CTS	CTSs
Fig. 3.4 (a)	4	(0 1 0 0 1), (1 0 0 0 0), (0 0 0 1 0), (0 0 1 0 0)
Fig. 3.4 (b)	5	(0 0 0 0 1), (0 0 0 1 0), (0 0 1 0 0), (0 1 0 0 0), (1 0 0 0 0)
Fig. 3.4 (c)	3	(0 0 1 0 1), (0 1 0 1 0), (1 0 0 0 0)

ii) Time Complexity Analysis

Theorem 1: Let an n -wire circuit consisting of m number of k -CNOT ($k \geq 2$) gates is given and ${}^nC_2 + {}^nC_3 + \dots + {}^nC_{n-1} = q$; ${}^qP_1 + {}^qP_2 + \dots + {}^qP_q = p$. (C and P stands for combination and permutation respectively). Then:

(a) There are total p number of different circuits C_i where each k -CNOT gate appears at most once, and then $\max(m_1, m_2, \dots, m_p) = q = O(2^n)$; m_i is the number of gates in circuit C_i ($1 \leq i \leq p$). (It is important to note that here we treat two gates as distinct only if their control inputs differ; we do not treat two gates as distinct if they have same control inputs but different inversion input. This assumption is inline with our assumptions throughout the whole book because we said earlier that the total book is *fixing* property oriented – so we are not concerned with the inversion input of a k -CNOT gate.)

(b) If we allow more than one occurrence of a gate that have same control input wires in a circuit, then $\max(m_1, m_2, \dots, m_p) = \infty$. ■

Proof: (a) As k -CNOT gate has maximum k control inputs and the number of wires is n , so $k \leq (n-1)$. We can have nC_2 number of 2-CNOT gates, nC_3 number of 3-CNOT gates, ..., ${}^nC_{n-1}$ number of $(n-1)$ -CNOT gates in a n -wire k -CNOT circuit with $k \geq 2$. So the total number of possible gates are ${}^nC_2 + {}^nC_3 + \dots + {}^nC_{n-1} = 2^n - n - 2 = q$. Now if we build a circuit where all of the above mentioned gates are used exactly once, then $m = q$. If some of these gates are not used but the used gates have exactly one occurrence as before, then $m < q$. Now we can permute these q gates to get a circuit, so the total number of circuits comprised of these q gates is: ${}^qP_1 + {}^qP_2 + \dots + {}^qP_q = p$.

(b) If a circuit is allowed to have more than one instance of any gate, then m may be raised to ∞ because we can append as many gates as we wish to build an arbitrary circuit, but use of some of these gates may be redundant. □

Lemma 1: We can confine the size of the *description_table* of ECTSG algorithm to maximum $q = O(2^n)$ ($q \neq \infty$). ■

Proof: We impose a condition that same ordering of values of the rows of the *description_table* should be adopted while taking input – for example, if (1, 3) is entered as description of a gate then (3, 1) is considered as an old gate and vice versa as these two descriptions are in fact same from ECTSG's viewpoint. The reasoning is: if the circuit has gates having same wires as control inputs, then they will have identical rows in the *description_table* because we are not storing the index of *inversion* input wire. As we are using these rows to fix the $TV_{(temp)}$ in ECTSG algorithm, once we have checked a row, no need to check another identical row. Hence we can keep only one row in case of duplicity, which completes the proof □

From the above discussion it is clear that ECTSG may not need to check all the m gates of a circuit. As each run of the ECTSG scans all the entries of the *description_table*, so the outcome of lemma 1 may save a lot of computation. So ECTSG is computationally feasible.

IV. EXPERIMENTAL RESULT

The result after the implementation of ECTSG is depicted in Table II. The designs of the benchmark circuits are taken from [8]. It is important to note that these machine-readable versions posted on [8] have a

characteristic: all the elements (discarding the *inversion* input of a gate) of a row are sorted in increasing order. So result shown in Table II produced by ECTSG algorithm is based on this structure, but actually ECTSG algorithm not necessarily needs the rows sorted. But the result may be different if different structures are used. Table II also contains a column that shows the result of ECTSG without sorting described in Section 3.2. From this column we notice that our conjecture mentioned in Section 3.2 holds correct for all the benchmark circuits as no result is better than the result found with sorting but converse is true for many cases.

Authors of [5] and [6] also used the benchmark circuits posted in [8]. As we said earlier that [5] showed that the size of CTS can be reduced to only 3 for any circuit consisting of k -CNOT gates ($k \geq 0$) but **always** needs modification of the circuit by adding one extra wire and replacing **every** k -CNOT gate by $(k + 1)$ -CNOT gate i.e. **always** needs to construct a DFT. As there may be hundreds of thousands of gates in a real-world circuit, it may be suitable to opt for ECTSG algorithm in this regard, which **may** need to construct DFT by adding maximum two extra wires and to replace **only** the k -CNOT gates having $k = 0$ and 1 by $(k + 2)$ -CNOT and $(k + 1)$ -CNOT gates respectively

TABLE II
EXPERIMENTAL RESULTS FOR DIFFERENT BENCHMARK CIRCUITS

Benchmark Circuits	Model	Bits of original circuit	Bits of DFT	Number of gates	Number of TVs by ECTSG without sorting*	Number of TVs by proposed ECTSG*	Number of TVs in [5]*	Number of TVs in [6]**
2of5	GT	6	8	18	2	2	8	4
4 49	GT	4	6	16	4	4	6	4
3 17	NCT	3	5	6	3	3	5	4
6sym	NCT	10	11	20	6	4	11	5
9sym	NCT	12	13	28	8	4	13	5
Cycle 10 2	GT	12	13	19	2	2	13	5
Cycle 17 3	GT	20	21	48	2	2	21	6
Gray code function	N/A	6	7	5	2	2	7	N/G
Ham 3	NCT	3	4	5	3	3	4	N/G
Ham7	GT	7	8	23	3	3	8	4
Ham 15	GT	15	16	132	5	5	16	12
Hwb4	GT	4	5	17	4	4	5	4
Hwb5	GT	5	6	55	6	6	6	7
Hwb6	GT	6	7	126	7	7	7	11
Hwb7	GT	7	8	289	8	8	8	13
Hwb8	GT	8	10	637	9	9	10	N/G
Hwb9	GT	9	10	1544	10	9	10	N/G
Hwb10	GT	10	11	3631	11	11	11	N/G
Hwb11	GT	11	12	9314	12	12	12	N/G
Mod5	NCT	5	6	8	3	3	6	N/G
Mod5adder	GT	6	8	21	3	3	8	4
Mod1024adder	GT	20	21	55	3	3	21	N/G
Mod 1048576adder	GT	40	41	210	3	3	41	N/G
Rd32	NCT	4	5	4	3	3	5	3
Rd53	GT	7	9	30	3	3	9	N/G
Rd73	NCT	10	11	20	7	4	11	4
Rd84	NCT	15	16	28	8	4	16	6
Xor5	NCT	5	6	4	2	2	6	N/G

* means this case works on DFT i.e. needs same number of bits as DFT.

** means this case works on original circuit i.e. needs same number of bits as original circuit.

N/G and N/A stands for “Not Given in this paper” and “Not Applicable” respectively.

That means if $k \geq 0$ in the original circuit, then we have to add 2 extra wires and all 0-CNOT and 1-CNOT gates are replaced by 2-CNOT gates, and if $k \geq 1$ in the original circuit, then we have to add 1 extra wire and all 1-CNOT gates are replaced by 2-CNOT gates, and if $k \geq 2$ then no need to construct DFT. Obviously, the less the number of control inputs in k -CNOT gate, the cheaper the gate is [9], so replacing 0-CNOT and/or 1-CNOT gates should not cost comparatively much. Moreover, in our case $|CTS|$ may be 2 i.e. optimal, but minimum $|CTS|$ of [5] is 3 for the same circuit at hand. The other method of [5] generates n TVs, namely, n weight-1 vectors, for n -wire k -CNOT ($k \geq 2$) circuit regardless the number of gates. In Table II we show that the number of TVs of the proposed is better than the method of [5] in most of the cases which generates exactly n TVs as CTS for an n -wire k -CNOT circuit ($k \geq 2$); and in the rest of the cases it generates equal number of TVs as [5].

Some other researchers, specifically, Patel et al. [4] and Allen et al [6] did not use *fixing property*. Both the algorithms of [4] and [6] outperform ECTSG in terms of the need for modification of the original circuit (i.e. hardware cost) as none of their technique requires to construct DFT like ECTSG (possibly i.e. in case of existing 0-CNOT and/or 1-CNOT gates); but as we said earlier that cost of 2-CNOT gates, which ECTSG requires to be added in a circuit where 0-CNOT and/or 1-CNOT gates exist, are nominal [9]. Moreover, ECTSG algorithm is superior to [6] in most cases, shown in Table II in terms of $|CTS|$ and equal in rest of the cases. The computational cost of [4] and [6] is presumably larger than ECTSG. And lastly, in [6], minimum $|CTS|$ is 3 in table II but it is reduced in ECTSG to 2.

V. CONCLUSION

Utilizing fixing property, we presented an algorithm that generates complete test set for stuck-at model for reversible circuits composed of k -CNOT gates with $k \geq 2$. The limitation $k \geq 2$ can be lifted because any n -wire k -CNOT circuit with $k \geq 0$ can be converted to a $(n + 2)$ -wire k -CNOT circuit with $k \geq 2$. We also provided some experimental results and comparison between the existing methods.

REFERENCES

- [1] C. Bennett, "Logical reversibility of computation", *IBM Journal of Research and Development* 17 (1973) 525–532.
- [2] R. Landauer, "Irreversibility and heat generation in the computational process", *IBM Journal of Research and Development* 5 (1961) 183–191.
- [3] <http://www.zyvex.com/nanotech/reversible.html>
- [4] K.N. Patel, J.P. Hayes, I. L. Markov, "Fault testing for reversible circuits", *VLSI Test Symposium*, 21 (2003) 410–416.
- [5] A. Chakraborty, "Synthesis of Reversible Circuits for Testing with Universal Test Set and C-Testability of Reversible Iterative Logic Arrays", *18th International Conference on VLSI Design* (2005) 249-254.
- [6] J.S. Allen, Jacob D. Biamonte, Marek A. Perkowski, "ATPG for Reversible Circuits using Technology-Related Fault Models", *International Symposium on Representations and Methodologies for Emergent Computing Technologies*, Tokyo, Japan, September (2005).
- [7] I. Shigeru, I. Yusuke, T. Satoshi, U. Shuichi, "On the Complexity of Fault Testing for Reversible Circuits", *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)* 105 (2005) 13-16.
- [8] D. Maslov, G. Dueck, and N. Scott, Reversible Logic Synthesis Benchmarks Page. <http://webhome.cs.uvic.ca/~dmaslov/>

Two-Level Dictionary-Based Text Compression Scheme

Md. Ziaul Karim Zia¹, Dewan Md. Fayzur Rahman², and Chowdhury Mofizur Rahman³

United International University, Dhanmondi, Dhaka, Bangladesh

Email: zia@yahoo.com¹, dfayzur@yahoo.com², cmr@uiu.ac.bd³

Abstract — In this paper a new dictionary and memory based text compression technique is presented called a Two-Level Dictionary Based Text Compression Scheme. The original words in a text file are transformed into codewords having length 2 and 3 using a dictionary comprising 73680 frequently used words in English language. Among these words most frequently used words use 2 length codewords and the rest use 3 length codewords for better compression. The codewords are chosen in such way that the spaces between words in the original text file can be removed altogether recovering a substantial amount of space. Another unique feature of our compression scheme is that we have recovered unused bit of ASCII character representation from each character to save one byte per 8 ASCII characters. Lastly a back end existing compression algorithm is used to finally compress the file. We have achieved about 75% (compression ratio of 2.01 bits per input character) reduction in size using our new compression strategy with gzip and bzip2.

Index Terms — Text compression, Dictionary based compression, Huffman code, word transformation.

I. INTRODUCTION

Text compression is typically used to save storage and communication costs. Dictionary-based text compression is a compression technique, which provides the most significant improvement in the text compression performance [1,19]. It is based on the notion of replacing whole words with shorter codes from the real text file. Usually large volume of words with their codes is preserved in the dictionary. In the recent past, the M5 Data compression group, University of Central Florida has developed a family of reversible Star-transformations [2] which is applied to a source text along with a backend compression algorithm, achieves better compression. The drawback of these compression algorithms is that execution time performance and runtime memory expenditure is relatively high compared with the backend compression algorithm such as bzip2 and gzip.

When transmitting text or a binary executable, one must be able to reconstruct every bit perfectly—hence the need for lossless data compression. Compression is usually broken into two steps: modeling and coding. With a perfect, concise model that describes the generation of the input source which is to be compressed, one could reproduce the data without transmitting the source data. In practice, one must approximate and construct an

approximate mathematical model for the data. In English text, for example, one can model the probability of a letter occurring as a probability conditioned on letters that have already been transmitted. This probabilistic model is transmitted with a description of how the data differs from the model. In a coding step, information is mapped to compact codewords. Obviously, a codeword must decode to a unique value so there can be no doubt of the original message. Prefix codes are used so that no codeword is the prefix of any other codeword. It has been proved [3] that, for any non prefix code that may be uniquely decoded; a prefix code can be found with the same codeword lengths. Often the modeling and coding steps are tightly coupled. For instance, Lempel-Ziv [6] codes can be constructed as an input source is parsed into a “dictionary” model.

In this paper a new dictionary and memory based text compression technique is presented called a Two-Level Dictionary Based Text Compression Scheme. The original words in a text file are transformed into codewords having length 2 and 3 using a dictionary comprising 73680 frequently used words in English language. Among these words most frequently used words use 2 length codewords and the rest use 3 length codewords for better compression. The codewords are chosen in such way that the spaces between words in the original text file can be removed altogether recovering a substantial amount of space. Another unique feature of our compression scheme is that we have recovered unused bit of ASCII character representation from each character to save one byte per 8 ASCII characters. Lastly a back end existing compression algorithm is used to finally compress the file. We have achieved about 75% (compression ratio of 2.01 bits per input character) reduction in size using our new compression strategy with gzip and bzip2. Fig. 1 represents our system. We have conducted experiments on a large number of corpora and vast volume of dictionary (up to 73680 words) to evaluate the effectiveness of our approach.

The remainder of this paper is organized as follows. Section 2 introduces overview of the existing text compression algorithms. Subsequently, why transformation is needed is presented in Section 3. In Section 4, we propose the compression and decompression of our proposed method. Section 5 describes our Physical

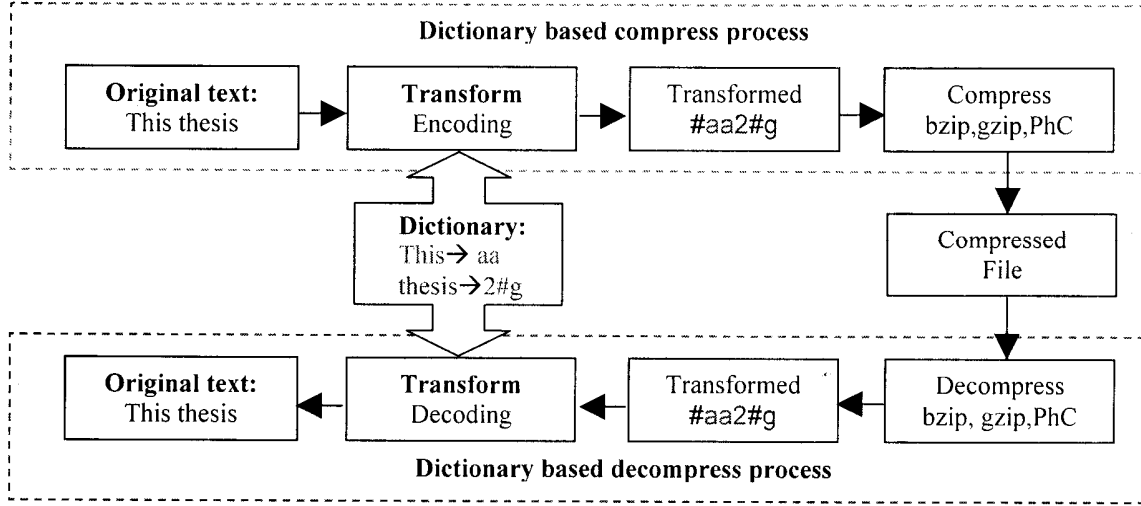


Fig. 1. Text Transformation Architecture

Compression (PHC) technique. Experimental results are given in Section 6 to verify the feasibility and effectiveness of the proposed method. Our concluding remarks are contained in Section 7.

II. OVERVIEW OF THE EXISTING TEXT COMPRESSION ALGORITHMS

A procedure known as static *Huffman coding* [14] can be used to build an optimal code in which the most frequently occurring symbols are given the shortest codewords. Huffman codes are established by storing the symbols of the alphabet in a binary tree according to their probability. As the tree is traversed from root to leaf, the code grows in length. When visiting the right child, a 0 is appended to the code. When visiting the left child, a 1 is appended. Thus, symbols which occur frequently are stored near the root of the tree and have the shortest codes.

In Arithmetic coding optimal compression ratio for a data source is traditionally described with respect to Claude Shannon's [4] definition of source entropy [15]; a measure of the source's information and therefore the average number of bits required to represent it. Sometimes the most frequently occurring symbol can contain so little information that it would be ideal to represent it with less than 1 bit. Huffman codes are restricted to using an integral number of bits per symbol, increasing the coding overhead. Arithmetic codes have been designed to support a fractional number of bits per symbol to bring the average length of a codeword much closer to the optimal. Knowing the probability of occurrence for each symbol, a unique identifier can be established for a series of symbols in the interval (0,1). Unlikely symbols narrow this interval so that more bits are required to specify it, while highly

likely symbols add little information to a message and require the addition of fewer bits as the interval refinement is coarser. As the fraction converges, the most significant bits become fixed, so the fraction can be transmitted most-significant-bit-first as soon as it is known.

The coding techniques described above are used in the algorithm families introduced below. There are two fundamental methods for constructing Lempel – Ziv [16] codes. Introduced in the late 1970s, these methods are known by the initials of their creators and the year of introduction: LZ77 and LZ78. LZ77 [5] maintains a current pointer into the source data, a search buffer, and a look-ahead buffer [5]. The search buffer is made up of symbols encountered prior to the current symbol, and the look-ahead buffer contains symbols which appear after the current symbol. Together, the buffers comprise a "window" which specifies the section of the input source under consideration. As the current pointer advances, the algorithm searches backward for the longest match in the search buffer. Instead of transmitting the matched symbols, they can be encoded with a triple: <offset from pointer, length of match, next codeword>. The "next codeword" is the codeword corresponding to the symbol in the look-ahead buffer following the match. The LZ78 [6] scheme was introduced to account for cases in which a nearby match cannot be found. Instead of the sliding search-buffer, LZ78 uses a separate dictionary, which also serves as a codebook. As each group of symbols is encountered, the dictionary is checked. An<index, code>pair is output where *index* corresponds to the longest prefix (if any) that matches the current input, and code is the unmatched symbol which follows. The pair is then added to the dictionary. A popular improvement to LZ78 is called LZW [7]. It seeds the dictionary with letters from

the source alphabet which eliminates the need to send the second element of the pair, shortening the number of bits that must be sent for a single character.

Prediction with Partial Match (PPM) uses Markov modeling followed by arithmetic coding. The Burrows-Wheeler Transform (BWT) reversibly permutes a block of source data so that it can easily be compressed. The fact that a certain string of symbols has appeared can aid in predicting which symbol will come next. For instance, if the letters *compr* appear in this article, there is a strong probability they will be followed by an *e* (*compress*). The PPM [8,18] scheme maintains such context information to estimate the probability of the next input symbol to appear. An arithmetic coder can use this stream of probabilities to code the source efficiently. *PPMd* [9] is a recent implementation of the PPM algorithm. The BWT [16] algorithm will naturally split the data into independent blocks of a predetermined size before the data is compressed. The blocks are then processed through the BWT algorithm. Since each block is independent, it should be possible to run the BWT algorithm on multiple blocks of data simultaneously and achieve speedup on a parallel machine. The separate blocks are then concatenated back together again to form the final compressed file. The sequential version of BWT will process the blocks in order. A parallel implementation will need to keep track of block ordering and write the compressed blocks back to disk in the correct order. The newest technique among those examined, the Burrows-Wheeler Transform, converts a block S of length a into a pair consisting of a permutation of S (call it L) and an integer in the interval $(0..n-1)$.

III. WHY TRANSFORMATION IS BENEFICIAL

There are three considerations that lead us to our transform algorithm. First, we gathered data of word frequency and length of words information from our collected corpora. It is clear that almost more than 60% of the words in English text have the lengths greater than three and more than 80% of the words in English text have the lengths greater than two. There exists a list of the 1000 most frequently used words in the English language. Read Naturally passages are made up of a high percentage of these words. The English vocabulary has increased greatly in more than 1500 years of development. The most nearly complete dictionary of the language, the Oxford English Dictionary (second edition, 1989), contains more than 600,000 words, including obsolete forms and variant spellings.

The second consideration is that the transformed output should be compressible to the backend compression algorithm. In other words, the transformed intermediate output should maintain some of the original context information as well as provide some kind of “artificial” but strong context. The reason behind this is that we choose BWT and PPM algorithms as our backend compression tools. Both of them predict symbols based on context information. Finally, the transformed code words can be treated as the offset of words in the transform dictionary. Thus, in the transform decoding phase we can use a hash function to achieve $O(1)$ time complexity for searching a word in the dictionary. Based on this consideration, we use a continuously addressed dictionary in our algorithm. In contrast, the dictionary is split into 22 sub-blocks in LIPT [2]. Results show that the new transformation is better than LIPT not only in time complexity but also in compression performance.

IV. COMPRESSION AND DECOMPRESSION OF PROPOSED METHOD

The concept of replacing words with shorter codewords from a given static dictionary has at least two shortcomings. First, the dictionary must be quite large, at least tens of thousands of words and is appropriate for a single language only (our experiments described in this paper concern English text only). Second, no “higher level”, e.g., related to grammar, correlations can be implicitly taken into account. In spite of those drawbacks, such an approach to text compression turns out to be an attractive one, and has not been given as much attention as it deserves. The benefits of dictionary-based text compression schemes are the ease of producing the dictionary (assuming enough training text in a given language), clarity of ideas, high processing speed, cooperation with a wide range of existing compressors, and last but not the least competitive compression ratios.

In this section, we present our modifications to dictionary based text preprocessor, as well as a few new ideas. In any text file there are spaces between words. It holds near about 20% spaces (by calculation when average word length is 4). So it will save space if we write text file without using spaces by implementing a novel dictionary based word replacing text compression scheme. The codewords have been designed in such a way that spaces between codewords can be removed altogether without any difficulty to retrieve the original words of the target text file. The result of our modifications with bzip2 and gzip can be observed in Table II. Rules for word replacing transform are given below:

TABLE I
LETTERS AND SYMBOLS USED TO FORM CODEWORDS

Code Type	Used Symbol	Code word example	Num. of possible codewords
2 length (Started with alphabet or key board symbol except punctuation and \$#<>)	a-z = 26 nos.	aa	6,440 (70x70+70x22) (For frequent used words)
	A-Z = 26 nos.	aA	
	@ % ^ & * () _ - + = ! \ / { } [] = 18 nos.	A\$	
	0-9 : ; ' ' , . ! ? \$ # < > = 22 nos.	A# Z/ &; %"	
3 length (Started with numeric number)	a-z = 26 nos.	1aa	67,240 (82x82x10)
	A-Z = 26 nos.	1Aa	
	@ \$ # % ^ & * () _ - + = ! \ : ; ' ' , . ! ? / { } [] < > = 30 nos.	2z\$	
	0-9 = 10 nos.	3{p 5} 9:?	
Control code	EoL	Except EoL	
	< >	End of Line	
	<< >> <<< >>>	To denote non dictionary word	
	\$ #	All capital 1 st capital	
Total			73,680

- Each word is converted to a fixed length codeword.
- The codewords are two type having length 2 and 3.
- Those words that are not present in dictionary are not converted to codewords and are written between '<' and '>' symbol.
- Two length codewords start with alphabet or keyboard symbol except punctuation and \$ # < > symbols. 3 length codewords start with number.
- High frequency and long length words are converted to 2 length codewords.
- There is no codeword for punctuation. So punctuation is not converted to codeword and no codeword begins with punctuation.
- Words started with capital letter are converted to their lowercase equivalent and to denote this change the flag '\$' is put in front of the respective codeword. Additionally it is worth using another flag to mark a conversion of a full uppercase word to its lowercase form with flag '#'. If all the letters of a word are in lower case then no flag (\$ #) is present before their codeword.

The dictionary used in this experiment is prepared in advance, and is shared by both the transformation encoding module and the transformation decoding

process. Most frequently used words are listed at the beginning of the dictionary with corresponding code words. The total dictionary is divided in to 26 files according to starting alphabet of words for improving time complexity of search. The letters and symbols used to form code words are shown in Table I.

In any compression process time complexity is a very important. For a fast compression and decompression process we made 26 dictionaries according to alphabetic groups and named the dictionary by same alphabet, e.g., 'Dictionary-A' contained the most frequent words that are started by 'a' and all the words of a dictionary maintain a sequence according to their frequency to achieve better compression performance. With the ordering specified above, each word in Dictionary is assigned a corresponding codeword. The first 6440 words in Dictionary are assigned two length codewords ("aa", "ab", ..., "Za" etc.). The next 67240 words are assigned three length codewords ("1aA", "9aB", ... , "99Z" etc.). Hence, the most frequently occurring words are assigned code words of two lengths. Using this mapping mechanism, a total number of 73680 words can be included in the Dictionary.

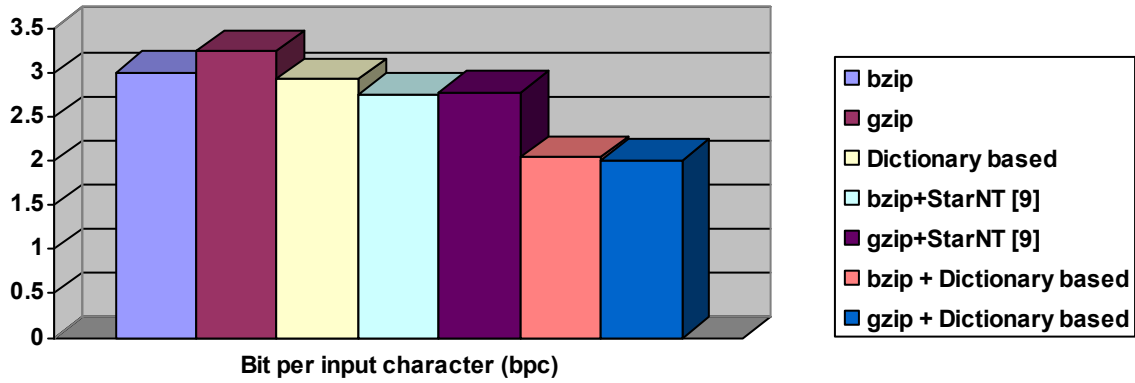


Fig. 2. Comparative Compression Ratio

TABLE II
COMPRESSION RESULT

Name of scheme	Bit per input character (bpc)	Average size (byte)
Bzip	3.01	11264
Gzip	3.25	11264
Dictionary based	2.93	11264
bzip+StarNT [9]	2.76	11954
gzip+StarNT [9]	2.78	11954
bzip + Dictionary based	2.05	11264
gzip + Dictionary based	2.01	11264

V. PHYSICAL COMPRESSION IN PROPOSED METHOD

At one time, every manufacture had a different code for characters which led to all kinds of confusion. Eventually, industry settled on an input-output code known as the American Standard Code for Information Interchange (ASCII). This code allows manufacturers to standardize I/O hardware such as keyboards, printers, video displays and so on. The ASCII code is a 7-bit code whose format is $X_6X_5X_4X_3X_2X_1X_0$ where each X is a 0 or 1. For instance, the letter A is coded as 1000001. Any text file consists of a set of characters that are represented by ASCII codes and we also know in memory each character holds 8-bit (1 byte) space instead of 7-bits, hence we loose 1 bit for storing a character in any memory device. In this way for 8 characters, we can recover 8 bits (1byte) if we store each character in just 7 bits. Actually (7x8) 56 bits of memory space is required for storing any 8 ASCII characters, as a result we need 56 memory Cells for storing 8 symbols that is equal to 7 byte. Thus 8 characters are stored in 7 bytes releasing 1 byte for storing one additional character. By

using this approach we recover 12.5% space from any text file. Here we use 8 character cluster for such compression process and recover a byte from each cluster. If the last cluster of a file contains less number of characters the compression algorithm takes additional null values to fill the cluster.

VI. EXPERIMENTAL RESULTS

The experiments are demonstrated on the text file which is 11KB (11264 byte) in length containing 10509 characters with space and new line. We compared our results with bzip+StarNT and gzip+StarNT. The results of bzip and gzip are also included in Table II and Fig. 2.

VII. CONCLUSION

In this paper we have presented a static dictionary-based preprocessor, called word replacing transformation, intended for text compression. Our ideas are significant extensions of the mechanisms presented in former algorithms, like StarNT. We have proposed several unique

approaches with results having excellent improvement in text data compression. First we have applied our dictionary based text transformation preprocessor and the result showed that compressed file uses 2.93 bits per character. We also achieve 2.05 bit per character with bzip and 2.01 with gzip as back end text compressor. The result will be more impressive by using BOA or RK compression scheme and if we increase the dictionary volume. The second scheme we have used is physical compression by utilizing the unused bit from the ASCII representation of a character. This scheme compresses a fixed percentage (12.5%) of space for any text file or preprocessed text file.

REFERENCES

- [1] R. Franceschini, H. Kruse, N. Zhang, R. Iqbal, and A. Mukherjee, "Lossless, Reversible Transformations that Improve Text Compression Ratio," *Project paper*, University of Central Florida, USA, 2000.
- [2] F. Awan and A. Mukherjee, "LIPT: A Lossless Text Transform to Improve Compression," *Proceedings of International Conference on Information and Theory: Coding and Computing*, IEEE Computer Society, pp. 452-460, April 2001.
- [3] R.J. McEliece, "The Theory of Information and Coding," *Encycl. Math. Appl.*, 3, Addison-Wesley, 1977
- [4] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379-423 and 623-656, July and October, 1948.
- [5] J. Ziv and A. Lempel, "A Universal Algorithm for Sequential Data Compression," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp.337-343, May 1977.
- [6] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Transactions on Information Theory*, vol.IT-24, no.5, pp.530-536, Sept. 1978.
- [7] Terry A. Welch, "A Technique for High Performance Data Compression," *IEEE Computer*, Vol. 17, pp. 8-19, June 1984.
- [8] J. Cleary and I. Witten, "Data compression using adaptive coding and partial string matching," *IEEE Transactions on Communications*, vol. 32, pp. 396-402, Apr. 1984.
- [9] W. Sun, A. Mukherjee, N. Zhang, "A Dictionary-based Multi-Corpora Text Compression System," In Storer JA, Cohn M, editors, *Proceedings of the 2003 IEEE Data Compression Conference*, IEEE Computer Society Press, pp. 448 - ,March 2003.
- [10] J. Seward, bzip2, Available : <http://www.spec.org/osg/cpu2000>
- [11] U. Manber, "A Text compression scheme that allows fast searching directly in compressed file," *ACM Transactions on Information Systems*, Vol.52, N0.1, pp.124-136, 1997.
- [12] Matthew V. Mahoney, "Adaptive Weighing of Context Models for Lossless Data Compression," *Florida Tech. Technical Report TR-CS-2005-16*, 2005
- [13] P. Skibi ski, Sz. Grabowski, S. Deorowicz, "Revisiting dictionary-based compression," *Software-Practice and Experience*, 2005; vol. 35, no. 15, pp. 1455-1476, 2005.
- [14] D. A. Huffman, "A method for the construction of minimum redundancy codes," *In Proc. IRE 40*, volume 10, pages 1098-1101, September 1952.
- [15] I. Witten, R. Neal, and J. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520-540, June 1987.
- [16] M. Burrows, and D.J. Wheeler, "A Block-sorting Lossless Data Compression Algorithm," *Digital Systems Research Center Research Report 124*, 1994, Available : <http://gatekeeper.dec.com/pub/DEC/SRC/research-reports/abstracts/src-rr-124.html>.
- [17] Nelson, Mark and Gailly, Jean-Loup, *The Data Compression Book*, second edition, M&TBooks, New York, 1995.
- [18] A. Moffat, "Implementing the PPM data compression scheme," *IEEE Transactions on Communications*, vol. 38, no. 11, pp. 1917-1921, November 1990.
- [19] F. Awan, N. Zhang, N. Motgi, R. Iqbal, and A. Mukherjee, "LIPT: A Reversible Lossless Text Transform to Improve Compression Performance," *Proceedings IEEE Data Compression Conference*, pp. 481-210, 2001.

Generation of Random Channel Specifications for Channel Routing Problem

Achira Pal [#], Tarak N. Mandal [@], Alak K. Datta [&], Debojit Kundu ^{\$} and Rajat K. Pal [®]

[#] Harinavi Subhasini Balika Sikshalaya, PO: Harinavi, Dist.: 24 Pgs. (South), Kolkata: 700 148, West Bengal, India

[@] IBM India Pvt. Ltd., 3rd Floor, Desk No.: 335, DLF IT Park, New Town, PO: Rajarhat, Kolkata: 700 156, India

[&] Department of Computer and System Sciences, Visva-Bharati, Santiniketan: 731 235, Birbhum, West Bengal, India

^{\$} HCL Technologies Ltd., Block: A8 & A9, Sector: 60, Noida: 201 301, Uttar Pradesh, India

[®] Department of Computer Science and Engineering, University of Calcutta, 92, A. P. C. Road, Kolkata: 700 009, India

Abstract — In this paper we develop algorithms for generating random channel specifications of channel routing problem in VLSI design. A channel is a rectangular routing region containing two sets of fixed terminals on two of its opposite sides and the other two opposite sides (of the rectangle) are open ends, may or may not contain any terminal of a net but the terminal position is not fixed before a routing solution is computed. Most of the problems in two-, three-, and multi-layer channel routing are beyond polynomial time computable. Hence for each of these problems it is unlikely to design a polynomial time deterministic algorithm. Developing heuristic algorithm might be a probable way out that hopefully provides good solutions for most of the instances occur in practice. Novelty of a heuristic algorithm is judged better if it works for a variety of large number of randomly generated instances of the problem. In fact, convergence of results of a heuristic algorithm is well established when the algorithm of a problem is executed for a huge number of randomly generated similar instances and the final result is computed making an average on all of them.

Index Terms — Channel routing problem, VLSI design, NP-hardness, Channel specification, Random generation, Simple channel instance, General channel instance, Algorithm.

I. INTRODUCTION

In this paper we are interested in generating random channel specifications for channel routing problem. To do so we first generate channel specifications having two-terminal nets only, where none of the channel specifications contains any vertical constraint. Such channel specifications that do not have any vertical constraint, are known as *simple channel specifications* [4, 5, 6]. Then we modify our requirement for computing general channel instances with multi-terminal nets, so that we can generate channel specifications containing two- or

more than two-terminal nets; also containing vertical constraints.

We know that the two-layer channel routing problem of area minimization is polynomial time solvable for simple channel specifications [1, 7, 8], but the problem of crosstalk minimization in two-layer channel routing is NP-hard even if the channel specifications are free from any vertical constraint [4, 5, 6]. So, it is unlikely that there would exist a polynomial time algorithm for computing two-layer minimum crosstalk channel routing solutions even for the *simple instances* of channel specifications. Incidentally, such channel instances are hardly available in literature, and generally the instances that are known as *benchmark channel specifications* contain both horizontal as well as vertical constraints [7], though such channel specifications are also very few.

A *channel specification* is usually obtained in the form of two m -element vectors TOP and BOTTOM indicating the *top* and *bottom terminal lists*, respectively, of a channel. We often may use *channel specification* and *channel instance* interchangeably. A *channel* is a rectangular routing region containing two sets of fixed terminals on two of its opposite sides, say top and bottom, and the other two sides, the left and right sides of the rectangle, are open ends. The left and right sides of the rectangle may contain any terminal of a net, but the terminal position is not fixed before having a routing solution. The fixed terminals on the top and at the bottom are usually aligned vertically in *columns*. The set of terminals that need to be electrically connected together is called a *net*. A channel specification with a set of eight two-terminal nets in a channel is shown in Figure 1; the length of the channel is 18.

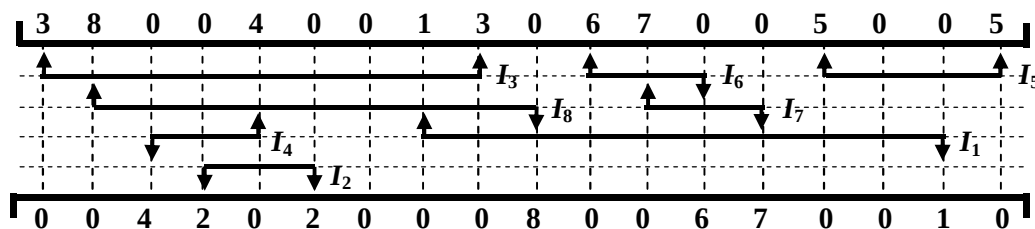


Figure 1. An example channel of eight nets; zeroes are non-terminals or vacant terminals, not to be connected. Intervals of the nets are placed in four different tracks. Terminals are vertically aligned along the columns of the channel. The length of the channel (i.e., the number of columns) is 18. Arrows indicate that the terminals to be connected, either on the top or at the bottom, to complete the required interconnection of all the nets present in the channel.

In order to obtain simple channel specifications, in this paper, we develop an algorithm for generating instances of such (simple) channel specification in random. There are several problems in channel routing that are hard in nature even if the channel specifications are free from any vertical constraints [4, 5, 6]; such randomly generated simple channel specifications can be used to compute desired routing solutions in executing heuristic algorithms designed for those problems [5, 6]. In the next section (Section II), we formulate the steps necessary for developing the algorithm in generating simple routing channel instances. Then we generalize the algorithm in generating general routing channel instances in Section III. We conclude the paper with a few remarks in Section IV.

II. GENERATION OF SIMPLE CHANNEL SPECIFICATIONS

A. An Overview

Let us assume, the number of nets be n in a channel that is being generated. So the channel length would be of $2n$, if no *blank column* (containing no terminal on the top as well as at the bottom) or *trivial column* (containing both the terminals of the same net) is introduced into the channel. In generating random channel specifications of simple in nature, we introduce a non-terminal (i.e., a vacant terminal) either on the top or at the bottom randomly, and the other side of the column would contain a terminal of a net. Our next task is to obtain the starting and the ending column positions of a two-terminal net, and on which side it would be (i.e., to obtain the terminal position of the net), on the top or at the bottom, randomly along the length of the channel. In this respect we like to fix up and maintain a set of criteria while generating the random channel specifications, as follows.

(a) Though it is needless to mention that the net numbers are nothing but symbols to differentiate themselves, we want to obtain the channel specifications where the net numbers would present randomly along the length of the channel, i.e., the nets are not sorted in succession based on their starting column positions from left to right or from right to left along the length of the channel, or something of that sort.

(b) Nets should be of different spans (or intervals) and they would present randomly along the length of the channel. If L_i (R_i) is the leftmost (rightmost) column position of net n_i , then $I_i = (L_i, R_i)$ is known as the *span* (or *interval*) of the net. As for example, the span of net 1 in Figure 1 is 9 and that of net 7 is 2. This criterion tells that all the smaller (or larger) nets are not accumulated on a side along the length of the channel.

(c) Generally, in practice, a channel contains smaller nets in large number and larger nets a few. Nets are defined smaller or larger based on their relative spans (or intervals) over the length of the channel. However, the number of nets with a fixed span (say, 5 nets of span 3 each, 2 nets of span 7 each, etc.) is not fixed; otherwise, the randomness of the channel to be generated may suffer.

Only we can say that, there would be a large number of smaller nets, and a few larger nets, but we do not precisely specify the exact spans of different nets or the number of nets of a fixed span. Obviously, the number of nets with some intermediate spans is neither more nor less. Actually, this criterion helps us in generating random channel instances where the number of nets gradually reduces as their spans increase along the length of the channel.

In order to generate the desired channel specifications, as guided by the criteria stated above, we now describe the channel specification generating procedure as follows. Let us assume the channel instance we want to generate is C that contains n nets, and its length is $m = 2n$ (as each of the n nets is a two-terminal net and we are not introducing any blank column or trivial column into C). We can simply generate two column numbers randomly for each net, and just place them on the top or at the bottom of the selected columns randomly. The problems that may arise in this method are that of generating the same number (as a net) more than twice (that means the net that has already been assigned) and the same column position c in C , $1 \leq c \leq m$, is accessed for assigning net j , which has already been occupied by net i , $j \neq i$, after starting the procedure (that can be considered as a *collision*). Finding out of a *free* column (where no net has yet been assigned) in some iteration might enhance the cost of the procedure (in terms of computational complexity) beyond some limit, and even if we use *linear probing technique* for such *collision resolution* (like *hashing*), it may result in some channel specification, which may not satisfy one or more criteria, mainly the third criterion described above.

So, generation of random column number for the nets is not a problem at all; even different methods for generating random numbers are easily available in literature, among which we can use any one of them. The problem gets apparent whenever a collision occurs, because we cannot predict anything about generation of random numbers. To get rid off this collision situation, we may think of the problem in a different way so that randomness remains unaffected.

B. Formulation of the Problem

In this subsection, we formulate the problem towards developing the algorithm in generating simple routing channel specifications, based on the overview of generating the same as discussed in the previous subsection. The channel specification may be thought as a linear *list* of two-tuple elements, where each element corresponds to a column in the channel (that is being generated) and the two tuples correspond to the top and the bottom terminal positions at each column. The length of the list is same as the length of the channel (say m), where n -number of nets are to be introduced when the channel is completely generated. For each list element we generate a net number; put it in one of the two tuples, i.e., either on the top or at the bottom, and the other side would contain a non-terminal.

At the very beginning, all the elements of the linear list of the two-tuple assumed above are initialized with zeros. In our algorithm, we replace $2n (= m)$ zeros in total in this two-tuple list, in order to generate a desired simple routing channel instance, where columnwise only one zero is randomly replaced by a net number. In order to generate such a channel specification, we assume another list of length m , which is same as the length of the channel specification that is being generated. Initially the list contains the natural numbers starting with zero onwards (i.e., 0 through $m-1$), i.e., each element i in the list contains the number $i-1$, $1 \leq i \leq m$.

We consider a singly linked linear list L of m nodes, and initialize the *information field* of node i ($L(i) \rightarrow info$) by $i-1$, where $1 \leq i \leq m$. In other words, the linear list L contains the column numbers 0 through $m-1$ of channel C to be generated. Initially all the columns are free, as we have not yet assigned any net to a column. If some net is assigned to column c , $0 \leq c \leq m-1$, then the node with information c is deleted from L , and keeps L as the list of free or available columns for the nets yet to be assigned. For each net i , $1 \leq i \leq n$, we randomly choose a node N from the existing list L of free nodes, retrieve its information part, which is a free column number in C , and store i in the free column. So, if N contains c in L , we replace c by i , and delete N from L . In effect, channel C would contain a terminal of net i at column c . Furthermore, in identifying the side (TOP or BOTTOM) of the channel, we choose another random number that signifies whether i is assigned on the top or at the bottom of c so that the other side of c is allotted for a non-terminal (i.e., 0). We treat this terminal as the position of the *initial terminal* of net i .

The problem of selecting a particular node in L may also arise for multiple times, if we do not delete the selected node from L . That is why we assign the current net to the column obtained by randomly selecting a node in L as stated above, and delete the node from L . In the same iteration (i.e., for net i) we generate the *final terminal* for net i too.

While generating the position of the *final terminal* of net i , $1 \leq i \leq n$, we have to keep in mind the last criterion of generating random channel specification stated above that is relating to spans of nets to be introduced in C . Here our intention is to generate many smaller nets and a few nets with larger spans. In order to obtain a random channel specification, we generate some smaller random numbers that would be added to or subtracted from column number c that is already selected randomly as the initial column position of net i . Initially we generate smaller random numbers and assign the nets with lesser spans along the length of the channel. Gradually L becomes sparse as the nets are assigned to C , and even for randomly generated smaller numbers nets with larger spans are evolved during later iterations.

As every time a new node with a new column number (as its information part) is selected, we can always be able to use the column successfully for a new net without any conflict, and in this way the problem of multiple occurrences of the same random number could be resolved.

C. The Complexity of the Algorithm

This is the way how a random channel C of n two-terminal nets is generated, where the length m of the channel specification is same as $2n$, and C contains no vertical constraint in it. One such randomly generated simple channel instance is shown in Figure 2(a). Figure 2(b) shows the horizontal constraint graph of this channel. In a *horizontal constraint graph*, n vertices are introduced corresponding to n nets belonging to a channel, and two vertices are connected by an edge if and only if the spans (or intervals) of the corresponding nets overlap in the channel [7, 10]. Table 1 shows how the nets in the channel generated are distributed over the length of the channel, as the net number increases. The table also shows the leftmost column position and the rightmost column position of each of the nets. From this table, it is clear that the spans of the nets also do not increase as the net numbers increase.

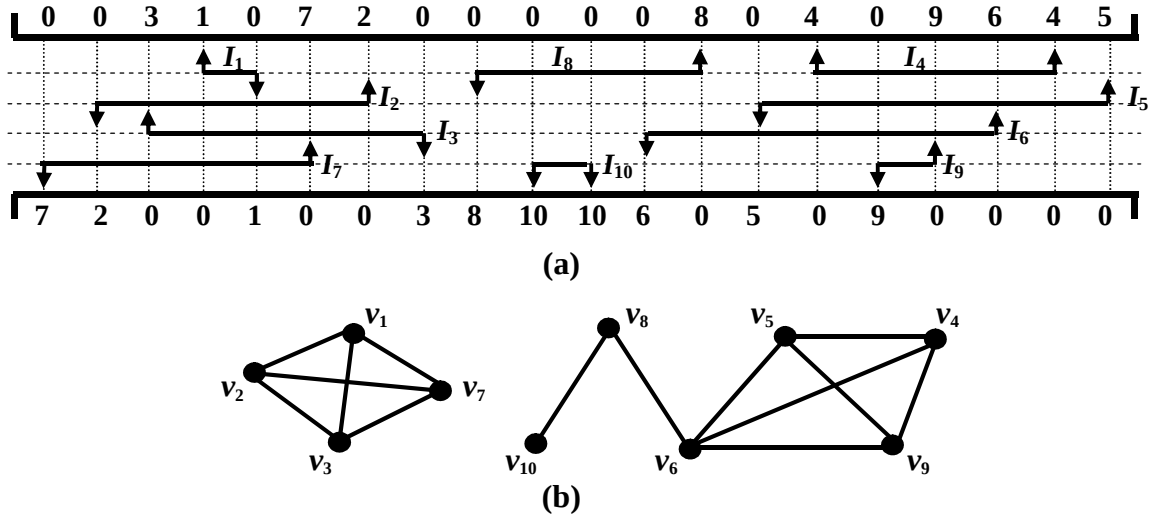


Figure 2. (a) The generated simple channel specification in this section. The channel contains 10 nets, and its length is 20. (b) The horizontal constraint graph of this generated simple channel specification. It consists of two components, as no net in this generated channel instance passes through both columns 8 as well as 9.

Net Number	Leftmost Column Position	Rightmost Column Position	Span (or Interval)
1	4	5	1
2	2	7	5
3	3	8	5
4	15	19	4
5	14	20	6
6	12	18	6
7	1	6	5
8	9	13	4
9	16	17	1
10	10	11	1

Table 1. Spans (or intervals) of different nets introduced into the randomly generated channel specification, shown in Figure 2(a).

Now we compute the computational complexities in generating such a random channel specification C as follows. In C , we have n nets in total. In addition, m is also of the order of n . Initially, we create a singly linked linear list L of order n . For each net i , $1 \leq i \leq n$, to be introduced into the channel, we randomly select a node N from the existing list L of free nodes and retrieve its information part to determine the initial terminal position of net i in a column in C . This linear search on the linked linear list L (of length $2n$) takes time $O(n)$. In the same iteration, we also determine the final terminal position of net i in a similar way, and that also takes time $O(n)$. So for each net i , $1 \leq i \leq n$, $O(n)$ time is required for finding out its terminal positions (in two different columns) in generating C . Hence for a channel of n nets to be introduced in C , the total time required is $O(n^2)$. Thus the time complexity in generating C is $O(n^2)$. Needless to mention that in generating C , we have used $O(n)$ dynamic space of computer memory. We summarize the complexity results in the following theorem.

Theorem 1. *Algorithm Simple_Random_Channel_Generator randomly generates simple channel specifications of only two-terminal nets without introducing any vertical constraint in it. The algorithm takes time $O(n^2)$, and space $O(n)$, where n is the number of nets introduced into the channel.*

III. GENERATION OF GENERAL CHANNEL SPECIFICATIONS

In the previous section we have described how to generate random channel instances, containing only two-terminal nets and there are no vertical constraints between any pair of nets. In general, (benchmark) channel specifications that are found in literature may have vertical constraints and nets may contain more than two terminals, i.e., there may be one or more than one terminals in between two end terminals of a net. So, the above-described procedure is applicable only on some special type of channel instances that are known as *simple channel specifications*. However, we can tune the procedure described above a bit, so that it becomes capable of generating channel instances having vertical constraints and nets may contain more than two terminals. In generating such general channel specifications, we

assume that there is an upper bound on the number of terminals and this number to be no more than six for a very few nets that are introduced into the channel being constructed.

Suppose we like to generate a random channel instance that consists of N nets ($N > 0$). So, the only input to be given here is the number of nets, i.e., N . Each net is assumed to be of k terminals ($2 \leq k \leq 6$). We need to place these N nets along the length of the channel. Also, there may be some non-terminals spread over the channel that should be incorporated in the specification. In order to generate a random general channel specification, first we have to estimate the length of the channel (L) in a way such that it is long enough to accommodate proposed terminals of all the nets as well as non-terminals. If k_i is the number of terminals of net i , then total number of terminals for all the nets can be computed by adding the number of terminals for each net. For example, if there are 10 nets in the channel then the total number of terminals for 10 nets is $K = k_1 + k_2 + \dots + k_{10}$. If NT is the number of non-terminals then total number of terminals is $P = K + NT$. These P terminals are evenly distributed over top and bottom positions of the channel specification along the length of the channel, in a random fashion.

Now we can calculate the estimated length of the channel as $L = \lceil P/2 \rceil$. As an illustration, say, there are 5 nets having 2 terminals each, 3 nets having 3 terminals each, and there are a 5-terminal and a 6-terminal net of a general channel specification to be generated of 10 nets. Also, say, number of non-terminals (NT) is 7. So the total number of terminals is calculated as follows:

$$P = K + NT = (2 \times 5 + 3 \times 3 + 5 \times 1 + 6 \times 1) + 7 = 30 + 7 = 37.$$

Hence, $L = \lceil P/2 \rceil = 19$.

So, the channel of length L is now capable of accommodating K terminals and NT non-terminals. In general, there are more nets having fewer terminals, and the random distribution of terminals should follow this constraint. Now, we have to locate the terminals of the nets at top / bottom positions in different columns in a random fashion. In order to place them properly, we have used three auxiliary lists, described as follows.

1. **NET_LIST**: The length of this list is also N . Initially the list is filled in with positive integers from 1 to N . An instance of the list is as follows.

NET_LIST: 1 2 3 4 5 6 7 8 9 10

This list tells that the nets to be introduced into the random general channel specification that is to be generated.

2. *TERM_COUNT*: The length of this list is N , i.e., same as the number of nets. Each location of this list is used to store the number of terminals (k) for a net. For example, if $N = 10$, then *TERM_COUNT* may be looked like as:

TERM_COUNT: 2 2 2 2 2 3 3 3 5 6

This means that, there are five nets of two terminals each, three nets of three terminals each and two nets having five and six terminals only.

3. *TERM_LIST*: The length of this list is $2L$, i.e., same as the sum of total number of top plus total number of bottom terminal positions. Initially the list is populated by positive integers started from 1 up to $2L$. For the above example, as $2L = 38$, so *TERM_LIST* must have positive integers from 1 to 38, as shown below.

TERM_LIST: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38

Now, it is the time to describe how channel specification C is generated with the help of these three auxiliary lists. At the very beginning of generating a desired random general channel specification, we initialize all the $2L$ top and bottom terminal positions of the channel by zeros. Then we start with generating the channel specification in N successive iterations. In each iteration, we select a net number (i) randomly and locate its terminals in different columns (c) in C , and again randomly determine the terminal position, either on the top or at the bottom at column c for net i . First we randomly select an element p from *NET_LIST* and remove the element from the list immediately; the length of *NET_LIST* is reduced accordingly. So, p is the net, whose terminals are to be assigned in different columns in C . Now, an element k from *TERM_COUNT* is selected randomly and the entry is deleted, and the size of this list is also reduced as before. So, p is considered as a k -terminal net. Say, we have selected the sixth entry from

NET_LIST and the ninth entry from *TERM_COUNT*; so $p = 6$ and $k = 5$, i.e., p is a 5-terminal net, whose terminal positions are to be determined now.

For each of these k terminals we generate a random number r and select the r -th element (which we denote by another number s) from *TERM_LIST* and delete the element as before, which results in a reduction of the length of *TERM_LIST*. The number s is now located at column $c = \lceil s/2 \rceil$; if s is odd (even), p (the net under consideration) is then placed at top (bottom) position at column c . When all the terminals of net p are assigned in different columns, we are left with *TERM_LIST* of length reduced by k , i.e., the length of *TERM_LIST* is same as $2L - k$, at the beginning of the second net / iteration under consideration.

In this way we extract elements from the three lists and using them we fill in the top and bottom positions of column c , and after N iterations we obtain a desired channel specification. Note that based on the initialization of the channel specification of length L (with all terminal positions containing only zeros), ultimately $2L - K$ terminal positions are left unassigned with some net terminals; so these terminal positions contain only non-terminals for the generated channel specification. One such randomly generated general channel instance is shown in Figure 3(a). The vertical constraint graph of this channel is shown in Figure 3(b). Vertical constraints among a set of nets determine a relative ordering over the nets along the height of the channel [7, 10]. The *vertical constraint graph* consists of n vertices for n nets present in a channel. If a column of the channel contains terminals of two different nets, say a and b , on the top and at the bottom, respectively, then in the vertical constraint graph, a directed edge is introduced from vertex a to vertex b .

Table 2 shows how the nets in the generated channel are spanned over its length, as the net number increases. Furthermore, terminal count, i.e., the number of terminals per net is also shown in the table. Data shows the randomness of the generated channel instance.

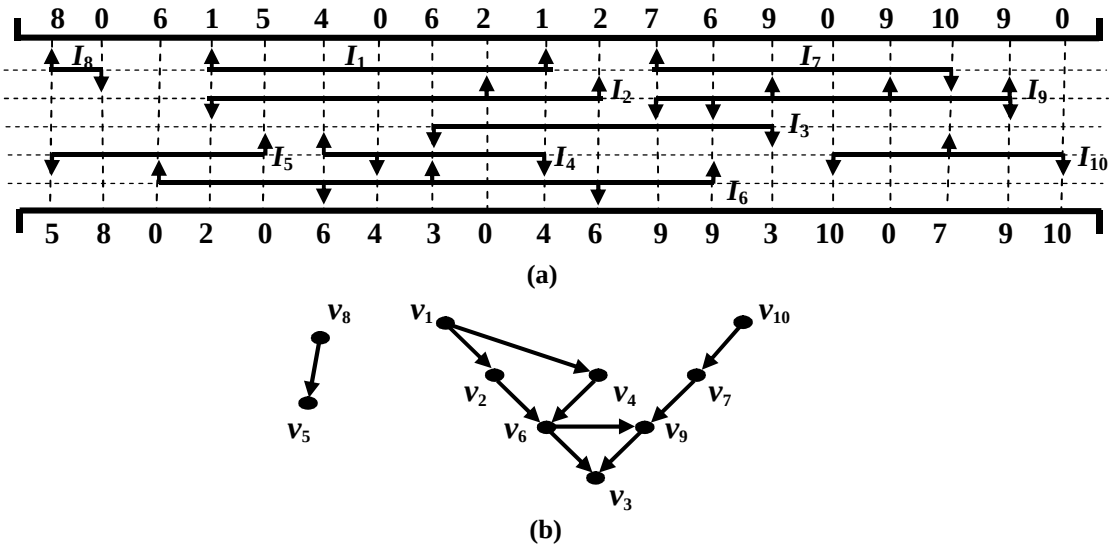


Figure 3. (a) The generated general channel specification in this section. The channel contains 10 nets, and its length is 19. (b) The vertical constraint graph of this generated general channel specification. Incidentally, it consists of two components, and does not contain any cyclic vertical constraint.

Net Number	Leftmost Column Position	Rightmost Column Position	Span (or Interval)	Terminal Count
1	4	10	6	2
2	4	11	7	3
3	8	14	6	2
4	6	10	4	3
5	1	5	4	2
6	3	13	10	5
7	12	17	5	2
8	1	2	1	2
9	12	18	6	6
10	15	19	4	3

Table 2. Spans (or intervals) of different nets introduced into the randomly generated general channel specification, shown in Figure 3(a). The number of terminals per net is also shown in this table in the rightmost column of Terminal Count.

We now analyze the computational complexities of algorithm *General_Random_Channel_Generator*. Suppose we have introduced N number of nets in total; so, N iterations are required. In each iteration, we need to search some random locations from the three auxiliary lists. Length of each of the lists *TERM_COUNT* and *NET_LIST* is same as the number of nets, i.e., N , and also the length ($2L$) of *TERM_LIST* is linearly dependent on N , i.e., L is $O(N)$. If we represent all these lists by singly linear linked lists, then the sequential search time for each list is dependent on their length. As the length of each of these (three) lists is $O(N)$, and as they are searched in a sequential manner, the total time to search the lists is $O(N)+O(N)+O(N)$, i.e., $O(N)$. So, for N iterations, as we like to introduce N nets into the channel, the total time complexity is $O(N^2)$. Also, space complexity is dependent on the length of each of all these lists, which is $O(N)$. So, space requirement is bounded by $O(N)$. We summarize the complexity results in the following theorem.

Theorem 2. *Algorithm General_Random_Channel_Generator randomly generates general channel specifications of multi-terminal nets that contain both horizontal as well as vertical constraints in it. The algorithm takes time $O(n^2)$, and space $O(n)$, where n is the number of nets introduced into the channel.*

IV. CONCLUSION

In this paper we have developed algorithms for generating random channel specifications of channel routing problem in VLSI design. We have developed two algorithms, one for generating random simple channel specifications and the other for generating random general channel specifications. Simple channel specifications contain only horizontal constraints but general channel specifications contain both horizontal as well as vertical constraints. General channel specifications may also contain cyclic vertical constraints. There are routing models, such as reserved two-layer no-dogleg Manhattan channel routing model [2, 7, 10], in which there is no feasible routing solution for general channel instances with cyclic vertical constraints. In that case an additional task of minor modification of the generated general channel specification is required in order to resolve vertical constraint violations.

In any case the algorithms developed in this paper are extremely useful when we require a large number of channel instances. As most of the problems in two-, three-, and multi-layer channel routing are computationally hard to solve [3, 4, 7, 9], for each of these problems it is unlikely to design a polynomial time deterministic algorithm. Rather, developing heuristic algorithms might be a probable way out that hopefully provides desired solutions for most of the instances occur in practice.

REFERENCES

- [1] A. Hashimoto and J. Stevens, "Wire Routing by Optimization Channel Assignment within Large Apertures," *Proc. of 8th ACM Design Automation Workshop*, pp. 155-169, 1971.
- [2] T.-T. Ho, S. S. Iyengar and S.-Q. Zheng, "A General Greedy Channel Routing Algorithm," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 10, pp. 204-211, 1991.
- [3] A. S. LaPaugh, *Algorithms for Integrated Circuit Layout: An Analytic Approach*, Ph.D. thesis, Lab. for Computer Sc., MIT, Cambridge, 1980.
- [4] A. Pal, A. K. Datta, T. N. Mandal and R. K. Pal, "Hardness of Crosstalk Minimization in Two-Layer Channel Routing," *Manuscript*, 2008.
- [5] A. Pal, D. Kundu, A. K. Datta, T. N. Mandal and R. K. Pal, "Algorithms for Reducing Crosstalk in Two-Layer Channel Routing," *Journal of Physical Sciences*, Vol. 10, pp. 167-177, Dec. 2006.
- [6] A. Pal, D. Kundu, A. K. Datta, T. N. Mandal and R. K. Pal, "Algorithms for High Performance Two-Layer Channel Routing," *Proc. of 22nd IEEE TENCON*, CD: Session: WeSC-O1.4 (Four pages), Oct. 30 – Nov. 02, 2007.
- [7] R. K. Pal, *Multi-Layer Channel Routing: Complexity and Algorithms*, Narosa Publishing House, New Delhi, 2000. (CRC Press, USA, 2000 and Alpha Sc. Int. Ltd., UK, 2000.)
- [8] R. K. Pal, A. K. Datta, S. P. Pal and A. Pal, "Resolving Horizontal Constraints and Minimizing Net Wire Length for Multi-Layer Channel Routing," *Proc. of 8th IEEE TENCON*, vol. 1, pp. 569-573, 1993.
- [9] G. A. Schaper, *Multi-Layer Channel Routing*, Ph.D. Thesis, Dept. of Computer Sc., Univ. of Central Florida, Orlando, 1989.
- [10] T. Yoshimura and E. S. Kuh, "Efficient Algorithms for Channel Routing," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 1, pp. 25-35, 1982.

Efficient Generation of Combinatorial Families

Md. Shariful Islam Bhuyan and Md. Saidur Rahman

Department of Computer Science and Engineering,
Bangladesh University of Engineering and Technology,
Dhaka, Bangladesh

Abstract—In this paper, we propose a unifying framework for efficient generation of combinatorial objects by giving recursive definition of an abstract combinatorial class. The definition can be instantiated to an array of specific combinatorial classes by specifying the framework parameters appropriately. Our framework defines a rooted spanning tree over a representative graph of that abstract combinatorial class where exhaustive generation can be done by a depth-first traversal. As an illustration, we show an instantiation for the combinatorial class of combinations as well as give a novel constant time generation algorithm for them.

Index Terms—Combinatorial Generation, Unifying Framework, Combinatorial Family, Combination, Combinatorial Algorithm

I. INTRODUCTION

The problem of generating all objects of a combinatorial class, each exactly once, so that successive objects differ in a pre-specified and usually small way, can be formulated as finding a Hamiltonian path/cycle problem in a graph; the vertices of the graph are the objects themselves. Two vertices are joined by an edge if they differ from each other in a pre-specified way. This graph is a representative graph of the corresponding combinatorial class (see Figure 1). A Hamiltonian path/cycle in this graph is called a combinatorial Gray code in the literature [1].

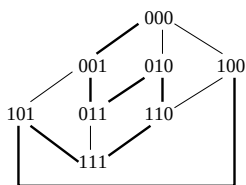


Fig. 1. A Hamiltonian path (indicated by the thick lines) over a representative graph of 3-digit binary number starting from 000.

Another approach that has been recently introduced in the combinatorial generation arena is genealogical tree or family tree approach. In this approach, the generation problem is formulated as finding a rooted spanning tree over a representative graph of the corresponding combinatorial class (see Figure 2). Every node of the tree, internal or external, is a combinatorial object and there is a unique child to parent relationship among the nodes. There is also a unique root object which has no parent. By traversing

this tree in a depth-first order, we can have a generation algorithm.

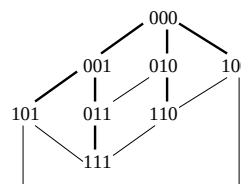


Fig. 2. A spanning tree (indicated by the thick lines) over a representative graph of 3-digit binary number starting from 000.

If we examine the two approaches mentioned above, we find that the goal of both approaches is to traverse every vertex of the representative graph of a particular combinatorial class exactly once. The goal is achieved by traversing a Hamiltonian path in the former and by traversing a rooted spanning tree in the later. The problem of determining whether a given graph has a Hamiltonian path or cycle is NP-complete. Moreover, it is not even guaranteed that a particular representative graph will have one. So this approach is not expected to work always. However, the best known algorithms for combinatorial generation are based on combinatorial Gray codes. On the contrary, intuitively speaking, finding a spanning tree is much easier.

In this paper, we propose a unifying framework for combinatorial generation by giving a recursive definition of an abstract combinatorial class. There are some framework parameters which are left unspecified in the definition. We can instantiate the recursive definition to define a concrete combinatorial class by specifying the framework parameters appropriately. Our unifying framework does not define any specific combinatorial class rather provides enough flexibility to define a broad range of combinatorial classes through framework parameters including n -tuple, combination as well as its variants, set partition, integer partition, binary tree and many more. Since a variety of combinatorial classes are unified under a common abstract recursive definition, we use the term 'Unifying'.

Previous attempts for combinatorial generation with minimal change using both the Gray code [2], [8] and genealogical tree approaches [5], [6], [7] are based on the intrinsic nature of each problem. This hardens the universal

adoption of a generation technique. Takaoka [4] formalizes the idea of combinatorial Gray code by a twisted lexicographic tree, and gives an iterative algorithm for efficient generation of certain combinatorial classes. In contrast, we formalize and unify the idea of genealogical tree where the formalization gives simpler constant-time generation scheme for larger classes of combinatorial family. Our idea is somewhat similar to [9]. However, we reduce the abstract reverse search procedure to instantiation of a formal framework with only two parameters depending on the specific combinatorial class.

Our framework recursively defines a rooted spanning tree over an abstract combinatorial class where the nodes of the tree are distinct objects of that class. By traversing that rooted spanning tree in a depth-first order we can visit every object. The point to focus is that we want to make a transition between two adjacent nodes in constant amount of time on average so that we can have an efficient generation scheme. We specify the parent to child relationship and child to parent relationship among the members of that class. We also present a generic algorithm for generation of every member. As an illustration of instantiation, we provide a specific example. We define the combination generation problem, instantiate our framework to the combinatorial class of combinations and prove the correctness of instantiation as well as present the specific algorithm.

The paper is organized as follows. In section 2 we present some definitions and concepts required for the rest of the paper. In section 3 we give the formal definition of our unifying framework. In section 4 we present an instantiation of our framework for the combinatorial class of combinations as well as give a novel constant time generation algorithm for them. In the conclusion we explore some future directions.

II. PRELIMINARIES

In this section, we define some terms used in this paper and follow the conventions in [3].

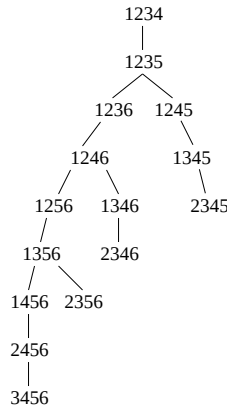


Fig. 3. A genealogical tree for combination of 4-element taken from a set of 6-element.

In this paper, we represent a *combinatorial object* of a combinatorial class $\mathcal{P}$ by a string $a_1 a_2 \dots a_n$ of length n such that each a_i is a value which is assigned to the i th position and taken from Σ as well as satisfies the constraint set $\mathcal{C}$ of $\mathcal{P}$. We denote the value of i th position using $v(i)$.

In the representation of a specific combinatorial object, if the value of i th position or $v(i)$ cannot be decremented due to the constraint set $\mathcal{C}$, then we call $v(i)$ the *minimal value* of the i th position of that combinatorial object. For a specific position i , this minimal value can vary from object to object. We call the minimum of all possible minimal values of the i th position as the *minimum value* of the i th position or i_{min} . In Figure 3, the minimal value of the 2nd position is 2 in 1456 but not in 2456. In 2456, minimal value of the 2nd position is 3. We cannot decrement the value of the 2nd position less than 3 with all other positions unchanged since it would violate the constraint set. However, the minimum value of the 2nd position is 2 since there are combinations that contain 2 in the 2nd position and no other combination has a value less than 2 in the 2nd position. We call a position a *minimum position* if the position contains the minimum value for that position. Otherwise we call the position a *non-minimum position*. In 1235, 1st, 2nd and 3rd positions are minimum positions and 4th position is a non-minimum position. We also define the concept of *maximal value* of the i th position, *maximum value* of the i th position or i_{max} , *maximum position* and *non-maximum position* in a similar way.

Next we give some tree-specific definitions. In the rooted spanning tree defined by our recursive definition, there is a unique root object T_r and every node $\mathcal{T}$ except the root, both internal and external, has a parent. In Figure 3, $T_r = 1234$. We denote the parent of a node $\mathcal{T}$ as $P(\mathcal{T})$ throughout the paper. Our idea of parent to child relationship is that the parent object and child object differ only in a single position r and by a single increment. The r th position is called the *change position* of $\mathcal{T}$ and a *generating position* of $P(\mathcal{T})$. According to an alternative notation, if $\mathcal{T}$ is a parent object, then $\mathcal{T}(k)$ is one of its children where k th position is a generating position in $\mathcal{T}$ and the change position in $\mathcal{T}(k)$. Note that, an object can have multiple generating positions but only a single change position. In a non-leaf object $\mathcal{T}$, the rightmost generating position is called the *start position* i and the leftmost generating position is called the *end position* j . In our framework, we assume that the generating positions are consecutive in an object. In Figure 3, 1235 is the parent of 1236. The 4th position is a generating position of 1235 as well as the change position of 1236. 1246 has a single change position namely the 3rd position but it has two generating position. The start position of 1246 is 3 and end position is 2.

III. FRAMEWORK DESCRIPTION

In this section, we give the recursive definition of our abstract combinatorial class.

A. Recursive Definition

Let $\mathcal{P}$ be the set of all objects of a combinatorial class. Let Σ be the alphabet and $\mathcal{C}$ be the constraint set defined for $\mathcal{P}$. Let $\mathcal{S}$ be the set of every object $\mathcal{T}$ that can be generated by the recursive definition given below. Let $\mathcal{T} \in \mathcal{S}$ be an object where $\mathcal{T} = a_1 a_2 \dots a_k \dots a_n$. Let $\mathcal{T}(k)$ be another object where $\mathcal{T}(k) = a_1 a_2 \dots a_k + 1 \dots a_n$. Alternatively, we can say that $\mathcal{T}(k)$ can be found by incrementing the k th position of $\mathcal{T}$ by one. We call $\mathcal{T}(k)$ the *child* of $\mathcal{T}$ and $\mathcal{T}$ the *parent* of $\mathcal{T}(k)$ if and only if $\mathcal{T}(k) \in \mathcal{S}$. We now give the formal definition of our unifying framework in definition 1.

Definition 1: The recursive definition has two steps.

Basis Step. Let $T_r \in \mathcal{S}$ be a unique object of length n , where $T_r \in \mathcal{P}$ and every position of T_r is set to the minimum. Then T_r is the unique root object.

Recursive Step. Let $\mathcal{T} \in \mathcal{S}$ be an object where $\mathcal{T} \in \mathcal{P}$. Let there be a start position i and an end position j in $\mathcal{T}$ such that $\mathcal{T}(k) \in \mathcal{P}$ for every k , where $j \leq k \leq i$. Then $\mathcal{T}(k) \in \mathcal{S}$. If there is no such pair (i, j) , then $\mathcal{T}$ is a leaf object and has no child.

Note that our recursive definition generates every object $\mathcal{T} \in \mathcal{P}$ where $\mathcal{P} = \mathcal{S}$. Specification of the root object T_r , or alternatively, specification of the *minimum value* for every position is the first parameter of our recursive definition. The choice of the root object T_r is problem specific. The start position i is specified by the parent to child relationship given below. Specification of the end position j is the second parameter of our recursive definition. The choice of the end position j is also problem specific.

B. Parent to Child Relationship

In this subsection, we define the parent to child relationship among the members of a combinatorial class that can be instantiated from our unifying framework. This relationship also serves as the child generation rule for a given object $\mathcal{T} \in \mathcal{S}$.

Definition 2: If $\mathcal{T} \in \mathcal{S}$, then $\mathcal{T}$ must have a *change position* r which is incremented in $P(\mathcal{T})$ to get $\mathcal{T}$. If $v(r)$ is not a maximal value then $i = r$. If $v(r)$ is a maximal value and r is not the leftmost position then $i = r - 1$. Otherwise, the position $r - 1$ does not exist and no start position i can be found. In that case $\mathcal{T}$ has no child. If $\mathcal{T} = T_r$ then $i = r$ where r is the rightmost position of $\mathcal{T}$.

The end position j depends on the constraint set $\mathcal{C}$ and is problem specific. By incrementing every k th position of $\mathcal{T}$ we can generate its children where $j \leq k \leq i$.

C. Child to Parent Relationship

In this subsection, we define the child to parent relationship among the objects $\mathcal{T} \in \mathcal{S}$ so that every non-root object $\mathcal{T}$ must have a parent and only one parent. In this way, each syntactically identical object is generated exactly once namely from its own parent. We need to define the constraint set $\mathcal{C}$ in such a way so that duplication never occurs. We ensure that the parent $P(\mathcal{T}) \in \mathcal{S}$ of an object $\mathcal{T} \in \mathcal{S}$ can be found by decrementing the change position of $\mathcal{T}$ so that the child to parent relationship remains the reverse of parent to child relationship. Our idea is, the parent $P(\mathcal{T})$ of an object $\mathcal{T} \in \mathcal{S}$ must be identical to $\mathcal{T}$ at every position except in its change position. We give our child to parent relationship by using the following two lemmas.

Lemma 1: For every non-root object $\mathcal{T} \in \mathcal{S}$, every position to the left of the change position r is a minimum position.

Proof: The proof is immediate from the definition 1. ■

Lemma 2: For every non-root object $\mathcal{T} \in \mathcal{S}$, the unique parent $P(\mathcal{T}) \in \mathcal{S}$ can be found by decrementing the leftmost non-minimum position.

Proof: Can be proved using Lemma 1. Detailed proof is omitted in this paper. ■

The Lemma 2 gives the definition of $P(\mathcal{T})$ and ensures that we can always find a unique parent of a non-root object $\mathcal{T}$. Next lemma ensures that there can be no duplication of objects in the tree generated by the recursive definition.

Lemma 3: For every non-root object $\mathcal{T} \in \mathcal{S}$, there is a unique sequence of decrement(s) in the leftmost non-minimum position(s) which transforms $\mathcal{T}$ into the root object T_r .

Proof: Can be proved using Lemma 2. Detailed proof is omitted in this paper. ■

Since there is a unique sequence of decrement(s) that transforms an object $\mathcal{T} \in \mathcal{S}$ into the root T_r , by reversing the decrement(s) we can generate that particular object $\mathcal{T}$, starting at the root. The path is unique, thus it will lead to a unique object.

D. Algorithm

In this subsection we give a straightforward implementation of an algorithm based on our parent to child relationship. Our algorithm generates every object in constant time on average if the determination of the end position j can be done in constant time. Thus we can have a *Constant Amortized Time* or *CAT* algorithm by carefully specifying the rule to determine end position. The algorithm is basically a depth first traversal over the rooted spanning tree defined by the recursive definition. We initialize the root object T_r and then traverse every object in a top down fashion. Here $\mathcal{T}$ is the current object

and r is the change position of $\mathcal{T}$.

Procedure Find-All-Child-Objects($\mathcal{T}, r$)
begin
 output $\mathcal{T}$;
 if ($v(r)$ is not a maximal value) **then** $i = r$;
 elseif ($r \neq 1$) **then** $i = r - 1$;
 else $\mathcal{T}$ has no generating position **return**;
 endif
 $j = \text{getEndPosition}()$; $k = i$;
 repeat
 Find-All-Child-Objects($\mathcal{T}(k), k$); $k = k - 1$;
 until ($k = j - 1$)
end;

Algorithm Find-All-Objects()
begin
 $\mathcal{T} = T_r$; $i = \text{rightmost minimum position } r$;
 Find-All-Child-Objects(T_r, i);
end;

E. Complexity Analysis

Before stating the space and time complexity we introduce some notation. Let n be the length of each object. Let $\mathcal{N}$ be the total number of objects in the combinatorial class $\mathcal{P}$. Let d_k be the difference between the maximum value and minimum value of the k th position where $1 \leq k \leq n$. Alternatively $d_k = k_{\max} - k_{\min}$. Let $d_{\max}$ be the maximum possible d_k where $1 \leq k \leq n$. We assume that the time complexity of the procedure getEndPosition is $O(\mathcal{F})$ where $\mathcal{F}$ is a placeholder for a complexity term.

Theorem 1: Algorithm **Find-All-Objects** uses $O(n * d)$ space and runs in $O(\mathcal{N} * \mathcal{F})$ time.

Proof: We give proof of space and time complexity separately.

To determine the worst case space complexity, we consider a special case. We set k th position of the root T_r to $k_{\min}$ for every k where $1 \leq k \leq n$. Let $\mathcal{T}_{\max}$ be an object with every position set to $k_{\max}$ where $1 \leq k \leq n$. In the worst case the value of $d_k = d_{\max}$ for every k where $1 \leq k \leq n$. So the path from T_r to $\mathcal{T}_{\max}$ contain a total of $n * d_{\max}$ increment and this path is the longest of all. As we traverse this tree, we must store the trace of the current path until we have reached to the leaf. So at the worst case we have the space complexity $O(n * d_{\max})$.

To determine the time complexity, we note that every transition from parent to child when we visit a child for the first time requires a single increment. In the same way every transition from child to parent when we trace back requires a single decrement. Given the change position both the increment and decrement can be done in $O(1)$. From the assumption, the complexity of selecting the end position j is $O(\mathcal{F})$. There is no need to find the end position when we are tracing back. So the traversal of

every edge of the rooted spanning tree can be done in $O(\mathcal{F})$ for forward direction and in $O(1)$ for backward direction. From the property of tree we know that a tree of $\mathcal{N}$ nodes has total $\mathcal{N} - 1$ edges. So the traversal takes $O((\mathcal{N} - 1) * \mathcal{F} + (\mathcal{N} - 1))$ which is $O(\mathcal{N} * \mathcal{F})$. We can readily see that if the end position j can be selected in $O(1)$ on average then we can have a constant amortized time or *CAT* algorithm. ■

IV. GENERATION OF COMBINATIONS

In this section we illustrate our framework with an example of the combinatorial class of combination. A *combination* of n things, taken k at a time, often called simply a k -combination of n things, is a way to select a subset of size k from a given set of size n . Set of all combinations essentially is a set of all subsets of a set of n elements or n -set with cardinality k . We use the term k -subset and combination interchangeably.

A. Framework Instantiation for Combination Generation

Let $\mathcal{P}$ be the set of all combinations. We represent a combination as $a_1 a_2 \dots a_n$ where a_i is the value of the i th position drawn from alphabet Σ . Here $\Sigma = [m]$ and $n \leq m$. Here the constraint set C has two rules. First, we do not allow same value in different positions simultaneously. Second, to avoid duplication we take the lexicographically sorted member. Combination of these two rules implies that values of the positions are kept sorted or $a_1 < a_2 < \dots < a_n$. $v(i)$ can be incremented when $v(i) + 1 < v(i + 1)$. Let $\mathcal{S}$ be the set of every combination $\mathcal{T}$ that can be generated by the instantiated recursive definition. All other definitions are as is.

To instantiate our recursive definition we follow a three-step procedure. The steps are given below:

- 1) The unique lexicographically smallest combination $\mathcal{T} \in \mathcal{P}$ of length n is defined as the root T_r where $T_r = 12 \dots n$. Every i th position of T_r is set to $i_{\min}$. Here $i_{\min}$ is the minimum value of i th position where $1 \leq i \leq n$. In this case $i_{\min} = i$. All other combinations contain an equal or a greater value at every position than T_r . Maximum value of i th position or $i_{\max}$ is $i + m - n$ where $1 \leq i \leq n$.
- 2) For root combination T_r the start position $i = n$ and end position $j = n$. For non-root combination, if the change position is r then the end position $j = r - 1$ given that $r \neq 1$. If $r = 1$ and $v(r)$ is not a maximal value then $j = r$. Otherwise, this is a leaf and has no generating position.
- 3) To prove the correctness of our instantiation we show that our parameter specifications enforce the set equality $\mathcal{P} = \mathcal{S}$ where $\mathcal{S}$ is the set of every combination $\mathcal{T}$ that can be generated by our recursive definition.

In the next few lemmas we prove several aspects of our instantiation.

Lemma 4: For every non-root combination $\mathcal{T} \in \mathcal{P}$, there is a unique combination $\mathcal{T}' \in \mathcal{P}$ where $\mathcal{T}'$ can be found by decrementing the leftmost non-minimum position.

Proof: Let r be the leftmost non-minimum position of the combination $\mathcal{T} \in \mathcal{P}$. Since $v(r) - 1 \in [m]$, the value of every position of $\mathcal{T}'$ is drawn from $[m]$. Decrementing the r th position retains the relation $v(r-1) < v(r) < v(r+1)$ in $\mathcal{T}'$. So $\mathcal{T}'$ satisfies the constraint set $\mathcal{C}$ hence $\mathcal{T}'$ is also a combination or $\mathcal{T}' \in \mathcal{P}$. ■

We call $\mathcal{T}'$ the parent of $\mathcal{T}$ or $P(\mathcal{T})$ since Lemma 4 is a restatement of the child to parent relationship in our unifying framework.

Lemma 5: For every non-root combination $\mathcal{T} \in \mathcal{P}$, there is a unique sequence of decrement(s) in the leftmost non-minimum position(s) which transforms $\mathcal{T}$ into the root combination T_r .

Proof: Can be proved using Lemma 4. Detailed proof is omitted in this paper. ■

Lemma 6: If there is a unique sequence of decrement(s) in the leftmost non-minimum position of a non-root combination $\mathcal{T} \in \mathcal{P}$ which transforms $\mathcal{T}$ into the root T_r , then $\mathcal{T} \in \mathcal{S}$.

Proof: Can be proved using Lemma 4. Detailed proof is omitted in this paper. ■

Now we state the required theorem to show that the instantiated version of our framework indeed generates every combination and no more. For this we have to show that, the set $\mathcal{S}$ of all combinations that can be generated by the recursive framework and the set $\mathcal{P}$ of all possible combinations are equal.

Theorem 2: For every combination $\mathcal{T}$, $\mathcal{T} \in \mathcal{P}$ if and only if $\mathcal{T} \in \mathcal{S}$.

Proof: *Case $n = m$.* When $\mathcal{T} = 12 \dots n$ it is the root $T_r \in \mathcal{S}$ as well as $T_r \in \mathcal{P}$. T_r is the only combination in $\mathcal{P}$. On the other hand, T_r has no generating position and the only combination in $\mathcal{S}$. So $\mathcal{P} = \mathcal{S}$.

Case $n < m$. To prove the necessary condition let $\mathcal{T} \in \mathcal{P}$ be a combination where $\mathcal{T} \in \mathcal{S}$. If $\mathcal{T} = T_r$, then the start position $i = n$ and end position $j = n$. So only the n th position of $\mathcal{T}$ is a generating position. If $\mathcal{T}$ is a non-root combination with the change position r , then $j = r - 1$ given that $r \neq 1$. If $v(r) + 1 < v(r + 1)$, then $i = r$. If $v(r) + 1 = v(r + 1)$, then $i = r - 1$. If $r = 1$ and $v(r)$ is not a maximal value, then $i = r$ and $j = r$. If $r = 1$ and $v(r)$ is a maximal value, then $\mathcal{T}$ is a leaf object. In all cases $\mathcal{T}(k) \in \mathcal{P}$, because incrementing the generating position k ensures that the value of every position of $\mathcal{T}(k)$ is drawn from $[m]$ and the relation $v(k - 1) < v(k) < v(k + 1)$ as well as the constraint set $\mathcal{C}$ is satisfied for $j \leq k \leq i$. So, if a object $\mathcal{T}$ is generated in our instantiated framework, then it is a combination or $\mathcal{T} \in \mathcal{P}$.

To prove the sufficiency condition we see that if $\mathcal{T} \in \mathcal{P}$

is a non-root combination, then using the Lemmas 5 and 6 it is immediate that $\mathcal{T} \in \mathcal{S}$. ■

B. Algorithm

In this subsection we give a straight forward implementation of our instantiated parent to child relationship. Our algorithm generates every combination in constant time on average since the determination of the end position j can be done in constant time. Thus we have a *CAT* algorithm for combination generation. We initialize the root combination T_r and then traverse every combination in a depth-first fashion. Here $\mathcal{T}$ is the current combination and r is the change position of $\mathcal{T}$.

Procedure Find-All-Child-Combinations($\mathcal{T}, r$)

```

begin
  output  $\mathcal{T}$ ;
  if( $v(r) \neq r + m - n$ ) then  $i = r$ ;
  elseif( $r \neq 1$ ) then  $i = r - 1$ ;
  else  $\mathcal{T}$  has no generating position return;
endif
 $j = r - 1$ ;  $k = i$ ;
repeat
  Find-All-Child-Combinations( $\mathcal{T}(k), k$ );  $k = k - 1$ ;
until( $k = j - 1$ )
end;
```

Algorithm Find-All-Combinations()

```

begin
   $\mathcal{T} = T_r$ ;  $i = n$ ;
  Find-All-Child-Combinations( $\mathcal{T}_r, i$ );
end;
```

Our algorithm starts with procedure *Find-All-Combinations*. This procedure set the start position i to the rightmost minimum position n . Next, algorithm calls the procedure *Find-All-Child-Combinations*. It first outputs the current combination and then checks the corresponding change position r whether $v(r)$ is a maximal value or not. If there is a room for increment, then $i = r$. If there is no room for increment and $r \neq 1$, then i is set to $r - 1$. Otherwise, this is a leaf combination and the procedure returns. If this is a non-leaf combination, then the end position $j = r - 1$. This is a *CAT* algorithm since computation of j takes constant time. Then a loop runs and increments every position from i to j and calls the procedure *Find-All-Child-Combinations* recursively. The correctness of our algorithm follows from Lemmas 5 and 6 and the complexity can be achieved from Theorem 1.

V. CONCLUSION

In this paper, we present a unifying framework for combinatorial generation. We are also working on the instantiation of other combinatorial classes. A theoretical

advancement might be capturing the core algebraic properties of the combinatorial classes those conform to this framework. Another interesting direction is to develop a random sampling scheme by making the traversal a random tree walk.

REFERENCES

- [1] J. T. Joichi, Dennis E. White, and S. G. Williamson: Combinatorial Gray codes. *SIAM J. on Computing*. 9(1) (1980) 130-141
- [2] F. Ruskey: Simple combinatorial Gray codes constructed by reversing sublists. *Proc. of ISAAC'93, Lecture Notes in Computer Science*. 762 (1993) 201-208
- [3] F. Ruskey: *Combinatorial Generation*. Book in preparation. (2003)
- [4] T. Takaoka: $O(1)$ time algorithms for combinatorial generation by tree traversal. *Comput. J.* 42(5) (1999) 400-408
- [5] S. Nakano: Efficient generation of plane trees. *Information Processing Letters*. 84 (2002) 167-172
- [7] S. Nakano: Efficient generation of tri-connected plane triangulations. *Computational Geometry Theory and Applications*. 27(2) (2004) 109-122
- [6] K. Yamanaka, S. Kawano, Y. Kikuchi and S. Nakano: Constant time generation of integer partitions. *IEICE Trans. Fundamentals*. E90-A (2007) 888-895
- [8] C. Savage: A survey of combinatorial Gray codes. *SIAM Review*. 39 (1997) 605-629
- [9] D. Avis and K. Fukuda: Reverse search for enumeration. *Discrete Applied Mathematics*. 65 (1996) 21-46

City Planning and Development Using Geographic Information Systems

Mustafa A. Kamal

Department of Computer Information Systems, University of Central Missouri
Warrensburg, MO 64093, USA

Abstract - In recent years, GIS technology is used from managing natural disasters such as landslides to habitat mapping through remote sensing. GIS system has also been used to address land use issues. Many local governments are adding GIS technology as part of their regular administrative management system. High cost of GIS system, shortage of appropriate GIS specialists, lack of understanding and policies for homogeneous integration of GIS into planning are making the system, in many public sector cases, less than successful. We worked with a city government in the middle USA and participated in evaluating their system. Although small in size (population – thirty thousand), it's location next door to a large metropolitan city made things complicated. The city incorporated a GIS system into many aspects of public sector development, offered web-based services to its community and published maps with search tools. The GIS based services gave the city government a high-tech reputation and a need-based analysis of public sector projects.

The system gave the city officials access to real and accurate data which were not previously available. This often added more policy debates on sensitive and long range issues. It added a new dimension to city development and helped to find a combination of solutions in city wide constructions, drainage system, traffic control, and other engineering applications. Overall, the GIS system improved the efficiency of the city government and their administration processes as well as improved quality of decisions made by the city planners and managers.

Index terms – GIS, planning, remote sensing, GPS

DEFINITION AND SCOPE OF GIS

A GIS (Geographic Information System) package is an organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display geographically referenced information [1-4]. The GIS database management system combines map layers and process data. Its capability of three-dimension interface can create a visual landscape or contour for infrastructure and engineering projects. For example, a GIS system can locate a traffic incident spot in a second, and call out the emergency

connections from the related database; the system can also create a 3D model of a park without physically building it. The majority of cases of desktop GIS software use are non-technical in nature. By using statistical analysis and GIS tools, practitioners can easily identify the target and build a what-if scenario. One of the most attractive point of GIS modeling is that it can quickly show the real-world results of a complex statistical manipulation [3]. However, at a local government level GIS uses is very limited to none.

CHALLENGES IN IMPLEMENTING GIS

GIS technology is not cheap. Cost come from hardware, software, infrastructure, people and training. Data acquisition and digitization also add to the cost. Integration of GIS into existing technologies, such as different operation systems, workstations, networking equipments, wireless networks, DBMS systems, management systems, are major issues. Many local governments lack funding for such implementation. With the complexity and limitation of the GIS and GPS collection, managers need to understand how to organize a GIS team. A GIS project needs GPS collectors and many specialists. For example, network specialists needed to transmit GIS data through the existing network without creating traffic overload; database specialists needed to merge and organize GIS data into database system. GIS specialist needed to process aerial photography and data digitization, store and retrieval. Website specialist may design the Internet Management System and GIS Internet tools to offer the service to the end users.

There can be impact on public policy issues too. Accurate and real time data may remove ambiguity from many public policy decisions. For example, GIS data may suggest that a planned commercial area may destroy the nature environment in that area or reduce the total area of parks in the city. Moreover, a GIS can also be a useful tool to find more ways to solve a conflict or existing problems. For example, a GIS-produced map is often used to explain zoning issues to the city council; a GIS map overlaid with different statistic data is used to solve city-wide policy issues

and conflicts. In Summit, parcel- specific maps are available on the internet for public uses and gives a common understanding of current patterns of development.

MAPPING DATA FOR POLICY MAKERS

Here are the general steps for a GIS manager to map data into a GIS project.

1. Process

- Manager needs to analyze the policy questions to identify the important features, and create different scenarios to match the basic requirement within the policy environment.
- Select visualization techniques, such as 3-dimensional GIS tools or a handy GIS software, which can appropriately emphasize the features that has been selected.
- Integrate different layers of the data and other required information, including statistical result, into a small number of displays.

2. Categorize the data

The manager decides what and how many attributes should be included in the data. For example, a street layer need parcel numbers, the name of property owners, name of the street, and

identification number etc. The manager needs to consider the future use of the data and its potential users.

3. Equal intervals

After the basic information has been built in the categories, the manager can combine it with statistical data and start to analyze the overall view of the questions. To clearly distinguish results, the data should be divided into several equal intervals and marked in different color or symbols. GIS will help the manager to create these tools and export the data to make a quick report.

4. Avoiding Visual Misperceptions

When analyzing the geographic data, visual misperception should be avoided. Misrepresentation or bias may cause mistakes in categorizing the data in analysis. For example, we may categorize houses around a lake as high income area which may not necessary be true for all units.

5. Analyze the policy questions

The analysis should involve a population large enough to allow for statistical analysis to be meaningful. Some other variables may also affect the result of analysis.

GIS Project – from planning to maintenance

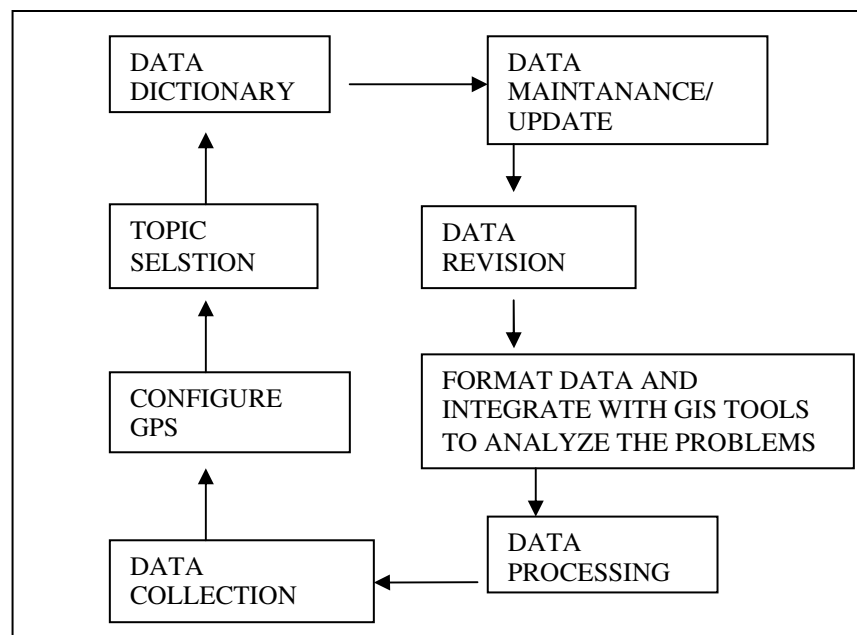
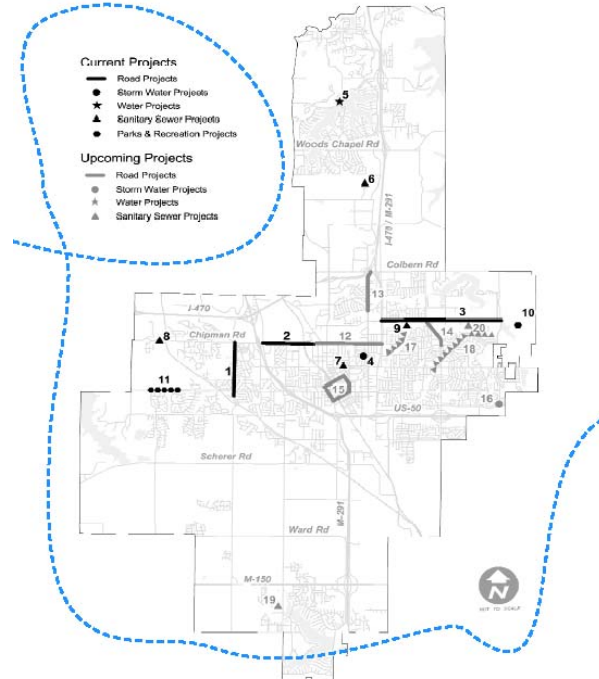


Figure 1: GIS project life cycle

CASE STUDY - GIS APPLICATION IN SUMMIT CITY

Summit City routinely applies their GIS system to evaluate current and proposed activities. For example, prior to citywide events GIS system is consulted to study impact on resources such as police, fire, evacuation plan, available housing, demand on water/sewer and on public health, recycling etc are looked at. Summit City is continuously developing and updating its GIS environment and compatible database. On a regular basis, GIS applications in Summit also include review of land use, impact on tax base, water and sewer infrastructure, availability of storm and sanitary sewer and drains, etc. The local government publishes upcoming project description and estimate completion date, performance expectations and availability to public. Here is an example of upcoming projects published by Summit City. This digitized map combined several different layers from different projects.



Project Description		Estimated Completion Date
CURRENT PROJECTS	1 Pryor Rd, 4th to Chipman	November 2003
	2 Chipman Rd, Phase I, Blue Parkway to Olive	December 2003
	3 Tudor/Scruggs, 291 hwy to Blackwell	May 2004
	4 Independence/Orchard Stormwater Improvements	July 2003
	5 Bowlin Road Pump Station Renovation	September 2003
	6 Maybrook Manhole Rehabilitation - Sanitary Sewer	November 2003
	7 Eastridge Sanitary Sewer Replacement	September 2003
	8 Edgewood Trail Sanitary Sewer Replacement	August 2003
	9 SA/5B Modifications - Sanitary Sewer	August 2003
	10 Legacy Community Center	Fall 2003
	11 Longview Loop - Greenway	Fall 2003
Project Description		Estimated Start Date
UPCOMING PROJECTS	12 Chipman Road, Phase II, Olive to M-291 Hwy	Spring 2004
	13 Independence Avenue, Deerbrook to Colbern	Winter 2004
	14 Todd George, Tudor to McKee	Winter 2004
	15 Downtown Improvements - Road, Water, Storm and Sanitary Sewer	Summer 2004
	16 Blackwell and Shenandoah - Stormwater	Spring 2004
	17 West Prairie Lee Sanitary Sewer Improvements, Phase I - Stormwater	Fall 2004
	18 South Prairie Lee and Bristol Sanitary Sewer	Winter 2003
	19 Middle Big Creek Excess Flow Basin - Sanitary	Fall/Winter 2004
	20 Scruggs Road Pump Station Improvements - Sanitary	Summer/Fall 2003

Figure2: Projects at Summit City

The city government offers an active on-line city map with simplified GIS tools to search the data for personal use. The GIS department updates their services on the website and its intranet. Due to ongoing construction of roads and utilities and construction of housing and new business the GIS quickly react to the changing events and its environments. The system allows staffs at different geographical location to work, store and retrieve data from the same database on real time. The system maintains its' communication with federal, state and other local governments by transmitting digitized map data through the networks. A summary of the process of land use application and its data flow is showing in Figure 2. Applicants submit paper plan to the planning department which will check land use plan in the current and in the future as well as the environment of that land from the database and give suggestions to the applicants. After the plan is approved and receives all required signatures, the draft plan is sent to the County office for its parcel number. The final draft comes back to the city government and is added into the database. The detail information of the land can be checked on County's Web site, including the records of tax payment, name of the land lord, the location, the size and the value of that land, etc. Customers can compare value of the property, transportation layout, school district, etc.

Before a piece of land is approved for public facility, such as schools, hospitals, or shopping mall, the city evaluates whether related constructions or improvements around that land is necessary. These may include building of new roads and/or upgrade existing roads, add traffic lights and add signs. The Summit City reviews impact on school boundaries from such constructions. GIS facilitates the process by redrawing boundaries in different ways to best fit the need. School Committee members of the city select which subdivisions they want to include within the new boundary and the GIS tags them with the school's code. The GIS knows the number of elementary students within each subdivision, so it can total the number of students within the proposed boundary. They can change the boundary by changing which subdivision they select, and GIS will automatically show the results of different combinations. The same method can also be used to analyze and create legislative districts.

GIS EVALUATION: HEALTH SYSTEMS HOSPITAL IN SUMMIT CITY

Health Systems Hospitals, consisting of eight area hospitals and 14 primary care practices provide a range of primary, acute, tertiary and chronic care services. The System proposed to build a 52-bed, \$57 million hospital in the Summit City. The health system is an organization in which the physicians and hospitals, functioning as an integrated unit, assume responsibility for the delivery of comprehensive, cost-effective, quality health care for people in the metropolitan area and the surrounding region.

The state is responsible for oversight of the alarming increase in healthcare costs and control of the healthcare constructions. The Health Facilities Review Committee is responsible for issuing a certificate-of-need (CON). The program is designed to contain health care costs by stopping unnecessary spending on facilities and high-dollar equipment. There is an inherent problem in CON denial that it may hinder competitions. The Summit City was responsible to conduct its' analysis of the proposal and submit their findings to the state CON committee. The Summit City extensively used its GIS system to come up with layers of relevant data and visual images and completed its' analysis. They looked at such issues as locations of nearly hospitals, layers of emergency route, street name, current traffic on streets, population per square mile, etc. They also looked at various "what-if" scenarios.

A. EVALUATIONS - POSITIVE:

1. The planned hospital is near a grid of interstate highways with easy access.
2. The project would include a \$22 million, 70,000-square-foot health care campus, expected to open two years after start of construction. It would include a 40,000-square-foot medical office building, a 24-hour Emergency-Center and a diagnostic center. It will incorporate state-of-the-art medical technology and treatments into a patient-friendly campus setting.
3. The new hospital will increase Summit City's medical resources and the delivery of needed healthcare choices.
4. It is estimated that the new hospital will employ 150-200 healthcare employee who will reside in the City. In addition, the new hospital will have an annual payroll exceeding \$12 million and an average annual salary of \$57,000 by 2009.

B. EVALUATION - NEGATIVE:

1. According to GIS analysis, location of the new hospital is too close to the other two main hospitals in the city, Summit Hospital and St. Murray Hospital. Although the hospital is easily accessible through a grid of freeways, the majority of the population does not reside near the expressway. The emergency vehicles will end up using the main thoroughfare rather than the expressways.
2. These hospitals have almost the same target patients. The increase in number of hospital beds far exceeds the foreseeable increase in projected population growth. The other two major hospitals in the area already experience less than 80% bed occupancy. Therefore, value added will not offset the cost of constructing a new hospital.
3. The addition of a new hospital will increase the healthcare cost.

FINAL DECISION

Based on the findings the Health Facilities Review Committee decided to deny the new hospital project. "We remain convinced there's not a need for a new hospital in the area. 'There are eight facilities within the 10-mile radius and two in the most immediate area,'" the committee reported.

Health Systems Hospitals claimed that the CON law is unconstitutional, and file an appeal to ask the state Supreme Court to repeal the CON law for hospitals and other acute-care providers. They think the state should regulate health facilities based on quality criteria instead of the number of the buildings. The Health Systems Hospitals won the lawsuit and a Certificate of Need was granted.

CONCLUSION

GIS is a handy tool with wide range of applications to analyze events and decision making. The system offers decision makers multiple layers of pertinent information to select. In the event of a decision making process such as a land-use conflict illustrated here, the users can analyze various layers such as roads, housing, land owners, traffic conditions at different time of day/night, future expansion of road plans, city boundary, neighbor areas, population density, and resource locations, etc.

Various options like 'what-if' conditions - move the blocks, turn the buildings or create a virtual road across the land can be explored. Most importantly, the form of digital data can give immediate feedback and thus save valuable time and resources. However, a powerful tool is still a tool. The use of the tool depends on the user and their geopolitical environment. GIS can offer the best tools to enhance decision making process. Currently many companies are marketing handheld GPS / PDA receivers for GIS users with sub-meter accuracy in real-time.

REFERENCE

1. The Guide to Geographic Information Systems: <http://www.gis.com/whatisgis/>
2. Geographic Information Systems : http://en.wikipedia.org/wiki/Geographic_information_system
3. Spatial Analysis And GIS (Technical Issues in Geographic Information Systems) by A S Fotheringham;
4. "Beyond Maps-GIS and decision making in Local Government." O'Looney. John. ESRI press, 2000, p-72

An Application Program Interface for vBulletin

Saumen Seal Sami, Tazilur Rahman, Khondker Shajadul Hasan, Javed Siddique

Department of Computer Science & Engineering, North South University
12 Kemal Ataturk Avenue, Banani, Dhaka-1213, Bangladesh

E-mail: saumen@gmail.com, tanzilurrahman@hotmail.com, {shajadul,jsiddique}@northsouth.edu

Abstract — vBulletin Calendar API is built on calendar feature of vBulletin, a commercial internet forum software produced by Jelsoft Enterprises Ltd. It acts as a remote agent for client applications intended to be developed by third party software developers. Such applications can be Facebook application, desktop application, IRC-bot and mobile applications for instant event notification and reminder. Important notices, announcements posted from departments of educational institutions often do not reach to target people, in time, due to various reasons. In this paper, we have proposed to develop an interface through which individual client applications monitor a common source of notice service center and notify users instantly.

Index Items — API, HTML, XML, vBulletin, MySQL, PHP-BB, SMS flag, DoS, URI.

I. INTRODUCTION

Notices work as a key point for any organization or institution as media of relaying information to the respective persons and/or sectors. Especially for the large institutions or organizations, it has a vital role. Currently these notices reach to target people by means of flyers, official websites and email. But the target people have to stop by flyers or browse websites or check emails on right time. It is often not possible for all of the notices to get updated right ahead of time. The goal of this development is to make notices and announcements available to maximum internet users instantly so that an internet user does not have to sit all the time in front of computer and browse for newest announcements.

vBulletin Calendar API processes queries and performs operations remotely requested by computer and mobile applications connected to internet. Our API provides requested data in XML format.

vBulletin Calendar API is intended to build for software application developers. Our product is like an interface or remote program that provides number of functionalities which can be called to provide data and process requests.

This development carries out the motto of increasing flexibility, efficiency of the notice posting system and distribution of these notices to the target people. This

development of the calendar feature of vBulletin is designed in such a way that it easily fits to any type organization including business, educational, corporate etc.

II. PRESENT INFRASTRUCTURE

Currently different organizations follow different policies or methods to inform all of their clients and employees about the latest notifications of information. Most of the educational institutions use the group emailing system to inform the specific groups. For example: administration group, faculty group and student group. In some places, students use official/unofficial forum to notify the latest event and news. In these cases they use online full featured forum software like vBulletin, phpBB etc. In Forum system, people request/register for an account and the forum admin setup an account, provides proper user access permission. After that, the user can log in and read threads and get informed.

So, to get informed of a notice a user must have to log on to the forum or at best they can be informed through email. Still the user has to be in touch to their mailbox. In short, the present infrastructure may play an important role but surely not the quickest and most effective. Among other two's, hospital mostly informs their clients through phone. Banks and other companies mainly use paper based notices, whereas software firms, telecom companies use SMS group notification, email to post to their employees and clients.

III. PROPOSED INFRASTRUCTURE

In this manner the proposed infrastructure has been designed considering the facts that lie currently in the present infrastructure. The intention is to dynamically change the view and developing a new structure which will automatically fit into any of the present system. Our structure uses the existing technology of vBulletin Calendar system as it is one of the most popular and mostly used commercial forum software. We set our target to develop a system to work as an intermediate way between the forum database and the third party client application which can be developed in anonymous production environment by any independent software developer by following some

defined protocols. The system verifies and processes requests, collects information from the vBulletin system and passes it to the third party applications on request.

The first thing is the organizations have to maintain a vBulletin forum database of their own. The users will have the accounts to that forum as it requires to any other forums. The vBulletin has a calendar mechanism that has been customized by us to fit for the API.

Services available for client application itself are defined as follows:

1. Marked Event List (user specific)
2. Modify SMS flag (event & user specific)
3. Latest Event

Services available for client application user are available as:

1. Authentication
2. Calendar List
3. Category list (calendar specific)
4. Event List (calendar & category specific)
5. Event Description
6. Event mark/unmark
7. Latest event (user specific)

Authentication is used for user classification and as the user identified all the services are ready for him according to his access level and they are grouped as: Calendar List, Category list, Event list, Event Description, Event Mark/Unmark, latest event information.

A. Routing (URI access protocol)

Rather than using the standard "query string" approach to URLs that is synonymous with dynamic systems, vBulletin Calendar API uses a segment-based approach. For example, the authentication URL is

http://{host}/api/authenticate/{user}:{pass}/

For this, we have used Apache Module mod_rewrite URL Rewriting Engine. This module uses a rule-based rewriting engine (based on a regular-expression parser) to rewrite requested URLs on the fly. It supports an unlimited number of rules and an unlimited number of attached rule conditions for each rule to provide a really flexible and powerful URL manipulation mechanism.

As you can see, first the URL is matched against the *Pattern* of each rule. When it fails mod_rewrite immediately stops processing this rule and continues with the next rule. If the *Pattern* matches, mod_rewrite looks for corresponding rule conditions. If none are present, it just substitutes the URL with a new value

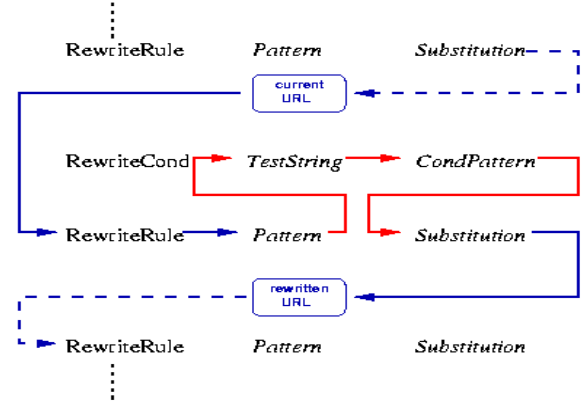


Fig. 1. Control flow through rewriting rule set structure

which is constructed from the string *Substitution* and goes on with its rule-looping. But if conditions exist, it starts an inner loop for processing them in the order that they are listed. For conditions the logic is different: we don't match a pattern against the current URL. Instead we first create a string *TestString* by expanding variables, back-references, map lookups, etc. and then we try to match *CondPattern* against it. If the pattern doesn't match, the complete set of conditions and the corresponding rule fails. If the pattern matches, then the next condition is processed until no more conditions are available. If all conditions match, processing is continued with the substitution of the URL with *Substitution*.

IV. DESIGN AND IMPLEMENTATION

The API does not come with any GUI; instead API provides XML response on request and this response is retrievable by any accessible network client application written in any programming language. No external hardware tools are required for any of the functionality. API requires vBulletin system installed and running on same web server. vBulletin works on a MySQL database consists of one hundred and sixteen (114) tables. We have added two additional tables for storing authentication tokens and event notifications. The API fetches data from this database and provides output in XML format.

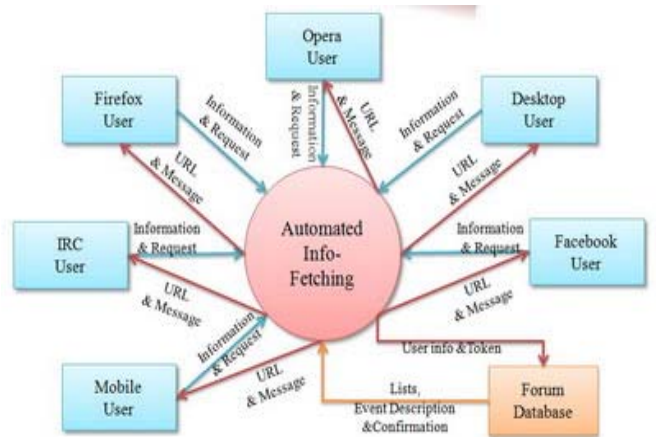


Fig. 2. A simple flow diagram of the structure

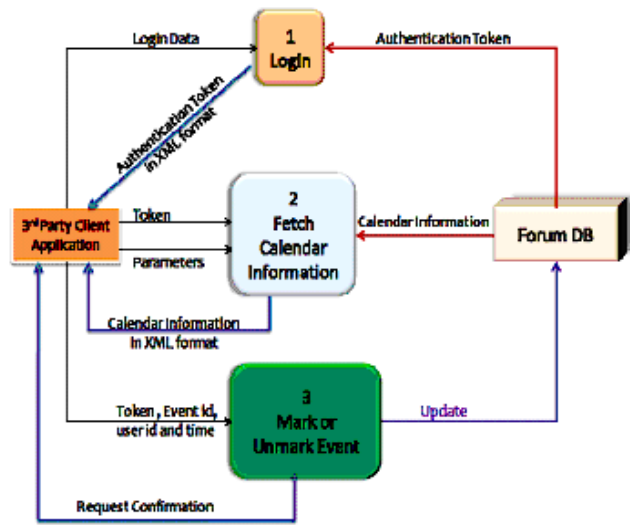


Fig. 3. Data flow diagram for the proposed structure

Service request and processing are performed in following order:

- Client application sends HTTP request to API with required parameters.
- Our API then prepares and sends SQL query to vBulletin database server.

- The communication between API and client application is committed through HTTP.

A. API service description: Authentication

According to our specification, every client application keeps separate authentication token for each user. This token is required to fetch calendar list, event list, event notification etc. via API.

- Application user identifies by providing username and password to client application.
- Client application calls API using our defined URI (described in Routing section).
- API on server validates the username and password and on success returns an authentication token.
- This authentication token is assigned to a user for a time period (This time period is configurable).

On the contrary, the API returns an error code when exception occurs.

This feature offers several advantages:

1. Performance is improved because login data is validated only once per session instead of with each all other service requests.

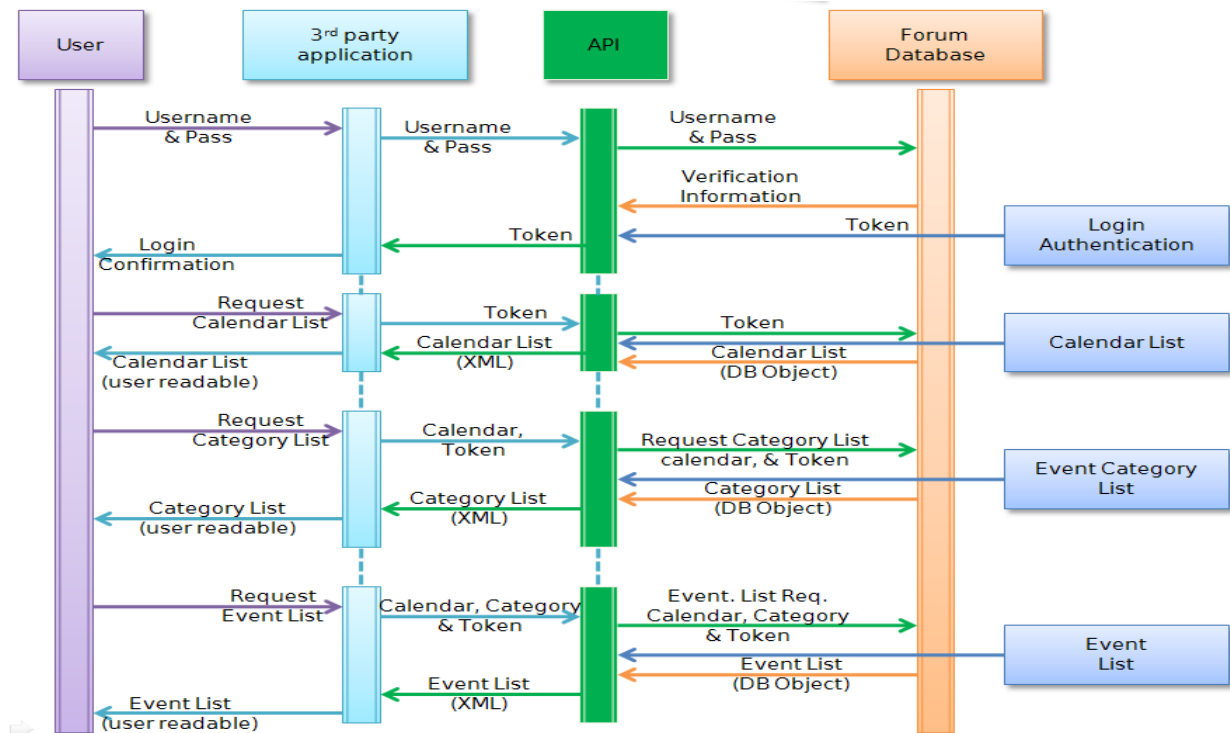


Fig. 4. Sequence Diagram for the proposed API structure

- In response, database server returns result set to API. API converts the result set into XML format and sends XML response to client application.

2. Security is tightened by minimizing the number of times login data is transmitted per session.

B. API service description: Fetch Calendar List

Gives list of calendars to which a user has at least read access. Client application can request for list of calendars and this list is filtered on user account permission.

C. API service description: Fetch Category List

Gives list of event categories for a specific calendar to which a user has at least read access. Client application can request for list of event categories and this list is filtered on user account permission.

D. Description of Other Core Methods

Here, 3rd party sends HTTP request to API which sends SQL query to forum DB in response forum DB sends back DB result that API converts into XML format and sends XML response to 3rd party Application 3rd party gets authentication token in XML format which it uses for further data retrieve like for fetch the calendar information it sends the token with necessary parameters and in response gets the calendar information in XML format. When a mark or unmark event occurs the data is updated in forum DB as well as an output of XML format of the data saved.

V. SECURITY AND SAFETY ISSUES

Our product is strong enough to handle security issues and does not hamper vBulletin system. Any big company or institution can integrate this API without concerning security and safety issues because we have ensured these issues during our design and implementation phase.

The API is written in PHP. Therefore, the web administrator can read API source code and see what's going on inside and make changes as and if required.

Rewrite Engine defines the rules of accessing URI resources in very strict way. Therefore, our API is SQL Injection protected because of strict URI access protocol.

The API uses Cache technology. This feature stands against DoS (Denial-of-Service) attack.

VI. CONCLUSION

We feel that the model we have proposed and implemented is not only a good one but also the best solution for the moment for instant notification of any notices made by a big organizations or institutions targeting to a particular group of people that include client or employee. This is because our API makes event announcements available to various platforms

including Web browser (extensible like Firefox), Desktop application, IRC Bot, Mobile application etc.

According to the requirement of organizations, they can develop various client side applications. As our API receives HTTP request and the response is in an XML format, so the client side application can be developed and used in any environment. The one and only fact is the availability of internet connection which is not an unattainable issue in this fastest growing world with technology.

REFERENCES

- [1] Deitel and Deitel, *Internet and world wide web: how to program*, 3rd edition, Prentice Hall 2003.
- [2] Silberschatz, Korth And Sudarshan, *Database System Concepts*, 5th Edition, Mc Graw Hill 2006.
- [3] Tim Converse and Joye Park Morgan, *PHP5 and MySQL Bible*, Wiley 2004.
- [4] "Denial-of-service attack", Wikipedia. 17 Apr. 2007 http://en.wikipedia.org/wiki/Denial-of-service_attack
- [5] "vBulletin® user guide", Online user manual 2008, <http://www.vbulletin.com/docs/>
- [6] Jeremy Frumkin, Jeremy Frumkin, "Definitions for the OCKHAM Reference Model" on Sep 15, 2004, <http://wiki.osuosl.org/display/OCKPub/ORMDefinitions>
- [7] "Apache Tutorials: Apache Server HTTP Docs" <http://httpd.apache.org/docs/1.3/howto/htaccess.html>
- [8] "Apache Tutorials: rule-based rewriting engine to rewrite requested URLs on the fly", *mod_rewrite*, http://httpd.apache.org/docs/2.2/mod/mod_rewrite.html

Structural Operational Semantics of Packages in Java

Abdullah Al Farook, Mohammad Shamsul Arefin and Md. Moshikul Hoque

Computer Science and Engineering Department, Chittagong University of Engineering & Technology, Bangladesh

Email(s): farookcanada@yahoo.com, sarefin_406@yahoo.com and moshikulh@yahoo.com

Abstract — A good understanding of programming language can be developed and understood in a machine independent manner called structural operational semantics. It is a mathematical model for specifying the meaning of a language by defining the effect of a running program in terms of its structure. The emphasis of this work is to present the structural operational semantics of several important aspects of packages in Java programming. With that we have showed how structural operational semantics is helpful for understanding the structure of packages.

Index Terms — Java, structural operational semantics, package, abstraction, environment.

I. INTRODUCTION

Java is unique among popular general purpose programming languages in that it provides a construct called package where classes, interfaces and sub packages defined in the source file will automatically be added to the package having the name specified by package name that are desirable in an Object Oriented Language. Package Structure helps the programmer to separate design part from coding part of the program and powerful capabilities to reuse code even outside the program.

Structural Operational Semantics (SOS) provides a framework for describing the meaning and specification of programming languages. SOS was introduced by G. D. Plotkin [1] as a logical means to defining operational semantics. The basic idea behind SOS is to define the behavior of a program in terms of the behavior of its parts, thus providing a structural, i.e., syntax oriented and inductive, view on operational semantics. An SOS specification defines the behavior of a program in terms of a set of transition relations. SOS specifications take the form of a set of inference rules, which define the valid transitions of a composite piece of syntax in terms of the transitions of its components. A growing number of programming languages from commercial to academic spheres have been given usable semantic descriptions by means of structural operational semantics. Because of its intuitive appeal and flexibility, structural operational semantics has found considerable application in the study of the semantics of concurrent processes. Moreover, it is becoming a viable alternative to denotational semantics in

the static analysis of programs and in proving compiler correctness.

Recently, structural operational semantics has been successfully applied as a formal tool to establish results that hold for classes of process description languages. This allows for the generalization of well-known results in the field of process algebra and for the development of a meta-theory for process calculi based on the realization that many of the results in this field only depend upon general semantic properties of language constructs.

A subset of Java has been taken to our work that would be termed as Mini Java [2] in the following sections of the paper. Semantics for declaration of packages have been given. The semantics is based on the environment and state. Two kinds of environment have been used – class-environment and working-environment. The semantics describes the changes in the environment and state at different steps of package execution, which is helpful for understanding the execution structure of packages.

II. LITERATURE REVIEW

Java semantics is an extensive research area. The first important research domain is the proof of the soundness of the Java type system. Drossopoulou and Eisenbach [3, 4, 5] are specifying semantics of several Java subsets in order to prove that the type soundness in these subsets (Javas, Javase, and Javar). They work on a large sequential of subset reduction theorem.

Attali et al. [6] presented a formal semantics of a large subset of Java, including inheritance, dynamic linking and multi-threading. They use a big-step semantic to describe the object-oriented features and use small-step semantics (structural operational semantics) to define the semantic of the concurrency. They are not concerned with typing. Their specification is on one side executable and on the other side it will be the basis for formal verification of Java programs.

M. S. Arefin et al. [7] showed structural operational semantics of Java interfaces. They showed how structural operational semantics is helpful for understanding the structure of Java interfaces.

P. D. Mosses [8] is working on Modular approach of Structural Operational Semantics (MSOS). They are trying to show MSOS for Java type sound.

Ryder et al. [9] conducted a study of the usage patterns of exception-handling constructs in Java programs. Their study offers an evidence to support my belief that exception-handling constructs are used frequently in Java programs and more accurate exception flow information is necessary. There are several research directions for exception constructs. The one is enabling a developer to make better use of exception mechanism, which includes analysis of uncaught exceptions, analysis of exception flow to facilitate understanding of the exception behavior. In Java, the JDK compiler ensures, by an interprocedural analysis, that clients of a method either handle the exceptions declared by that method, or explicitly redeclare them. Robillard and Murphy [10] have developed Jex: a tool for analyzing exception flows in Java. They describe a tool that extracts the flow of exceptions in a Java program, and generates views of the exception structure.

B. Chang et al. [11] proposed inter procedural exception analysis that estimates uncaught exceptions independently of programmer's declared exceptions. They compared their analysis with JDK-style analysis by experiments on large realistic Java programs, and have shown that their analysis is able to detect uncaught exceptions, unnecessary catch and throws clauses effectively.

In this paper the structural operational semantics of Packages in Mini Java is defined. In comparing with other semantics (denotational, axiomatic) the structural operational semantics seem to be simpler to understand. It may even be argued that the denotational one is a great deal more complex, because it requires an understanding of fixed points.

III. MINI JAVA

Mini Java is a subset of Java that contains some important properties of Java. Below we have provided the abstract syntax [2] of Mini Java on which we will work. Nonterminals appear in italics, identifiers with the suffix Id (e.g. ClassId) indicate the identifiers of newly declared entities whereas identifiers with the suffix Name (e.g. ClassName) are entities that have been previously declared. The keyword appears in bold.

```

n          ::= an integer
Class      ::= class ClassId [extends ClassName]
              {ClassVarDec (MethodDec)*}
MethodDec  ::= AccessMode ReturnTyp MethID
              (Para) {MethodBody}
MethodBody ::= VarDec Command ; return
              [Expr]

```

```

ClassVarDec ::= ε
              | ClassVarDec; ClassVarDec
              | AccessMode int VarId
              | AccessMode boolean VarId
              | AccessMode ClassName VarId
VarDec      ::= ε
              | VarDec; VarDec
              | int VarId
              | boolean VarId
              | ClassName VarId
AccessMode  ::= public | private | protected
ReType     ::= int | boolean | void |
              ClassName
Para        ::= ε
              | para, para
              | int VarId
              | boolean VarId
              | ClassName VarId
Commands    ::= ε
              | Commands; Command
Command     ::= Var = Expr
              | Var = BExpr
              | if BExpr Command
Expr        ::= Value
              | Var
              | Expr. MethName (Expr*)
              | new Classname ( )
              | Expr NOP Expr
Value       ::= n
              | null
Var         ::= VarName
              | Var. VarName
NOP         ::= + | - | * | /
Bop         ::= > | < | >= | <= | =
BValue      ::= true | false
BExpr       ::= BValue
              | Var
              | !BExpr
              | Expr BOp Expr
              | BExpr || BExpr
              | Bexpr && Bexpr

```

IV. ENVIRONMENT AND STATES

The environment of a program contains all mapping from identifier to denotable value known to the phrase at an instant. The state (σ) of a program contains all mapping from location to some value, which is storable or undefined. It is a snapshot of the contents of the program memory of the abstract machine. Thus, the state provides the means of storing temporarily the values that needed to be manipulated

IV. HOW JAVA PROGRAM ELABORATED AND INTERPRETED

The environment can be classified into two parts, class environment (η) and working environment (ρ). Here is how the Java program deals with them.

Step 1: Before the execution of the main method, class declarations are elaborated. The result is a class environment (η). Class environment of a program contains all mapping from identifier to class abstraction. Class abstraction of a class will contain all needed information (field attributes, method abstraction and super class name) of that class. It can be defined as follows:

```
ClassAbn ::= { { SuperClassName FieldAttr*
MethodAbn* } }
FieldAttr ::= (AccessMode Type FieldName)
MethAbn ::= (AccessMode MethodName ArgumentType
ReturnType MethodBody)
Type ::= int | boolean | ClassName
AccessMode ::= private | public
ArgumentType ::= () | (T1f1, ..., Tzfz)
Where z ≥ 0 and fi is a name of a formal parameter
(i ≤ z) And Ti is a Type (i ≤ z)
ReturnType ::= void | Type
```

Step 2: The execution of the main method starts and working-environment (ρ) will be created by declarations in the main method.

IV. DEVELOPING SEMANTIC FUNCTION FOR PACKAGE DECLARATION

In this section, we have provided the detail developing procedure of SOS of packages in Java. For this purpose, we have provided the required grammar and an example based on which we have developed the semantic function.

A. Package Declaration

A package declaration introduces a new technique where each class and interface defined in the source file will automatically be added to the package having the name specified by package name. A Package consists of a number classes, interfaces and sub packages. General form of package declaration is given below.

```
PackageDeclaration ::= SystemPackageopt package
PackageName Nestedopt PackageBody
SystemPackageopt ::= import System package name
Nestedopt ::= import NestedpackageName
packageBody ::= packageMember Declarations
packageMember Declaration ::= ( InterfaceDeclaration*,
ClassDeclaration * )
IfaceDeclaration ::= IfaceModifieropt interface IfaceName
ExtendsIFacesopt IfaceBody
```

```
IfaceModifier ::= public
ExtendsIFacesopt ::= extends SuperInterfaceName+
IfaceBody ::= { IfaceMemberDeclaration }
IfaceMemberDeclaration ::= MethodDeclaration*
ConstDeclaration*
MethodDeclaration ::= AccessModeopt ReturnType
MethodName (para)
ConstDeclaration ::= AccessModeopt Type Final-
varName= value
AccessMode ::= public
ClassDeclaration ::= ClassModifieropt class Class Name
superopt Interfaceopt ClassBody
Superopt ::= extends SuperClassName
Interfaceopt ::= implements InterfaceName1,
.....,InterfaceNameN
ClassBody ::= { ClassMemberDeclaration }
```

Below we have provided our own package definition, which will be used for developing semantic functions in later section. In the following example keywords are written in bold.

```
import java.util.* //System Package
package pack_01 //User Defined Package
import pack_01.nested1.* //Nested Package
public class class_01
{
    private int a ; float b ;
    protected int c;
    public void display() {MethodBodyOfdisplay} ;
    public void getdata() {MethodBodyOfgetdata} ;
}
public interface IFace_1
{
    void meth1 ();
    int meth2 (int i);
    int Y=1; int N =0; }
public interface IFace_2
{ void meth3 (); }
public interface IFace_3 extends IFace_1, IFace_2
{ void meth4 (); }
public class class_02 extends class_01
{
    double z;
    public int display1() {MethodBodyOfdisplay1}; }
public class class_03 implements IFace_1, IFace_2
{ int p ; private float q;
    public void meth1() {meth1Body};
    public int meth2 (int i) {meth2Body};
    void nonIfaceMethod () {nonIfaceMethodBody}
}
package pack_01.nested1
public class class_05
```

```
{
  int m; float n;
  public void display2( ){MethodBodyOfdisplay2};
}
```

B. Package Abstraction

Package abstraction of a package will contain all needed information of that package. When a package declaration is found in a Java Program, the information about that package will be kept as a package abstraction.

```
Package Abn ::= { SystemPackage PackageName
NestedPackage InterfaceAbn ClassAbn }
InterfaceAbn ::= {SuperIFaceName MethodAbn*
ConstAttr* }
MethodAbn ::= (AccessMode MethodName
ArgumentType ReturnType Body )
ConstAttr ::= ( Access Type VarDeclarator ConstValue)
ClassAbn ::= (SuperClassName FieldAttr* MethodAbn*)
FieldAttr ::= (AccessModes Type FieldName )
MethodAbn ::= (AccessMode MethodName ArgumentType
ReturnType Body )
```

C. Semantic Functions

Several semantic functions are necessary to define in order to process and keep important information about packages.

CreateFieldAttr is a semantic function that will take variable declaration in a class declaration as an argument. It will return field attribute as a result. Several variable declarations will be processed one by one.

CreateMethodAbn is a semantic function that will take method declarations in a class declaration as an argument. It will return method abstraction as a result.

CreateClassAbn, *CreateExtendClassAbn* are semantic functions that will take class declarations in a package declaration as an argument. As every class is an extension of class object, so first function will return object as super class name, which is called class abstraction. The second function will take super class name, variable declarations and method declarations as argument.

v range over variable declaration in a class
m range over method declaration in a class
s range over super class name
I range over interface names

$$\frac{\text{CreateFieldAttr}(v) = \text{VarAttr}^* \quad \text{CreateMethodAbn}(m) = \text{MethodAbn}^*}{\text{CreateClassAbn}(v, m) = \text{ClassAbn}}$$

$$\frac{\text{CreateFieldAttr}(v) = \text{VarAttr}^* \quad \text{CreateMethodAbn}(m) = \text{MethodAbn}^*}{\text{CreateExtendClassAbn}(s, v, m) = \text{ExtendClassAbn}}$$

CreateInterfaceAbn and *CreateExtendsInterfaceAbn* are semantic functions that keep necessary information about Interfaces declared in a package declaration.

w range over constant declarations in an interface
n range over method declaration in an interface
declaration

$$\frac{\text{CreateConstAttr}(w) = \text{ConstAttr}^* \quad \text{CreateMethodAbn}(n) = \text{MethodAbn}^*}{\text{CreateInterfaceAbn}(w, n) = \{ \{ _, \text{ConstAttr}^* \} \text{MethodAbn}^* \}}$$

I range over super interface names

$$\frac{\text{CreateConstAttr}(w) = \text{ConstAttr}^* \quad \text{CreateMethodAbn}(n) = \text{MethodAbn}^*}{\text{CreateExtendIfaceAbn}([I_k]_1^n, w, n) = \{ \{ [I_k]_1^n \text{ConstAttr}^* \text{MethodAbn}^* \} \}}$$

IsSameSignature is a semantic function that takes method declaration in the interface definition and method declarations in the class declaration. If all the methods declared in the interface exist exactly in the class declaration (fulfilling all the requirements mentioned above), and then it returns **true**, otherwise it returns **false**.

IsSameSignature(methDeclarationsInIface,
methDeclarationsInClass) = **true** | **false**

CreateClassAbnImplement is a semantic function, which takes interface names (the interfaces that are implemented by that class), variable declarations and method declarations of the class as its arguments. It will return *ClassAbn* as a result.

CreatePackageAbn and *CreateNestedPackageAbn* are semantic functions that will take class declarations and interface declarations as arguments. These functions will return package abstractions as a result. The first function imports extra argument the system package name. The creation of package abstraction is explained with the help of example 1 and example 2.

Example 1: Creation of package abstraction of package pack_01.

```
importSystemPackageName    import Java .util.*
importNestedPackag        import pack_01.Nested1
```

CreateClassAbn

$$\left\{ \begin{array}{l} \text{(private int a) (float b)} \\ \text{(protected int c)} \\ \text{public void display ()} \\ \quad \{MethodBodyOfdisplay\} \\ \text{public void getdata ()} \\ \quad \{MethodBodyOfgetdata\} \end{array} \right\}$$

=

$$\left\{ \begin{array}{l} \text{Object (private int a)} \\ \text{(_float b)(protected int c)} \\ \text{public display () void} \\ \quad \{MethodBodyOfdisplay\} \\ \text{public getdata () void} \\ \quad \{MethodBodyOfgetdata\} \end{array} \right\}$$

CreateExtendClassAbn

$$\left\{ \begin{array}{l} \text{class_01, _ double z,} \\ \text{public void display1 ()} \\ \quad \{MethodBodyOfdisplay1\} \end{array} \right\}$$

=

$$\left\{ \begin{array}{l} \text{class_01, _ double z} \\ \text{(public display1() void)} \\ \quad \{MethodBodyOfdisplay1\} \end{array} \right\}$$

CreateInterfaceAbn

$$\left\{ \begin{array}{l} \text{(void meth1 (); int meth2} \\ \text{(int i); int Y =1; int N = 0;)} \end{array} \right\}$$

=

$$\left\{ \begin{array}{l} \text{_, ConstAttrOfIface_1,} \\ \text{MethodAbnOfIface_1} \end{array} \right\}$$

CreateExtendIfaceAbn

$$\left\{ \begin{array}{l} \text{([IFace_1, IFace_2]} \\ \text{(void meth4 ())} \end{array} \right\}$$

=

$$\left\{ \left\{ \begin{array}{l} \text{[IFace_1, IFace_2] (public} \\ \text{meth4 () void) } \end{array} \right\} \right\}$$

CreateClassAbnImplement

$$\left\{ \begin{array}{l} \text{IFace_1,} \\ \text{int p; private float q;} \\ \text{public void meth1()} \\ \quad \{meth1Body\} \\ \text{public int meth2 (int i)} \\ \quad \{meth2Body\} \\ \text{void nonIfaceMethod ()} \\ \quad \{nonIfaceMethodBody\} \end{array} \right\}$$

=

$$\left\{ \begin{array}{l} \text{Object (public int Y 1) (public} \\ \text{int N 0) (_ int p) (private float q)} \\ \quad \text{MethodAbnOf_meth1} \\ \quad \text{MethodAbnOf_meth2} \\ \quad \text{MethodAbnOf_nonIfaceMethod} \end{array} \right\}$$

CreatePackageAbn

$$\left\{ \begin{array}{l} \text{import Java .util.*,} \\ \text{import Pack_01.Nested1,} \\ \\ \left\{ \begin{array}{l} \text{(private int a) (float b)} \\ \text{(protected int c)} \\ \text{public void display ()} \\ \quad \{MethodBodyOfdisplay\} \\ \text{public void getdata ()} \\ \quad \{MethodBodyOfgetdata\} \end{array} \right\} \\ \\ \left\{ \begin{array}{l} \text{class_01, _ double z,} \\ \text{public void display1 ()} \\ \quad \{MethodBodyOfdisplay1\} \end{array} \right\} \\ \\ \text{(void meth1 (); int meth2} \\ \text{(int i); int Y =1; int N = 0;)} \\ \\ \left\{ \begin{array}{l} \text{([IFace_1, IFace_2]} \\ \text{(void meth4 ())} \end{array} \right\} \\ \\ \left\{ \begin{array}{l} \text{IFace_1,} \\ \text{int p; private float q;} \\ \text{public void meth1()} \\ \quad \{meth1Body\} \\ \text{public int meth2 (int i)} \\ \quad \{meth2Body\} \\ \text{void nonIfaceMethod ()} \\ \quad \{nonIfaceMethodBody\} \end{array} \right\} \end{array} \right\}$$

$$\left\{ \left\{ \begin{array}{l} \text{Java, _util, pack01.nested1} \\ \text{Object (private int a)} \\ \text{(_float b)(protected int c)} \\ \text{public display () void} \\ \quad \{MethodBodyOfdisplay\} \\ \text{public getdata () void} \\ \quad \{MethodBodyOfgetdata\} \end{array} \right\} \right\}$$

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{class_01, _ double z} \\ \text{(public display1() void)} \\ \quad \{MethodBodyOfdisplay1\} \end{array} \right\} \\ \\ \left\{ \begin{array}{l} \text{_, ConstAttrOfIface_1,} \\ \text{MethodAbnOfIface_1} \end{array} \right\} \\ \\ \left\{ \left\{ \begin{array}{l} \text{[IFace_1, IFace_2] (public} \\ \text{meth4 () void) } \end{array} \right\} \right\} \end{array} \right\}$$

$$\left\{ \begin{array}{l} \text{Object (public int Y 1) (public} \\ \text{int N 0) (_ int p) (private float q)} \\ \quad \text{MethodAbnOf_meth1} \\ \quad \text{MethodAbnOf_meth2} \\ \quad \text{MethodAbnOf_nonIfaceMethod} \end{array} \right\}$$


```

ConstAttrOffface_1 = (public int Y 1) (public int N 0)
MethodAbnOffface_1 = (public meth1 ( ) void) (public
                    meth2 (int i) int)
MethodAbnOf_meth1   = (public meth1 ( ) void
                    meth1Body)
MethodAbnOf_meth2   = (public meth2 (int i) int
                    meth2Body)
MethodAbnOf_nonIfaceMethod = ( _ nonIfaceMethod ( )
                    void nonIfaceMethodBody)

```

$$\text{CreateClassAbn} = \left\{ \begin{array}{l} \text{int m; float n;} \\ \text{public void display2(} \\ \quad \{ \textit{MethodBodyOf display2} \}; \end{array} \right\}$$
$$\begin{aligned} \text{CreateNestedPackageAbn} & \left[\begin{array}{l} \text{int m; float n;} \\ \text{public void display2()} \\ \quad \{ \textit{MethodBodyOfdisplay2} \}; \end{array} \right] \\ = & \left\{ \begin{array}{l} \text{pack_01, object, (_int m)} \\ \quad \quad \quad (_ \text{float n}) \\ \text{public display2 () void)} \\ \quad \{ \textit{Method body of display2} \} \end{array} \right\} \end{aligned}$$

Package abstraction that we have got by using the functions *CreatePackageAbn* and *CreateNestedPackageAbn* will be added to the package environment.

$$\frac{\text{CreatePackageAbn}(t, u, x, y, z) = p}{\langle \text{package } d, t, u, x, y, z, \eta, \rho, \sigma \rangle \Rightarrow_d \langle _, \eta', \rho, \sigma \rangle}$$

Here, $\eta' = \eta[p/d]$

We have given a formal description of structural operational semantics for a solid subset of Java called Mini Java. Structural operational semantics of several important aspects of packages in Java programming were clearly written down. With that we have showed how structural operational semantics is helpful for understanding the structure of packages. It can also be concluded that SOS of this important property of Java can serve as a basis of understanding and reasoning how programs behave. Moreover, these semantics can be used for various kinds of analysis and verification. These semantics may be provided for reference –potentially even for standardization.

- [1]. Gordon D. Plotkin, "A Structural Approach to Operational Semantics", Tech. Rep. DAIMI FN-19, Computer Science Department, Aarhus University, Aarhus, Denmark, 1981.
- [2]. C. Tayapiwattana, "Operational Semantics of Java", Master's Thesis, Asian Institute of Technology, Bangkok, Thailand, 1999.
- [3]. S. Drossopoulou and S. Eisenbach, "Is the Java type system sound?", In *Fourth International Workshop on Foundations of Object-Oriented Programming*, 1997.
- [4]. S. Drossopoulou and S. Eisenbach, "Java is type safe-Probably", In *Proceedings of the European Conference on Object-Oriented Programming*, 1997.
- [5]. S. Drossopoulou, S. Eisenbach, and S. Khurshid, "Is Java Sound?" In *Theory and Practice of Object Systems*, 1998.
- [6]. I. Attali, D. Caromel, and M. Russo, "A formal Executable Semantics for Java", 1998.
- [7]. M. S. Arefin, M. A. Rahman, S. Chakraborty and M.R.Siddique, "Structural Operational Semantics of Interface in Java", In *the proceedings of ICCIT, Dhaka Bangladesh*, pp. 532-536, December 2002.
- [8]. P. D. Mosses, "Modular Structural Operational Semantics", *The Journal of Logic and Algebraic Programming*, pp. 195-228, 2004.
- [9]. R. K. Chatterjee, B. G. Ryder, and W. A. Landi, "Complexity of concrete type-inference in the presence of exceptions", *Lecture notes in Computer Science*, vol. 1381, pp. 57-74, April 1998.
- [10]. M. P. Robillard and G. C. Murphy, "Analyzing exception flow in Java programs", in *Proc. of '99 European Software Engineering Conference and ACM SIGSOFT Symposium on Foundations of Software Engineering*, pp. 322-337, Springer-Verlag, 1999.
- [11]. B. M. Chang, J. Jo, K. Yi, and K. Choe, "Interprocedural Exception Analysis for Java", *Proceedings of ACM Symposium on Applied Computing*, pp 620-625.

Automated Digital Archive for Land Registration and Records

Mabruka Toaha*, Shahriar Khan**

* Department of CSE, North South University, Kemal Ataturk Ave, Banani, Dhaka
mabruka.toaha@gmail.com

** School of Engineering and Computer Science, Independent University, Bangladesh,
14/2 A, Baridhara North Road, Dhaka - 1212, khandhaka@gmail.com

Abstract - Land management is an important issue for Bangladesh, where the land registration process is a major barrier to the economic growth of the country. Technology and IT can greatly facilitate the creation of a proper land management system. This paper focuses on various issues and problems related to the current system of land management. It offers an alternative management system using IT features and shows a comparison between the existing system and the proposed system. It is shown that a digitized land management system in Bangladesh can bring positive outcomes in the entire process by minimizing hassle, expenditure, time, and staff dishonesty.

Index Terms - Land, Records, Bangladesh, Automated

I. INTRODUCTION

Land dominates the life of almost every person in Bangladesh, particularly in rural areas. A proper recording of land rights, along with the need for a sound land management system is important for a largely agrarian economy under heavy population pressure. The existing land management system is very ineffective and cumbersome at best. The governmental system of land trading is very time consuming and problematic. Taking the advantage of this intricate system, a section of personnel are grabbing the properties of the weak. As a result, the number of lawsuits, illegal possessions, strife, scuffles, and other crimes are increasing day by day. The present superannuated land management system is responsible to a great extent for the deteriorating law and order situation of the country.

II. PRESENT SYSTEM

The directorate of Land Records and Surveys (DLRS) is entrusted with the responsibility of carrying out cadastral surveys and preparation and maintenance of Record Of Rights (ROR) in respect of every parcel of land in the country. With the passage of time, the organizational structure of the Directorate changed considerably to meet new requirements.

There are two parts in the preparation of the record of rights. The first part is to draw a revised mouza map showing changes in location, area and characteristics of land. The second part is to prepare the Khatian (cadastre)

including the ownership, area and character of land. These two parts together are called record of rights. The phases followed in the survey and preparation of record of rights are traverse survey, plot to plot survey, Kistwar (demarcation of boundary), Khanapuri (preliminary record writing), Bujharat (local explanation), attestation, draft publication, hearing of objections, appeal, final Junch (checking), printing and publication of Khatian.

A. Land Sale Registration

One fundamental aspect of land registration in Bangladesh is that it is merely the record of an isolated transaction. It only shows that a particular transaction took place but it does not establish whether the parties are legally entitled to enter into such a transaction. In other words registration takes place without any critical examination of relevant records. This kind of registration of deeds in Bangladesh has been in force since the Registration Act of 1908.

The amount of tax levied on the immovable property transferred is 10 percent of the total value of the property. When the buyer and the seller of a parcel of land come to an agreement on the transfer of land, a deed is prepared on the stamp of required value with the help of a deed writer, and then it is presented before the Sub-Registrar for registering it and for putting the official seal on it. Both the buyer and the seller have to appear before the Sub-Registrar who takes cognizance of the land transfer and puts his signature with the official seal. The registered deed is copied in the relevant volume. It takes about 9-12 months to get the registered deed. The deed prepared by the deed writer contains a lot of unnecessary words and sentence, which could be avoided. At present the land records kept in the Tehsil office are not checked to find out whether the seller is the real owner of the land or not, or if any other persons have got any interest in the plot of land in question. There are instances of false registration in both urban and rural areas. The incidence of false registration is much higher in urban areas. A plot of land may be sold to different people leading to decades- long litigation.

A. Detailed Mapping

Information on the area and location of land can be collected through four different approaches:

- Ground Survey,
- Aerial/Satellite Photography,
- GPS based data collection system and
- Existing information of lands in the land offices.

Bangladesh is well suited for cadastral mapping by vertical air photographic techniques. Its primary advantage is its flatness and air visibility of plot boundaries. A scale of 1:12,000 will be sufficient enough in most areas. The main steps in aerial survey will consist of:

- i) taking of photographs,
- ii) rectification (1:12,000 scale),
- iii) correction in site,
- iv) providing a survey number,
- v) area calculation,
- vi) preparation of draft Khasra,
- vii) preparation of map with internal details,
- viii) public review of draft Khasra and map
- ix) and the final publication.

Using aerial photography methods will have three main advantages:

First, contrary to the prevailing notions it will be less costly and will have less man-management problems to deal with. Second, it will cut down the time taken for conventional cadastral settlement operations of about ten years by at least three years. And finally it will yield much more accurate results. However, this method will require substantial initial investment in terms of purchase of equipment and aircraft and training of officials.

B. Creating Effective Database

In the existing procedure, the land records stored in the Tehsil office are not checked to find out whether the seller is the real owner of the land or not. Retrieval of any information is very difficult which raises a question to the purpose the volumes of registers will serve. Computerization of registration records would therefore make the job much simpler and facilitate easy retrieval of registration records. The new system will handle all the manual activities of land records and tax using computerized database system. It will have database with a simpler format with the necessary information for the conveyance of the property. The database will contain information such as names and occupations of both buyer and seller, with their present and permanent addresses, the particulars of the land (lot number, JL number, name of mouza), the mode by which the seller became the owner

of the plot (i.e. purchase, inheritance, etc.), the classification of land, area of land, boundary demarcating, the value at which the land is being transferred, tax etc. In case of property inheritance the buyer field will be replaced by present owner and the seller field will be replaced by previous owner.

The effectiveness objectiveness of the database will include:

- Ensuring that the stored data can be used for a variety of applications like: record keeping, land registration, mutation, property tax and capital gain collection and resolving dispute.
- Maintaining data that are both accurate and consistent and avoiding data redundancy.
- Ensuring purposeful information retrieval system so that all the data required for current and future applications will be readily available.
- Ensuring efficient upgrading system etc.

C. Database Organization (Logical and Physical Views)

The database is to be shared by many users. Users at different levels will have different user schema. Examining the required views an overall logical model of the database will be developed. Finally the logical model of the database will be transformed into a corresponding physical database design. The physical design will involve fundamental information on how data will be stored and related as well as how they will be accessed. The figure below shows how the user reports and user schema will be related to the logical model and physical design.

D. Features of the Proposed System

The Land Information Management Systems software will be an integrated system wherein the user will be provided with the facility to navigate and perform all operations from Data Entry to Data Validation to Records Display through one single Interface. The main modules available in the LIMS will be as follows:

Data Entry Module: The Data Entry Module will have interfaces to enter data of the registers that are maintained for revenue purposes. Facility to view and take prints for each and every register will be provided along with the option to view and list the contents of the registers.

Security Module

Considering the confidential and sensitive nature of the land records, the issue of security is to be given the highest priority while designing the software. The important features of the Security Setup will be:

Restricted user access to records wherein only authorized land functionaries have permission to edit records in the database.

* Digital signature for authentication; All relevant employees will have the option to encrypt their digital signatures in the document that they will handle.

* A separate module for tracking each and every transaction that takes place in the database along with the name of the user responsible for such action.

* Backup facility to help in regeneration of past records in case of accidental or deliberate deletion.

The security of the software along with the entire system will be maintained at three separate levels:

Logical Level:

The system will permit the user with correct passwords or authorization codes to enter the system or a particular part of the database. There will be at least three different level of security access:

1) Level 1: Employed personnel at the supreme level will be allowed to view all the stored land records. They will be allowed to add delete and edit land records related to buying, selling and mutation of land. They will also be allowed to view the taxation records on a certain piece of land.

2) Level 2: Employees at the second level will be allowed to view, add and edit the existing data, i.e. they will be able to store land sell and mutation records. However, to delete any land record from the database, they will require special permissions or they will have to hand over the word to the level 1 user.

3) Level 3: The database will have a limited portion of the each entry as “read only to all”. Any one interested to view the buying and selling records on any piece of land will be able to do so by paying certain government fees.

Graphical Module

This module will be developed using VC++ and Active X components and will have a linkage with the GIS software. This is primarily the map display module of the system. It will provide all the tools that are available in any of the Graphic Viewer Software of international repute. The noteworthy features of the graphical module will be:

- 1) Map Display
- 2) Display the maps within a predefined window area.
- 3) Zoom In, Zoom Out, Pan View, Fit View, Refresh View, and Window Zoom facilities.
- 4) Facility to view legends like schools, masjids, graves etc.
- 5) Edge-matched maps within the villages so that one village will correspond to one map in the digital environment. However, in the real world, a village map may well be distributed over multiple sheets.
- 6) The spatial database (maps) generated on a true scale thus keeping the dimensions accurate.

Query Capabilities

The system will provide facility to query on the owner name, Khatian number in addition to random selection of parcels from the map and querying on them. Since the spatial and the textual data will be integrated, all queries have spatial components, which will give the user the flexibility to pinpoint locations on the map. In addition, a number of Reports will be designed keeping the daily requirements of the district administration in view. The reports to be available at proposed system will include:

- 1) List of Government owned lands
- 2) List of revenue assigned land's
- 3) Status of mutation transactions
- 4) Mutation types.

Solution Architecture to the Proposed System

The Solution provided for proposed network system will have a decentralized structure. A computer at the District level will act as the local server whereas all the other machines installed in the Thana Headquarters and other focal points will be designated as clients. Connectivity between the clients and the local server will be established by means of an appropriate method like Wi-Max. The local servers at the District levels can be connected to the main server via the optical fiber backbone of Bangladesh Railway.

The software will be designed following the client/server architecture in order to improve the usability, flexibility, interoperability and scalability of the solution as compared to a centralized system. Here, a client will be defined as a requester of services and a server as the provider of services.

Strategic Approach to the Proposed System

Ascertaining the overall feasibility of the IT implemented land management system relates to clarifying the objectives of the new system and determining if the system moves towards its objectives in some way. The objectives of the proposed system include:

- * Developing a digital archive and retrieval system,
- * Automating manual procedures to improve in various ways like reducing errors, increasing speed and accuracy and cutting down employee time required,
- * Smoothing the land purchase and sale registration process,
- * Reducing law suits related directly or indirectly to land,
- * Introduce transparency in the government thereby reducing corruption.

Technical Feasibility: The proposed system requires around 534 computers and 138 printers that are to be distributed to all the corners of the country.

Number of computers:

Client computers at the Thana levels : $460 \times 1 = 460$
 Local servers at the District levels : $64 \times 01 = 64$
 Main server and backups at the Capital : $5 + 5 = 10$
 Total : 534 computers

Number of Printers:

Printers with local servers : $64 \times 2 = 128$
 Printers with Main servers : 05
 Additional printers : 10
 Total : 138 printers

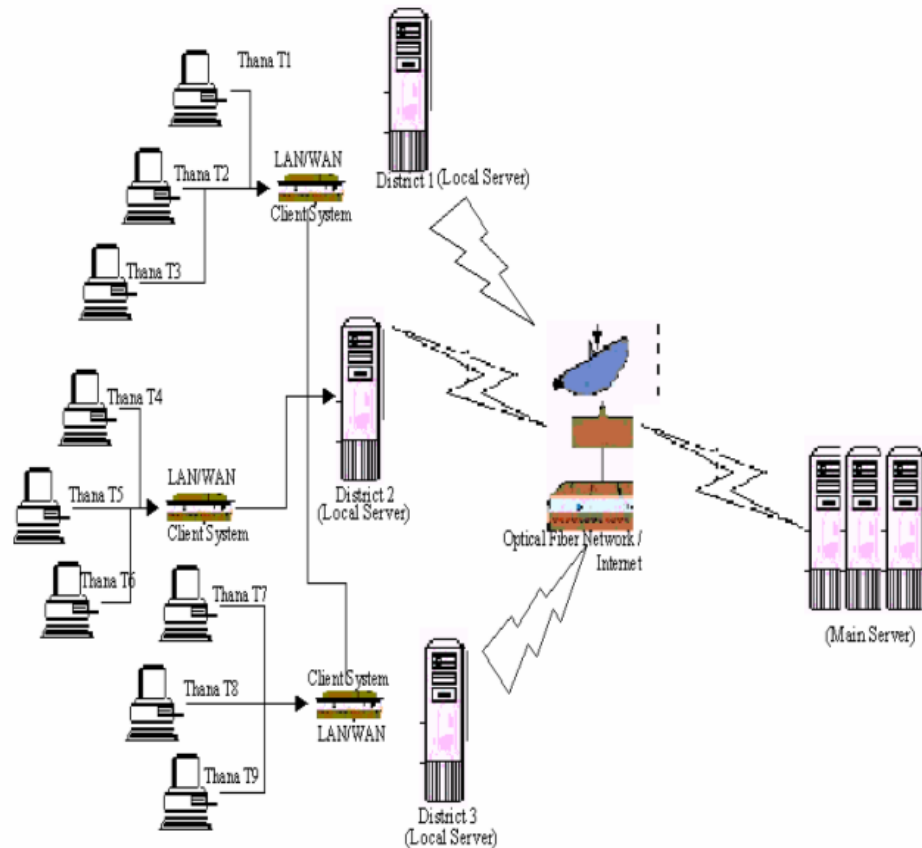


Fig. 2 Architecture of the Proposed System

V. DIFFICULTIES OF CHANGE:

There will be several obstacles to moving forward from the present land management system to the proposed system. Some of the difficulties the new system has to overcome are:

- * The Directorate of Land Records and Survey has earned notoriety for corruption and inefficiency.
- * The low level stuffs are temporary and they try to make as much as they can through corrupt practices.

* Powerful land owners try to induce officers to do undue favors.

* Corruption exits at the stage of attestation and objection hearing.

If the new system is implemented, not only will the corrupt system come down, the previous records of the employees will be exposed to the public. Strong opposition from existing employees is to be expected.

TABLE 1
COMPARISON OF PRESENT AND PROPOSED SYSTEMS

Function	Process	Present System	Proposed System
Land Record System	Updating	Highly time consuming. About 46 years will be needed for total updating	Updating will be done immediately after the change. So no extra time will be needed.
	Intricacy	High intricacy	No intricacy
	Objection/Appeal	Takes several months to years for completion	A few days will be needed
	Manpower	DLRS has a total of 9098 stuffs for survey on payroll	Not more than one thousand people will be needed for the job
	Printing time	At least 28 years are required for printing all Khatians with current setup	It may take a few days for a whole Thana & the job can be done at several Thanass simultaneously
	Printing Cost	About 400/- per Khatian	About 2-3/- per Khatian
Land Sale Registration	Backlog	High backlog	No backlog
	Amount of Printing	10 copies per Khatian	As much as needed, whenever needed.
	Integrity	No linkage between DLRS & Collectorate setup	A linkage will be developed through computerized networked
	Information Dissemination	No information dissemination. The record is kept in the office where the transaction occurred	Information will be disseminated through computer network
	Immoral Behavior of Land Buyers	Tendency of understating actual land value for paying less stump duty	Due to stored up-to-date price no opportunity of understating.
	Deed Registration	Takes 9-12 months	1-7 days will be required

IV. COMPARISON BETWEEN PRESENT AND PROPOSED SYSTEM

The proposed Land Information Management System will help to improve efficiency and processing and. It can be an effective database, providing accurate information to both officials and the public. It will give a secure access of data and also check unauthorized manipulation. The system will speed up the present system significantly, introducing transparency in the government, and reducing corruption. Having a Relation Database LIMS can be integrated to other e-Governance Applications.

V. CONCLUSION

The present process of land registration process is a barrier to the economic growth of our country, and creates financial overheads on the trading of land. The problems occur mainly due to the backdated land record and storing system. The proposed system will greatly simplify the tasks, and become an effective replacement to the existing system.

REFERENCES

- [1] Acharya, Ballabh Prasad (1987): *Policy of Land Acquisition and Development - Analysis of an Indian*

- Experience*. Third World Planning Review, vol. 9, No. 2, Liverpool University Press, United Kingdom.
- [2] Brammer, Hugh (1984): *Land Use Planning in the Tropics*. ADAB News, vol. XI, No. 1, Dhaka.
- [3] Marulanda, Liliana and Florian Steinberg (1991): *Land Management and Guided Land Development in Jakarta*.
- [4] Paper based on work by Soetojo H., N. Rukmana and F. Steinberg (1990) entitled "Problems and Issues of Land Development for Jakarta's Low-Income Families" and Triaco on the Guided Land
- [5] J. Brownlee, Managing information to national standards, 1998.
- [6] Local Government Management Board (LGMB) Information society news, Vol. 1, issue 1, pp 9 – 11.
- [7] Burrough, P.A. and McDonnell, A., *Principles of Geographical Information Systems*, Oxford University Press, New York, ISBN 0-19-823366-3, 1998.

Development of a Highly Optimized Preemptive Real Time Operating System (pRTOS)

Md. Al Mehedi Hasan, Shamim Ahmad*

Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology

*Department of Computer Science & Engineering, University of Rajshahi, Rajshahi-6205,
Bangladesh E-mail: mehedi_ru@yahoo.com, shamim_cst@yahoo.com

Abstract— A real-time operating system (RTOS) is software which ensures that time critical events are processed as efficiently as possible. In this paper, an attempt has been taken to develop a real time operating system, named Preemptive Real Time Operating System (pRTOS), in which all of the important issues regarding to a real time application have been considered. In this pRTOS, strictly preemptive scheduling algorithm has been used. This scheduling policy makes sure that important tasks are handled first and the less important later. The Bitmap technique has been used to find out the highest priority task from the unsorted ready list. The complexity of this technique for selecting the highest priority task is $O(1)$, which is much faster than the linear search technique having complexity of $O(n)$. This pRTOS can support 64 priority levels ranges from 0 to 63. In addition with this, it is a highly configurable RTOS. Moreover, it can be adopted in a board range of hardware platform, say, Intel x86, MIPS, Hitachi SH, Power PC and Strong ARM processors. This RTOS has been tested on Intel x86 and from the obtained result, it has been found that our developed pRTOS can be used for various application, say, for automated industrial systems, control-systems, high-tech electronics/electrical products and home applications.

Index Terms — RTOS, Scheduling, Context Switching, Priority Resolution, Task Synchronization, Preemptive Real Time Operating System (pRTOS)

I. INTRODUCTION

Real-time computing is where system correctness not only depends on the correctness of logical result but also on the result delivery time. So the operating system should have features to support this critical requirement to render it to be termed a Real-time operating System (RTOS). The RTOS should have predictable behavior to unpredictable external events. "A good RTOS is one that has a bounded (predictable) behavior under all system load scenario i.e. even under simultaneous interrupts and thread execution [1]." A true RTOS will be deterministic under all conditions. These operating systems occupy little space from 10 KB to 100KB as compared to the General Operating systems which take hundreds of megabytes [2]. Many commercial embedded operating systems (RTOSs) are available on the market [3, 4, 5] and most of these

RTOSs have primitive real-time capabilities that are adequate for only very simple real-time applications. On the other hand, some real-time systems often require sophisticated scheduling capabilities to handle applications with high performance requirements.

With this background, in this paper, an attempt has been taken to develop a real time operating system, named Preemptive Real Time Operating System (pRTOS), in which all of the important issues regarding to a real time application have been considered. This operating system (pRTOS) has been developed to provide the optimized approach to handle the critical issues of scheduling and synchronization, response-time constraints, task priorities, latencies and deadlines, and inter-task communications. It also provides some additional services, say management of the resources, memory, tasks, and time.

II. BASIC REQUIREMENTS OF RTOS

The following are the basic requirements of an RTOS.

- Multitasking and Preemptability
- Task Priority Level
- Task Synchronization & Inter-task Communication
- Predefined latencies
 - Task switching latency
 - Interrupt latency
- Portability and Modularity
- Scalable
- Scheduling Policy

III. ARCHITECTURE OF THE DEVELOPED pRTOS

Fig. 1. illustrates the architecture of the developed RTOS. The aim of this work is to design a Real Time Operating System mainly with the following characteristics

1. Low overheads in terms of time, code size.
2. Ease of portability to other CPU architectures.

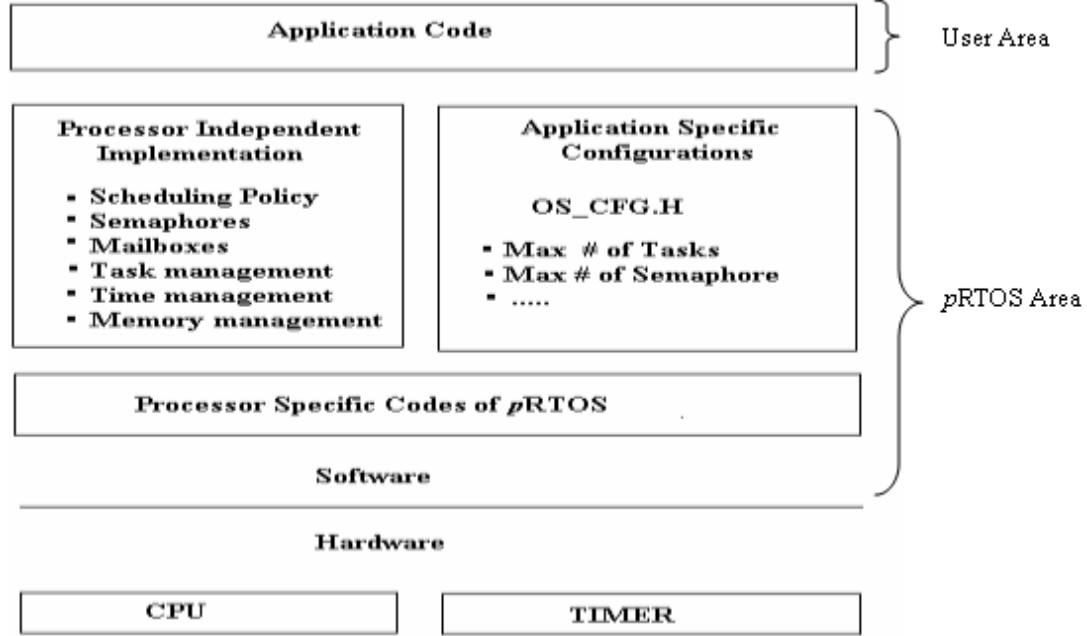


Fig. 1. Architecture of the pRTOS

From the point of view of portability, since it is very important to have a modular design, in which the hardware (CPU/board) dependent code must be clearly isolated from the hardware independent code, as a result, in developing this pRTOS, the design phase has been divided into two parts, these are,

- Processor independent part
- Processor dependent part

In this RTOS, the processor independent part has been written in C language. The processor dependent part has been written in assembly language. The necessities of processor dependent part is that the operations, such as, context switching involves of storing of the CPU register contents of the currently running process and loading of the parameters of new task into the CPU registers.

A. Scheduling Policy

The heart and soul of any operating system is its scheduler [6, 7]. We have used strictly preemptive scheduling algorithm in our developed pRTOS. In this policy, at any instance of time, only the highest priority task that is ready to run executes. If a task T_i runs, it means that all tasks T_j with priorities lesser than T_i are blocked. This scheduling policy makes sure that important tasks are handled first and the less important later. For example, in an aircraft cruise control system, the fight controller task

will have more priority than a task that controls the air-conditioning system.

Algorithm: Preemptive Scheduling

1. Task T_i with priority i is running.
 2. If (task T_i has finished){
 - a) Select the next highest priority task is ready to run.
- ```

}
Else{
 if (a task T_j ($j > i$) is in the ready list){
 a) Select the task T_j which is ready to run.
 }
 Else{
 Go to step 1.
 }
}

```

#### B. Priority Resolution Technique

Priority based task switching may kill considerable amount of time when searching of highest priority task is a great challenge. From this view point, a robust technique, called Bitmap technique, has been used in this RTOS which ensures single time find out the highest priority task [8, 9]. The complexity of this technique for selecting the highest priority task is  $O(1)$ . This is more faster than the linear search technique to find the highest priority task from the ready list whose complexity is  $O(n)$ .

### C. Context Switching and Interrupt Handling

Context switching means saving the context of the interrupted routine or task and retrieves or load the new context of the called routine or task. It becomes the responsibility of the RTOS to service this context switching. An interrupt is a signal that the hardware can send when it wants the processor's attention. It also becomes the responsibility of the RTOS to service these interrupt as immediately as it can at a "safe" time.

Two types of context switching method are developed in this RTOS.

**Task level context switch:** To perform task level context switch the following functions are developed:

- void pOSStartHighRdy(void);
- void OS\_FAR pOSCtxSw(void);

**Interrupt level context switch:** To perform interrupt level context switch the following function are developed:

- void pOSIntCtxSw(void);

### D. Task Synchronization & Inter-Task Communication

Semaphores have been used to achieve synchronization of tasks and resources. When semaphore is created, it is associated with a specific of tokens. Once a semaphore is created, tasks can request a specific number of tokens from the semaphore. When the requested number of token are available, the number of tokens in the semaphore is decremented, and given to the task. If the number of tokens exceed the number of available tokens, the tasks blocks. We have developed semaphore management API. Mailbox is an inter-task communication mechanism. A task posts a message in the mailbox and another task will read the message. This pRTOS has mailbox functions.

We have developed a structure (Semaphore or Mailbox Control Block) to maintain the Semaphore and Mailbox functionality.

Semaphore or Mailbox Control Block:

```
typedef struct sema_mailbox {
 UBYTE Wait_Group;
 UBYTE Wait_Table[8];
 UWORD SemaCnt;
 void *Pointer;
} OS_SEMA_MAILBOX;
```

**Wait\_Group:** Group corresponding to the tasks waiting for an event (Semaphore or Mailbox) to occur

**Wait\_Table[8]:** List of tasks waiting for event (Semaphore or Mailbox) to occur

**SemaCnt:** It is used to count when this structure is used in Semaphore

**Pointer:** It is used to point to a message in the mailbox and also used to make chain of SemaMailbox Control block.

## IV. EVALUATION OF INVOLVED FEATURES IN THE DEVELOPED pRTOS

The features which we have included and developed are listed below

1. To support multiple tasks in real-time applications, the pRTOS is multi-tasked and preemptible.
2. pRTOS supports 64 tasks out of which one are system Idle Task. pRTOS is a pre-emptive priority base scheduler with 64 priority levels within 0 to 63 assigned during task creation.
3. pRTOS is a portable operating system. Its code can be ported on many processors that are commonly used in the designing of real time application.
4. pRTOS has task synchronization and inter-task communication functions. It uses Semaphore, Message Mailbox for task synchronization and inter-task communication. pRTOS has semaphore functions, which are usable like event flags, resource acquiring keys or counting semaphores.
5. pRTOS handles multiple levels of interrupts i.e., the pRTOS is preemptible at task level and at the interrupt level as well.
6. pRTOS is extremely modular. It has a clear set of functions defined for doing jobs like task, semaphore, scheduling management and context switching.
7. pRTOS is scalable. The functions which are needed in our application can be predefined in a configuration file included in the application codes.
8. pRTOS uses strictly pre-emptive scheduling. This scheduling policy makes sure that important tasks are handled first and the less important later.

## V. TEST METRICS AND MEASUREMENT RESULTS

RTOS has strict timing constraints and it should respond, under all possible load condition, in a very

predictable way. Real time operating systems are required to be very deterministic. This means different latencies of the systems should be predictable. Best operating system should have minimum latency and least amount of standard deviation.

#### A. Context Switching Time

The operation of context switching has been developed in assembly language. The context switching time can be optimized by only using proper coding and it has been done in this work. It is included in the processor dependent part of the pRTOS. It can be recalled that context switching involves storing the CPU register contents of the currently running process and loading the new task parameters in the CPU registers. We have calculated the required bus cycle in executing the assembly code for context switching which is shown in Table I.

TABLE I  
MEASUREMENT RESULTS

| Test Metrics                   | CPU cycles |
|--------------------------------|------------|
| Task Level Context Switch      | 225        |
| Interrupt Level Context Switch | 177        |

#### B. Scheduling Complexity

The scheduling complexity characterizes the time required by the RTOS to select a ready task. This duration depends not only on the scheduling algorithm, but also on the number of ready tasks when the algorithm runs. The scheduling algorithm of pRTOS uses Bitmap technique to find the highest priority task from the unsorted ready list. The complexity of this technique for selecting the highest priority task is  $O(1)$ . This is more faster than the linear search technique to find the highest priority task from the ready list whose complexity is  $O(n)$ .

#### C. Space Complexity

A true RTOS should occupy little space from 10 KB to 100KB as compared to the General Operating systems which take hundreds of megabytes. RTOS should be scalable. Only the OS functions that are necessary part of

the applications codes, thus having reduced memory requirements. The functions needed for task servicing, inter-process communications, etc. must be predefined in a configuration file included in the user codes. This pRTOS is highly scalable. It occupies only 25 KB memory if we consider all of its functionality and its kernel.

## VI. CONCLUSION

There is a need of real-time operating system (RTOS), which can provide the solution of the complex issues of real time-time system and can provides some services of different functionalities. Our developed pRTOS can be used for various applications, say, for automated industrial systems, control-systems, high-tech electronics/electrical products and home applications.

## REFERENCES

- [1] Dedicated Systems Experts, "What makes a good RTOS. Brussels", Dedicated Systems Experts, Belgium, 2001
- [2] Brian Santo, "Embedded Battle Royale," IEEE Spectrum, pp.36-41, December, 2001.
- [3] VxWorks Programmer's Guide, 5.1, Wind River Systems, 1993.
- [4] L. M. Thompson, "Using pSOS+ for embedded real-time computing," in COMPCON, pp. 282-288, 1990.
- [5] D. Hildebrand, "An architectural overview of QNX," in Proc. Usenix Workshop on Micro-Kernels and Other Kernel Architectures, Apr. 1992.
- [6] Sriram V Iyer, Pankaj Gupta, "Embedded Realtime Systems Programming", Tata McGraw-Hill, 2004.
- [7] Andrew S. Tanenbaum, Albert S. Woodhull, "Operting Systems Design and Implementation", Second Edition, Prentice-Hall, 2001.
- [8] Jean J. Labrosse, iC/OS: The Real-Time Kernel, R&D Publications, Lawrence, 1993.
- [9] S.-J. Oh, et al., "Deterministic Task Scheduling for Embedded Real-Time Operating Systems," IEICE Trans. Inf. & Syst., Vol. E87-D, No. 2, pp. 123-126, Feb. 2004.

# Resource Capability Discovery and Description Management System for Bioinformatics Data and Service Integration - An Experiment with Gene Regulatory Networks

Emdad Ahmed

Integration Informatics Laboratory  
Department of Computer Science  
Wayne State University  
Detroit, MI 48202, USA  
Email: emdad@wayne.edu

**Abstract**—Traditional legacy HTML based web sites/ page can be thought of as *web services* because the dynamic web pages can take user input argument via web forms and response to user query. The ability of agents and services to automatically locate and interact with unknown partners is a goal for Web based Data Integration system. This "serendipitous interoperability" is hindered by the lack of an explicit means of describing what web pages are able to do and in order to do it what input it takes and what output it produces, that is what is their *capabilities* [1]. The tremendous success of the WWW is countervailed by the efforts needed to search and find relevant information. For tabular structures embedded in HTML documents, typical keyword or link-analysis based search fails. The next phase envisioned for the WWW is automatic ad-hoc interaction between intelligent agents, web services, databases and semantic web enabled applications. A large amount of information available on the Web is formatted in HTML tables, which are mainly presentation oriented and are not suited for database applications. As a result, how to capture information in HTML tables semantically and integrate relevant information is a challenge. We are envisioning another layer of web abstraction where user can query intra web document table like structure. Our prototype application is based on *WebFusion* and an ad hoc query language *BioFlow* [2], [3], [4], [5], [6] a software agent that can simulate a person interacting with web search forms and extracting information from the resulting pages by means of an API. We need to develop a framework which is able to query search web forms and the web page tables in a SQL way. In this context we also report a Java based implementation for integrating Flybase and AlignACE site.

**Index Terms**—Table structure, Table modeling, Web mining, Ontology generation, Semantic Web, Intelligent Wrapper, Web Information Extraction, HTML forms, Web Data Integration, extraction ontology, Hidden Web, Web Automation.

## I. INTRODUCTION

Turning the current Web into a Semantic Web requires automatic approaches for annotation of existing data since manual annotation approaches will not scale in general. By "capability" we mean an operation offered by the provider of a data source to a user to extract data from the services

data store and structure it for transfer to the customer [7]. By Web resource we mean any information, data structure, representation which can be queried by some language. Example of web resource are: URLs, links, Web Forms which can be filled for query, outgoing URLs, Keyword, attribute which can be found inside Meta Tag (this conforms to key string matching problem) or anywhere in web page; Repeated structure found in web page (e.g Table, list). Information extraction from the Web is the process of extracting data from a set of Web pages in order to fill a given database schema. Web information extraction presents two particular features: (a) the input documents are semi-structured: they are neither raw data (full text) nor strictly typed (relational tables), but a structure is provided (typically by HTML tags) (b) the information sources are heterogeneous, i.e. the same kind of information is often presented in different formats. A Web IE system receives two kinds of input: a DB schema that represents the information the user wants to retrieve; a set of IS (Web sites) containing the data to be extracted.

Such a system aims at populating the database with data extracted from the sources and must provide mechanisms for dealing with semi-structured documents and for accessing heterogeneous information sources. The presence of heterogeneous information sources, whose models are not known a priori, makes scalability one of the main features for these systems. In other words, the system must be easily programmed for dealing with a large number of different information sources. An important design element for these systems is the representation techniques used for modelling the database schema and the information sources. Indeed, this issue permits a classification of approaches to Web Information Extraction, which are as follows: *Manual extraction*, *Procedural extraction*, *Declarative extraction* and *Automatic extraction*.

### A. Motivating Example

**Example 1.** Suppose a user want to do the following series of activities on a site:

- Go to [www.amazon.com](http://www.amazon.com).
- Type "Ruby" into the search text field.
- To narrow down the results, click "Books",
- then for further narrowing click "Computers & Internet" in the left sidebar.

By fetching the web page from the URL and parsing and analyzing the links and auto form filling, the above tasks can be automated by our system. We call it *Web User Activity Automation*. Web task automation can be done by two different ways: *Recorded Macro* and *Programmable API*. We will go for the second one.

**Example 2.** Suppose a client want to search for Book from two sites [www.amazon.com](http://www.amazon.com) and [www.bn.com](http://www.bn.com) at once and see the combined result. Both the site offer advanced search web form. Traditionally the client has to visit both the sites iteratively filling the form and gather data manually. If a Web Data Ingetration System as envisioned by ours is present, the client has to submit the query once and the system will do the rest automatically on the fly.

**Example 3.** Let us consider the following SQL query for book search from amazon site.

```
SELECT Title, Rating
FROM amazon.com
WHERE Title like "Discrete Mathematics"
AND Rating=5;
```

By analyzing the advanced search option of amazon, we see that *Title* attribute can be mapped to Interface Schema, whereas *Rating* attribute can be mapped to Result Schema. For the above example with the Interface that are available in amazon site, user query can be satisfied only for the *Title* attribute, not for *Rating* attribute. Consequently, user will see a lot of tuples containing all the *Rating* returned in response to query, whereas he/she is only interested to see record corresponds to Rating 5. Our system can solve these types of user query by efficiently mapping user query to site.

**Technical Challenges:** The web forms of the two sites look the same but how to automatically fill the form is much more difficult as will be clear from the following discussion. The client will have a GUI, the query will be submitted from there. The query need to be rewritten as per the mappings of the respective site. Our proposed web resource capability description/ discovery method will take care of the mappings. Each site's capability (*form information*) will be captured and stored in resource description repository. The envisioned overall architecture is shown in the following figure 1. In response to the form, we will have tabular data. Semantic enrichment of the tabular data will be made semi-automatically and through ontology and human knowledge. Tabular schema will also be stored

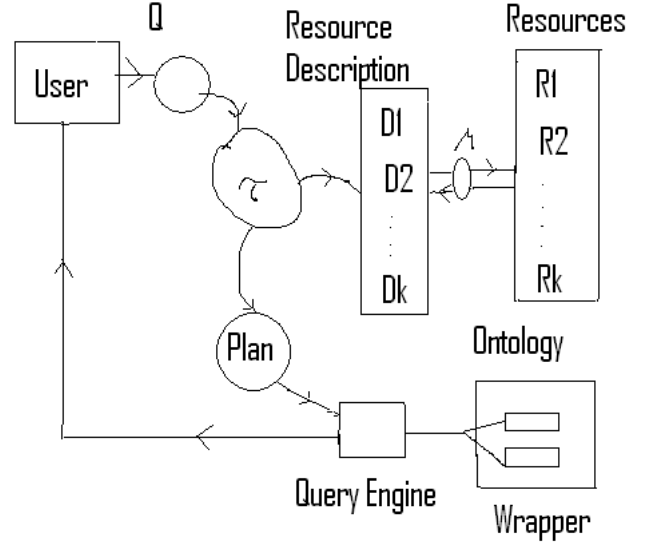


Fig. 1. Resource Description /Discovery Mechanism

in the resource description repository. Thus our *resource capability description* will be focused toward handling web form and representing table schema, i.e column name identification and label assignment. The rest of the paper deals with based on the above two ideas. The current research is our ongoing work where we proposed and identified a model for web data integration. The current focus of our research is to describe a web page so that some body else can later use to query source. We have developed prototypes for bioinformatics web data integration. To the end, our system can solve Web Data Intergration problem in an ad hoc manner as follows:

```

Define function RegulatoryRegion
Extract 800bp sequence
Using wrapper FlyBaseToACE in ontology GeneMapping
From URL Flybase
Submit FBgnNum string;

Define function motif
Extract 6..8bp sequence
Using wrapper FlyBaseToACE in ontology GeneMapping
From URL AlignACE
Submit 800bp string;
```

### B. Related Work

Turning the current web into a semantic web requires automatic approaches for annotation of existing data since manual approaches will not scale in general. [8] presented a method for automatically filling in forms to retrieve the associated dynamically generated pages. They consider the task of duplicate record elimination and uses a stratified sampling approach to decide when a particular source has been fully extracted. Using their approach automated agents can begin to systematically access portions of the "Hidden Web". The Hidden Web Exposer (HiWE) project at Stanford proposed a way to extend crawlers beyond

the Publicly Indexable Web (PIW) by giving them the *capability* to fill out Web forms automatically. The related work references are the following: [9], [8], [10], [11], [12],[13], [14]. [15] presented **DeLa**, which reconstructs (part of ) a "hidden" back-end database. It does this by sending queries through HTML forms, automatically generating regular expression wrappers to extract data objects from the result pages and restoring the retrieved data into an annotated (labelled) table. The present research work is our ongoing work where we have identified and proposed ontology based and declarative workflow query language for ad hoc web data integration on the fly [16], [17]. Our experimentation on Bioinformatics Web Data and Service Integration has also been reported in another work [18].

## II. DEEP WEB QUERY MODEL

Structured Web databases can be queried via query forms or through Web service interfaces. We uniformly refer to both access methods as "query interfaces". Through query interfaces, data consumers (e.g end users) are able to express their information needs by imposing selection conditions on certain attributes of interest. Our system views a Web database as a single relational table  $DB$  with a set of queriable attributes  $A_q = \{attr_{q1}, attr_{q2}, \dots, attr_{qn}\}$  (interface schema) and a set of result attributes  $A_r = \{attr_{r1}, attr_{r2}, \dots, attr_{rm}\}$  (result schema). Each  $attr_{qi} \in A_q$  represents the queriable attribute through the query interface, while the result attribute  $attr_{rj} \in A_r$  corresponds to the attributes displayed in the result pages. Each query operation can be modelled using SQL syntax as:

```
SELECT {attrr1, attrr2, ..., attrrm}
FROM DB
```

```
WHERE attrq1 = valq1, attrq1 = valq1, ..., attrqn = valqn
```

where  $val_{qi}$  is the corresponding attribute value filled into the query form.

We model the web site as  $S \subseteq Q \times R$ , where  $Q$  is the query interface schema and can be represented as  $Q \subseteq F \times P$  and  $R$  is the result schema, can be represented as  $R \subseteq L \times V$ . The semantics of the above definition is as follows: we have a set of form label  $F$  and a set of corresponding parameter  $P$ . In response to the web form, we will have tabular data, which have a set of values  $V$ . The label  $L$  of the values  $V$  may or may not be in the web page. One of the major fundamental research problem is how to find  $L$ , given  $F, P, V$ . This is a separate research topic and we address the problem somewhere else.

## III. ONTOLOGY DEFINITION

Ontologies are a convenient way to describe relationships among the terms in the universe being modeled for the purpose of heterogeneous information integration on the web. It is also a convenient way to describe extraction

rules for web resources in the form of wrappers. So, in BioFlow, we allow schema mapping relationships and data extraction rules for web sites be included for the use by the query engine. An ontology mainly captures three types of relationship: *synonym*, *isa* and *partof* [19]. The general structure of ontology definition is shown below - for a given ontology  $o$ , its definition contains a set of schema mapping statements and a set of wrapper statements. The syntax of the mapping statements and wrapper statements are shown next.

define ontology  $o$

schema mapping statements/wrapper statements;

end ontology  $o$ ;

mapping  $m : URL\ u \rightarrow a_1 : b_1, \dots, a_k : b_k$

$c_1 \# d_1, \dots, c_m \# d_m, e_1 @ f_1, \dots, e_n @ f_n$ ;

wrapper  $w : URL\ u \rightarrow a_1 :: r_1, \dots, a_k :: r_k$ ;

where  $a_i, b_j, c_k, d_l, e_m, f_n$  are terms (including nested terms), and  $r_p$  are regular expressions corresponding to XPath expression of the table in an HTML page, and where  $a : b$  means  $a$  is a synonym of  $b$  (and vice versa),  $c \# d$  means  $c$  is a homonym of  $d$  (and vice versa),  $e @ f$  means  $e$  is an antonym of  $f$  (and vice versa), and finally  $a :: r$  means  $a$  is an attribute located at  $r$ .

As a simple example of ontology, we present here a table pertaining to Flybase site. We call the table ontology *GeneMapping*. The above ontology shows mapping between

TABLE I  
GENEMAPPING

| CG Number | FBgn Number | F/R | Range                 |
|-----------|-------------|-----|-----------------------|
| CG12055   | FBgn0001091 | R   | 2R:3679403..3680885   |
| CG6871    | FBgn0000261 | F   | 3L:18815706..18821294 |
| CG4769    | FBgn0035600 | F   | 3L:5353290..5355898   |
| CG17903   | FBgn0000409 | F   | 2L:16719874..16722659 |
| CG7070    | FBgn0003178 | F   | 3R:18193234..18198464 |
| CG4581    | FBgn0025352 | R   | 2R:19754124..19756139 |
| CG7176    | FBgn0001248 | R   | 3L:8349503..8354628   |
| CG10120   | FBgn0002719 | R   | 3R:8538819..8548267   |
| CG3476    | FBgn0031881 | F   | 2L:7027594..7028959   |
| CG2107    | FBgn0035383 | F   | 3L:2981023..2983551   |
| CG6050    | FBgn0024556 | F   | 2R:9395063..9396838   |

a CG number to a corresponding FBgn Number gene, its direction, forward or reverse as well as its corresponding range in the *chromosome arm*. We went further in *geneontology* site to see synonymous name for the *EjTum* gene. The following three names showed up: CG6050 ; l(3)L4569 and mtEF-Tu. We also get the Flybase Database entry field as FBgn0024556. We report in the above table last row only one mapping (CG6050, FBgn0024556), whereas user can pose query using any one of the set  $\{CG6050, l(3)L4569, mtEF - Tu, FBgn0024556\}$ . A system is required which can rewrite user query variable to the site. We need to use a most Generalized Term as query variable to be submitted into the site. For the Flybase site, the Generalized Term happens to be the FBgn number. We store all the synonymous relationship in relational table

as multivalued dependency [20] as follows: GO ontology

TABLE II  
SYNONYM GENE MAPPING

| Generalized Term | Synonym Term                               |
|------------------|--------------------------------------------|
| FBgn0024556      | EtTum                                      |
| FBgn0024556      | CG6050                                     |
| FBgn0024556      | l(3)L4569                                  |
| FBgn0024556      | mtEF-Tu                                    |
| FBgn0001091      | GAPDH                                      |
| FBgn0001091      | BEST:GH12586                               |
| FBgn0001091      | CG12055                                    |
| FBgn0001091      | G3-P dehydrogenase                         |
| FBgn0001091      | GA3PDH                                     |
| FBgn0001091      | GADPH                                      |
| FBgn0001091      | Gapd                                       |
| FBgn0001091      | GAPDH I                                    |
| FBgn0001091      | GAPDH-1                                    |
| FBgn0001091      | GAPDH                                      |
| FBgn0001091      | Gapdh43E                                   |
| FBgn0001091      | gh12586                                    |
| FBgn0001091      | glyceraldehyde 3 phosphate dehydrogenase I |
| FBgn0001091      | Glyceraldehyde-3-phosphate dehydrogenase I |

which is a more a nomenclature and taxonomy, than a formal ontology, is highly successful and extensively used. Although GO is technically a nomenclature rather than an ontology it has been successfully used to annotate large volumes of data and consequently support interoperability and integration from heterogeneous data sets. *Applied to Bioinformatics, an ontology is a "controlled vocabulary for the description of the molecular functions, biological process and cellular components of gene products"*. Current data integration approaches should be enhanced by ontological knowledge and ISA relationships in ontologies should be explored. [21]

#### IV. WEB FORM FILLING

Information transmitted to the server in a CGI request is fundamentally just a list of (name, value) pairs with appropriate URL encoding string. Thus we can characterize a form with  $n$  controls as a tuple  $F = \langle U, (N_1, V_1), (N_2, V_2), \dots, (N_n, V_n) \rangle$ , where  $U$  is the URL to which the encoded CGI request is sent, and the  $(N_i, V_i)$  are (name, value) pairs to be sent. [10] described 5 heuristics to deal with the from filling problem. There are two ways to submit a form for CGI processing. HTTP POST method submit the form in the body of the request. HTTP GET method submit the form as part of the URL. A question mark (?) separates the base URL and action path from the encoded names and values. This query is then sent directly to the Web site of interest. It has the same effect as that of a user clicking the Go button without selecting or typing anything on the Web form. For our experiment with *Flybase* and *AlignACE* site form filling, the following Java code snippet captures those functionality:

```
System.out.println("Enter Flybase ID ");
try {
 inputStr = mystdin.readLine();
} catch (IOException e) {
 System.out.println("IOException: " + e);
}
url1String.append("http://www.flybase.org/reports/");
url1String.append(inputStr);
url1String.append(".html");
```

```
url1 = url1String.toString();
FlyBaseToACE fetcher1 = new FlyBaseToACE(url1);
if (HEADER.equalsIgnoreCase(option)) {
 pageHeader1 = fetcher1.getPageHeader();
}
else if (CONTENT.equalsIgnoreCase(option)) {
 pageContent1 = fetcher1.getPageContent();
 rangeIndex = pageContent1.indexOf("Sequence location");
 if (rangeIndex != -1) {
 ourRange = pageContent1.substring(rangeIndex+26,
 rangeIndex+51);
 checkForwardReverse = pageContent1.substring
 (rangeIndex, rangeIndex+60);
 }
 computedRange = deleteComma(ourRange);
 url2String.append("http://www.flybase.org/cgi-bin
 /getseq.html?range=");
 url2String.append(computedRange);
 url2String.append("&source=dmel&format=
 DecoratedFasta&adrange=");

 if (checkForwardReverse.indexOf("+") != -1)
 url2String.append("-2000"); // the gene is forward
 else
 url2String.append("+2000"); // the gene is reverse

 url2 = url2String.toString();
 FlyBaseToACE fetcher2 = new FlyBaseToACE(url2);

 if (checkForwardReverse.indexOf("+") != -1) {
 System.out.println("THE GENE IS FORWARD,
 add 2K upstream");
 pageContent2 = fetcher2.getPageContent();
 output800bp = process2Kup(pageContent2);
 }
 else {
 System.out.println("THE GENE IS REVERSE COMPLEMENTED,
 add 2K downstream");
 pageContent2 = fetcher2.getPageContent();
 output800bp = process2Kdown(pageContent2);
 }
 url3 = "http://atlas.med.harvard.edu/cgi-bin/alignace.pl?
 name='Drosophila'&nc=10&xp=10&gc=0.38&seq=";
 url3 = url3 + output800bp;
 FlyBaseToACE fetcher3 = new FlyBaseToACE(url3);
 pageContent3 = fetcher3.getPageContent();
 log (pageContent3);
}
```

#### A. Need to consider Web Form DOM

In spite of having a tag in the HTML specification called **label** for the declaration of a label, it is almost not used. Form labels refers to which is visually adjacent to the left or on the top of a form field. Labels in Web Form always appear in HTML page, but DOM do not catch it, programmers have to parse the page and extract label. Based on the above discussion, we propose to take care of the web form label. DOM for a typical web form look like:

document.form.name.element.value (*Existing*)  
document.form.**label**.name.value (*our proposal*)

#### B. HTML tables - location and extraction problems

Data on the Web in HTML tables is mostly structured, but we usually do not know the structure in advance. Thus we cannot directly query for data of interest. [9] proposed a solution based on document-independent extraction ontologies. DeLA uses four heuristics to assign label for the tabular data. As for their second heuristics, when data

is represented in voluntary labels in table headers such as  $\langle TH \rangle$ ,  $\langle THEAD \rangle$ , then the label assignment is trivial. We take whatever we find inside the heading. We experimented with the following site: [http://www.mrc-lmb.cam.ac.uk/genomes/FlyTF/quite\\_sure\\_TFs.html](http://www.mrc-lmb.cam.ac.uk/genomes/FlyTF/quite_sure_TFs.html) to test the functionality and materialize the web response in MySQL relational DB. We use MySQL as it outperform other relational DB for Data Intensive Web Application (DIWA) [22]. On the fly materialization of web page table need to take care of the following delimiter characters and clean those, otherwise they will conflict with the DB system's DDL and DML.

String delims = " ' ! @ # % & \* ( ) - \_ = + ] [ | ' ; : / ? . > , < " ;

More heuristics for column label assignment are as follows: Inner Label Assignment (extract repeating prefix as label of column); Inter Label Assignment (consider consecutive two columns, first one is the label, second one is the data); Column Splitting (first portion is the column label, second portion is the data)

We wish to add more heuristics such as any consecutive four character in the set "ACGT" will not be anything other than a sequence (of course exception is there like TATA, which is a company name as well as a gene sequence) We also wish to add that only *syntactic labeling* is not enough. Suppose we get X as the repeating prefix character in a column of all the data. Then according to our heuristics it will be labeled as X. But X can be *chromosome* as well as *geometric coordinate*. In order to mitigate the confusion, we plan to use web knowledge source such as Wikipedia, Google etc to resolve the concept extraction, thereby enabling *semantic labeling*.

## V. INTEGRATING FLYBASE TO ALIGNACE SITE

In this section, we briefly discuss our experience with implementation of a Bioinformatics Web Data Integration system. Flybase is a site about a database for drosophila genetics and molecular biology. We search the site for gene, say *sin3*. In response to the quick search, a *overview page* shows the attribute name and values e.g sequence location, gene direction etc. We extract those from the *overview page*. Then we get the genome region, the result is given in *detail page* as *Decorated Fasta* format. The Flybase site has limited query processing capability such as it can take 2K - 5K sequence to update region display. But in our program, we allow user to add arbitrary region, then also allow to add arbitrary up and down sequences. Our target is to extract say 800 bp from start sequence from *sin3* gene. The whole process is automated as a sequence of *web access as scientific workflow*. Then the extracted sequence is automatically submitted in the AlignACE site to extract sequence motif, say 6 to 8 bp. We have plan to extend the functionality to integrate from AlignACE to TRANSFAC as well.

## VI. IMPLEMENTATION ARCHITECTURE

The high level architecture of the system has been shown in Figure 1. Here we implement the systems as a layered architecture consist of four layers as depicted in the following Figure 2. The *first layer* is the entire

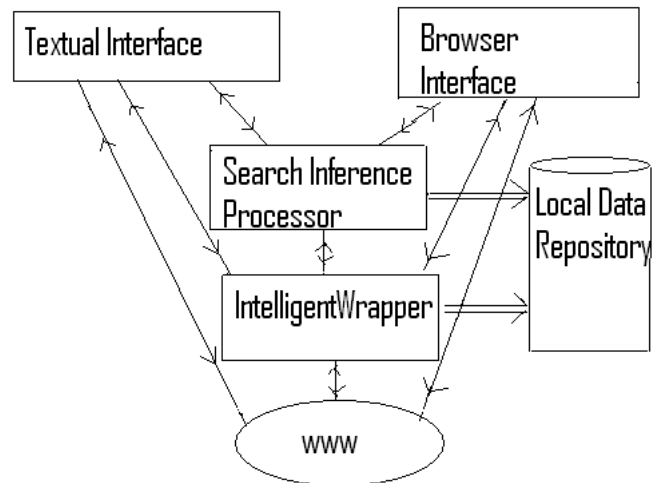


Fig. 2. Overall System Architecture

world-wide web on the Internet. Because HTML documents are not only generated automatically by specific application programs such as *FrontPage*, *DreamWeaver*, but also written by human manually, it is common that some documents have syntactic errors. Tools like *JTidy* can be used to fix the HTML tag errors. The *second layer* is the Intelligent Wrapper that will implement our proposed Label Assignment method. It accesses the world-wide web through the Internet (Java URL and URLConnection class), fetches HTML documents and converts them into web objects as nested table. In order to speed-up query processing, the Intelligent Wrapper will also cache web objects in the local *resource description repository* and builds proper indexes on these web objects. It will allow the user to adjust web objects by adding, adjusting or deleting attributes or objects and it can also learn the pattern from the user's interaction and then process similar HTML documents accordingly. *Not only the proposed Wrapper will extract tabular data from web, but also it will map the data into attribute-value construct so that the system can later query as ordinary relational database.* That is why it is called the Intelligent Wrapper. The *third layer* of the system is the Search and Inference Processor that will implement *BioFlow*. It is mainly in charge of query processing. It obtains queries to be processed from the user interface layer and checks whether or not web objects involved are in the local resource description repository. If web objects involved are not in the local resource description repository, then it invokes



the Intelligent Wrapper to fetch them. The *last layer* is the Intelligent User Interface. Two kinds of interfaces will be provided: textual interface and browser interface (*we have already implemented using Java JEditorPane*). They will provide different kinds of environment for the user to issue commands and express queries, perform syntactical analysis, and pass valid commands in internal format to the Search and Inference Processor. They also display query results generated by the Search and Inference Processor to the user in different format, the web objects generated by the Intelligent Wrapper, and original web document like Internet Explorer.

## VII. CONCLUSION AND FUTURE WORK

In this paper we presented an implementation of a procedural and logical approach to web resource capability description/ discovery. By capability we mean what a web site can do for us and in order to perform its tasks what input it needs and what output it produces. In this connection we have identified web form to be the gateway. In response to web form, we assume a tabular output. By our *Intelligent Wrapper*, we convert the repeated structure into relational model, we can identify attributes and values in the table by our label assignment heuristics, then we can apply normal DB SQL like queries on it.

The main contribution of this paper is that we have presented a new top-down approach how to query HTML pages declaratively, which even Google search engine is missing i.e *Google can NOT search within a table of HTML document*. We have proposed a novel architecture consisted of four layers which can modularize the activities. We have proposed to take care of the Web Form DOM as it will enable greater web information integration as discussed in our paper. Our Java based implementation used the classes URL, URLConnection, JEditorPane, HTML Swing Parser etc. To the best of our knowledge, such a integrated solution is not available anywhere although some other systems can only extract/export region of interest web data. Ours is an integrated solution, which we plan to extend further by the concepts as discussed in the paper.

## ACKNOWLEDGMENT

The work is partially supported by Wayne State University, Computer Science Department conference, travel fund.

## REFERENCES

- [1] P. Oaks, A. H. M. ter Hofstede, and D. Edmond, "Capabilities: Describing what services can do," in *ICSOC LNCS 2910*. Springer-Verlag Berlin Heidelberg, December 2003, pp. 1–16.
- [2] G. Z and H. Jamil, "Streamling biological data analysis using bioflow," in *International Symposium on Bio-Informatics and Biomedical Engineering*. IEEE, March 2003, pp. 258–262.
- [3] C. L. H. Jamil, and W. N., "Automatic composite wrapper generation for semi structured biological data based on table structure identification," *ACM SIGMOD*, vol. 33, no. 2, pp. 58–64, June 2004.
- [4] C. L. and H. Jamil, "Supporting remote user defined functions in heterogeneous biological databases," in *International Symposium on Bio-Informatics and Biomedical Engineering*. IEEE, November 2001, pp. 144–152.
- [5] C. L., H. Jamil, and W. N., "Automatic wrapper generation for semi structured biological data based on table structured identification," in *DEXA International Workshop on Biological Data Management*. IEEE, September 2003, pp. 1–10.
- [6] H. Jamil, "Achieving interoperability of genome databases through web mediators," in *International Symposium on Bio-Informatics and Biomedical Engineering*. IEEE, November 2000, pp. 118–125.
- [7] J. Yang and L. Xu, "Describing capabilities of internet data sources for information discovery and sharing," in *12th Australasian Database Conference*. IEEE, March 2001, pp. 1–8.
- [8] S. W. Liddle, S. H. Yau, and D. W. Embley, "On the automatic extraction of data from the hidden web," in *LNCS 2465*. Springer-Verlag London, UK, December 2001, pp. 212–226.
- [9] D. W. Embley, C. Tao, and S. W. Liddle, "Automating the extraction of data from html tables with unknown structure," *Data and Knowledge Engineering*, vol. 54, no. 1, pp. 3–28, November 2005.
- [10] J. P. Lage, A. S. da Silva, P. B. Golgher, and A. H. F. Laender, "Automatic generation of agents for collecting hidden web pages for data extraction," *Data and Knowledge Engineering*, vol. 49, no. 1, pp. 177–196, December 2004.
- [11] M. Liu and T. W. Ling, "A conceptual model for the web," in *19th International Conference on Conceptual Modeling (ER 2000)*. LNCS Springer-Verlag, October 2000, pp. 1–14.
- [12] —, "A data model for semistructured data with partial and inconsistent information," in *7th International Conference on Extending Database Technology (EDBT 2000)*. LNCS 1777 Springer-Verlag, March 2000, pp. 1–15.
- [13] S. Li, Z. Peng, and M. Liu, "Extraction and integration information in html tables," in *4th International Conference on Computer and Information Technology (CIT'04)*. IEEE Computer Society, September 2004, pp. 1–6.
- [14] C. Rivero and D. Ruiz, "An advanced query capability description framework implementation," in *International Conference on Enterprise Information System*. IEEE, April 2002, pp. 1–8.
- [15] J. Wang and F. H. Lochovsky, "Data extraction and label assignment for web databases," in *WWW*. ACM, 2003, pp. 1–13.
- [16] E. Ahmed and H. M. Jamil, "A survey on bioinformatics data and service integration using ontology and declarative workflow query language," in *PhD Qualifying Survey Report*. Department of Computer Science, Wayne State University, March 2007, pp. 1–60.
- [17] E. Ahmed, H. M. Jamil, and L. Pile, "Resource capability discovery and description management system for bioinformatics data and service integration - an experiment with gene regulatory networks," in *PhD Prospectus Technical Report*. Department of Computer Science, Wayne State University, November 2008, pp. 1–70.
- [18] E. Ahmed, "Bioinformatics web data and service integration - an experiment with gene regulatory networks," in *5th International Conference on Electrical and Computer Engineering*. ICECE, Technical Sponsor IEEE Bangladesh Section, December 20–22 2008.
- [19] V. Sugumaran and V. C. Storey, "Ontologies for conceptual modeling: their creation, use and management," *Journal of Data and Knowledge Engineering*, vol. 42, no. 1, pp. 251–271, June 2002.
- [20] A. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*. USA: McGraw Hill, 2005.
- [21] E. Ahmed, "Use of ontologies in software engineering," in *17th International Conference on Software Engineering and Data Engineering (SEDE)*. ISCA, Los Angeles, USA, June 30 - July 2 2008, pp. 145–150.
- [22] E. Ahmed, K. M. I. Asif, and M. Rahman, "Performance analysis of data intensive web application (diwa) - a case study," in *7th International Conference on Computer and Information Technology*. ICCIT, December 2004, pp. 226–231.

# A Data Hiding Capacity Estimation Technique for Biometric Images

Ghazaleh Sarbishei<sup>\*</sup> Saied Hossaini Khayat<sup>\*\*</sup>

<sup>\*</sup>Ferdowsi university of Mashad, Mashhad, Iran, *Ghsarbishei@gmail.com*

<sup>\*\*</sup>Ferdowsi university of Mashad, Mashhad, Iran, *shk@ieee.org*

**Abstract-** In this paper an estimation technique for the data hiding capacity in biometric images is presented. We consider the QSWT algorithm for data hiding in biometric images and investigate the effect of message strength in increasing capacity in the presence of different types of attacks.

**Index terms-** data hiding, image watermarking, DWT, QSWT, channel capacity.

## I- Introduction

Data hiding refers to embedding information within a host data such as text, audio, image or video. In watermarking applications, the hidden data represents authorship information, a time stamp or copyright information [1]. Data hiding in biometric images has been proposed to make the biometric systems secure and resilient to deliberate manipulations and attacks.

Data hiding capacity is a measure of the amount of information that can be hidden in a digital image while satisfying requirements such as robustness and invisibility. Estimating the capacity can set an upper-bound to the amount of data that can be hidden in an image, therefore it can help in the design of efficient data hiding algorithms. There has been ongoing research in data hiding capacity estimation in recent years. In [2], each pixel is considered as an independent channel and the capacity is calculated based on the theory of parallel Gaussian channels. Reference [3] presents an information theoretic model for data hiding and studies the capacity problem under several types of attacks. In [4], zero-error information hiding capacity in JPEG-compressed domain is presented. Reference [5], presents an analysis of watermarking capacity based on the content of wavelet sub-bands and uses the concept of wavelet quantization matrix and Noise Visibility Function.

The factors determining data hiding capacity include (1) the statistical model used for the host image, (2) the distortion constraints placed on the data hider and the attacker, and (3) the information available to the data hider, attacker and decoder. In fact, the content of an image has great influence on its data hiding capacity.

Data hiding capacity can also be influenced by the watermark strength, but high strength does not necessarily imply high capacity, as will be shown in simulation results. In previous works [2-5], capacity is estimated for data hiding in wavelet domain. In those works, the watermark strength is constrained based on the content of wavelet sub-bands.

In this paper, a framework for estimating the capacity of data hiding in biometric images based on QSWT algorithm is presented. Automated biometrics can provide accurate and reliable user authentication method. Watermarking biometric images can guarantee secure transmission of acquired images from intelligence agencies to a central image data base [6]. We assume that these watermarked biometric images are used for automatic user recognition and authentication. Since authentication will be done automatically, considerations of Human Visual System in watermarking can be ignored to some extent. In other words, in this proposed scheme, watermark (message) strength can be distributed equally among different wavelet sub-bands, regardless of their frequency contents.

Also in this paper, we compare the data hiding capacity in the presence of different attacks and discuss the effect of increased watermark strength on the capacity.

This paper is organized as follows. In section II, a mathematical model used for image watermarking is presented. Section III describes the QSWT approach for data hiding. Section IV gives capacity results for Gaussian channels. In section V, simulation results for biometric images are presented.

## II- Mathematical Model for Data Hiding

In data hiding schemes, a host image is considered as a communication channel for transmitting messages. Therefore, watermarking capacity can be estimated using traditional information theory. Fig. 1 shows a common model for watermarking in transform domain. In this model, the forward transform block decomposes the image into its coefficients of L bands. This forward transform can be either discrete cosine transform (DCT)

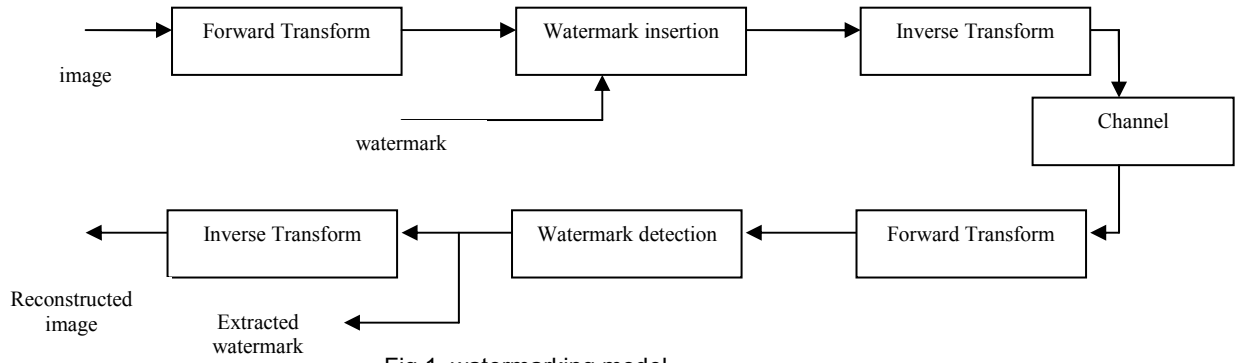


Fig 1. watermarking model

or discrete wavelet transform (DWT). Then, a watermark (message) is added to each band. The watermarked image is reconstructed using an inverse transform block. This image undergoes some certain type of processing (such as compression, addition of noise, median filtering) to yield the image. At the receiving end, watermark is extracted by decomposing this image into its  $L$  bands (using the same forward transform block) and the hidden message is extracted from each band. In this paper, we aim to estimate the capacity of such communication channel which is used to hide messages in biometric images such as fingerprint and iris images. In [9] a model for data hiding channel is presented, as shown in Fig.2 where,  $S$  is a watermark (message) to be transmitted through the channel. The channel noise is modeled as a combination of two sources:  $I$  the noise due to the host image and  $P$ , the noise due to processing. In this paper, we estimate the capacity for the case that image undergoes median filtering, JPEG compression, Gaussian filtering, sharpening and addition of Gaussian noise.

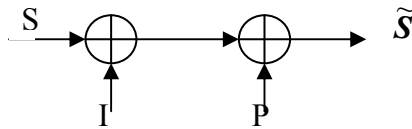


Fig 2. Data hiding channel [9]

### III- Channel Specification

In the above model for data hiding, we assume that the data hiding channel is a combination of  $L$  parallel sub-channels. The decomposition of the image is performed by applying a forward DWT transform. DWT transform is identical to a hierarchical sub-band system where the sub-bands are logarithmically spaced in frequency domain. The image is first decomposed into 4 parts of high, middle and low frequencies ( $LL1$ ,  $HL1$ ,  $LH1$ ,  $HH1$ ). The sub-bands  $LH1$ ,  $HL1$  and  $HH1$  are the finest scale wavelet coefficients. Sub-band  $LL1$  is further decomposed and this process is repeated several times. In some data hiding schemes, the watermark message is hidden in these sub-bands according to the magnitude of their coefficients. So each sub-band of the wavelet decomposition can be

considered as one of the sub-channels defined in our model.

In some schemes, the coefficient selection method is based on the Qualified Significant Wavelet Tree (QSWT) [7]. This method takes the relationships of DWT coefficients and spatial information into consideration and thus achieves more robustness. The principles of QSWT are as follows:

A parent-child relationship can be defined between wavelet coefficients at different scales corresponding to the same location. Excluding the highest frequency sub-bands, every coefficient at a given scale can be related to a set of coefficients at the next finer scale of similar orientation. The coefficient at the coarse scale is called the *parent*, and all coefficients corresponding to the same spatial location at the finer scales are called *children*. In other words, a family tree is constructed by grouping  $n$  levels of wavelet coefficients in the same spatial direction. The parent is a wavelet coefficient in one of the three highest frequency sub-bands  $HL_n$ ,  $LH_n$  or  $HH_n$  and its children are the ones located along the same spatial location in the next higher scales [8]. For a given parent, the set of all coefficients of all finer scales corresponding to the same orientation are called *descendants*. In Fig. 3 a wavelet tree consisting of all descendants of one coefficient in  $LH3$  is shown.

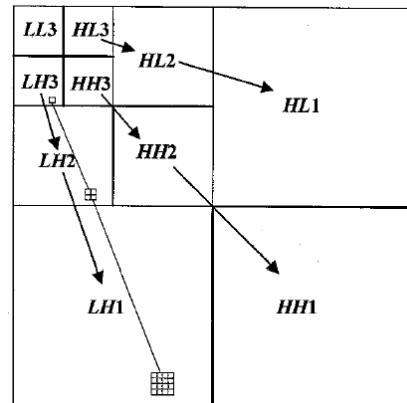


Fig. 3. DWT decomposition and all descendants of one coefficient in  $LH3$ .

Choosing an appropriate coefficient as a parent for data hiding is done based on some amplitude thresholds.

If a wavelet coefficient  $x_n(i, j) \in D$ , where  $D$  is a sub-band labeled  $HL_n, LH_n, HH_n$ , at the coarsest scale is a parent of  $x_{n-1}(p, q)$ , satisfying  $|x_n(i, j)| > T_1$ ,  $|x_{n-1}(p, q)| > T_2$  and,  $p = 2i - 1, \dots, 2i$ ,  $q = 2j - 1, \dots, 2j$ ,  $n, i, j > 1$ , for a given threshold  $T_1$  and  $T_2$ , then  $x_n(i, j)$  and its children are called a QSWT [7].

So in this case the elements of each sub-channel are the coefficients of the QSWT.

In this paper, we will show the effect of appropriate data hiding schemes in increasing the data hiding capacity. As will be shown in section V, we have estimated the data hiding capacity for two cases and compared estimated capacities.

#### IV- Channel Capacity

The capacity of an AWGN channel can be calculated according to

$$C = W \log_2(1 + \sigma_x^2 / \sigma_n^2)$$

in which  $W$  is the channel bandwidth,  $\sigma_x^2$  is the variance of the message signal to be transmitted and  $\sigma_n^2$  denotes the noise variance. In this work we assume that message is transmitted through  $L$  bands of transformed (DWT) image. So in an  $M \times N$  pixel image, we have  $(M \times N) / L$  coefficients. According to the Nyquist sampling theory, sampling points should be  $2W$ , so the bandwidth of an image will be  $M \times N / 2$  [5].

The decomposition of an image into  $L$  sub-bands results in  $L$  parallel sub channels with two noise sources in each sub-channel. Since we have no knowledge about the distribution of these noises, we assume the worst Gaussian distribution [10]; we assume that these two noises are Gaussian and independent. So the channel noise can be replaced by one noise with variance  $\sigma_i^2 + \sigma_p^2$  where  $\sigma_i^2$  is the variance of image noise and  $\sigma_p^2$  is the variance of processing noise. Thus the total capacity of  $L$  parallel channels is given by [9]:

$$C = \frac{MN}{2L} \sum_{j=1}^L \log_2 \left( 1 + \frac{\sigma_{x_j}^2}{\sigma_{i_j}^2 + \sigma_{p_j}^2} \right) \quad (1)$$

In most watermarking applications,  $\sigma_x^2$  is calculated according to perceptibility requirements. For instance in [9]  $\sigma_{x_j}^2$  is replaced by the visual threshold of band  $j$ . In other words,  $\sigma_{x_j}^2$  is the maximum message signal energy permitted in band  $j$  based on its perceptual quality effects. This parameter also depends on the magnitude of the particular coefficient. Due to the properties of human visual system, a coefficient with larger magnitude can be altered to a larger extent than a coefficient with smaller magnitude. Our goal is to have an estimate of the average energy that can be added to a particular band. Since the human visual system is more

sensitive to lower frequency bands than higher ones, and lower frequency bands have higher variance, a reasonable model for the visual threshold can be [9]

$$\sigma_{x_j}^2 = k \sigma_{i_j}^{2\alpha} \quad (2)$$

where  $0 \leq \alpha \leq 1$ , and  $k \ll \sigma_{i_j}^2$ , and  $\sigma_{i_j}^2$  is the variance of the coefficients in band  $j$ .

In some scenarios such as automatic biometric user authentication, where recognition is done automatically, human perception will be of less importance. In these cases, the energy of message signal can be distributed equally among different bands regardless of their variances. In this paper, we set  $\alpha = 0$  and calculate the capacity for different values of  $k$ .

Channel noise is the sum of two noise sources, image noise with variance  $\sigma_i^2$  and processing noise with variance  $\sigma_p^2$ . In order to model the image noise, we assume that  $f_{i_j}(i_j)$  is the distribution of the  $j^{\text{th}}$  sub-band of the host image. The image noise is split into its components in  $L$  sub-bands which are modeled as random variables with variance  $\sigma_{i_j}^2$ . The next step is to obtain their entropy equivalent Gaussian variances (or the Gaussian random variable that has the same entropy as these random variables). This is achieved by plotting a histogram of the coefficients of each band and calculating the entropy. If  $\Delta x$  is the width of  $n$  bins of the histogram  $g_i(m), m = 1, 2, \dots, n$  and  $p$  is the total number of coefficients in band  $j$ , the entropy and the equivalent Gaussian variance  $\sigma_{i_j}^2$  are obtained as

$$H_j = - \sum_{i=1}^n \frac{g_j(i)}{p \Delta x} \log_2 \left( \frac{g_j(i)}{p \Delta x} \right) \Delta x \quad (3)$$

$$\sigma_{i_j}^2 = 2^{2H_j / \ln 2} / 2\pi e$$

Thus the image noise in sub-band  $j$  can be replaced by a Gaussian noise of variance  $\sigma_{i_j}^2$  [9]. Note that this noise source can be omitted in non-blind approaches, since the image noise can be subtracted from the received image. Therefore, we expect such schemes to have higher capacity than blind watermarking approaches.

The processing noise is estimated as the variance of an equivalent additive noise which substitutes the actual (nonlinear) processing noise sources [9], such as JPEG compression and gamma correction. We define processing noise as the equivalent additive noise which accounts for the reduction in correlation between the transform coefficients of the original image and the transform coefficients of the image obtained after processing. We assume the processing noise in each sub-band as  $\sigma_{p_j}^2, j = 1, \dots, L$ . In order to obtain this variance, we apply the processing to  $n_i$  test images and decompose both the original and processed images into  $L$  sub-bands. Thus we will obtain  $MNn_i / L$  samples for each sub-band. Let  $c_{j_k}, k = 1, \dots, (MNn_i) / L$  be the coefficients of band  $j$  of test images and

$\tilde{c}_{j_k}, k=1, \dots, (MNn_i)/L$  be the corresponding coefficients of the images subjected to processing. The variance of processing noise is obtained as follows:

$$\frac{\langle c_j, \tilde{c}_j \rangle}{|c_j| |\tilde{c}_j|} = \frac{\langle c_j, (c_j + n_j) \rangle}{|c_j| |c_j + n_j|} = \rho_j \quad (4)$$

$$\sigma_{p_j}^2 = \left( \frac{1}{\rho_j^2} - 1 \right) |c_j|^2$$

In the next section, image noise and processing noise are estimated for fingerprint and iris images based on the above equations by means of simulations.

## V- Experimental Results

We estimate data hiding capacity for two databases of fingerprint images (364×328 pixels) and iris images (576×768 pixels). The images were decomposed into  $L=10$  sub-bands by applying two dimensional wavelet transform. We used 4-tap Daubechies filters for wavelet decomposition of images. We have assumed that the watermarked image may be subjected to different types of attacks including median filtering, JPEG compression, Gaussian filtering, sharpening and addition of Gaussian noise. We estimated the noise due to these attacks and figured out that in order for the images to resist these attacks how many bits can be hidden in them. In other words, acceptable capacities for these images were estimated in order to be robust against different attacks. Since we have focused on data hiding for automatic user authentication schemes, we are able to eliminate the effect of human visual system on designing watermarking algorithms. Thus we choose  $\alpha = 0$  in Equation (2) so that message energy is distributed equally among different sub-bands. The constant  $k$  is chosen to be smaller than the variance of the coefficients of each sub-band.

Also since we are targeting user authentication applications, we use non-blind data-hiding, as a result we have omitted the variance of image noise  $\sigma_i^2$ . The variance of processing noise  $\sigma_p^2$  is estimated according to Equaion (4). So the capacity of data hiding channel is given by

$$C = \frac{MN}{2L} \sum_{j=1}^L \log_2 \left( 1 + \frac{k}{\sigma_{p_j}^2} \right) \quad (5)$$

The estimated capacity for different values of  $k$  in the presence of different kind of attacks is shown in Fig.4 and Fig.5 for the following two cases:

- Case A: We have assumed that all coefficients of each sub-band can be used for data hiding.
- Case B: We have assumed only the coefficients of QSWT are considered as data hiding channels.

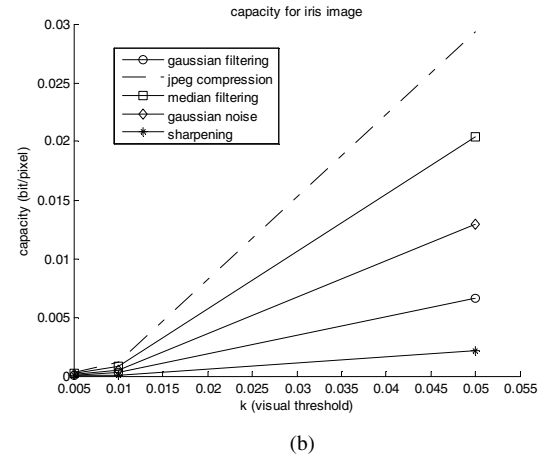
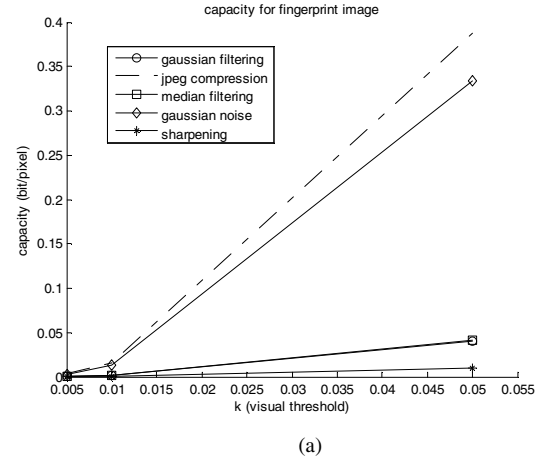


Fig 4. Capacity estimates (case A) a) fingerprint images, b) iris images

It is shown that in the first case, fingerprint images provide more capacity for data hiding in comparison with iris images. In our simulations, constant  $k$  indicates the energy of the message to be hidden. It is shown that increasing  $k$  has no effect in increasing capacity in the presence of sharpening attacks. In other words, increasing message energy is not appropriate for achieving higher capacity in these cases. This is true because sharpening is a kind of high pass filtering and we assumed that the data is hidden in mid and high frequency bands. Also by using QSWT algorithm, the total capacity will increase. This is more obvious for iris images.

## VI- Conclusion

We have presented a technique for estimating the data hiding capacity of biometric images for user authentication applications. Since human perception is of less importance in such scenarios, the message energy can be distributed equally among different sub-bands of wavelet transform. Our experimental results show that increasing message energy will increase data hiding capacity, although other properties such as visual

perception may decrease. Also it was shown that the QSWT algorithm will increase the capacity.

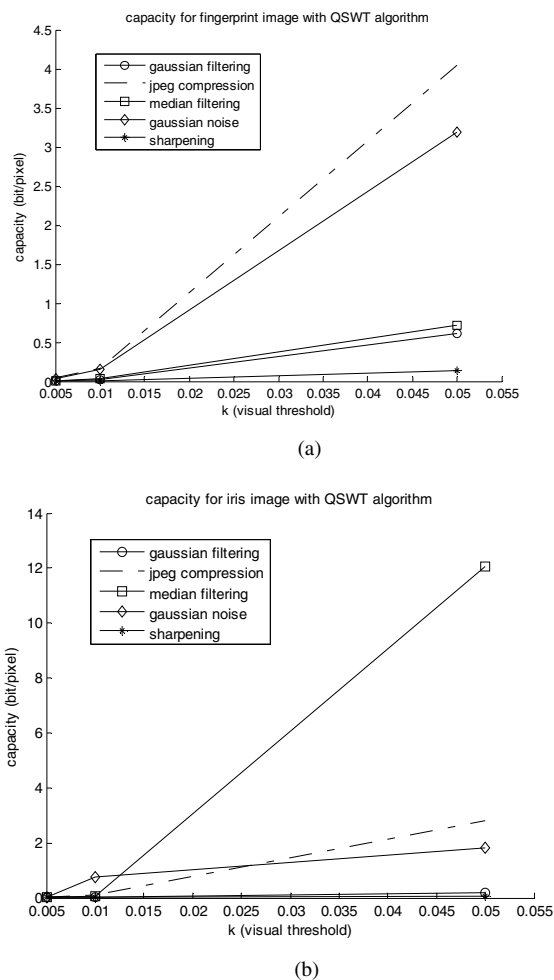


Fig. 5. Capacity estimates (case B) for a) fingerprint image b) iris image

Also increasing message energy in the presence of compression and Gaussian noise will cause a significant increase in the capacity but in the case of sharpening and filtering message energy does not have any significant effect on data hiding capacity.

#### References

- [1] P. Moulin, M. K. Mihcak, "A framework for evaluating the data-hiding capacity of image sources", *IEEE Trans. Image Processing*, Vol. 11, No. 9, September 2002.
- [2] S. D. Servetto, C. I. Podilchuk, K. Ramchandran, "Capacity issues in digital image watermarking", *IEEE Intl. Conf. on Image processing*, 1998, volume 1.
- [3] P. Moulin, M. K. Mihcak, "A framework for evaluating the data hiding capacity of image sources", *IEEE Trans. Image processing*, vol. 11, no. 9, 2002.

- [4] C. Y. Lin, S. F. Chang, "Zero-error information hiding capacity of digital images", *IEEE Intl. conf. on image processing*, vol.3 2001.
- [5] Z. Fan, Z. Hongbin, "Wavelet domain watermarking capacity analysis", *proc. SPIE Vol. 5637*, 2005.
- [6] N. K. Ratha, J. H. Connel, R. M. Bolle, "Enhancing security and privacy in biometric-based authentication systems", *IBM Systems Journal*, Vol. 40, No. 3, 2001.
- [7] M. Hsieh, D. Tseng, Y. Huang, "Hiding digital watermarks using multiresolution wavelet transform", *IEEE Trans. On industrial electronics*, vol. 48, No. 5, Oct 2001.
- [8] B. McKinnon, X. Qi, "Adaptive wavelet-based family tree quantization for digital image watermarking",
- [9] M. Ramkumar, A. N. Akansu, "Capacity estimates for data hiding in compressed images", *IEEE Trans. Image Processing*, Vol. 10, No. 8, August 2001.
- [10] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, 2<sup>nd</sup>, Ed. New York, Wiley, 1991.

# Regression Diagnostics in Large and High Dimensional Data

A. A. M. Nurunnabi<sup>†</sup>, and Mohammed Nasser<sup>‡</sup>

<sup>†</sup>School of Business, Uttara University, Dhaka-1230, Bangladesh,

E-mail: pyal1471@yahoo.com

<sup>‡</sup>Department of Statistics, Rajshahi University, Rajshahi-6205, Bangladesh,

E-mail: mnasser.ru@gmail.com

**Abstract** — “Learning methods” play a key role in the fields of statistics, data mining, and artificial intelligence, intersecting with areas of engineering and other disciplines. These methods for analyzing and modeling data come in two flavors: supervised and unsupervised learning. Regression analysis and classification are two well known supervised learning techniques. To get an effective model from regression analysis it is necessary to check and preprocess the data set in astronomy, bio-informatics, image analysis, computer vision etc, especially when the data sets are large and high dimensional. In these industries large or fat data appear with unusual observations (outliers) very naturally. Checking raw data for outliers in regression is regression diagnostics. Most of the popular diagnostic methods are not good enough for large and high dimensional data. The aim of this paper is to provide a new measure for identifying influential observations in linear regression for large high dimensional data.

**Index Terms** — Data mining, high dimensional data, influential observation, outlier, regression diagnostics, supervised learning.

## I. INTRODUCTION

Data and knowledge mining/discovery is learning from data. Learning means extraction of the regularity from the data. Learning methods for analyzing and modeling data are divided into two: “supervised learning” and “unsupervised learning” to the pattern recognition and machine learning community. Classification, decision trees, neural networks, and regression analysis are well known supervised learning methods. In recent years regression analysis has drawn much attention for its wide range of applicability to the user of learning community. Regression can be performed using many different types of techniques, including classification, genetic algorithms, neural networks and optical flow estimation.

Supervised learning requires input data that has both predictor (independent) variables and target (dependent/response) variable, whose value is to be estimated. By various means, the process “learns” how to model (predict) the value of the target variable based on the predictor variables. Regression analysis is one of the most important supervised learning methods that have the

objective to specify a relationship between the target and explanatory (independent) variables. Supervised learning requires that the target variable is well defined and that a sufficient number of its values are given. Data contamination is a practical problem, introduced either due to unreliable data sources or during data transmission. Noise (data contamination) can greatly impair data learning. It is often more severe problem in the context of streams (large and high dimensions) because it interfaces with change adaptation. Regression analysis is not an exception to this kind of problem and regression diagnostics is a remedy. Problem is that most of the identification techniques are performed well only for small and few dimensional data.

In this paper we have a brief introduction of outliers in large and high-dimensional data in Section II, Section III focused on regression diagnostics. We propose a new diagnostic measure for identifying multiple influential observations for high dimensional large data set in Section IV. Numerical demonstrations are thereafter followed by conclusion in Section VI.

## II. OUTLIERS DETECTION IN LARGE AND HIGH DIMENSIONAL DATA

Outlier detection has been suggested for numerous applications, such as credit and fraud detection, clinical trials, voting irregularity analysis, network intrusion, severe weather prediction, geographic information system, and other data mining tasks [1], [2]. Outlying observations do not inevitably ‘perplex’ or ‘mislead’; they are not necessarily ‘bad’ or ‘erroneous’, and the experimenter may be tempted in some situations not to reject an outlier but to welcome it as an indication of new and important findings.

Outlier detection methods can be divided into two, univariate and multivariate methods. Another fundamental taxonomy of outlier detection methods is between parametric and nonparametric methods that are model-free (e.g., [3]). Statistical parametric methods either assume a known underlying distribution of the

observations or, at least, they are based on statistical estimates of unknown distribution parameters. These methods flag as outliers those observations that deviate from the model assumptions. They are often unsuitable for high-dimensional data sets and for arbitrary data sets without prior knowledge of the underlying data distribution [4]. Within the class of non-parametric methods one can set apart the data mining methods, also called distance-based methods. These methods are usually based on local distance measures and are capable of handling large databases (*e.g.*, [3], [5], [6]). Another class of outlier detection methods is founded on clustering techniques, where a cluster of small sizes can be considered as clustered outliers [7], [8].

### III. REGRESSION DIAGNOSTICS

The purpose of regression (model function) is to map a data item to a real valued function prediction variable, classical model is

$$Y = X\beta + \varepsilon. \quad (1)$$

We predict the output  $Y$  with least squares method via the model,

$$\hat{Y} = X\hat{\beta} \text{ (in vector-matrix form);} \quad (2)$$

where  $\hat{\beta} = (X^T X)^{-1} X^T Y$  and the corresponding

residual vector,  $r = Y - \hat{Y} = (I - H)\varepsilon$ ; in scalar form,

$$r_i = \varepsilon_i - \sum_{j=1}^n h_{ij} \varepsilon_j; \quad i = 1, 2, \dots, n$$

where  $H = X(X^T X)^{-1} X^T$  is the leverage matrix and its  $j$ th diagonal element is  $h_{jj}$ . Usually  $r_i$  and  $h_{ij}$  along with their different transformed versions are used for the diagnosis of outliers and leverage points respectively.

In regression, outliers can deviate into three ways (i) the change in the direction of response ( $Y$ ) variable (ii) the deviation in the space of explanatory variable(s), deviated points in  $X$ -direction called leverage points, and (iii) the other is change in both the directions. It is known that as with outliers, high leverage points need not be influential, and influential observations are not necessarily high-leverage points. According to Belsley *et al.* [9], influential observation is one which either individual or together with several other observations has a demonstrably larger impact on the calculated values of various estimates than is the case for most of the other observations. Robust regression and regression diagnostics are two complementary approaches to study with unusual observations. Robust regression first wants to fit a regression to the majority of the data and then to discover outliers. On the contrary, field diagnostics is a combination of graphical and numerical tools. It is

designed to detect and delete (if necessary recheck/correct) and then fit the good data by classical methods. It tries to iteratively detect possibly wrong data and reject/refit them through analysis of globally fitted model.

A large body of literature is now available (see [9-12]) for the identification of outliers and influential observations. Among them Cook's distance [13], DFFITS and DFBETAS [9] become very popular with the practitioners for identifying influential observations. These are based on single case deletion and unable to cope with multiple unusual observations. There are some group deletion measures (see [14]) but they are failed in case of high dimensional large data sets.

High dimensional data introduces several problems to traditional statistical analysis. Computation time increases more rapidly with number of explanatory variable than the number of observations. Another problem for group deletion diagnostics is to find group of outlying observation  $d$ , the deletion of which produces the largest reduction to the residual sum of squares.

### IV. NEW DIAGNOSTIC MEASURE

We introduce a two-step diagnostic measure originated from Pena's [15] idea attaching with formal and/or informal group deletion methods (see [14]). Group deletion helps us to reduce the maximum disturbance by deleting the suspect group of influential cases at a time and 'the deletion of which produces the largest reduction in the residual sum of squares. It helps to make the data more homogeneous. We look how every point in a sample is influenced by the specific ( $i$ th) observation. In our method, after deleting the suspect group at a time we again delete specific sample point from each of the observations one after another and measure the difference between the forecast with and without the specific one. Group deletion attaching with Pena's [15] idea improves the measure and show nice results in presence of masking and/or swamping phenomena. Graphical displays like index plot and character plot of explanatory and response variables could give us some ideas about the influential observations, but these plots are not useful for higher dimension of regressors. There are some suggestions in the literature for using robust regression techniques like LMS, LTS, reweighted least squares (RLS) (see [10]) for finding the group of suspect influential observations. We consider one of the two proposed procedures of Hadi and Simonoff [14], where they suggested dividing the data set in two initial subsets: a 'basic' subset that contains the first  $k+1$  clean observations and a 'nonbasic' subset that contains the remaining observations.

At the first step we find out all suspect influential (unusual) cases. We suggest using graphical display and/or regression diagnostic measures and/or robust regression methods



whatever can helps to find the suspect group with maximum number of influential cases.

In the second step, we assume that  $d$  observations among a set of  $n$  observations are suspected as influential observations and to be deleted. Let us denote a set of cases 'remaining' in the analysis by  $R$  and a set of cases 'deleted' by  $D$ . Hence without loss of generality, assume that these observations are the last  $d$  rows of  $X$  and  $Y$  so that

$$X = \begin{bmatrix} X_R \\ X_D \end{bmatrix} \quad Y = \begin{bmatrix} Y_R \\ Y_D \end{bmatrix}.$$

After formation of the deletion set indexed by  $D$ , we would compute the fitted values  $\hat{Y}^{(-D)}$ . Let  $\hat{\beta}^{(-D)}$  be the corresponding vector of estimated coefficients when a group of observations indexed by  $D$  is omitted. We define the vector of difference between  $\hat{y}_j^{(-D)}$  and  $\hat{y}_{j(i)}^{(-D)}$  as

$$t_{(i)}^{(-D)} = (\hat{y}_1^{(-D)} - \hat{y}_{1(i)}^{(-D)}, \dots, \hat{y}_n^{(-D)} - \hat{y}_{n(i)}^{(-D)})^T \quad (3)$$

$$= (t_{1(i)}^{(-D)}, \dots, t_{n(i)}^{(-D)})^T \quad (4)$$

and

$$t_{j(i)}^{(-D)} = \hat{y}_j^{(-D)} - \hat{y}_{j(i)}^{(-D)} = \frac{h_{ji} \hat{\epsilon}_i^{(-D)}}{1 - h_{ii}}, j=1, 2, \dots, n \quad (5)$$

where

$$h_{ji} = x_j^T (X^T X)^{-1} x_i \text{ and } \hat{\epsilon}_i^{(-D)} = y_i - \hat{y}_i^{(-D)}.$$

Finally, we introduce our new measure as squared standardized norm,

$$M_i = \frac{t_{(i)}^{(-D)T} t_{(i)}^{(-D)}}{kV(\hat{y}_i^{(-D)})}, \quad (6)$$

where

$$V(\hat{y}_i^{(-D)}) = s^2 h_{ii} \quad \text{and} \quad s^2 = \frac{\hat{\epsilon}^{(-D)T} \hat{\epsilon}^{(-D)}}{n - k}.$$

Using (3) to (6), we obtain

$$M_i = \frac{1}{ks^2 h_{ii}} \sum_{j=1}^n h_{ji}^2 \frac{\hat{\epsilon}_i^{(-D)2}}{(1 - h_{ii})^2}. \quad (7)$$

We use a confidence bound type cut-off value for it, and consider the  $i$ th observation to be influential if it satisfies the rule

$$|M_i| > \text{median}(M_i) + 4.5 \text{MAD}(M_i), \quad (8)$$

where

$$\text{MAD}(M_i) = \text{median}\{|M_i - \text{median}(M_i)|\} / 0.6745.$$

## V. NUMERICAL EXAMPLES

We consider here two artificial large data sets, one is multi-structural clustered (step) and the other is high-dimensional, for the identification of influential observations. Such data sets are frequently used in classification, cluster analysis and optical flow estimation. We want to show the performance of our new measure and compare with the existing popular diagnostic methods.

### A. Multi Structural Step Data

Our first example uses an artificial data set that is taken from (see [16]) a paper of optical flow estimation (computer vision). In this experiment, we investigate how a multistructural clustered data is analyzed which is contaminated by number of outliers (noise). We generate the step data set, since every line is corrupted by Gaussian noise with zero mean and different variances. The  $i$ th line has  $n_i$  samples,  $n_0$  data points are randomly distributed in the range of (0,100). The data points are defined as follows.

Steps are:  $X = (0, 55)$ ,  $Y = 30$ ,  $n_1 = 120$ ,  $\sigma = 1.5$  ('+');

$X = (20, 100)$ ,  $Y = 60$ ,  $n_2 = 50$ ,  $\sigma = 2$  ('-');  $n_0 = 50$  ('o').

First we plot the points with LS and robust (LMS and LTS) line, it is clearly shown that data sets are clustered on two lines and some observations do not hold any pattern. As the common definition, observations those are assigned 'o' and '-' in Fig.1 can be treated as outliers (apart from the bulk of the data). Judging from the LMS line it is worth mentioning that last 100s are outliers. Now we apply our newly proposed diagnostic technique and form the deletion set  $D$  by the hundreds. Index plots (Fig. 2. (b) and (d)) of residuals and leverage points do not show any indication of heterogeneity.

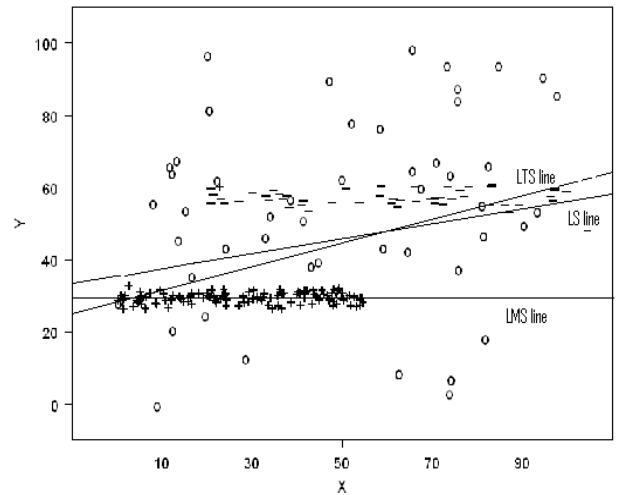


Fig.1. Scatter plot of multi structural step data.

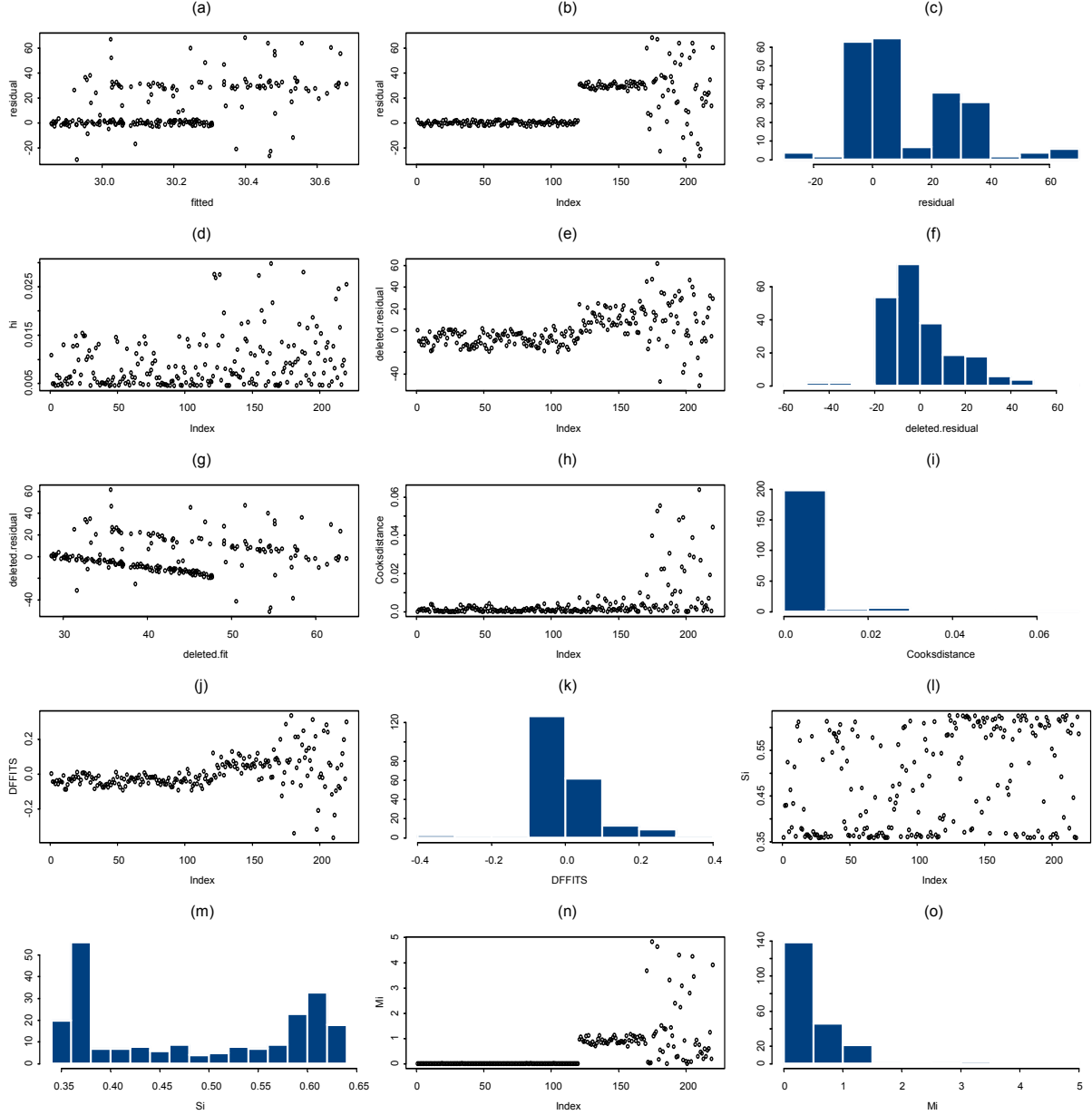


Fig. 2. Regression diagnostics for multistructural step data; (a) Plot of residual versus fitted value (b) Index plot of residuals (c) Histogram of residuals (d) Index plot of leverage,  $h_i$  values (e) Index plot of deleted residuals (f) Histogram of deleted residuals (g) Deleted residual versus deleted fit (h) Index plot of Cook's distance (i) Histogram of Cook's distance (j) Index plot of DFFITS (k) Histogram of DFFITS (l) Index plot of  $S_i$  (Pena) (m) Histogram of  $S_i$  (n) Index plot of  $M_i$  and (o) Histogram of  $M_i$ .

We compute the diagnostic measures Cook's distance and DFFITS along with Pena's measure ( $S_i$ ) and  $M_i$  (new measure) and use them to draw index (dot) plot and histogram. Without  $M_i$ , rest 3 are failed to indicate

influential cases properly. Index plot of  $M_i$ , Fig. 2 (n) shows that last 100 observations are influential with a

few masking observations. Histogram of  $M_i$ , Fig. 2 (o) clearly indicates the presence of three different patterns of observations.

### B. Large and High Dimensional Data

We consider an artificial data set that is generated in a similar fashion described by Pena [15]. The data set shows the presence

of heterogeneous variances in the sample points and a mixture of continuous and categorical variables, generated by the model

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{20} X_{20} + \beta_{21} Z + \varepsilon. \quad (9)$$

We generate 500 observations, where  $X$ 's have 20 dimensions and they are independent random drawings from uniform distributions. The first 400 observations for each of the  $X$  variables are generated from Uniform (0, 10)

and last 100 observations (401-500) from Uniform (9, 10) that makes the presence of heterogeneous variance in the data set. For the categorical explanatory variable  $Z$ , the first 400 observations are set at  $z = 0$  and the last 100 observations are set at  $z = 1$ . For the null model we generate errors from Normal (0, 1). The parameter values have been chosen as

$\beta_0 = \beta_1 = \dots = \beta_{20} = 1$ , and  $\beta_{21} = -100$  so that the standard diagnostics do not show any evidence of heterogeneity.

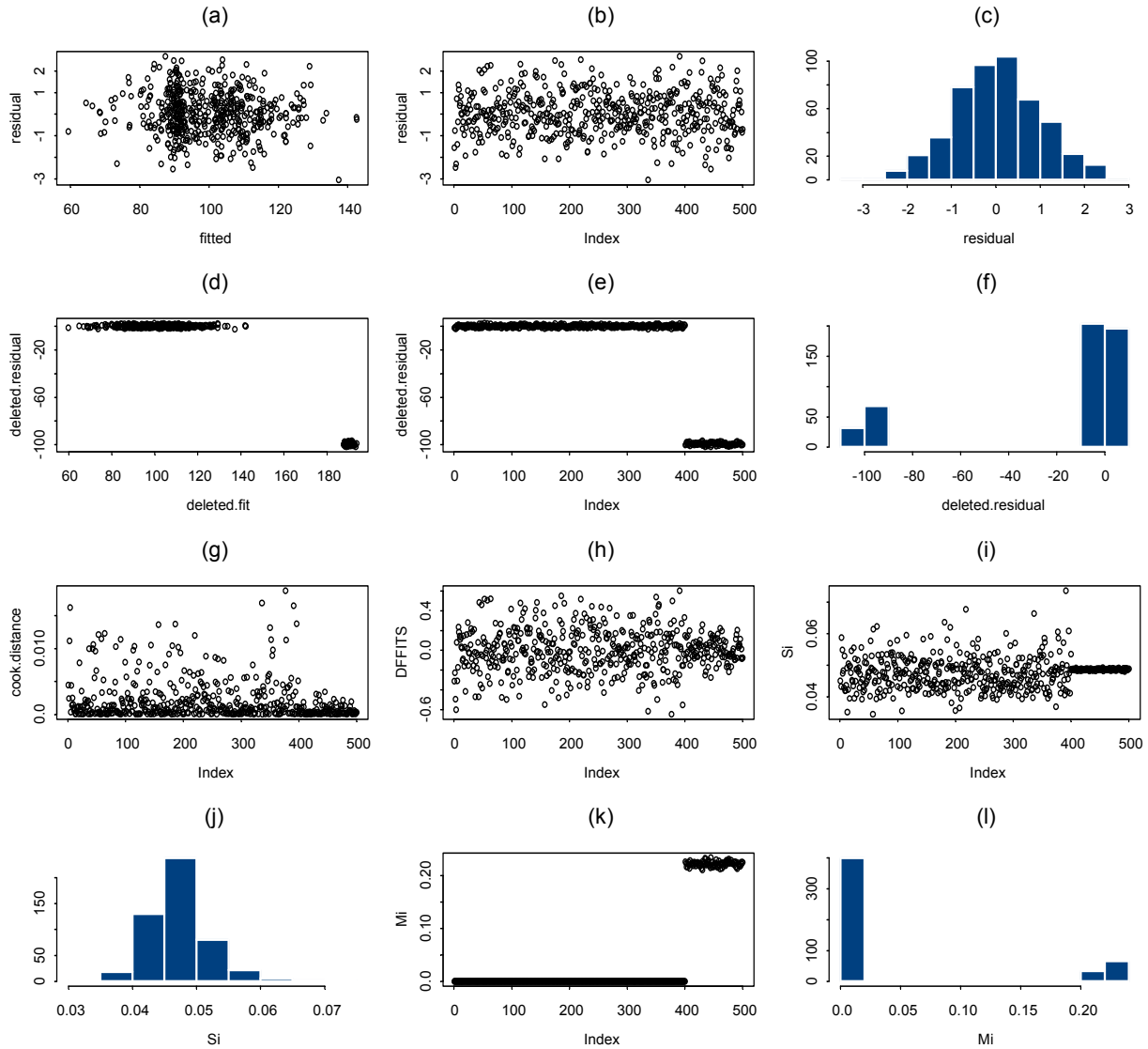


Fig. 3. Influence analysis of the large and high dimensional data; (a) Plot of residual versus fitted value (b) Index plot of residuals (c) Histogram of the residuals (d) Plot of deleted residual versus deleted fitted value (e) Index plot of deleted residuals (f) Histogram of the deleted residuals (g) Index plot of Cook's distance (h) Index plot of DFFITS (i) Index plot of  $Si$  (j) Histogram of  $Si$  (k) Index plot of  $Mi$ , and (l) Histogram of  $Mi$ .

Results that we obtain from this artificial data are shown in Fig. 3. The residual versus fitted value plot, Fig. 3 (a)

shows no indication of heterogeneity. When the last 100 observations are deleted, the deleted residual versus deleted fitted value plot and histogram of deleted residual ( Fig. 3. (d), (f) ) show clear indications of the presence of heterogeneity. We consider above mentioned three measures to identify the 100 influential cases but their index plots (Fig. 3. (g), (h), and (i)) show they are totally failed to identify the influential cases. In index plot, Fig. 3. (k) we see that proposed  $M_i$  can successfully identify all influential cases. Fig. 3. (l), histogram of  $M_i$  also shows the heterogeneous variances in the data set that is absent in histogram of  $S_i$ .

## VI. CONCLUSION

In this article, we propose a group deletion diagnostic measure for identifying influential observations (outliers) in linear regression that performs well for large and high dimensional data. The technique also gives some clear indications of presence of influential cases in a multidimensional step (cluster) data when the existing popular diagnostic methods are not well enough. To get real data involving thousands of cases and hundreds of dimensions is not easy and also expensive. Time and space constrains do not permit us to perform the tasks this time; we shall try to do it as a further work.

## ACKNOWLEDGEMENT

The authors are thankful to the anonymous reviewers for their valuable suggestions and to the organizing and publication committee of the conference.

## REFERENCES

- [1] T. Fawcett, and F. Provost, "Adaptive fraud detection," *Data Mining and Knowledge Discovery*, Vol.1, No. 3, pp. 291-316, 1997.
- [2] K. L. Penny, and I. T. Jolliffe, "A comparison of multivariate outlier detection methods for clinical laboratory safety data," *The Statistician*, Vol. 50, No. 3, pp. 295-308, 2001.
- [3] G. L. Williams, R. A. Baxter, H. X. He, S. Hawkins, and L. Gu, "A comparative study of RNN for outlier detection in data mining," *IEEE International Conference on Data Mining (ICDM'02)*, Macbashi, Japan, 2002.
- [4] S. Papadimitriou, H. Kitawaga, P. G. Gibbons, and C. Faloutsos, "LOCI: Fast outlier detection using the local correlation integral," *International. Research Laboratory Technical report*, No. IPR=TR-02-09, 2002.
- [5] S. Hawkins, H. X. He, G. J. Williams, and R. A. Baxter, "Outlier detection using replicator neural networks," in *Proceedings of the Fifth International Conference on Data Warehousing and Knowledge Discovery*, France, 2002.
- [6] E. Knorr, R. Ng, and V. Tucakov, "Distance based outliers: algorithms and applications," *VLDB Journal: Very Large Data Bases*, Vol. 8(3-4), pp. 237-253, 2000.
- [7] E. Acuna, and C. Rodriguez, "A meta analysis study of outlier detection methods in classification," Technical Paper, Department of Mathematics, University of Puerto Ricoat Mayaguez, <http://www.academic.uprm.edu/~eacuna/paperout.pdf>, in *Proceedings, IPSI*, 2004.
- [8] S. Shekhar, C. T. Lu, and P. Zhang, "Detecting graph based spatial outlier," *Intelligent Data Analysis: An International Journal*, Vol. 6, No. 5, pp. 451-468, 2002.
- [9] D. A. Belsley, E. Kuh, and R. E. Welsch, *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York: J. Wiley & Sons, 1980.
- [10] P. J. Rousseeuw, and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: J. Wiley & Sons, 1987.
- [11] S. Chatterjee, and A. S. Hadi, *Sensitivity Analysis in Linear Regression*, New York: J. Wiley & Sons, 1988.
- [12] A. C. Atkinson, and M. Riani, *Robust Diagnostic Regression Analysis*, New York: Springer, 2000.
- [13] R. D. Cook, "Detection of influential observations in linear regression", *Technometrics*, Vol.19, pp.15-18, 1977.
- [14] A. S. Hadi, and J. S. Simonoff, "Procedure for the identification of outliers in linear models," *Journal of American Statistical Association*, Vol. 88, pp. 1264 - 1272, 1993.
- [15] D. Pena, "A new statistic for influence in linear regression," *Technometrics*, Vol. 47, pp.1-12, 2005.
- [16] J. Colliez, F. Dufrenois, and D. Hamad, Robust regression and outlier detection with SVR: application to optical flow estimation, <http://www.macs.hw.ac.uk/bmvc2006/papers/446.pdf>, access 8<sup>th</sup> March, 2008.

# Comparisons of Maximum System Lifetime in Diverse Scenarios for Body Sensor Networks

<sup>1</sup> Md. Nazrul Islam Mondal <sup>2</sup> Md. Shahid Uz Zaman, <sup>3</sup> Abdullah Al Masud, <sup>4</sup> Md. Jahangir Alam

<sup>1</sup> Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Bangladesh, Email: [nimbd@yahoo.com](mailto:nimbd@yahoo.com), Ph. & Fax: +880-721-750838

<sup>2</sup> Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Bangladesh, Email: [szaman22@yahoo.com](mailto:szaman22@yahoo.com), Ph. & Fax: +880-721-750838

<sup>3</sup> Department of Electronic and Telecommunication, International Islamic University Chittagong, Bangladesh, Email: [iammasud@yahoo.com](mailto:iammasud@yahoo.com)

<sup>4</sup> Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Bangladesh, Email: [jahangir01326cse@yahoo.com](mailto:jahangir01326cse@yahoo.com), Ph. & Fax: +880-721-750838

**Abstract** — In this paper, the main target is to compare the system lifetimes in diverse scenarios based on our new energy dissipation model in the body sensor network as well as maximize the system lifetime of sensor nodes when they will make communication among body sensors and personal communication unit. Nowadays, the ultra low energy consumption is the very much important challenge for medical applications. The best compression technique like LPC is selected for energy saving based on some calculations. This research work studies and analyzes the transceiver energy consumption for different compression algorithms and select the best technique like LPC for energy saving. Formulation of a linear programming problem is also the important part of this research work, where is to maximize the system lifetime which is equivalent to the time until the first node runs out of battery. Maximum system lifetimes are calculated by MATLAB optimization technique using and without using efficient compression algorithm like LPC in various environments. Results show that maximum system lifetimes calculated in different scenarios using efficient compression technique like LPC is better than without using compression technique.

**Index Terms** — WBSN, transmitting energy and receiving energy, linear prediction coding (LPC), maximum system lifetime, linear problem, low energy consumption.

## I. INTRODUCTION

Consider a wireless body sensor network of static nodes as depicted in Fig.1, where each node operates on limited battery energy. Assume that each node generate constant bit rate and send or receive information to or from neighbors with and without compression technique. Also assume that the patient's body sensor (one sensor node for every patient to detect heart beat signal) is fixed and its distribution is shown in the Fig.1. In this paper, we select the best compression algorithm like LPC and calculate the transmitting and receiving energy according to the new energy dissipation model shown in Fig.4. Then we formulate the linear programming problem where

objective is to maximize the system lifetime. Finally compare the results with the help of optimization technique and show that maximum system lifetimes calculated in various scenarios using efficient compression technique like LPC is better than without using compression technique. Most of the previous works deal with low energy consumption for WSN but they did not maximize the system lifetime with the help of optimization technique using energy efficient LPC algorithm. In [1], the authors described the application of WSN just to monitor the battle field. In [2], the authors described the WBAN architecture for broader telemedical system using energy efficient local communication protocols such as ZigBee, Wi-Fi.

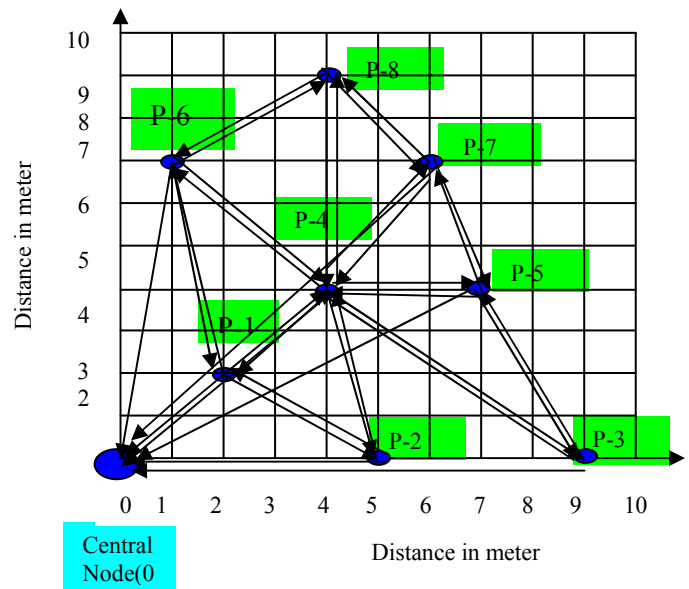


Fig.1. Application scenario for maximum system lifetime

The authors in [3] explained the energy aware network layer protocols like LEACH, CICADA, and S-MAC for

maximize the battery lifetime. The work in [4] considered multi destination case instead of single destination case for maximizing the system lifetime. In [5], instead of minimizing the total consumed energy, they tried to maximize the network lifetime till first node dies.

## II. OBJECTIVES OF THE RESEARCH

The objectives of the research work are written in below:

- To study the energy consumption parameters in the sensor networks.
- To analyze the different compression techniques to select the best one for energy consumption in the sensor nodes
- To formulate the problem as a linear programming problem where the objective is to maximize the minimum system lifetime of sensor nodes
- To create an application scenario and calculate the energy consumption parameters for various environments
- To calculate and compare the maximum system lifetime using the best energy efficient algorithm like LPC and without using compression algorithm with the help of optimization technique for a fixed bit rate

## III. MAXIMUM SYSTEM LIFETIME

Maximum system lifetime which is equivalent to the time until the first node runs out of battery for a particular network.

### A. Formulation of linear problem

The formulation of maximum system lifetime routing problem is described by the author for the case where information generation rates are fixed [4], [6]. Since  $T_{sys}$  in equation (1) is not a linear function of flow variables,  $x_{ij}^k$ , to formulate as a linear programming problem, the author defined  $T_{sys}^{-1}$  as an inverse system lifetime shown in equation (6) [6]. Linear programming problem is formulated considering some parameters described below:

### B. Given information for the network

The parameters that are considered to formulate the problem for the network as follow [6]:

- $\mathcal{V}$  : set of nodes (including node 0 which is central node)
- $\mathcal{V}'$  : set of nodes (excluding node 0)
- $\mathcal{V}_i$  : set of nodes that can be reached by node  $i$
- $t^k$  : traffic rate from sensor node  $k$  to central node
- $E_i$  : initial battery energy of sensor node  $i$

- $\alpha_{ij}^t$  : energy consumed by sensor node  $i$  to transmit a packet (total bits) to node  $j$
- $\alpha_i^r$  : energy consumed by sensor node  $i$  to receive a packet (total bits)

So, the *Lifetime* of node  $i$  is written by

$$T_i = \frac{E_i}{\sum_{k \in \mathcal{V}'} \sum_{j \in \mathcal{V}_i} \alpha_{ij}^t x_{ij}^k + \sum_{k \in \mathcal{V}'} \sum_{j \in \mathcal{V}_i} \alpha_i^r x_{ji}^k} \quad (1)$$

### C. Decision variables, system lifetime and their relations

- $x_{ij}^k$  : traffic flow in bps from node  $i$  to node  $j$  for data of sensor node  $k$ .

The *Lifetime* of node  $i$ , denoted by  $T_i$  can be computed as in (1) follows [5], [6]. But, our target is to maximize the system *lifetime*, called MLR (Maximum Lifetime Routing), denoted by  $T_{sys}$ . The system lifetime under the flow  $x_{ij}^k$  can be defined as the length of time until the first node dies among all nodes in the network, which is the same as the minimum lifetime over all nodes, shown in equation (2).

$$T_{sys} = \min_{i \in \mathcal{V}} T_i \quad (2)$$

So, maximizing  $T_{sys} = \min_{i \in \mathcal{V}} T_i$  can be rewritten by

$$\text{Minimizing } T_{sys}^{-1} = \max_{i \in \mathcal{V}} T_i^{-1} \quad (3)$$

To maximize the system lifetime, the author needs to find the flow under flow conservation condition and satisfaction of traffic demands as follows [6]:

Flow conservation and satisfaction of traffic demands:

$$\forall i \in \mathcal{V}, k \in \mathcal{V}', \sum_{j \in \mathcal{V}_i} x_{ji}^k - \sum_{j \in \mathcal{V}_i} x_{ij}^k = \begin{cases} -t^k, & i = k \\ t^k, & i = 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$\forall k \in \mathcal{V}', i \in \mathcal{V}, j \in \mathcal{V}_i, x_{ij}^k \geq 0 \quad (5)$$

### E. Overall optimization problem

The overall problem is to minimize the cost function mentioned in equation (2) subject to the above set of constraints. From equation (1), it is obvious that the cost function is not yet a linear function of decision variables but it should be linear as considered as linear programming problem. So, the overall linear optimization problem is as follows [5], [6]:

$$\text{minimize } T_{sys}^{-1} \quad (6)$$

subject to

$$\sum_{k \in \mathcal{V}'} \sum_{j \in \mathcal{V}_i} \alpha_{ij}^t x_{ij}^k + \sum_{k \in \mathcal{V}'} \sum_{j \in \mathcal{V}_i} \alpha_i^r x_{ji}^k - T_{sys}^{-1} E_i \leq 0 \quad (7)$$

$$\forall i \in v, k \in v', \sum_{j \in v_i} x_{ji}^k - \sum_{j \in v_i} x_{ij}^k = \begin{cases} -t^k, i = k \\ t^k, i = 0 \\ 0, \text{otherwise} \end{cases} \quad (8)$$

$$\forall k \in v', i \in v, j \in v_i, x_{ij}^k \geq 0 \quad (9)$$

#### IV. SELECTION OF AN EFFICIENT ALGORITHM

The energy consumption of the transceiver depends on the both system and circuit level design. If the CMOS circuits for designing the sensor nodes increase, total energy consumption must be increased. It is given that 20k and 10k CMOS gate circuits are needed for compression and decompression respectively [7], [8]. Every 1k gate needs 24  $\mu$ W powers for compression and decompression. So, computation power depends on the hardware of the sensor nodes. Another one is the transmission power and it is related to data rate minimization, distance and frequency. This data rate minimization can be achieved by local processing at the sensor node like data compression algorithms. We compare the different data compression algorithms by the following equation (10), [7], [8] and select **LPC** as the best algorithm for our purpose based on some calculations.

$$E_{tot} = P_{TX} \times (T_{ON-TX} + T_{START-TX}) + P_{RX} \times (T_{ON-RX} + T_{START-RX}) + P_{OFF} \times T_{OFF} \quad (10)$$

$P_{TX/RX}$  = compression and decompression power consumption of the transceiver

$T_{ON-TX/RX}$  = transmitter/receiver on-time

$T_{START-TX/RX}$  = start up time of the transceiver

$P_{OFF}$  = power consumption of the transceiver in the standby/off state

$T_{OFF}$  = transceiver off-time =  $1 - [T_{ON-TX} + T_{START-TX} + d \times (T_{ON-RX} + T_{START-RX})]$

$d$  = ratio of received packets to transmitted packets

##### A. Energy consumption in different methods varying transceiver ON time and Data Rate

We calculate the energy consumptions in different methods varying ON time of the transceiver with help of equation (10) and finally compare to select the best algorithm like LPC that are shown in the following Fig.2 and Fig.3.

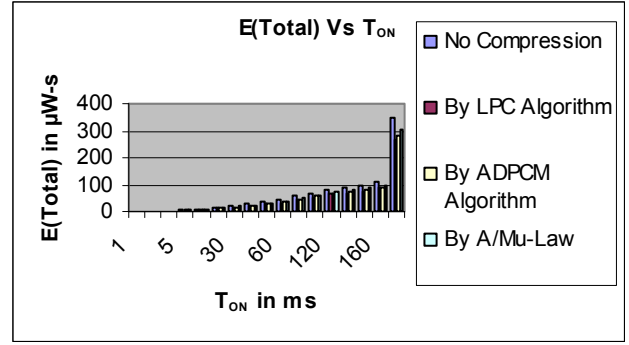


Fig.2. Energy consumption in different methods varying transceiver ON time

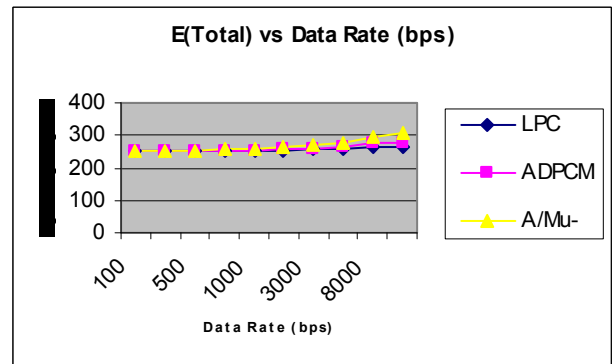


Fig.3. Energy consumption in different methods varying transceiver data rate

#### V. NEW ENERGY DISSIPATION MODEL

We used our new energy dissipation model for calculating the transmitting and receiving energy consumption of the transceiver which is shown in the following Fig.4 [9].

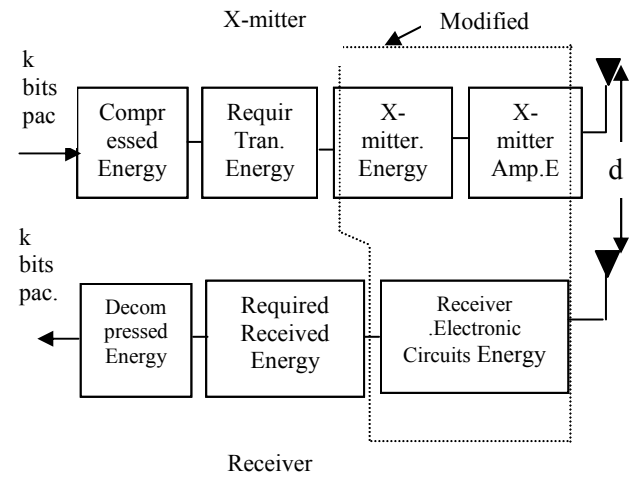


Fig.4. New energy dissipation model for calculating the energy consumption parameters

### Case I: Without compression algorithm

We calculate transmit and receive energy consumption according to proposed radio energy dissipation model for our work as follows [9]:

Transmit energy consumption,  

$$\alpha_{ij}^t = E_{total} + P_{out} \times T_{on} + P_{off} T_{off} + E_{amp} \times d_{ij}^n \quad (11)$$

Where,  $E_{total}$  = total electronic circuit energy consumption (i.e ADC, filtering etc),  $P_{out} \times T_{on}$  = required transmission energy consumption,  $P_{off} T_{off}$  = transmitter off energy consumption,  $E_{amp}$  = amplifier energy consumption

$d_{ij}^n$  = distance between node i and j where n is the path loss exponent

Received energy consumption,  

$$\alpha_i^r = E_{total} + P_{rec} \times T_{on} + P_{off} T_{off} \quad (12)$$

Where,  $E_{total}$  = total electronic circuit energy consumption (i.e ADC, filtering etc),  $P_{rec} \times T_{on}$  = required received energy consumption,  $P_{off} T_{off}$  = receiver off energy consumption

### Case II: Using compression algorithm

We also calculate transmit and receive energy consumption according to proposed radio energy dissipation model for our work as follows [9]:

Transmit energy consumption,  

$$\alpha_{ij}^t = E_{total} + P_{TX} \times (T_{ON-TX} + T_{START-TX}) + P_{off} T_{off} + E_{amp} \times d_{ij}^n \quad (13)$$

Where,  $E_{total}$  = total electronic circuit energy consumption (i.e ADC, filtering etc),  $P_{TX} \times (T_{ON-TX} + T_{START-TX})$  = compression circuit energy with start-up time of compressor,  $P_{out} \times T_{on}$  = required transmission energy consumption,  $P_{off} T_{off}$  = transmitter off energy consumption,  $E_{amp}$  = amplifier energy consumption

$d_{ij}^n$  = distance between node i and j where n is the path loss exponent

Received energy consumption,  

$$\alpha_i^r = E_{total} + P_{RX} \times (T_{ON-RX} + T_{START-RX}) + P_{rec} \times T_{on} + P_{off} T_{off} \quad (14)$$

Where,  $E_{total}$  = total electronic circuit energy (i.e ADC, filtering etc),  $P_{RX} \times (T_{ON-RX} + T_{START-RX})$  = decompression circuit energy consumption with start-up time of decompressor,  $P_{rec} \times T_{on}$  = required received energy consumption,  $P_{off} T_{off}$  = receiver off energy consumption.

## VI. COMPARISONS OF THE SYTEM SYSTEM LIFETIME

We compare the system lifetime of our system by the following ways:

### Case I: For free space model (When n = 2)

TABLE I  
MAXIMUM SYSTEM LIFETIME WITH VARYING BATTERY ENERGY USING COMPRESSION TECHNIQUE

| Node Battery Energy (μW-s) |                   |                   |                   |                   |                   |                   |                   | System LT(s) |
|----------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------|
| Node 1                     | Node 2            | Node 3            | Node 4            | Node 5            | Node 6            | Node 7            | Node 8            |              |
| 1×10 <sup>4</sup>          | 1×10 <sup>4</sup> | 1×10 <sup>4</sup> | 1×10 <sup>4</sup> | 1×10 <sup>4</sup> | 1×10 <sup>4</sup> | 1×10 <sup>4</sup> | 1×10 <sup>4</sup> | 000.02       |
| 5×10 <sup>4</sup>          | 5×10 <sup>4</sup> | 5×10 <sup>4</sup> | 5×10 <sup>4</sup> | 5×10 <sup>4</sup> | 5×10 <sup>4</sup> | 5×10 <sup>4</sup> | 5×10 <sup>4</sup> | 0000.1       |
| 1×10 <sup>5</sup>          | 1×10 <sup>5</sup> | 1×10 <sup>5</sup> | 1×10 <sup>5</sup> | 1×10 <sup>5</sup> | 1×10 <sup>5</sup> | 1×10 <sup>5</sup> | 1×10 <sup>5</sup> | 0000.2       |
| 5×10 <sup>5</sup>          | 5×10 <sup>5</sup> | 5×10 <sup>5</sup> | 5×10 <sup>5</sup> | 5×10 <sup>5</sup> | 5×10 <sup>5</sup> | 5×10 <sup>5</sup> | 5×10 <sup>5</sup> | 0001.1       |
| 1×10 <sup>6</sup>          | 1×10 <sup>6</sup> | 1×10 <sup>6</sup> | 1×10 <sup>6</sup> | 1×10 <sup>6</sup> | 1×10 <sup>6</sup> | 1×10 <sup>6</sup> | 1×10 <sup>6</sup> | 0002.2       |
| 5×10 <sup>6</sup>          | 5×10 <sup>6</sup> | 5×10 <sup>6</sup> | 5×10 <sup>6</sup> | 5×10 <sup>6</sup> | 5×10 <sup>6</sup> | 5×10 <sup>6</sup> | 5×10 <sup>6</sup> | 0010.7       |
| 1×10 <sup>7</sup>          | 1×10 <sup>7</sup> | 1×10 <sup>7</sup> | 1×10 <sup>7</sup> | 1×10 <sup>7</sup> | 1×10 <sup>7</sup> | 1×10 <sup>7</sup> | 1×10 <sup>7</sup> | 0021.4       |
| 5×10 <sup>7</sup>          | 5×10 <sup>7</sup> | 5×10 <sup>7</sup> | 5×10 <sup>7</sup> | 5×10 <sup>7</sup> | 5×10 <sup>7</sup> | 5×10 <sup>7</sup> | 5×10 <sup>7</sup> | 0107.1       |
| 1×10 <sup>8</sup>          | 1×10 <sup>8</sup> | 1×10 <sup>8</sup> | 1×10 <sup>8</sup> | 1×10 <sup>8</sup> | 1×10 <sup>8</sup> | 1×10 <sup>8</sup> | 1×10 <sup>8</sup> | 0197.3       |
| 5×10 <sup>8</sup>          | 5×10 <sup>8</sup> | 5×10 <sup>8</sup> | 5×10 <sup>8</sup> | 5×10 <sup>8</sup> | 5×10 <sup>8</sup> | 5×10 <sup>8</sup> | 5×10 <sup>8</sup> | 1070.7       |
| 1×10 <sup>9</sup>          | 1×10 <sup>9</sup> | 1×10 <sup>9</sup> | 1×10 <sup>9</sup> | 1×10 <sup>9</sup> | 1×10 <sup>9</sup> | 1×10 <sup>9</sup> | 1×10 <sup>9</sup> | 2141.4       |



TABLE II  
MAXIMUM SYSTEM LIFETIME WITH VARYING BATTERY ENERGY USING WITHOUT COMPRESSION  
TECHNIQUE

| Node Battery Energy ( $\mu\text{W-s}$ ) |                 |                 |                 |                 |                 |                 |                 | System LT(s) |
|-----------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------------|
| Node 1                                  | Node 2          | Node 3          | Node 4          | Node 5          | Node 6          | Node 7          | Node 8          |              |
| $1 \times 10^4$                         | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | 000.0017     |
| $5 \times 10^4$                         | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | 000.0083     |
| $1 \times 10^5$                         | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | 000.0200     |
| $5 \times 10^5$                         | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | 000.0800     |
| $1 \times 10^6$                         | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | 000.2000     |
| $5 \times 10^6$                         | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | 000.8000     |
| $1 \times 10^7$                         | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | 0001.700     |
| $5 \times 10^7$                         | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | 008.3000     |
| $1 \times 10^8$                         | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | 016.6000     |
| $5 \times 10^8$                         | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | 082.8000     |
| $1 \times 10^9$                         | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | 165.6000     |

*Case II: For multipath model (When  $n = 4$ )*

TABLE III  
MAXIMUM SYSTEM LIFETIME WITH VARYING BATTERY ENERGY USING COMPRESSION TECHNIQUE

| Node Battery Energy ( $\mu\text{W-s}$ ) |                 |                 |                 |                 |                 |                 |                 | System LT(s) |
|-----------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------------|
| Node 1                                  | Node 2          | Node 3          | Node 4          | Node 5          | Node 6          | Node 7          | Node 8          |              |
| $1 \times 10^4$                         | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | 0000.02      |
| $5 \times 10^4$                         | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | 0000.09      |
| $1 \times 10^5$                         | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | 0000.20      |
| $5 \times 10^5$                         | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | 0000.80      |
| $1 \times 10^6$                         | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | 0001.60      |
| $5 \times 10^6$                         | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | 0008.20      |
| $1 \times 10^7$                         | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | 0016.30      |
| $5 \times 10^7$                         | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | 0081.70      |
| $1 \times 10^8$                         | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | 0163.40      |
| $5 \times 10^8$                         | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | 0817.10      |
| $1 \times 10^9$                         | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | 1634.10      |

TABLE IV  
MAXIMUM SYSTEM LIFETIME WITH VARYING BATTERY ENERGY USING WITHOUT COMPRESSION  
TECHNIQUE

| Node Battery Energy ( $\mu\text{W-s}$ ) |                 |                 |                 |                 |                 |                 |                 | System LT(s) |
|-----------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------------|
| Node 1                                  | Node 2          | Node 3          | Node 4          | Node 5          | Node 6          | Node 7          | Node 8          |              |
| $1 \times 10^4$                         | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | $1 \times 10^4$ | 00.0009      |
| $5 \times 10^4$                         | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | $5 \times 10^4$ | 00.0044      |
| $1 \times 10^5$                         | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | $1 \times 10^5$ | 00.0090      |
| $5 \times 10^5$                         | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | $5 \times 10^5$ | 00.0400      |
| $1 \times 10^6$                         | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | $1 \times 10^6$ | 00.0900      |
| $5 \times 10^6$                         | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | $5 \times 10^6$ | 00.4000      |
| $1 \times 10^7$                         | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | $1 \times 10^7$ | 00.9000      |
| $5 \times 10^7$                         | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | $5 \times 10^7$ | 04.4000      |
| $1 \times 10^8$                         | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | $1 \times 10^8$ | 08.8000      |
| $5 \times 10^8$                         | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | $5 \times 10^8$ | 44.0000      |
| $1 \times 10^9$                         | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | $1 \times 10^9$ | 87.9000      |

*The graph for free space model (Case I):*

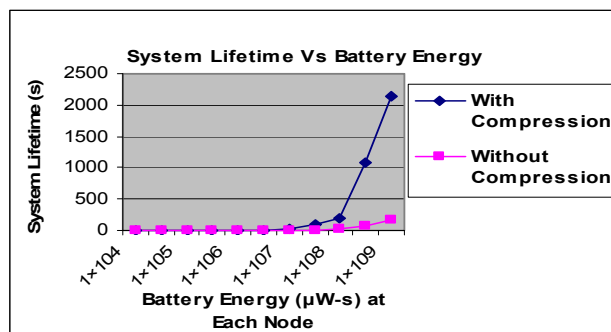


Fig.5. Comparison of the system lifetimes in 8-node body sensor networks for free space

*The graph for free multipath model (Case II):*

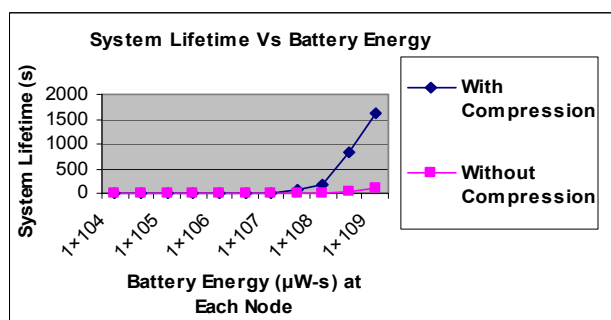


Fig. 6. Comparison of the system lifetimes in 8-node body sensor networks for multipath

## VII. CONCLUSION

In wireless body sensor networks, the nodes operate on limited battery energy. Our major target to support uninterrupted operation by enhancing the battery lifetime of the body sensor nodes for diagnosis the heart beat signal. Therefore, the main objective of this paper is to study and analyze the different algorithms for selecting the best one for low energy consumption in the body sensor nodes and then apply this efficient technique to maximize the system lifetime for a particular application scenario. In this paper, a new energy dissipation model is proposed and applied this best technique for this improved model to calculate the energy consumption parameters like transmitting and receiving energy consumption with respect to a particular application environment. MATLAB optimization technique is used to find out the system lifetimes with respect to our application scenario where we assumed our problem as a linear programming problem. Finally, we compared our results (system

lifetimes) using and without using compression technique in two different environments like free space model where  $n = 2$  and multipath model where  $n = 4$ . Results show that the maximum system lifetime using compression technique is definitely better in both cases.

## REFERENCES

- [1] Pearson Matthew, Beckford Jamal & Akin John, "Wireless sensor Network Proposal, ECE290", 2007. URL: [www.engr.uconn.edu/ece/SeniorDesign/projects/ecesd83/Proposal.pdf](http://www.engr.uconn.edu/ece/SeniorDesign/projects/ecesd83/Proposal.pdf).
- [2] Otto Chris, Milenkovic Aleksandar, Sanders Corey & Jovanov Emil, "System Architecture of a Wireless Body Area Network for Ubiquitous Health Monitoring". Journal of Mobile Multimedia, Vol. 1, No 4, 307-326, 2006.
- [3] Estrin D., Heidemann J, & Ye W, "Energy Efficient MAC Protocol for Wireless Sensor Networks", In Proceedings of the IEEE Infocom, USC/Information Sciences Institute. New York, NY, USA: IEEE June, pp. 1567-1576, 2002
- [4] Chang Jae-Hwan & Tassiulas L., "Energy Conserving Routing in Wireless Ad-hoc Networks,". In Proceedings of IEEE INFOCOM, Tel Aviv, Israel, Mar, pp.22-31, 2000.
- [5] Chang Jae-Hwan & Tassiulas L., "Maximum Lifetime Routing in Wireless Sensor Networks". IEEE/ACM Transactions on Networking, Vol.12, No.4, August, 2004.
- [6] Saengudomlert P., "Lecture Note on Optimization for Communications and Networks". AT77.9011, P-32-36, Dept. of Telecommunications, School of Engineering & Technology, Asian Institute of Technology, August,2007. URL:[http://203.159.18.14/tcweb/AT77\\_9011\\_Optimization](http://203.159.18.14/tcweb/AT77_9011_Optimization).
- [7] Toumaz Technology Limited, "Consideration for the Design of an Ultra-Low Power Wireless Transceiver for Sensor Network Applications". UbiMon First Deliverable, 85 Milton Park Abingdon, UK, March 2005.
- [8] Toumaz Technology Limited, "Trade-Off in Communication versus Computation for Body Sensor Network Nodes". UbiMon Second Deliverable, 85 Milton Park Abingdon, UK, April, 2005.
- [9] Heinzelman Wendi B., Chandrakasan Anantha P. & Balakrishnan Hari, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks". IEEE Transactions on Wireless Communications, Vol. 1, No. 4, October, 2002.

# Cluster Analysis of Microarray data of *Shewanella oneidensis*

Sharmin Nilufar and Mst. Jannatul Ferdousi Ara

Department of Computing Science, University of Alberta, Edmonton, Canada T6G 2E8  
Department of Computer Science and Engineering, Dhaka City College, Dhaka, Bangladesh

**Abstract**— This paper presents a comprehensive analysis of a novel temporal dataset of *Shewanella oneidensis*. Here we propose to cluster the temporal gene expression data of *Shewanella oneidensis* to define its molecular response at different time intervals following acidic pH and basic pH exposure and to find a relation of these temporal data at different environmental conditions. A mapping between those clusters is also defined to describe the interaction and reaction of those genes in different time in two separate conditions. Several conceptually variant clustering algorithms are applied to find the actual structure within data and the results of clustering algorithms are evaluated by several robust internal indices.

**Index Terms**— Clustering, temporal data, *Shewanella oneidensis*, validation techniques.

## I. INTRODUCTION

Recent advents of technologies in the area of bioinformatics allow us to study thousands of genes by analyzing microarray gene expression data qualitatively. The cluster analysis of gene expression data facilitates to identify and group those genes with similar expression pattern. Clustering can be considered as the important initial steps of categorization and investigation of genes in most of the microarray data analysis research. Preliminary validation of the data quality can also be achieved by the visualization of the clustering results. Clustering also helps to explain the inter-correlation structure and expression changes across the samples.

In this paper we propose to analysis a novel temporal dataset of *Shewanella oneidensis*. *Shewanella oneidensis* is a metal-reducing bacterium. It has diverse respiratory capability [16]. It is very challenging and imperative task to define the genetic basis and regulatory mechanisms underlying the ability of *S. oneidensis* to survive and adapt to various environmentally relevant stresses. Gao *et al* [19] examines temporal gene expression profiles in cells subjected to heat stress by using whole-genome DNA microarrays for *S. oneidensis* to define this organism's molecular response to elevated growth temperatures. The molecular response of *Shewanella oneidensis* MR-1 to variations in extracellular pH was investigated by Leaphart *et al* [24] based on genomewide gene expression profiling. They discovered that cells elicited both general and specific transcriptome responses when challenged with environmental acid (pH 4) or base (pH10) conditions over a 60-min period.

This paper endeavors to analysis the microarray data of *Shewanella oneidensis*. There is not enough insight

information on the underlying molecular mechanisms and interaction enabling this bacterium to cope with extracellular acid and alkaline environments is available at present. To define this *Shewanella oneidensis*'s molecular response at different time intervals following acidic pH and basic pH exposure and to find a relation of these temporal data at different environmental conditions we propose to cluster the temporal gene expression data of *Shewanella oneidensis* and find a mapping between those clusters to describe the interaction and reaction of those genes in different time in two separate conditions. Since the actual number of clusters in this dataset is unknown we propose to use a number of conceptually variant clustering algorithms to find the actual structure within data. The result of clustering algorithm can be evaluated by several internal indices. The rest of the paper is organized as follows. Different aspects of gene clustering are described in section 2. A number of challenges of clustering methods including validation of the clustering results and estimation of the proper number classes are discussed in section 3 and section 4 respectively. Section 5 gives a brief idea about the dataset used in this research and section 6 discusses experimental setup and results. Finally concluding remarks are given in section 7.

## II. CLUSTERING

Current experiments on microarray data contain  $10^3$  to  $10^4$  genes. It is actually important and meaningful to cluster both gene and samples (also called features). Gene based clustering aims to group co-expressed genes together by considering genes as objects and samples as features. On the other hand sample based clustering cluster samples by treating gene as sample where each sample based group may represents some macroscopic phenotype, such as clinical syndromes or subtypes of disease [1] [2]. Another very powerful approach known as bi-clustering or subspace clustering designed to cluster subset genes that have similar expression pattern only in a subset of samples or time periods [13]. The results of these clustering techniques provide a good initial investigation in most of the microarray data before actually examining into specific pathways or genetic mechanisms. In the literature, there are a large number of statistical and computational approaches proposed and applied for gene clustering. The most popular clustering algorithm for gene expression data analysis are hierarchical clustering [11], K-means [6], partitioning around medoids [1], self-organizing maps (SOM)

[9], Self-Organizing Tree Algorithm (SOTA) [7] [8], fuzzy c-means, finite mixture models. These clustering algorithms are relatively equivalent in the sense that one method does not dominate all others and their performance depends on the structure and quality of the data.

K-means is probably the most widely used general clustering algorithm. This algorithm starts by assigning the input data points into  $k$  initial sets, either at random or using some heuristic data. The centroid of each set is then calculated. It reassigns each data point to the cluster with the closest centroid. The cluster centers are recalculated for the new clusters, and algorithm iterate these two steps until convergence. This algorithm is very popular because of its quick convergence. Different kinds of distance measures are used in kmeans and centroid of clusters are calculated differently for the different supported distance measures. Widely used distance measures are squared Euclidean distance where each centroid is the mean of the points in that cluster, cosine distance where data points are treated as vectors, correlation distance where objects are considered as sequences of values and hamming distance which calculates percentage of bits that differ is actually suitable for binary data

Hierarchical algorithms find successive clusters using previously established clusters. Hierarchical algorithms can be agglomerative or divisive. Agglomerative algorithms begin with each element as a separate cluster and merge them into successively larger clusters. Divisive algorithms begin with the whole set and proceed to divide it into successively smaller clusters. Single linkage, complete linkage and average linkage based methods are applied to calculate distances between new cluster and each of old clusters. Single-linkage considers the distance between one cluster and another cluster to be equal to the shortest distance from any member of one cluster to any member of the other cluster. Complete-linkage computes the distance between one cluster and another cluster to be equal to the greatest distance from any member of one cluster to any member of the other cluster. Average-linkage calculates the distance between one cluster and another cluster to be equal to the average distance from any member of one cluster to any member of the other cluster. Although the standard hierarchical clustering performs well for gene expression data, it is comparatively slow and not very robust when analyzing thousands of gene expression profiles [6] [4].

Self-Organising Maps (SOM) is based on the unsupervised neural networks and offer a more robust framework when clustering huge amounts of noisy microarray data [7] [8] [9]. The process can be summarized in two phases. In the first phase called training phase a map is built, the neural network organizes itself, using a competitive process. A large number of input vectors are given to the network. During the next phase the mapping phase a new input vector may quickly be given a location on the map, it is automatically classified or categorized. There will be one

single *winning* neuron: the neuron whose weight vector lies closest to the input vector. However, the SOM has some inherent problems. Firstly, since SOM is a topology-preserving, the number of clusters has to be fixed at the beginning of the method. Secondly, the clusters obtained in this technique are not proportional since training of the network depends on the number of items in each cluster.

In fuzzy clustering which is based on fuzzy subset, each point has a degree of belonging to clusters, rather than belonging completely to crisp cluster. The concept of fuzzy c-means clustering is very similar to k-means clustering where the number of clusters should be fixed before starting the clustering process. Main challenge of apply fuzzy clustering or microarray data is the choice of the fuzziness parameter [20]. Some new approaches for example model-based clustering and tight clustering have been proposed to handle scattered genes which contains noisy gene [1] [12]. Good reviews of various gene clustering methods and their advantages and limitations are discussed in [1] and [4].

### III. CLUSTER VALIDATION MEASURE

Since microarray data are very high dimensional and usually contain high levels of noises and many missing values, it is very challenging to detect the appropriate clusters. Moreover clustering approaches are biased towards particular clustering results which match with their own clustering properties. Most of the clustering techniques produce some kind of clusters even in the absence of natural clusters. To detect these limitations, cluster validation techniques are essential tools to verify the clustering result. Cluster-validation techniques provide a logical and systematic evaluation of the amount and type of structure captured by a clustering and should therefore be an important measurement in the analysis of clustering results [3].

A number of different validation techniques proposed by several researchers. Validation techniques are mainly divided into two board categories, external and internal validation measures. External validation measures are applied to obtain an entirely objective evaluation when the prior knowledge of the correct class levels is available. On the other hand internal validation measures do not use any prior information of the correct class level and are appropriate for the clustering result evaluation of new dataset where the actual class levels are not known. This kind of measures assesses the quality of the clustering result by employing information intrinsic to the data [3] [4] [22].

Popular external indices are rand index and Jaccard coefficient which used contingency table to assess the consensus between a partitioning and ground truth [22]. The rand Index determines the similarity between two clusters as a function of positive and negative agreements in pairwise

cluster assignments while in Jaccard coefficient only positive agreements are rewarded. *Internal validation measures* can be categorized into several groups based on the properties they considered for validation for example compactness, connectedness and separation. Most of the popular internal validation indices employ a linear or nonlinear combination of several properties for example Dunn Index, the Davies–Bouldin Index or the Silhouette Width [21] assess both intra-cluster homogeneity and inter-cluster separation, and compute a ultimate score as the linear or non-linear combination of these measures. Some other external indices are also proposed which are based on the stability of the clustering results and compliance between partitioning and distance information of the original data. An extensive review of the different validation techniques can be found in [3].

#### IV. ESTIMATING OF NUMBER OF CLUSTERS

Estimating the actual number of clusters is very difficult in the real microarray data, since the underlying genetic interactions in an organism are so complex that the definition of gene clusters and the exact number of clusters are vague. Many methods for estimating the number of clusters have been proposed in the literature. One of the popular and vigorous techniques is to perform a comprehensive comparison of internal validation measures with different number of clusters and pick the cluster number which has the maximum performance under internal measure [3] [22]. This type of use in class number selection has been the most common application of internal validation measures in bioinformatics. Most of the internal measures can be used to estimate the number of clusters in a dataset. A robust method is to compute clustering results for a range of different numbers of clusters and plot of the performance under the internal measure as a function of the number of clusters. The best number of clusters can often be identified as a ‘knee’ in the resulting performance curve.

#### V. DATASET

The dataset we used in this project are obtained from Leaphart et al [24]. Datasets contains gene expressions of 4692 genes after espousing them to pH 4(acid) and pH10 (base) environment for 60 min time. Expressions are recorded after 30 min and 60 min. We apply filtering technique to eliminate all genes with missing values and obtain 3560 genes which have value in all time points in both conditions. We divide the data sets as acid dataset which contains the expression values of genes in 30 min and 60min time point in acidic condition and base dataset which contains the gene expressions in 30 and 60 min after exposing them in base environment.

### VI. EXPERIMENTS AND RESULTS

#### A. Experimental Setup

Most of the experiments are performed in MATLAB 7. Several special matlab toolboxes are used to perform cluster analysis and validation for example self organizing map toolbox and fuzzy clustering toolbox.

#### B. Clustering and cluster validation

When clustering a novel biological dataset, cluster validation plays a very different role. Although a completely objective validation of cluster quality is truly not possible, the application of cluster validation at different steps during the clustering process can help us to improve the quality of results, and enhance the confidence in the obtained final result. Several conceptually distinct clustering algorithms are applied to *Shewanella oneidensis* dataset. We have applied five different clustering algorithms namely kmeans, self organizing map (SOM), single linkage, average linkage and complete linkage hierarchical methods to find the actual structure of the dataset. The clustering results are obtained for different number of clusters. The performances of the clustering methods are evaluated by computing the silhouette width which is one of the effective internal measures. Figure 4 shows the plot of the performance of clustering techniques for different number of clusters. The clustering results on the acid datasets are given here. From the figure 1 we can see that most of the algorithms generate highly similar partitions. Since several conceptually different algorithms generate highly similar partitions we can say that the actual structure of the data has been discovered.

From the graph in Figure (1) we can see that kmeans outperforms all other clustering method for almost all of the cases. The complete and average linkage algorithms also obtained similar results. The worst performance is obtained from single linkage hierarchical algorithm especially for cluster size 2 to 8. Single linkage which employs the concept of connectedness is not appropriate for the structure of our dataset.

#### C. Kmeans with different distance measures

We have applied different types of distance measure in kmeans to find clusters which match with the meaningful biological clusters. Figure 2 shows the clusters retrieved with kmeans for different types of distance measures. From the visualization of clustering results and their silhouette widths we can see that cosine distance is performed best and able to find the well defined clusters which are also biologically meaningful. The correlation distance gives the worst performance.

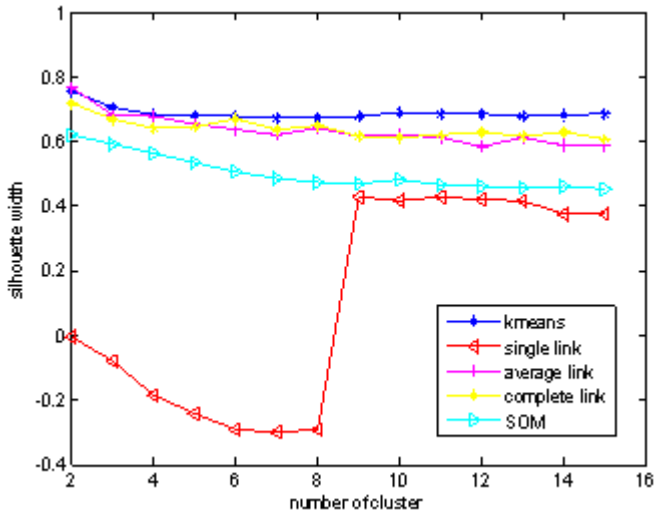


Figure 1: Performance of several clustering algorithms

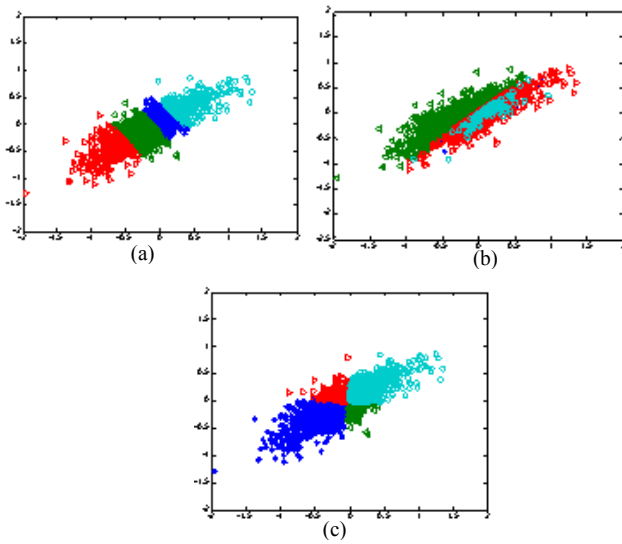


Figure 2: Clusters with kmeans using (a) Euclidian distance (Silhouette width= 0.6327), (b) Correlation distance (Silhouette width= -0.0386) and (c) Cosine distance (Silhouette width= 0.766322)

#### D. Number of Clusters

To select optimum number of clusters ( $k$ ) we compute several internal indices. Plot in figure 3 shows the performance curves of several internal indices for various values of  $k$ .

Most of internal indices agreed that cluster number 4 and cluster number 8 give good results comparing to other values of  $k$ . When considering the two most robust internal measures: Davis Bouldin index, silhouette width we can easily see that 4 and 8 are promising numbers of clusters as Davis Bouldin index achieve minimum values and silhouette width get maximum values in those cases. As a result, we have selected cluster number 4 and 8 for further experiments. From the several internal measures we can conclude that 4 is the best number of clusters for two datasets. We perform kmeans with four clusters. The clustering results we obtained from kmeans for 4 clusters using cosine distance can be defined

biologically. We define each cluster based on the nature of most gene expression in two time points of that clusters. For example figure 4 shows four clusters namely induced-induced, repressed-induced, repressed-repressed, induced-repressed. Most of the genes or in some cases all of the genes in those clusters have the same properties as their cluster names. For example, all of the genes in the cluster induced-induced are induced in both time points (i.e. at T1 and T1).

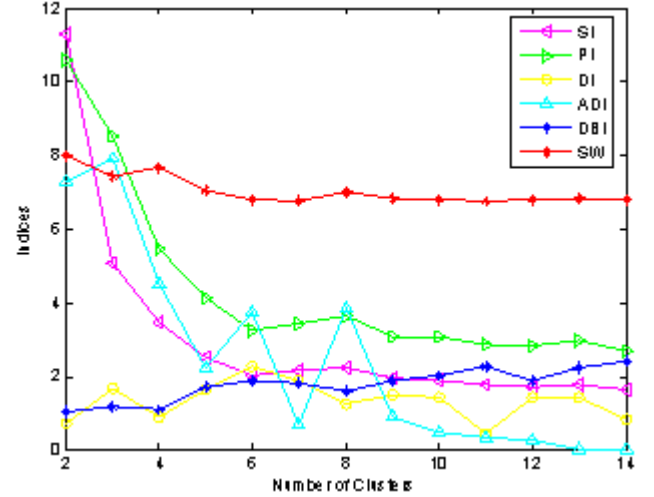


Figure 3: Clustering result with 8 clusters in Base condition

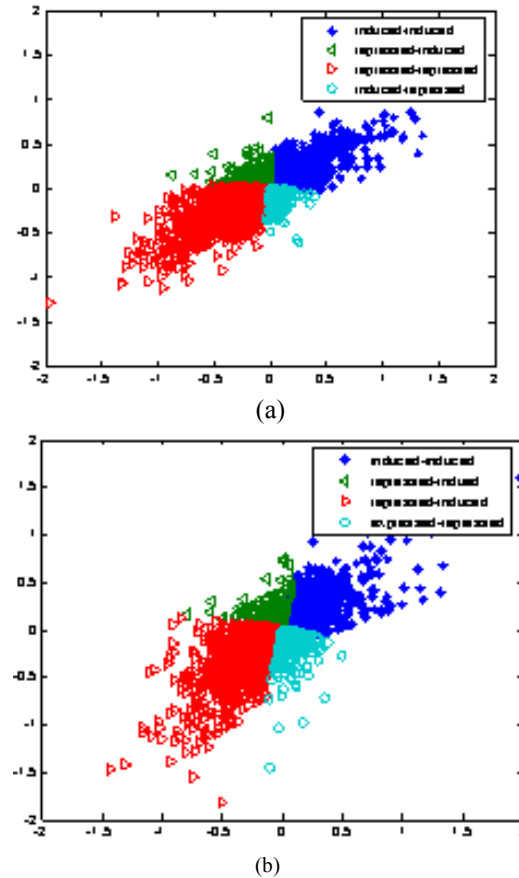
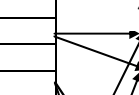
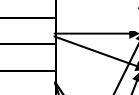


Figure 4: Visual examples for four clusters for (a) Acid dataset and (b) Base dataset

The sizes of similar clusters in two conditions are very close. For example, the size of cluster induced-induced in acidic condition is 1391 and in base condition is 1244. Although the number of genes shared in these two conditions are only 446. Also the number of genes that are induced in both time points in acid condition and repressed in both time points in base condition is 530. Thus we can say that many of the genes behave very differently in two different conditions

Table 1  
CLUSTER BEHAVIOUR IN (a) ACID AND (b) BASE CONDITION

| Acid (pH4)  | T1        | T2        | Number of genes |                                                                                   |
|-------------|-----------|-----------|-----------------|-----------------------------------------------------------------------------------|
| Cluster1    | induced   | induced   | 1391            |  |
| Cluster2    | repressed | induced   | 441             |                                                                                   |
| Cluster3    | repressed | repressed | 1456            |                                                                                   |
| Cluster4    | induced   | repressed | 302             |                                                                                   |
| (a)         |           |           |                 |                                                                                   |
| Base (pH10) | T1        | T1        | Number of genes |                                                                                   |
| Cluster1    | induced   | induced   | 1244            |  |
| Cluster2    | repressed | induced   | 536             |                                                                                   |
| Cluster3    | repressed | repressed | 1363            |                                                                                   |
| Cluster4    | induced   | repressed | 417             |                                                                                   |
| (b)         |           |           |                 |                                                                                   |

Considering the plot of internal indices in figure 3 we can also find a pick when the number of clusters is 8. Figure 5 shows the clustering results for two conditions using 8 clusters. From the clustering results for acid condition we see that the two big clusters induced-induced and repressed-repressed, each of them are divided into 3 different clusters based on the difference of the inducement and repression in two time points. For example the clusters1 we obtained are sub clusters of i-i(L-H) which are slightly induced in first time point and largely induced in the second time points. On the other hand i-i(simi) is induced similarly in both time points and cluster 4 i-i(H-L) are induced highly in first time point and decrease inducement in second time point. When considering the base condition we see that the cluster induced-induced are divided into two clusters: i-i(simi) and i-i(H-L). Similarly the cluster repressed-induced are divided into two clusters: r-i(simi) and r-i(L-H). On the other hand the cluster repressed-repressed in divided into 3 different clusters. To compare the reaction of genes in two different environments we calculate the intersection between classes. There are only 5 genes sharing the cluster 7 (i-r) in acid condition and cluster 5 (r-i) in base condition. On the other hand maximum number of genes are found in the intersection of cluster 3 (i-i(simi)) in acid condition and cluster 6 (i-i(simi)) in base condition.

## VII. CONCLUSION

In this paper, we employ several clustering methods and validations techniques to detect the actual structure of a novel dataset of *Shewanella oneidensis*. It is in fact not easy to provide an unbiased and experimental comparison of different

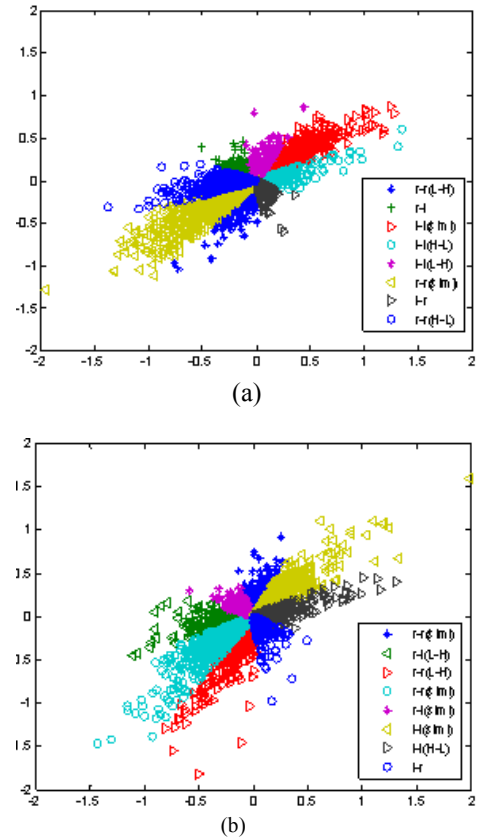


Figure 5: Visualization of Clustering result with 8 clusters in (a) Acid and (b) Base conditions respectively

clustering approaches. The original true clustering assignment in a novel dataset is generally unknown and the concept of a cluster is not mathematically well-defined in unsupervised learning. After explaining several validation indices we found that kmeans with cosine distance perform best for this dataset. The cluster number 4 and 8 able to find biologically meaningful clusters. Genes of *Shewanella oneidensis* have strong response to acid condition than base condition. More genes are induced and repressed in both time points in acidic condition than the base condition. In both conditions more genes are repressed than induced. Although the results shows similar distribution of the genes in four classes, the intersection between similar clusters in two different conditions contain small number of genes which imply many of the genes behave completely different in two conditions. Clustering result obtained with 8 clusters give more meaningful small clusters than 4 clusters. Detail study of dataset using different number of clusters and draw a complete mapping between clusters in two different conditions are left for our future work.

## ACKNOWLEDGMENT

This work is supported by Natural Science and Engineering Research Council of Canada (NSERC), Alberta Ingenuity Fund.

# REFERENCES

- [1] A. Thalamuthu , I. Mukhopadhyay , X. Zheng , and G. C. Tseng, Evaluation and comparison of gene clustering methods in microarray analysis' *Bioinformatics*, 22(19): 2405--2412. October 1, 2006
- [2] J. Dopazo, Microarray data processing and analysis. *Microarray data analysis II. Kluwer Academic. Publication*, 43--63.
- [3] Julia Handl, Joshua Knowles and Douglas B. Kell . Computational cluster validation in post-genomic data analysis. *Bioinformatics* 21: 3201--3212, Aug 1, 2005.
- [4] S. Datta and S. Datta. Comparisons and validation of statistical clustering techniques for microarray gene expression data. *Bioinformatics* 19(4): 459--466, 2003.
- [5] N. Bolshakova and F. Azuaje. Machaon CVE: cluster validation for gene expression data. *Bioinformatics*. 19(8): 2494 --2495, 2003.
- [6] F. X. Wu, W. J. Zhang, and A. J. Kusalik. A genetic k-means clustering applied to gene expression data. *Canadian Conference on AI*, 520-526, 2003.
- [7] H.C. Wang, J. Dopazo, J.M. Carazo. Self-organizing tree growing network for classifying amino acids. *Bioinformatics* 14(4): 376-377, 1998.
- [8] H. C. Wang, J. Dopazo, L.G. de la Fraga, Y.P. Zhu, J.M. Carazo. Self-Organizing Tree-growing Network for the classification of Protein Sequences. *Protein Science* 7:2613-2622, 1998.
- [9] T. Kohonen. Self-Organizing Maps. *Series in Information Sciences*, 30. Springer, Heidelberg. Second ed. 1997.
- [10] K. Y. Yeung, C. Fraley, A. Murua, A. E. Raftery and W. L. Ruzzo. Model-based clustering and data transformation for gene expression data. Technical Report no. 396, *Bioinformatics* 17:977-987 2001.
- [11] Michael B. Eisen, Paul T. Spellman, Patrick O. Brown, and David Botstein. Cluster analysis and display of genome-wide expression patterns', *PNAS*, 1995
- [12] G. C. Tseng, W. H. Wong. Tight Clustering: A Resampling-Based Approach for Identifying Stable and Tight Patterns in Data. *Biometrics* 61 (1), 10--16, 2005.
- [13] Q. Sheng, Y. Moreau, and B. D. Moor. Biclustering microarray data by Gibbs sampling. *Bioinformatics*. 19(2003), 196--205.
- [14] C. Spieth, J. Supper, F. Streichert, N. Speer and A. Zell. Systems biology JCell—a Java-based framework for inferring regulatory networks from time series data. *Bioinformatics*, June 16, 2006
- [15] M.J.L. de Hoon, S. Imoto, K. Kobayashi, N. Ogasawara, S. Miyano. Inferring Gene Regulatory Networks from Time-Ordered Gene Expression Data of Bacillus Subtilis Using Differential Equations. *Pacific Symposium on Biocomputing*. 8: 17--28, 2003.
- [16] X. F. Wan, N. C. VerBerkmoes, L. A. McCue, D. Stanek, H. Connelly, L. Wu, X. Liu, T. Yan, A. Leaphart, R. L. Hettich, J. Zhou, and D. K. Thompson,. Transcriptomic and proteomic characterization of the fur regulon in the metal-reducing bacterium Shewanella oneidensis. *Journal of Bacteriology*. 186: 8385--8400, 2004.
- [17] Ziv Bar-Joseph. Analyzing time series gene expression data. *Bioinformatics*. 19: 2004.
- [18] J. Ernst , G. J. Nau , and Z. Bar-Joseph. Clustering short time series gene expression data. *Bioinformatics* 21: i159-i168, May 6, 2004.
- [19] H. Gao, Y. Wang, X. Liu, T. Yan, L. Wu, E. Alm, A. Arkin, D. K. Thompson, and J. Zhou. Global Transcriptome Analysis of the Heat Shock Response of *Shewanella oneidensis*. *Journal of Bacteriology*, 186(22): 7796--7803, 2004.
- [20] D. Demb  l   , and P. Kastner. Fuzzy C-means method for clustering microarray data. *Bioinformatics* 19: 973--980, 2003.
- [21] P. J. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20, 53--65.
- [22] Gat-Viks.. 'Scoring clustering solutions by their biological relevance. *Bioinformatics*. 19: 2381--2389, 2003.
- [23] T. T. Vu and J. Vohradsky. Nonlinear differential equation model for quantification of transcriptional regulation applied to microarray data of *Saccharomyces cerevisia*. *Nucleic Acids Res*. January, 2007.
- [24] A. B. Leaphart, D. K. Thompson, K. Huang, E. Alm, X.-F. Wan, A. Arkin, S. D. Brown, L. Wu, and J. Zhou. Transcriptome analysis of *Shewanella oneidensis* gene expression in response to acidic and alkaline pH stress. *Journal of Bacteriology*. 188, 1633--1642, 2006.



# Performance Evaluation of DODE of a Voice/Data Integrated Wireless Mobile Network

H. M. Rahat<sup>1</sup>, Runa Laila<sup>2</sup>, Sabiha Rahman Juthy<sup>3</sup>, Md. Imdadul Islam<sup>4</sup>, and M. R. Amin<sup>5</sup>

<sup>1,2,3,5</sup>Department of Electronics and Communications Engineering, East West University, Mohakhali, Dhaka, Bangladesh, e-mail: juthy@ewubd.edu, ramin@ewubd.edu

<sup>4</sup>Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh, e-mail: imdad@juniv.edu

**Abstract** — An analytical model has been developed to improve the performance of a wireless mobile communication network with voice and data integrated traffic system. To reduce the probability of no channel available, which is an important quantity for performance evaluation, delays to the last data end user is introduced to the system. It has been found that by applying a very small delay to the last data end user, the probability of no channel available can be drastically reduced compared to the usual erlang case.

**Index Terms** — Last data end user, performance evaluation, probability of no channel available, voice/data integration.

## I. INTRODUCTION

As the wireless communications industry moves toward 3G and 4G networks, one of the important objectives is the use of a single wireless system for multimedia applications to support a variety of communications services, including voice, data, and voice in various forms and combinations [1]. A key technical problem in such integrated systems is that of multi-user access. As we know that an access method that efficiently supports one category of service may be unsuitable for another category of service. In the 1G and 2G networks, the wireless industry evolved around two separate paths for voice- and data-oriented applications [2]. If a data service can be efficiently integrated with a voice service, transmission resources that are otherwise wasted (because there is no voice transmission) can be used for data, which typically do not have stringent delay requirements. First, the voice-oriented networks evolved into supporting data. More recently with the popularity of voice over IP over the Internet and PSTN, supporting voice in data-oriented WLANs has become attractive as well [3].

In a packet communication environment, voice and data have different requirements. Voice packets can tolerate errors and even packet losses of 1 to 2 percent of voice packets has insignificant effect on the perceived quality of reconstructed voice [3], whereas data packets are sensitive to loss and errors but can generally tolerate delays. Also, the rate at which information is transmitted is constant in the case of voice, thereby making circuit-switching a viable and

efficient approach, whereas information generated for data transmission is very bursty. As a result, voice- and data-oriented networks use different multiple access methods. In a wireless environment, the simplest approach is to assign different frequency bands to isochronous (voice) and asynchronous (data) packets. However, integration in one frequency band will result in a more efficient use of the bandwidth, a simpler radio interface, and an environment that provides a better control for synchronizing voice and video [3].

Let us consider a mixed service environment with data and voice. For bandwidth detection, such a mixed service can further classified into the forms of circuit switching, burst or packet switching and hybrid switching which support both circuit and packet switching [4]. To support the maximum number of voice users, the best case would be that all available channels are shared between voice users. However, in cases where voice users occupy all available channels any data presented to the system is buffered and the mean delay for the data becomes high. This situation is defined to be the scalar quantity ‘bursty queue length’. It is clear that in this worse data case the buffer occupancy quickly becomes elevated and queuing time rapidly increase. Conversely, a better data case would be that all channels are available for data at any time. In this case the earlang capacity of voice user becomes low as voice users are less able to occupy channels. Trivially, it is clear that improvement in traffic for data could be made if bursty queue lengths could be ameliorated [5].

We organize the paper as follows. Section II describes the traffic model for the derivation of the stationary probability distribution of the network states. Expressions for the probability of no channel available and blocking probability have been obtained in this section. Section III describes the results of the investigation by showing the variation of the above mentioned two important quantities with respect to the offered traffic of both voice and data and also with respect to the delay of data end user. Finally, Section IV concludes the paper.

## II. TRAFFIC MODEL

We consider a hybrid traffic system with voice and data traffic. We assume that the voice has a high priority and that data has low priority.

For a queuing system with  $n$  servers and an infinite number of users consisting of both voice and data traffic according to Poisson process with the same arrival rate  $\lambda$  and service rate  $\mu$ , then the system can be denoted by  $M/M/n/n$ , and is known as  $n$ -server loss system [6]. The servers can be considered as channels. The steady state distributions of the probability states can then be written as

$$P_x = \frac{a^x}{x!} P_0; \quad x = 0, 1, 2, \dots, n \quad (1)$$

where  $a \equiv \lambda / \mu$  is known as the offered traffic and  $P_0$  can be obtained from the normalization that the sum of all the stationary probability states is unity. This is the usual Erlang theory [6]. The quantity  $P_n$  is an important quantity, called the probability of no channel available. It is the fraction of the time that the system is full, or the fraction of the time that customers are lost in the long run. It is implied that the system is saturated with the users and the mean duration is given by  $m_n = (n\mu)^{-1}$ . In general,  $m_n$  is in the order of many seconds and consequently bursty length of queue occurs. Now if we consider that the state of the system be at  $n-1$ , the next user is given the last available channel. A new protocol called delay of voice end user (DOVE) has been proposed in [5]. When the  $n$ -th user presents, this call is delayed for an arbitrary time. We redefine that the state space to be  $\{0, 1, 2, \dots, n-1, D, n\}$ . Fig. 1 shows the queuing model of DOVE.

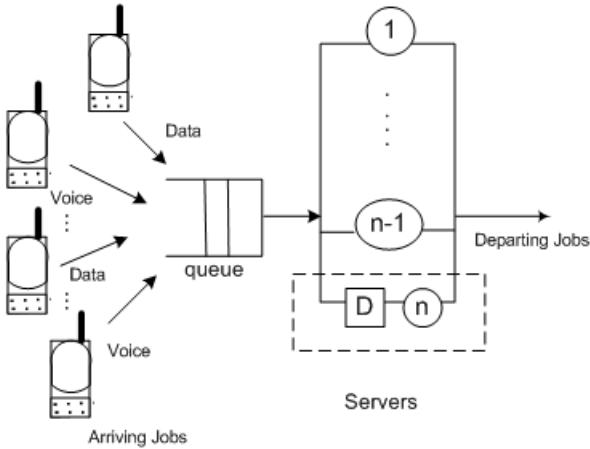


Fig. 1. Model of exchange with DOVE.

Implementation of DOVE is straightforward for mobile networks and would be triggered by a last user success condition [5]. Let  $d$  be the duration of time that the new call experiences in delay before being allowed a channel. In this

situation we can either have that  $d$  expires. In which case the delayed voice user is given permission to occupy the channel, or a channel becomes available before the expiration of the  $d$  and the voice user is given permission to this other channel. By delaying the  $n$ -th user, the duration for which all of the channels are occupied is reduced and the system stays in the state  $n-1$  for a greater proportion of time. Detail analysis is described in [5].

Since voice traffic should have higher priority than the data traffic, instead of DOVE we intend to study the delay of data end user (DODE) in this paper. Further, we do not assume the same arrival rate of the voice and data traffic. Instead, we assume that voice traffic and data traffic have different arrival rates and these are respectively as  $\lambda_v$  and  $\lambda_d$ . The corresponding Markov chain is shown in Fig. 2.

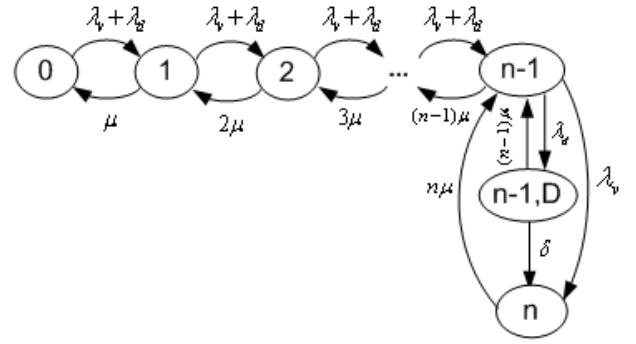


Fig. 2. Markov chain for the delay of data end user.

The steady state probability distributions are obtained from the following systems of equations:

$$(\lambda_v + \lambda_d)P_0 = \mu P_1, \quad (2)$$

$$(\lambda_v + \lambda_d + \mu)P_1 = (\lambda_v + \lambda_d)P_0 + 2\mu P_2, \quad (3)$$

$$(\lambda_v + \lambda_d + 2\mu)P_2 = (\lambda_v + \lambda_d)P_1 + 3\mu P_3, \quad (4)$$

$$\vdots$$

$$\begin{aligned} \{\lambda_v + \lambda_d + (n-1)\mu\}P_{n-1} \\ = (\lambda_v + \lambda_d)P_{n-2} + (n-1)\mu P_{n-1,D} + n\mu P_n, \end{aligned} \quad (5)$$

$$\{\delta + (n-1)\mu\}P_{n-1,D} = \lambda_d P_{n-1}, \quad (6)$$

$$n\mu P_n = \delta P_{n-1,D} + \lambda_v P_{n-1}. \quad (7)$$

The stationary probability distributions up to state  $n-1$  are still given by (1) with  $a \equiv (\lambda_v + \lambda_d) / \mu$  and the remaining  $P_{n-1,D}$  and  $P_n$  are obtained as

$$P_{n-1,D} = \left( \frac{\lambda_d}{\delta + (n-1)\mu} \right) \frac{a^{n-1}}{(n-1)!} P_0, \quad (8)$$

and

$$P_n = \left( \frac{a\delta + (n-1)\lambda_v}{\delta + (n-1)\mu} \right) \frac{a^{n-1}}{n!} P_0, \quad (9)$$

where

$$P_0 = \left( \sum_{x=0}^{n-1} \frac{a^x}{x!} + \left( \frac{a\delta + n\lambda_d + (n-1)\lambda_v}{\delta + (n-1)\mu} \right) \frac{a^{n-1}}{n!} \right)^{-1}. \quad (10)$$

The probability of no channel available and the blocking probability are obtained as

$$P_N = P_n = \left( \frac{a\delta + (n-1)\lambda_v}{\delta + (n-1)\mu} \right) \frac{a^{n-1}}{n!} P_0, \quad (11)$$

and

$$\begin{aligned} P_B &= P_N + P_{n-1,D} \\ &= \left( \frac{a\delta + n\lambda_d + (n-1)\lambda_v}{\delta + (n-1)\mu} \right) \frac{a^{n-1}}{n!} P_0, \end{aligned} \quad (12)$$

where the constant  $P_0$  is given by (10).

### III. RESULTS AND DISCUSSIONS

For the numerical appreciation of our results, we have displayed in the forms of graphs, the variation of the probability of no channel available and the blocking probability as functions of the arrival rates of the voice and data traffic as well as with respect to the delay of the data end user. Figs. 3 and 4 are displayed for the probability of no channel available with respect to the arrival rates of voice and data traffic respectively taking the delay  $d$  of the data end user as a parameter; whereas in Figs. 5 and 6 we show the variation of the probability of no channel available and the blocking probability respectively with respect to the delay  $d$  taking the service rate  $\mu$  as a parameter [7].

From Figs. 3 and 4 we notice that the behavior of the probability of no channel available is almost same, particularly for the lower values of the arrival rates of the voice and data traffic, whereas for the higher values of the arrival rates, the two cases differ markedly if we increase the delay  $d$ . We also observe that as we increase the delay  $d$ , the probability of no channel available drops noticeably. Since delay is applied to data traffic, slope of  $P_N$  against  $\lambda_d$  changes with respect to  $d$  but it does not have impact on slope of  $P_N$  versus  $\lambda_v$  curve (Figs. 3 and 4).

Fig. 5 shows the variation of the probability of no channel available with respect to the delay  $d$  of the data end user. It is interesting to note from this figure that we need only to give a very small value of the delay  $d$  to reduce  $P_N$  up to 67% from the Erlang value (with no delay to the data or voice end user). This is a remarkable result of the paper.

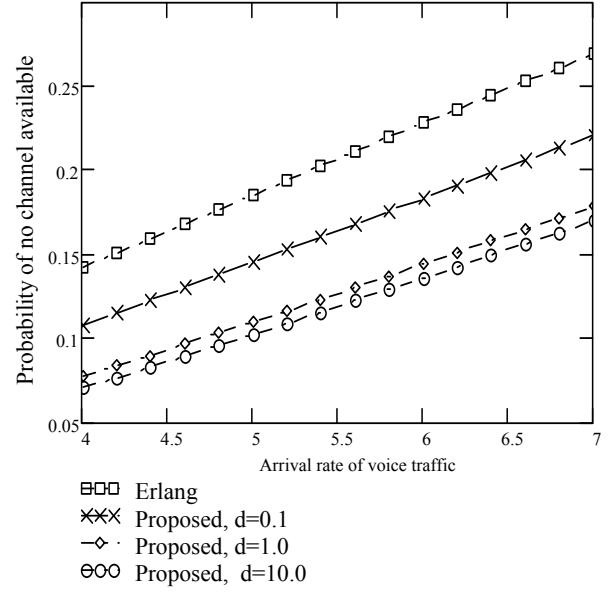


Fig. 3. The probability of no channel available as a function of the arrival rate of the voice traffic  $\lambda_v$  under different delay  $d = 1/\delta$  ( $n=7$ ,  $\mu = 1.5$ ,  $\lambda_d = 4$ ).

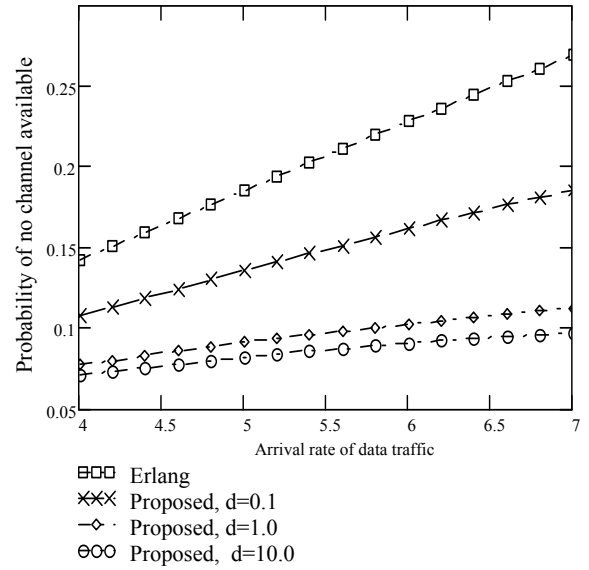


Fig. 4. The probability of no channel available as a function of the arrival rate of the data traffic  $\lambda_d$  under different delay  $d = 1/\delta$  ( $n=7$ ,  $\mu = 1.5$ ,  $\lambda_d = 4$ ).

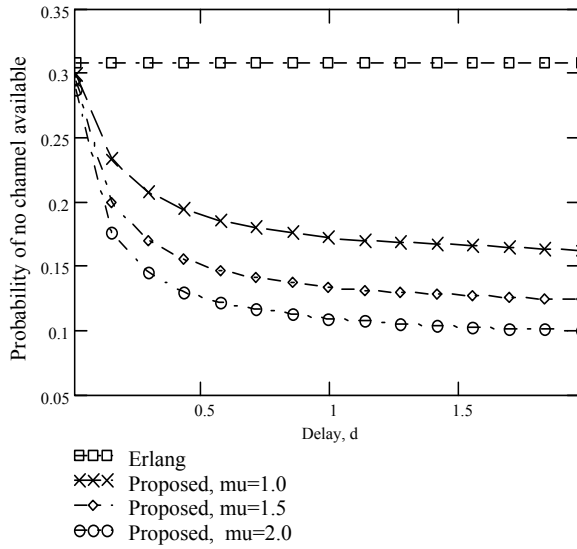


Fig. 5. The probability of no channel available as a function of the delay  $d = 1/\delta$  for different  $\mu$  ( $n=7$ ,  $\lambda_v = 4$ ,  $\lambda_d = 4$ ).

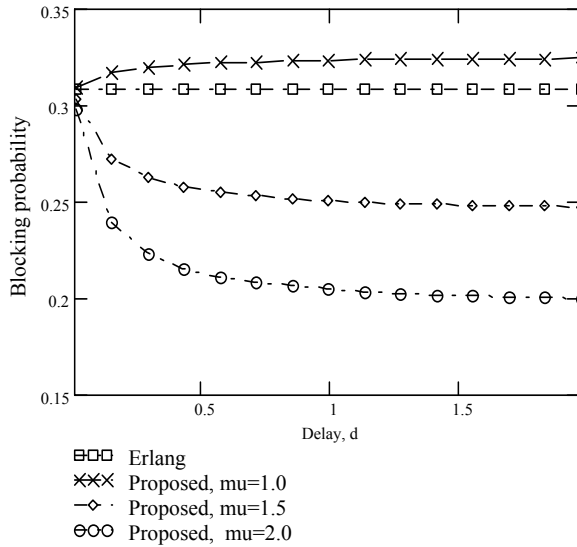


Fig. 6. The blocking probability as a function of the delay  $d = 1/\delta$  for different  $\mu$  ( $n=7$ ,  $\lambda_v = 4$ ,  $\lambda_d = 4$ ).

Fig. 6 shows the variation of the blocking probability  $P_B$  with respect to the delay  $d$  for different values of  $\mu$ . We observe that the blocking probability is somewhat increased (negligibly) when we include the delay of data end user for the data traffic in the system for small values of  $\mu$ . From the user's point of view, as to handle a voice request is more important than the blocking of data traffic, this negligible increase in the blocking probability for small value of  $\mu$  has no effect on the Erlang capacity for the users of the voice

calls. However, it is interesting to observe that as we increase  $\mu$ , the blocking probability decreases from the Erlang case. Furthermore, for large holding time (small  $\mu$ ) of call, the impact of delay is very small as is visualized from Fig. 6.

## VII. CONCLUSION

An analytical model has been developed to improve the probability of no channel available, which is a very important quantity for performance evaluation of a wireless mobile communication system. Delay of data end user has been introduced instead of delay of voice end user as the voice traffic is impatient traffic (customer retry until getting free channel) should have higher priority than the data traffic. It has been found that by applying a small delay to the data end user, the probability of no channel available can be drastically reduced. We have in this analysis incorporated different arrival rates of the voice and data traffic but same service termination rate.

## REFERENCES

- [1] K. K. Leung, W. A. Massey, W. Whitt, "Traffic Models for Wireless Communication Networks," *IEEEJSAC*, vol. 12, no. 8, pp.1353-1364, October 1994.
- [2] D. Bear, *Principle of Telecommunication Traffic Engineering*, London: Peter Peregrinus, 1980.
- [3] Data Communication Magazine, *Integrating Voice & Data McGraw Hill/Data Communications Book Series*, New York: McGraw-Hill, 3<sup>rd</sup> Ed., vol. 3, 1990.
- [4] N. M. Mitrou and G.L. Lyberopoulos, "Voice and data integration in the air-interface of a micro cellular mobile communication system," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 1-13, February 1993.
- [5] M. Mahdavi, R. M. Edwards, and S. R. Cvetkovic, "Policy for Enhancement of Traffic in TDMA Hybrid Switched Integrated Voice/Data Cellular Mobile Communications Systems," *IEEE Commun. Lett.*, vol. 5, no. 6, June 2001.
- [6] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory*, New York: John Wiley & Sons, 3<sup>rd</sup> Ed., 1998.
- [7] M. I. Islam and S. Hossain, "Impact of channel reservation on call blocking and forced termination in mobile cellular network", *Jahangirnagar Univ. J. Sci.*, vol. 26, pp. 123- 132, December 2004.

# A New Weighted Centroid Localization Algorithm in Wireless Sensor Networks

Yahya Forghani

Toos Institute of Higher Education, Mashad, IRAN

**Abstract** — Nodes in a sensor network are often randomly distributed. To assign measurements to locations, each node has to determine its own position. Algorithms for positioning in wireless sensor networks are classified into two groups: approximate and exact. In this paper, we propose a range-based approximate positioning approach which is almost the combination of WCL and EBTB. Then, compare it with two other approximate positioning approaches (WCL with time complexity of  $O(n)$  and EBTB with time complexity of  $O(n^2)$  and an exact positioning approach (QR Factorization with time complexity of  $O(n^3)$ ). Finally, it will be shown that EWCL (with time complexity of  $O(n)$ ) is the best localization algorithm with respect to the three other localization algorithms when the noise is high and its accuracy is close to the accuracy of QR when the noise is medium.

**Index Terms** — range-based; approximate; positioning; wireless sensor networks; noisy.

## I. INTRODUCTION

The increasing miniaturization of electronic components and advances in modern communication technologies lead to the development of extreme small, cheap, and smart sensor nodes [1]. These nodes consist of sensors, actuators, a low power processor, small memory, and a communication module. Nodes measure conditions of the environment, precalculate, aggregate, and transmit data to a base station. Thousands of these nodes form a large wireless sensor network to monitor huge inaccessible terrains.

Processor performance and available energy of each sensor node are highly limited by its physical size. Therefore, intensive communication and computation tasks are not feasible. Thereby, algorithms in sensor networks are subject to strict requirements covering reduced memory consumption, communication, and processing time.

As a result of the stochastic distribution of all nodes in the deployment phase, a determination of the node's position is required. Then, every measurement can be assigned to its physical position. Basically, a global position can be approximately determined with state-of-the-art positioning systems such as the Global Positioning System (GPS) or the Global System for Mobile Communication (GSM) [2], [3]. These systems are not suitable for tiny sensor nodes because of their physical size and costs as well as their high power consumption

resulting from the communication overhead [4]. A better approach is to equip only few nodes with a global positioning system. Then, these nodes determine their own position very precisely and provide this position to other nodes in the sensor network. Based on this approach, we present a new approximate positioning approach for localization in wireless sensor networks.

This paper is structured as follows. In Section II, we give a basic overview of the approaches for positioning in the wireless sensor networks and in Section III, BTB algorithm is described. Next, In Section IV, we describe our new positioning approach. Then, in Section V, we describe our simulation environment and its parameters. After discussing our simulation results in Section VI, finally we conclude the paper in Section VII.

## II. RELATED WORK

For the above-mentioned reasons, existing positioning techniques (e.g. GPS) cannot be integrated on all sensor nodes. The number of nodes with known position has to be limited. These nodes are referred to here as beacons, with the remaining nodes classed as sensor nodes. For the positioning of the sensor nodes we distinguish between approximate and exact approaches [5].

### A. Approximate Positioning

Many approximate approaches for the determination of sensor nodes exist in literature. These approaches are often resource-efficient but also result in higher positioning errors than exact positioning approaches, especially when the input parameters have low deviation of their true value. Examples of such approaches are the hybrid approaches [6], the Coarse Grained Localization (CGL) [7], the Approximate Point in Triangulation-algorithm (APIT) [8], the Weighted Centroid Localization (WCL) [9], the EBTB algorithm [10] and the Binary algorithm [11].

WCL is simple, has low complexity, requires a few resources, has rather good accuracy (especially in high noise situations) and is range-based. APIT has more complexity, requires more resources and has lower accuracy than WCL. But, it is range-free. The resource requirement of EBTB is more than WCL and less than APIT and its accuracy is often better than WCL and APIT

but it is range-based. CGL is range-free and requires fewer resources than the two previous approaches, but has low accuracy. The resource requirement of Binary is more than EBTB and less than APIT; its accuracy is often better than them and is range-free.

### B. Exact Positioning

In contrast to approximate approaches, exact approaches use the known beacon-positions and the distances to the sensor nodes in order to calculate their coordinates through the solution of non-linear equations. Using a minimum of three beacons (in two dimensions), the coordinates of the sensor nodes may be determined by their intersection. The use of more than three beacons gives more information in the system and allows the refinement of the position and the detection and removal of outlying observations. The least squares method (LSM) is used for the solution of the simultaneous equations. The LSM produces accurate results, however it is complex and resource-intensive and therefore it is not feasible on resource-limited sensor nodes. So, this approach is implemented distributed with some communication efforts [12]–[17].

In this paper, we introduce a new Weighted Centroid Localization algorithm called Extended WCL. Then, we compare it with QR Factorization [17] (which is a LSM) and WCL and EBTB (which are approximate positioning approaches).

## III. ESTIMATING THE POSITION OF A NODE USING TWO BEACONS (BTB ALGORITHM)

When  $C$  is a sensor node on the connecting line segment of the beacons  $A$  and  $B$ , and the range between the beacon  $A$  and the sensor node  $C$  is  $AC$  and the range between the beacon  $B$  and the sensor node  $C$  is  $BC$ , we have

$$AB = AC + BC. \quad (1)$$

Also, we know that parametric equation of the line segment  $AB$  is

$$\begin{aligned} x(u) &= x_A + u\Delta x \\ y(u) &= y_A + u\Delta y. \end{aligned} \quad (2)$$

where

$$\begin{aligned} \Delta x &= x_B - x_A \\ \Delta y &= y_B - y_A. \end{aligned}$$

So, with replacing  $u = \frac{AC}{AB}$  in (2), the coordinate of the point  $C$  is obtained.

If the sensor node  $C$  is not on the line segment  $AB$ , but is near to it, then

$$AB = AC + BC - \epsilon. \quad (3)$$

Whatever the sensor node  $C$  is nearer to the line segment  $AB$ ,  $|\epsilon|$  becomes less and when the sensor node  $C$  is on the line segment,  $\epsilon$  becomes equal to zero. As it is seen in Fig. 1, if the coordinates of the beacons  $A$  and  $B$  and the ranges between them and the sensor node  $C$  is determinate, the position of the sensor node  $C$  will be the point  $C_1$  or  $C_2$ . But, if  $|\epsilon|$  is small, the intersection of the line segment  $AB$  and the line segment  $C_1C_2$  (which is the point  $F$ ), will be a good approximation of the position of the sensor node  $C$ .

The point  $F$  (the approximate position of the sensor node  $C$ ) is approximately in middle of the two points  $E$  and  $D$ , where the coordinate of the point  $D$  (the intersection of a circle with the radius  $AC$  and the line segment  $AB$ ) and the coordinate of the point  $E$  (the intersection of a circle with the radius  $BC$  and the line segment  $AB$ ) are obtained easily from the parametric equation of the line segment  $AB$  (see [10]).

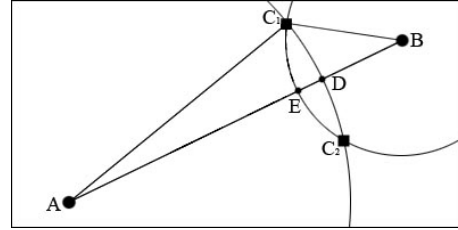


Fig. 1. Positioning in a sensor network

As it was said, if  $|\epsilon|$  is small, the point  $F$  will be a good approximation of the position of the sensor node  $C$ . So, we can select two beacons with minimum  $|\epsilon|$  and then estimate the position of the node  $C$  using the selected two beacons (BTB algorithm).

## IV. NEW APPROACH

### A. First Improvement of BTB

Almost like EBTB, in a noisy environment we should select two beacons with minimum  $|\mu|$ , where

$$\mu = \epsilon \cdot (aC + Cb), \quad (4)$$

because the accuracy of estimated range between the node  $C$  and the beacons “a” and “b” has inverse relation with the true range. In the other word, when the true range between the node  $C$  and the beacon “a” becomes more, the estimated range “aC” probably has more bias from the true range. So, in a noisy environment, we should select two beacons “a” and “b” that their  $\epsilon$  with respect to the node  $C$  is relatively minimum and these two beacons should be relatively the nearest beacons in the range of the node  $C$ .

So, instead of selecting two beacons with minimum  $|\epsilon|$ , we select two beacons with minimum  $|\mu|$ .

### B. Second Improvement

In BTB and EBTB algorithm and the prior subsection, we used only some couples of beacons which the vertical line from the node C on the connecting line of those couples of beacons was positioned on the connecting line segment of those couples of beacons (Fig. 1). But we can extend the algorithm to use the other couples of beacons. The Fig. 2 shows such a couple of beacons. As it is seen in this figure, the vertical line from the node C on the connecting line of the couple of beacons doesn't position on the connecting line segment of the couple of beacons.

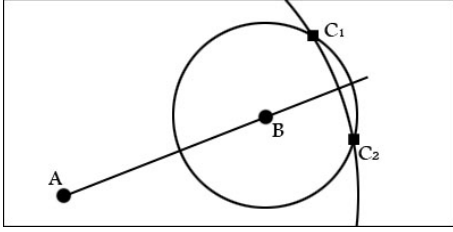


Fig.2. Positioning in a sensor network

### C. Third Improvement

Unfortunately, even with the prior improvements the described algorithm doesn't have a good accuracy in a noisy situation because the node C estimates its position using only two beacons with minimum  $|\mu|$ , while we can use the other couples of beacons to obtain a better estimation of the position of node C. In the other word, we can obtain an estimation of the position of the node C using each couples of beacons and then combine the estimated positions with different weights using the following equation to obtain a better estimation of the position of the node C (like WCL algorithm)

$$(x_N, y_N) = \frac{\sum_{i=1}^n \sum_{j=1, j \neq i}^n w_{ij} \cdot (x_{ij}, y_{ij})}{\sum_{i=1}^n \sum_{j=1, j \neq i}^n w_{ij}}, \quad (5)$$

where  $i$  and  $j$  are two beacons,  $n$  is the number of beacons,  $(x_{ij}, y_{ij})$  is the obtained estimation of the position of the node C using  $i$  and  $j$ ,  $(x_N, y_N)$  is the final estimation of the position of the node C. With regard to (4), if  $\epsilon_{ij}$  of two beacons  $i$  and  $j$  become close to zero,  $\mu_{ij}$  of them becomes close to zero, too. So, two problems occur.

First problem: In this case or in general when

$$\frac{\epsilon_{ij}}{iC + Cj} \quad (6)$$

become close to zero, the distance from the node C to the each of the beacons  $i$  and  $j$  will not affects much on  $\mu_{ij}$  which is not desirable and we want in this situation that the node C has positioned close to the connecting line of the couple of beacons,  $\mu_{ij}$  become close to zero only if the sum of the distances from the node C to each of the beacons  $i$  and  $j$  is a small value. In the other words, if the sum of the distances from the node C (which is close to the connecting line of the couple of beacon) to each of the beacons  $i$  and  $j$  is a big value, we should not rely much on the estimated position of the node C using this couple of beacons. But unfortunately in this case,  $\mu_{ij}$  become close to zero, so the described algorithm relies much on estimated position of the node C using this couple of beacons.

Second Problem: when  $\mu_{ij}$  of two beacons  $i$  and  $j$  become close to zero,  $w_{ij}$  of them become so big that  $(x_N, y_N)$  become close to a  $(x_{ij}, y_{ij})$  (the estimated position of the node C using this two beacons) and the effect of the other estimated positions of the node on  $(x_N, y_N)$  become too low which is not desirable (see (5)).

So, we should skip the couples of beacons which their  $\epsilon_{ij}$  is close to zero. In the other words, if

$$\frac{\epsilon_{ij}}{iC + Cj} < \delta, \quad (7)$$

We should skip the estimated position of the node C using the couple of beacons  $i$  and  $j$ .

## V. SIMULATION

To study the accuracy of EWCL and to compare it with WCL, EBTB and QR, we did many simulations in different situations in a square area with sides of 100 m. At the start of each simulation, some beacons (between 10 to 100 beacons, as an input parameter of simulation) and 10,000 nodes were randomly placed with a uniform distribution within the area. Then, the range between connected nodes was blurred by drawing a random value from a normal distribution having a normalized standard deviation with respect to the true range and having the true range as the mean. Maximum transmission range of each beacon was assumed between 20 m to 100 m.

Then, with a given input parameters, the position of each node was estimated using EWCL, EBTB, WCL and QR. Then, the number of the sensor nodes which could not estimate their position and the average of positioning error of sensor nodes (excluding the sensor nodes which

can not estimate their positions) and many other output parameters were registered. To account for the randomness in generating topologies and range errors we repeated each experiment 50 times with a different seeds, and report the averaged results.

Meanwhile, in the simulation, we found that 0.02 is a good value for  $\delta$  in (7). So, we present only the results of simulations which were done by this value of  $\delta$  and maximum transmission range of 50 m.

## VI. RESULTS

Fig. 3,4,5 show the accuracy of the proposed algorithm with respect to the three other algorithms, in three situation: when there is not any noise, when the noise is medium and when the noise is high. As it is seen in Fig. 3, in a noiseless situation, EWCL doesn't have a good accuracy. But in a real situation, when there is some noise in the environment, it has rather good accuracy and when AH becomes more than about 40, it becomes the best algorithm with the aspect of accuracy (Fig. 4). In high noise situation (Fig. 5), EWCL has the best accuracy with respect to the three other positioning algorithms.

Fig. 6,7,8,9 shows the sensitivity of each of the four localization algorithm to the noise. As it is seen, EWCL and WCL have low sensitivity to the noise, while QR and EBTB have high sensitivity to the noise.

## VII. CONCLUSION

As it was said, that WCL is with time complexity of  $O(n)$ , EWCL and BTB are with the time complexity of  $O(n^2)$  and QR is with time complexity of  $O(n^3)$ . So, EWCL is in the middle of the three other localization algorithms with respect to the resource requirement. But as it was seen, EWCL is the best localization algorithm when the noise is high and its accuracy is close to the accuracy of QR when the noise is medium.

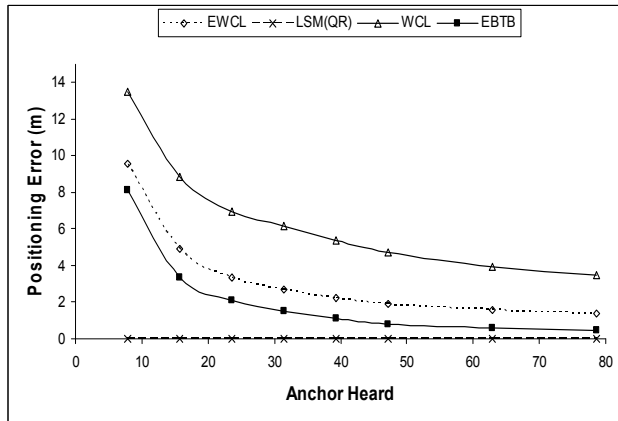


Fig. 3. Positioning error in a noiseless situation

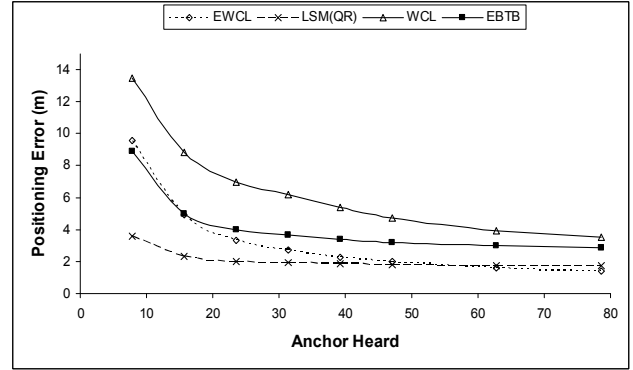


Fig. 4. Positioning error when the standard deviation of estimated range is about 0.033

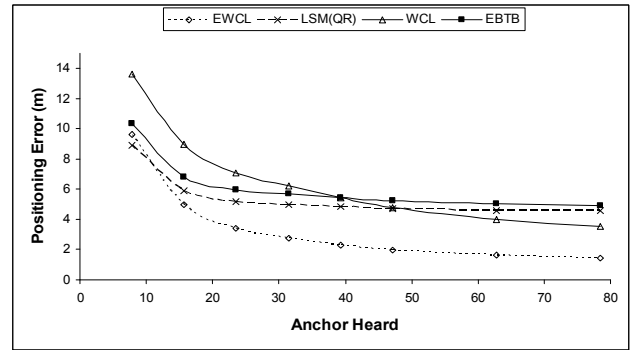


Fig. 5. Positioning error when the standard deviation of estimated range is about 0.083

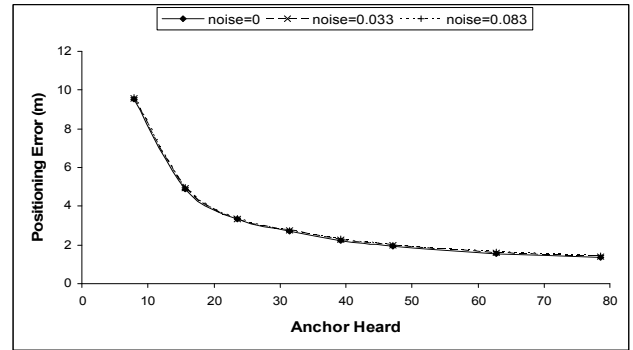


Fig. 6. Positioning error of EWCL

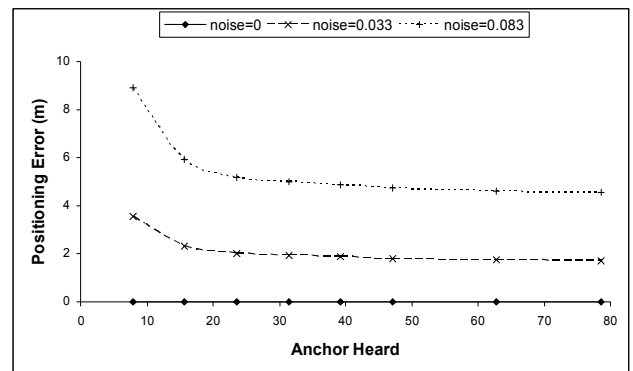


Fig. 7. Positioning error of QR



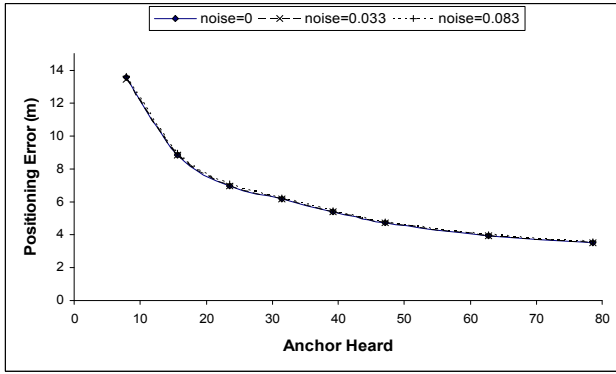


Fig. 8. Positioning error of WCL

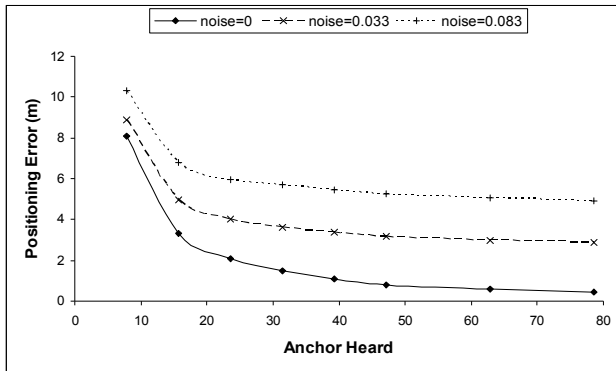


Fig. 9. Positioning error of EBTB

## REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramanian, E. Cayirci, "Wireless sensor networks: a survey," *Elsevier J. Computer Networks*, vol. 38, no. 4, pp. 393-422, 2002.
- [2] R. Bill, C. Cap, M. Kohfahl, T. Mund, "Indoor and outdoor positioning in mobile environments a review and some investigations on WLAN positioning," *Geographic Information Sciences J.*, vol. 10, pp. 91-98, 2004.
- [3] J. Gibson, *The mobile communications handbook*. CRC Press, 1996.
- [4] R. Min, M. Bhardwaj, S. Cho, A. Sinha, E. Shih, A. Wang, A. Chandrakasan, "Low-power wireless sensor networks," *Int. Conf. VLSI Design*, pp. 205-210, 2001.
- [5] F. Reichenbach, D. Timmermann, "Indoor Localization with Low Complexity in Wireless Sensor Networks," *IEEE Int. Conf. Industrial Informatics*, pp. 1018-1023, 2006.
- [6] C. Savarese, J. Rabaey, K. Langendoen, "Robust positioning algorithms for distributed ad-hoc wireless sensor networks," *USENIX Tech. Annu. Conf.*, pp. 317-327, 2002.
- [7] N. Bulusu, "Gps-less low cost outdoor localization for very small devices," *IEEE Personal Communications Mag.*, vol. 7, pp. 28-34, 2000.
- [8] H. Tian, H. Chengdu, B. M. Brian, S. A. John, A. Tarek, "Range-free localization schemes for large scale sensor networks," *9th Annu. Int. Conf. Mobile computing and networking*, pp. 81-95, 2003.
- [9] J. Blumenthal, F. Reichenbach, D. Timmermann, "Precise positioning with a low complexity algorithm in ad hoc wireless sensor networks," *PIK - Praxis der Informationsverarbeitung und Kommunikation J.*, vol. 28, pp. 80-85, 2005.
- [10] Y. Forghani, "A new approximate positioning approach in wireless sensor networks," *IEEE Int. Networking and Communications Conf.*, pp. 138-143, 2008.
- [11] Y. Forghani, "A binary approach for range-free localization," *IEEE National Conf. on Telecommunication Technologies*, 2008, in press.
- [12] A. Savvides, C. C. Han, M. B. Srivastava, "Dynamic fine grained localization in adhoc networks of sensors," *7th Annu. ACM/IEEE Int. Conf. Mobile Computing and Networking*, pp. 166-179, 2001.
- [13] Y. Kwon, K. Mechitov, S. Sundresh, W. Kim, G. Agha, "Resilient localization for sensor networks in outdoor environments," *25th IEEE Int. Conf. Distributed Computing Systems*, pp. 643-652, 2005.
- [14] A. A. Ahmed, H. Shi, Y. Shang: Sharp, "A new approach to relative localization in wireless sensor networks," *Second Int. Workshop on Wireless Ad Hoc Networking*, pp. 892-898, 2005.
- [15] T. C. Karalar, S. Yamashita, M. Sheets, J. Rabaey, "An integrated, low power localization system for sensor networks," *First Annu. Int. Conf. Mobile and Ubiquitous Systems: Networking and Services*, pp. 24-30, 2004.
- [16] K. Langendoen, N. Reijers, "Distributed localization in wireless sensor networks: A quantitative comparison," *Elsevier J. Computer Networks*, vol. 43, pp. 499-518, 2003.
- [17] F. Reichenbach, A. Born, D. Timmermann, R. Bill, "A distributed linear least squares method for precise localization with low complexity in wireless sensor network," in *Distributed Computing in Sensor Systems*, vol. 4026, Berlin, Springer, 2006, pp. 514-528.

# Performance evaluation of a wireless Orthogonal Frequency Division Multiplexing system under various concatenated FEC channel-coding schemes

Md. Dulal Haque<sup>1</sup>, Shaikh Enayet Ullah<sup>2</sup>, and Md.Razu Ahmed<sup>3</sup>

<sup>1</sup>Department of Computer& Communication Engineering, International Islamic University  
Chittagong, Chittagong-4203, Bangladesh. E-mail:dulal\_ape83@yahoo.com

<sup>2</sup>Department of Applied Physics And Electronic Engineering, Rajshahi University  
Rajshahi-6205, Bangladesh. E-mail:enayet67@yahoo.com

<sup>3</sup>Department Of Computer& Communication Engineering, International Islamic University  
Chittagong, Chittagong-4203, Bangladesh. E-mail:mdrazu@ete.iu.ac.bd

**Abstract** — In this paper, we study the effect of various concatenated Forward Error Correction (FEC) codes on the performance of a wireless Orthogonal Frequency Division Multiplexing (OFDM) system. In FEC concatenated channel code, the OFDM system incorporates Reed-Solomon (RS) encoder of (255,239,8), Cyclic Encoder of (15,11), Bose-Chadhuri-Hocquenghem (BCH) encoder of (127,64) with Convolution encoder of  $\frac{3}{4}$  and  $\frac{1}{2}$ -rated codes under different combinations of digital modulation (QPSK, 8PSK, 32-QAM and 64-QAM). The simulation study is made with the development of a computer program written in MATLAB source code on the processing of recorded audio signal under additive white Gaussian noise (AWGN) channel. The simulation results of estimated Bit error rate (BER) show that the implementation of concatenated RS(255,239,8) code with  $\frac{3}{4}$ -rated Convolutional code under QPSK modulation technique is highly effective to combat inherent interference in the communication system. Due to constraint in data handling capability of the MATLAB Editor, a segment of the recorded audio signal is used for analysis. The transmitted audio message is found to have retrieved effectively under noisy situation.

**Index Terms** — Orthogonal Frequency Division Multiplexing, Cyclic encoding, Reed-Solomon encoding, Convolution encoding (CC), Bose-Chadhuri-Hocquenghem (BCH) Bit error rate (BER) and Additive White Gaussian Noise (AWGN).

## I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is a multi carrier modulation technique well suited to overcome adverse effects in hostile transmission environment. This technique provides a reliable reception of signals affected by multipath propagation and selective fading and has been used in broadcast media such as European terrestrial digital video broadcasting (DVB-T) and digital audio broadcasting (DAB) and in the IEEE 802.11a(local area network, LAN) and the IEEE

802.16a(metropolitan area network, MAN) standards. OFDM is also being pursued for dedicated short-range communications (DSRC) for roadside to vehicle communications. With the advent of next generation (4G) broadband wireless communications, the combination of multiple-input multiple-output (MIMO) wireless technology with orthogonal frequency division multiplexing (OFDM) has been recognized as one of the most promising techniques to achieve high data rate and provide more reliable reception compared with the traditional single antenna system [1, 2]. However, in the present study, an effort has been made merely to concatenate the various channel encoding codes to improve the reliable reception performance of an OFDM wireless communication system under different digital modulation schemes such as QPSK, 8PSK, 32-QAM, 64-QAM. In OFDM, generally, the transmitted bit stream is divided into many different sub streams and sent over many different sub channels. Typically the sub channels are orthogonal under ideal propagation conditions. The data rate on each of the sub-channels is much less than the total data rate, and the corresponding sub-channel bandwidth is much less than the total system bandwidth. The number of sub streams is chosen to insure that each sub channel has a bandwidth less than the coherence bandwidth of the channel, so the sub channels experience relatively flat fading with insignificant amount of ISI effect. In almost all applications of multi-carrier modulation, satisfactory performance cannot be achieved without the addition of some form of channel coding. In wireless systems subjected to fading, extremely high signal-to-noise ratios are required to achieve reasonable error probability. In addition, interference from other wireless channels is frequently severe. On wire-line systems, large constellation sizes are commonly employed to achieve high bit rates. Coding in this case is essential

for achieving the highest possible rates in the presence of crosstalk and impulsive and other interference. Channel coding in OFDM systems can be implemented in time and frequency domain such that both dimensions are utilized to achieve better immunity against frequency and time selective fading. A concatenated coding scheme constituted from a combination of two 8-bit channel codes (block and convolution) along with proper time/frequency interleaving has been used in the present study. [3,4].

## II. TRANSMISSION SIMULATION MODEL

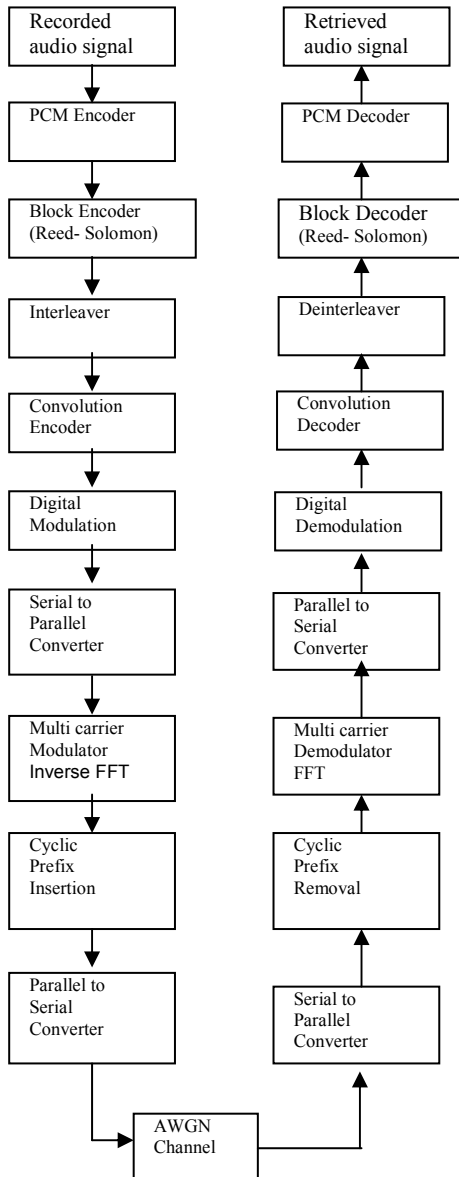


Fig. 1. A block diagram of an OFDM communication system with interleaved concatenated channel coding.

The block diagram of the simulated system model is shown in Figure 1. The input binary data stream obtained from a segment of recorded audio signal is ensured against transmission errors with Forward Error Correction codes (FECs) and interleaved. A block Reed Solomon (255, 239,8) code based on the Galois field GF (28) with a symbol size of 8 bits is chosen that processes a block of 239 symbols and can correct up to 8 symbol errors calculating 16 redundant correction symbols. The block formatted (Reed Solomon encoded) data stream is passed through a convolution interleaver of depth 12. A convolution code (CC) is applied to the interleaved symbols. Its rate  $R = m/n$ , where  $m$  is the number of input bits and  $n$  is the number of output bits- is equal to  $\frac{1}{2}$  and  $\frac{3}{4}$  and the constraint length,  $K$  of 5 and 7. The convolutionally encoded bits are interleaved further prior to converted into each of the either four complex modulation symbols in QPSK 8PSK, 32-QAM and 64-QAM modulation and fed to an OFDM modulator for transmission. The simulated coding and modulation schemes used in the present study are shown in Table 1.

TABLE I  
SIMULATED CONCATENATED CODING AND MODULATION SCHEMES

| FEC # | Modulation | RS code     | CC code rate  | BCH code    | Cyclic Code |
|-------|------------|-------------|---------------|-------------|-------------|
| 1     | QPSK       | (255,239,8) | $\frac{2}{3}$ |             |             |
| 2     | 8PSK       | (255,239,8) | $\frac{2}{3}$ |             |             |
| 3     | 32QAM      | (255,239,8) | $\frac{2}{3}$ |             |             |
| 4     | 64QAM      | (255,239,8) | $\frac{2}{3}$ |             |             |
| 5     | QPSK       | (255,239,8) | $\frac{3}{4}$ |             |             |
| 6     | 8PSK       | (255,239,8) | $\frac{3}{4}$ |             |             |
| 7     | 32QAM      | (255,239,8) | $\frac{3}{4}$ |             |             |
| 8     | 64QAM      | (255,239,8) | $\frac{3}{4}$ |             |             |
| 9     | QPSK       |             | $\frac{2}{3}$ | BCH(127,64) |             |
| 10    | 8PSK       |             | $\frac{2}{3}$ | BCH(127,64) |             |
| 11    | 32QAM      |             | $\frac{2}{3}$ | BCH(127,64) |             |
| 12    | 64-QAM     |             | $\frac{2}{3}$ | BCH(127,64) |             |
| 13    | QPSK       |             | $\frac{3}{4}$ | BCH(127,64) |             |
| 14    | 8PSK       |             | $\frac{3}{4}$ | BCH(127,64) |             |
| 15    | 32-QAM     |             | $\frac{3}{4}$ | BCH(127,64) |             |
| 16    | 64-QAM     |             | $\frac{3}{4}$ | BCH(127,64) |             |
| 17    | QPSK       |             | $\frac{2}{3}$ |             | CC(15,11)   |
| 18    | 8PSK       |             | $\frac{2}{3}$ |             | CC(15,11)   |
| 19    | 32QAM      |             | $\frac{2}{3}$ |             | CC(15,11)   |
| 20    | 64QAM      |             | $\frac{2}{3}$ |             | CC(15,11)   |
| 21    | QPSK       |             | $\frac{3}{4}$ |             | CC(15,11)   |
| 22    | 8PSK       |             | $\frac{3}{4}$ |             | CC(15,11)   |
| 23    | 32QAM      |             | $\frac{3}{4}$ |             | CC(15,11)   |
| 24    | 64QAM      |             | $\frac{3}{4}$ |             | CC(15,11)   |

In OFDM modulator, the digitally modulated information symbols are transmitted in parallel on sub carriers through

implementation as an inverse discrete Fourier transform (IDFT) on a block of information symbols followed by an analog-to-digital converter (ADC). To mitigate the effects of intersymbol interference (ISI) caused by channel time spread, each block of IDFT coefficients is typically preceded by a cyclic prefix [5,6]. At the receiver side, the received signal is OFDM demodulated, de-mapped, de-interleaved and then de-decoded in order to recover the data transmitted.

### III. SIMULATION RESULTS

BER performance of the wireless OFDM system under three concatenated FEC channel coding and different modulation schemes listed in the Table 1 have been evaluated. We first investigate the use of interleaved concatenated RS (255,239,8) and CC codes with code rates of  $\frac{1}{2}$  and  $\frac{3}{4}$  under QPSK, 8PSK, 32-QAM, and 64-QAM modulations. The BER performance is investigated and results are shown in Fig. 2.

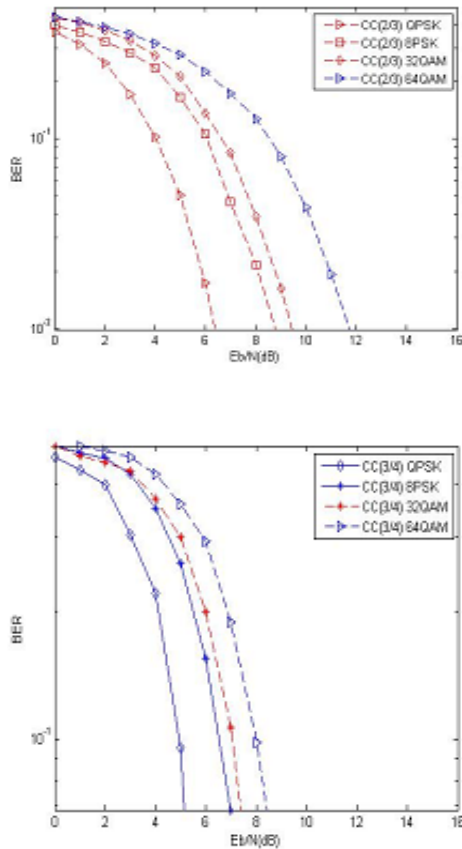


Fig. 2. System performance under different modulation schemes for RS (255,239,8) concatenated with CC (code rate of 2/3, 3/4)

The use of concatenated RS ((255,239,8) and CC with code rate  $\frac{1}{2}$  in 64-QAM modulation degrades the performance of the system. Next, we compare the BER performance of the OFDM system when channel coding is BCH and CC. The results are given in Fig. 3.

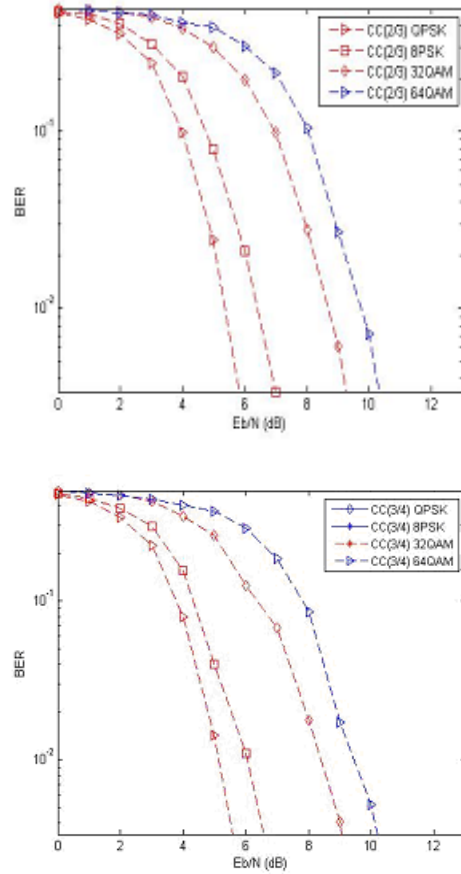


Fig. 3. System performance under different modulation schemes for BCH (127,64) concatenated with CC (code rate of  $\frac{1}{2}$  and  $\frac{3}{4}$ )

Finally; we compare the BER performance of the OFDM system when channel coding is Cyclic and CC. The results are given in Fig. 4. The simulation study gives an insight into ways of concatenating 8-bit Reed Solomon Code with  $\frac{2}{3}$  and  $\frac{3}{4}$  -rated Convolution code in 64-QAM modulations perform better than other concatenated codes and modulations. Fig. 2 through Fig. 4 show the bit error rate (BER) for different values of energy per bit to noise ratio ( $E_b/N$ ). Simulation results in Figure 2 show the advantage of considering a low ( $\frac{2}{3}$  and  $\frac{3}{4}$ ) convolution-coding rate for each of the four considered digital modulation schemes (QPSK, 8PSK, 32-QAM and 64-QAM). The performance of transmitted audio message is

found to have retrieved effectively under noisy situation that is shown in Fig. 5.

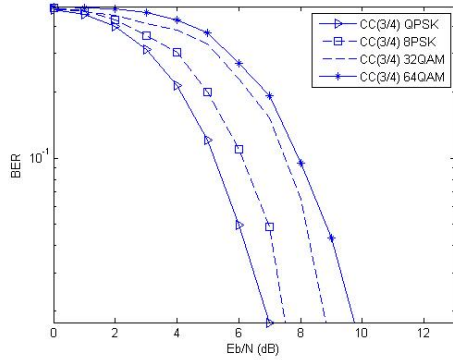
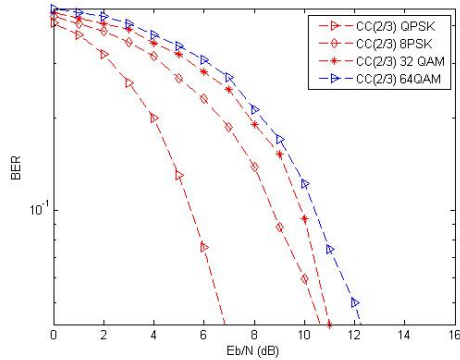


Fig. 4. System performance under different modulation schemes for Cyclic (15,11) concatenated with CC (code rate of  $\frac{1}{2}$  and  $\frac{3}{4}$ )

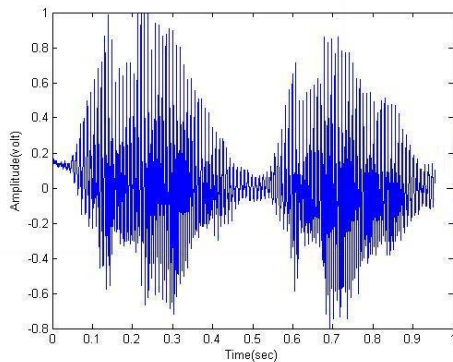


Fig. 5. (a) Transmitted audio signal

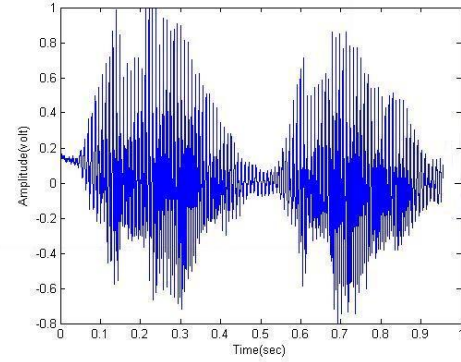


Fig. 5. (b) Retrieved audio signal

## VII. CONCLUSION

In this contribution, we have studied the performance of a wireless Orthogonal Frequency Division Multiplexing (OFDM) system adopting various concatenated FEC channel coding schemes and digital modulations. A range of system performance results highlights the impact of both digital modulation and concatenated channel coding schemes. In the context of system performance, it is observed that the implementation of interleaved Reed Solomon code with  $\frac{3}{4}$ -rated Convolution code under QPSK modulation technique provides satisfactory result among the three concatenated FEC channel coding schemes.

## REFERENCES

- [1] Gordon L. Stüber, Steve W. McLaughlin and Mary Ann Ingram, 2004: Broadband MIMO-OFDM wireless communications, Proceedings of the IEEE vol. 92, no. 2,, pp. 271-294
- [2] Wei Zhang, Xiang-Gen Xia and Khaled Ben Letaief, 2007: Space-time/frequency coding for MIMO-OFDM in next generation broadband wireless systems, IEEE Wireless Communications, pp. 32-43.
- [3] Andrea Goldsmith, 2005: Wireless Communications, Stanford University, Cambridge University Press.
- [4] Ahmad R. S. Bahai and Burton R. Saltzberg, 2002: Multi-Carrier Digital Communications, Theory and Applications of OFDM, Kluwer Academic Publishers, New York, USA.
- [5] T. Kratochvil, 2003: Utilization of Matlab for digital Image transmission Simulation Using the DVB Error Correction Codes, Radio Engineering, Vol.12, no.4, pp. 31-37.
- [6] L. J. Cimini, Jr. 1985: Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing, IEEE Trans. Commun., vol. COM-33, pp. 665-675.

# A Cost Effective Multi-granular Location Information Strategy for MANET

Koushik Majumder<sup>1</sup>, S. R. Biradar<sup>2</sup>, Subir Kumar Sarkar<sup>3</sup>

<sup>1</sup>Dept. of CSE, West Bengal University of Technology, Kolkata, India

<sup>2</sup>Department of Computer Engineering, SMIT, Sikkim, India

<sup>3</sup>Dept. of ETCE, Jadavpur University, Kolkata, India

koushikmail@yahoo.com

**Abstract** — In the literature there are several location based routing schemes available for MANET. The effective management of location information has the maximum impact on the performance of this type of routing protocols. In this paper first we have made a study of the different categories of existing location services. Then we have proposed a new location management scheme which effectively implements the major location management functions; namely location update and location query. In this scheme we have incorporated the multi-granular location information strategy. This is particularly aimed at reducing the orthogonal costs associated with location update and location query which we have clearly shown in the analysis part of this paper.

**Index Terms** — Location Query Cost, Location Service, Location Update Cost, Mobile Ad-Hoc Networks, Multi-Granular Location Information, Position Based Routing Protocols.

## I. INTRODUCTION

Position based routing protocols (such as GLS [1], SLURP [2], SLALoM [3], DLM [4]) use information about the physical position of the participating nodes, hence eliminating some of the limitations of topology based routing algorithms. Commonly, each node determines its own position through the use of GPS. The routing decision at each node is then based on destination's position and the position of forwarding node's neighbors.

In position based routing, the nodes have neither to store routing tables nor to transmit messages to keep the routing table up to date. The overhead is also reduced, since the establishment or maintenance of routes is not required in a protocol that uses location information for routing. Another advantage is scalability. And lastly it uses localized algorithm.

## II. REVIEW OF LOCATION SERVICES

### A. Flooding-based Location Services

Location services can be broadly divided into flooding-based and quorum-based approaches. Flooding-based protocols follow two different approaches: proactive and reactive. In case of proactive approach, each (destination)

node periodically floods its location information to other nodes in the network, which maintain location tables to keep track of this information. DLS is an example of proactive flooding-based location services.

### DREAM Location Service (DLS)

In the DREAM Location Service (DLS) [5] each location packet (LP) contains the coordinates of the source node based on some reference system, the speed of the source node and the last transmission time of the LP. Each mobile node in the ad hoc network transmits an LP to other nodes. The rate of transmission to nearby nodes is higher than to faraway nodes. The faraway nodes *appear* to move more slowly than nearby mobile nodes, and, therefore, it is not necessary for a mobile node to maintain up-to-date location information on faraway nodes. The overhead of location packets is limited by differentiating between nearby and faraway nodes. In DLS, the accuracy of the location information in the network is controlled via the frequency with which the LPs are transmitted and how far the LP is allowed to travel.

In case of reactive (on-demand) flooding-based approaches (e.g. RLS [5]), if a sender node cannot find a recent location of a destination, it floods a scoped query in the network in order to search the destination.

### B. Quorum-Based Location Services

In case of quorum-based protocols, all the nodes in the network agree, implicitly or explicitly, upon a mapping that associates each node's unique identifier to one or more other nodes in the network. The mapped-to nodes act as the location servers for that node. The location servers will store periodical location updates. Location queries will be routed to and looked up at these servers. As long as the lookup reaches one of the servers that was updated about the location corresponding to a given identifier, the correct location information will be found. In other words, the quorum of location servers updated on each location change needs to intersect the quorum of location servers consulted with in a lookup.

### C. Explicit Quorum-based Location Services

In the explicit quorum-based approach, each location update of a node is sent to an explicitly defined subset of available nodes. This subset forms the *update quorum*. On the other hand, a location query for that node is sent to a potentially different subset known as the *query quorum*. The two subsets are designed such that their intersection is non-empty, and thus the query will be satisfied by some node in the update quorum.

### D. Flat Quorum-based Location Services

In case of flat quorum-based location services, the operation is based on the notion of quorums. However, in these approaches, all quorum servers are symmetric in their roles. In particular, no hierarchical structure exists in such quorums. The Column-row quorum-based protocol [6] is an example of flat geographical quorum-based location service.

### E. Hierarchical Geographical Quorum-based Location Services

In case of hierarchical explicit geographical quorum-based location services a hierarchical structure of the quorum nodes is deployed. LLS [7] is an example of hierarchical explicit geographical quorum-based location service.

#### The Locality aware Location Service (LLS) [7]

The Locality aware Location Service (LLS) [7] uses hierarchical spiral structure in order to guarantee the *locality aware* properties of the lookup/publish algorithms. A location service has a *locality aware lookup* algorithm if the cost of locating destination  $t$  from source  $s$  is proportional to the cost of the minimal cost path between  $s$  and  $t$ . A location service has a *locality aware publish* algorithm if the cost of updating the location service due to the movement of a node from  $x$  to  $y$  is proportional to the distance between  $x$  and  $y$ .

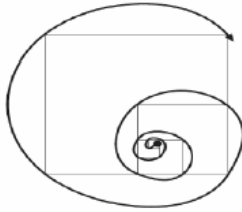


Fig. 1. Basic structure of LLS

Fig. 1 shows the basic structure of LLS. The issue of mobility is addressed in LLS by using virtual coordinates within the squares for storing information on  $t$  rather than searching for certain pre-designated nodes.

LLS achieves worst case lookup cost of  $O(d^2)$  and an average case linear cost of  $O(d)$ , where  $d$  is the minimal path length between source and destination. Additionally,

when a node moves distance  $d$ , the average cost of publishing its new location is  $O(d \log d)$ . The inherent locality of LLS makes it fault tolerant both to nodes failures and to network partitions.

### F. Implicit (Hashing-based) Quorum-based Location Services

In case of implicit quorum-based protocols, location servers are chosen via a hashing function, either in the node identifier space or in the location space.

### G. Flat Hashing-based Location Services

In flat hashing-based protocols (for example, SLALoM [3]), a well-known hash function is applied to map each node's identifier to a *home region* consisting of one or more nodes within a fixed location in the network area. All nodes in the home region act as location servers. They maintain the location information for the node and can reply to queries for that node. Typically, the number of location servers in the home region is independent of the total number of nodes in the network, and thus effectively each node's location is disseminated to  $O(1)$  location servers.

#### SLALoM [3]

SLALoM combines the strengths of SLURP [2] and GLS [1]. In this scheme each node is assigned multiple home regions distributed uniformly over the area in which the nodes move about. (The nodes in these home regions act as location servers for the node.) It is assumed that the nodes move about in a square region of area  $A$ . According to SLALoM the square is divided into  $G$  unit regions called *order-1 squares*. It then combines  $K^2$  of the order-1 squares to form *order-2 squares*. A node's home region will consist of an order-1 square. With some exceptions, every node has a home region in each order-2 square. Hence, every node has  $O(A/K^2)$  home regions.

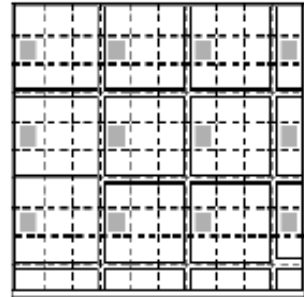


Fig. 2. Network hierarchy of SLALoM

### H. Hierarchical Hashing-based Location Services

In the hierarchical hashing-based protocols, the area in which nodes reside is recursively divided into a hierarchy of smaller grids. For each node, one or more nodes in each grid

at each level of the hierarchy are selected as its location servers. Location updates and queries traverse up and down the hierarchy. A major advantage of maintaining a hierarchy is that when the source and destination nodes are nearby, the query traversal is limited to the lower levels of the hierarchy.

The hierarchical grid structure is very efficient in distributing the load evenly across the network and in achieving efficiency and scalability of the location service. Since the height of the hierarchy is usually  $O(\log(N))$ , effectively each node's location is disseminated to  $O(\log(N))$  location servers.

The examples of hierarchical hashing-based protocols are Grid Location Service (GLS) [1], Distributed Location Management (DLM) [4].

#### GLS [1]

Grid Location Service (GLS) divides an area containing the ad hoc network into a hierarchical grid of squares. The largest square is called the level-H square. The level-H square is then recursively divided into four level-(H-1) squares until level-0 squares are reached, forming a so-called quad-tree. In each level- $i$  square (for  $i > 0$ ), node A selects three location servers, one in each level-( $i-1$ ) square that A is not in. The structure of GLS is shown in Fig. 3.

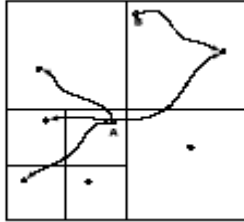


Fig. 3. Network structure of GLS

GLS selects the location servers based on the node ID in each server's service area (i.e., a quadrant). For a node C to be node A's location server, C has the smallest ID that is larger than A's ID in that quadrant, i.e.,  $C = \min \{x | \text{node } x \text{ is in the quadrant, } ID(x) > ID(A)\}$

Each node updates its location servers with its *exact* location after it moves a threshold distance  $\delta$ . To query for a particular node A, a node B sends the query to the node that is closest to A for which B has location information, and so on. Eventually the query would reach one of A's location servers.

### III. PROPOSED SCHEME

In our proposed scheme we have assumed that each node is equipped with GPS enabled devices through which the node can acquire location information about its current

geographic position. We also assume that each node has a transmission range of  $r_t$ .

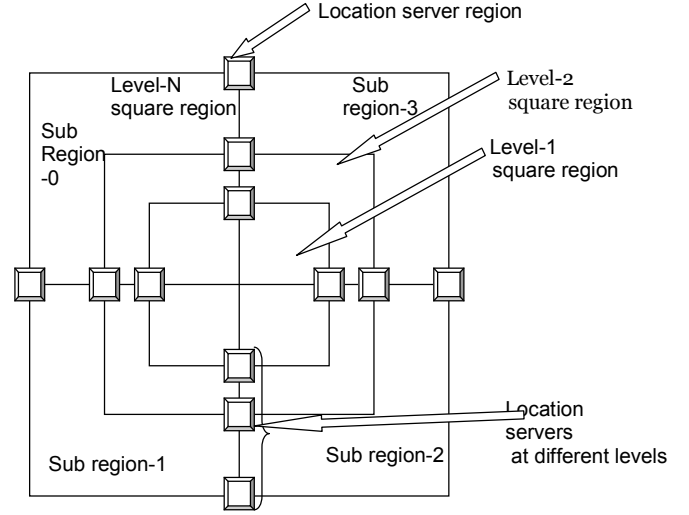


Fig. 4. Complete network structure of our proposed scheme

In our scheme we have divided the entire network area into  $N$  level of square regions. The arrangement is such that each level  $k$  square region encapsulates the level  $(k-1)$  square region and is encapsulated by level  $(k+1)$  square region. Each square region has a side length of  $2.2^l s$ , where  $l$  denotes the level number and  $s$  depends on the node density. The innermost region is the level-1 square region and the outermost region is the level- $N$  square region.

An implicit characteristic of MANET is the high mobility of nodes and frequent changes in network topology. As a result, the cost for location update becomes a major burden. If we keep track of only the fully qualified location information of the nodes, then there is a possibility of this information becoming stale quickly and thus requiring to invoke expensive location update routines. This increases the location update cost tremendously. In order to reduce this cost, in our scheme, we apply the concept of multi-granular location information strategy. We have assumed that the location information can be of two types – fully qualified location information and relative location information.

The fully qualified location information of a node includes the following parts.

|                    |
|--------------------|
| Node -id           |
| x-coordinate       |
| y-coordinate       |
| Location server-id |

Fig. 5. Fully qualified location information



On the other hand relative location information has two components.

|                    |
|--------------------|
| Location server id |
| Node id            |

Fig. 6. Relative location information

The location servers within a location server region are responsible for keeping track of the fully qualified location information of only those nodes that are currently within its assigned sub-region. On the other hand the location servers within a particular location server region also keep track of the relative location information of all the nodes that are currently within other sub-regions of the same level  $i$  square region. Therefore, when a node moves within a particular sub-region it needs to notify only the single location server region- that is currently in charge of that sub-region, regarding the change in its fully qualified location information, thereby reducing the location update cost.

#### A. Location Update

Location update has three components.

I. Location update for node movement within sub-region:

When a node A moves within the sub-region in which it is currently residing it needs to notify only the location servers currently in charge of that particular sub-region.

II. Location update for node movement between sub-regions:

On reaching its new sub-region the node A probes its neighbors to get information about its new location server region. Once it collects this information, it sends its current  $x$  and  $y$  coordinate positions and the node id to the new location server region. Now the node A is under the direct supervision of this new location server region. Therefore, this new location server region needs to update the location information regarding node A from relative to fully qualified. This new location server region then sends the new relative location information of the node to other location server regions within the same level  $i$  square region. Each of these other location server regions need to modify the location information about the node.

III. Location update for node movement between square regions at different levels:

After reaching its new sub-region within its new square region, node A probes its neighbors to get information about its new location server region. Once it gets this information, it sends its previous fully qualified address and the current  $x$  and  $y$  coordinate positions to this new location server region. The new Location server region can know the

previous level no. of node A from its previous fully qualified address. This new location server region then sends the new relative address of the node, (i.e., current location server id and node id) to the location server region in the previous level. This information is then relayed to all the other location server regions in the previous level. These location server regions after analyzing the current relative address of the node, delete the entry corresponding to node A from their database to reflect the fact that the node A is no longer at this level.

As the new location server region is in a square region which is at a level that is different than the level of its previous square region, therefore, it must make a new entry in its location information database about the new fully qualified location information of node A. This new location server region then needs to send the relative location information of node A to other location server regions within the new square region. These other location servers being in a square region at a level which is different than the previous square region, had no location information about node A and therefore, they will need to make new entries about the new relative location information of node A.

#### B. Location Query

Suppose a source node A wants to send a data packet to a destination node B but node A does not know the location information of node B. Under this scenario three situations can evolve.

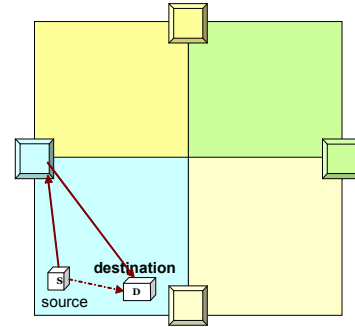


Fig. 7. Destination B is within same sub-region at same level as of source A

I. Destination B is within same sub-region at same level as of source A:

In this case the location server region that is in-charge of the sub-region contains the fully qualified address of node B. The source node A sends the data packet to the location server region. The location server region extracts the current  $x$  and  $y$  coordinate position of node B from its fully qualified address and sends the data packet to node B at that position.

II. Destination B is within other sub-region at the same level as of source A:

In this case the source node A sends the data packet to the assigned location server region of its sub-region. But the destination B is within a different sub-region. Therefore, the location server region of node A contains only the relative location information about destination B. From this information, the location server region of node A can find which location server region is currently containing the fully qualified address of node B. The location server region of node A then sends the data packet forwarded by node A to that particular location server region. This new location server region ultimately sends the data packet to the destination node B.

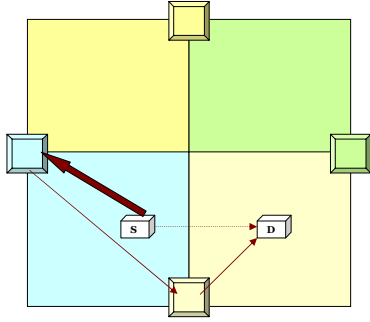


Fig. 8. Destination B is within other sub-region at the same level as of source A

III. Destination B is within other square region at different level than that of source A:

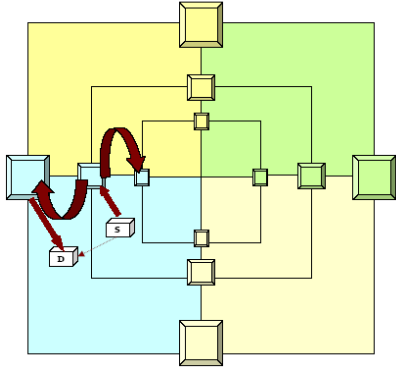


Fig. 9. Destination B is within other square region at different level than that of source A

The location server region now sends the data packet to the location server region of the square region that is encompassing the current level square region. It also forwards the packet to the location server region of the square region that is contained by the current level square region. The location server regions at other levels now follow the previously mentioned steps for location query. This process is continued until the destination node B is

found or the network boundary is reached. Thus, if the destination node falls within the network boundary, the data packet is propagated from the source node S to the destination node B through the intermediate location server regions.

#### IV. ANALYSIS OF LOCATION MANAGEMENT COSTS

Basically there are two types of costs, which are important for any location management scheme. They are - cost for location update and cost for location query. In the following sections we analyze these two types of costs.

##### A. Location Updation Cost [ $Cost_{update}$ ]

The cost for location update can be divided into three parts - i>Cost for location update for node moving within sub-region ( $Cost_{update-intra-subregion}$ ) ii>Cost for location update for node moving between sub-regions ( $Cost_{update-inter-subregion}$ ) iii>Cost for location update for node moving between square regions at different levels( $Cost_{update-inter-level}$ ).

Thus we can write,

$$Cost_{update} = Cost_{update-intra-subregion} + Cost_{update-inter-subregion} + Cost_{update-inter-level}$$

The cost for location update depends upon the amount of forwarding load. Thus the forwarding load, and as a result the cost will be higher for a packet traveling a greater distance. Cost for location update for node moving within sub-region ( $Cost_{update-intra-subregion}$ ) is basically the product of updation frequency and the cost of updation of one location server region. The cost of updation of one location server region is proportional to the average number of hops an update packet takes to reach the assigned location server region. We denote this cost by  $Cost(1)$ . We can approximate this cost by considering the distance  $D=\sqrt{2}.2^l.s$ ; where  $l$  denotes level number.

Let us denote  $z$  as the average progress for each forwarding hop where  $z$  is a function of the radio transmission range  $r_t$  and the node density ( $\gamma$ ) [8]. We assume both  $r_t$  and  $\gamma$  are constants. Therefore,  $z$  is also a constant. It is possible to derive the average number of hops an update packet takes by  $D/z$ . If we consider the average velocity of a node as  $v$ , and the transmission range of a node as  $r_t$ , then the updation frequency is  $v/r_t$ . Thus,

$$Cost_{update-intra-subregion} = v/r_t \cdot Cost(1)$$

$$\begin{aligned} \text{and } Cost(1) &\propto \sum_{l=0}^N \sqrt{2}.2^l.s/z \\ &\approx \sqrt{2}.s.N/z. \end{aligned}$$

If we assume  $S$  as the side length of the square region at the maximum level, i.e.  $N^{\text{th}}$  level square region,

then,  $S \propto 2^N$ . Thus,  $N \propto \log S$ . Since,  $S \propto \sqrt{M}$ , ( $M$ =Total Number of nodes in the network), we have  $N \propto \log \sqrt{M}$ .

$$\text{Thus, Cost}_{\text{update-intra-subregion}} = O(v \cdot \log \sqrt{M}). \quad (1)$$

Cost for location update for node moving between sub-regions ( $\text{Cost}_{\text{update-inter-subregion}}$ ) is the product of the boundary crossing rate ( $\Omega$ ) and the cost for updating the four location server regions ( $\text{Cost}(4)$ ).

$$\text{So, Cost}_{\text{update-inter-subregion}} = \Omega \cdot \text{Cost}(4).$$

The boundary-crossing rate is proved [9] to be proportional to  $v$ . The cost of updating four location server regions can be approximated by  $4(D_1)/z$ . Thus

$$\begin{aligned} \text{Cost}(4) &\propto \sum_{l=0}^N 4 \cdot \sqrt{2} \cdot 2^l \cdot s/z \\ &\approx 4\sqrt{2} \cdot s \cdot N/z. \end{aligned}$$

$$\text{Therefore, Cost}_{\text{update-inter-subregion}} = O(v \cdot \log \sqrt{M}). \quad (2)$$

Similarly we can formulate  $\text{Cost}_{\text{update-inter-level}}$  as

$$\text{Cost}_{\text{update-inter-level}} = \Omega \cdot \text{Cost}(8).$$

We can approximate the cost of updating eight location server regions by  $4(D_1 + D_{l-1})/z$ . Thus

$$\begin{aligned} \text{Cost}(8) &\propto \sum_{l=0}^N 6 \cdot \sqrt{2} \cdot 2^l \cdot s/z \\ &\approx 6\sqrt{2} \cdot s \cdot N/z. \end{aligned}$$

$$\text{Therefore, Cost}_{\text{update-inter-level}} = O(v \cdot \log \sqrt{M}). \quad (3)$$

Thus from “(1)”, “(2)” and “(3)” we have

$$\begin{aligned} \text{Cost}_{\text{update}} &= \text{Cost}_{\text{update-intra-subregion}} + \text{Cost}_{\text{update-inter-subregion}} + \\ &\quad \text{Cost}_{\text{update-inter-level}} = O(v \cdot \log \sqrt{M}). \end{aligned}$$

### B. Location Query Cost [ $\text{Cost}_{\text{query}}$ ]

If a source node has some data to send to a destination node, the source node must first query a location server region to get the current location information of the destination node. The cost for this activity of querying the location information is known as location query cost ( $\text{Cost}_{\text{query}}$ ). In order to calculate  $\text{Cost}_{\text{query}}$ , we have to measure the expected number of forwarding hops traveled by a query packet from the source node to its assigned location server region, which can be approximated by  $D/z$ . Therefore, the expected query cost is,

$$\begin{aligned} \text{Cost}_{\text{query}} &= \sum_{l=0}^N \sqrt{2} \cdot 2^l \cdot s/z \\ &\propto H \\ &= O(\log \sqrt{M}). \end{aligned}$$

## V. CONCLUSION

In this paper, we have discussed some of the existing location service schemes and have proposed a new scheme for location based routing among the nodes in mobile ad hoc network. The use of geo-location information allows us to avoid network wide flooding of data and control packets and reduces the total normalized overhead. This reduction is significant as it saves a lot of time as well as energy which are serious constraint issues one needs to keep in mind while designing the network ensuring its smooth functioning so that an optimum performance is tenable. In our scheme by dividing the entire network area into  $N$  level of square regions and using multi-granular location information, we have been able to provide a unique way to reduce the cost associated with both location update and location query. Further investigation on performance analysis of this scheme in different network scenarios can be taken as extended work.

## REFERENCES

- [1] J. Li, J. Jannotti, D. De Couto, D. Karger, and R. Morris. A scalable location service for geographic ad-hoc routing. In *Proceedings of ACM MobiCom*, pages 120.130, August 2000.
- [2] Seung-Chul M. Woo and Suresh Singh. Scalable routing protocol for ad hoc networks. *Wireless Networks*, 7(5):513-529, 2001.
- [3] Christine T. Cheng, H. L. Lemberg, Sumesh J. Philip, E. van den Berg, and T. Zhang. SLALoM: A scalable location management scheme for large mobile ad-hoc networks. In *Proceedings of IEEE WCNC*, March 2002.
- [4] Y. Xue, B. Li, and K. Nahrstedt. A scalable location management scheme in mobile ad-hoc networks. In *Proceedings of the IEEE Conference on Local Computer Networks (LCN '01)*, 2001.
- [5] T. Camp, J. Boleng, and L. Wilcox. Location information services in mobile ad hoc networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pages 3318-3324, 2001.
- [6] I. Stojmenovic. A scalable quorum based location update scheme for routing in ad hoc wireless networks. Technical Report TR-99-09, University of Ottawa, September 1999.
- [7] I. Abraham, D. Dolev, and D. Malkhi. LLS: a Locality Aware Location Service for Mobile Ad Hoc Networks. In *Proc. of the Joint Workshop on Foundations of Mobile Computing (DIALM-POMC)*, pages 75–84. ACM Press, 2004.
- [8] Seung-Chul M. Woo and Suresh Singh. Scalable routing protocol for adhoc networks. *Wireless Networks*, 7(5):513-529, 2001.
- [9] Yinze Yu, Guor-Huar Lu, and Zhi-Li Zhang, “Enhancing location service scalability with HIGH-GRADE,” Dept. of Comp. Sci. & Eng., University of Minnesota, Technical Report TR-04-002, 2004

# QoS based Fair Load-Balancing: Paradigm to IANRA Routing Algorithm for Wireless Networks (WNs)

Ayyoub Akbari Moghanjoughi <sup>\*,1</sup>, Sabira Khatun<sup>\*</sup>, Borhanuddin M. Ali<sup>\*</sup>, Raja S. A. R. Abdullah<sup>\*</sup>

<sup>\*</sup> Wireless Network Laboratory, Department of Computer & Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor Darul Ehsan, Malaysia.

**Abstract** — In this paper, a new algorithm, Intelligent agent AntNet based Routing Algorithm (IANRA) is proposed to enhance load balancing strategy in Wireless Networks (WNs). IANRA is based on Ants behaviour with some important factors such as: the specific self-organizing behaviour of ant colonies, the shortest path discovery and the related framework of Ant Colony Optimization (ACO). The main focus in IANRA is to find optimum and near optimum route by means of Genetic Algorithm (GA) using breeding capability of ants. Here, ants can produce a number of generations with the target to discover an optimized route. Hence IANRA is able to prevent of the difficulties which exist in existing routing algorithms, such as Ad hoc On-Demand Distance Vector (AODV), AntNet etc. The obtained results show that the efficiency of IANRA algorithm is better than AODV and AntNet or any other related algorithm. In addition, the proposed algorithm (IANRA) is able to reduce the end-to-end delay and increase the packet delivery ratio significantly.

**Index Terms** — IANRA, Routing Algorithm, Ant Colony Optimization (ACO), Load Balancing, Wireless Communication Networks, QoS.

## I. INTRODUCTION

In the past few years, a lot of research interest has been seen for routing in Wireless Networks (WNs). As mobile hosts and wireless network equipments have become widely available, an entirely new class of wireless application has been created, which wired network infrastructure can not support. With the rapid growth and evolution of wireless computer networks, not only their topologies are changing and progressing in continuous form, but also manner and type of utilization are permanently changing for various goals. Present routing algorithms, used in wireless computer networks are not suitable to coordinate with the instant growth. Hence, new algorithms for routing and load balancing are needed [1, 2, 3].

New mobile Ad-hoc networks (MANETs) are networks in which all nodes are mobile and communicate with each other via wireless connections. Nodes can join or leave a specific network domain at any time. There is no fixed infrastructure, all nodes are equal and there is no centralized control. There are no designated routers; any

node can serve as router for others, and data packets are forwarded from node to node in a multi-hop fashion [1]. In other hand WNs consist of static wireless routers and gateways, which are directly connected to the wired infrastructure. User stations are connected to the wired infrastructure via wireless link.

The main goal of routing algorithm is to maintain the Quality of Service (QoS) and to prevent the traffic overload in some parts of the networks [2, 3, 4]. First Dorigo and Di Caro proposed AntNet algorithm by considering only the delay in their routing tables [4, 5, 6]. It has been observed that, ants in a colony can focus on moving over the shortest path among various paths connecting their nest to source of food [7, 8]. The main catalyst of this colony-level shortest path behaviour and discovery is the use of a volatile chemical substance called pheromone. Usually shorter paths are discovered quickly and more frequently by the ants, and therefore can be marked with higher pheromone intensity [9].

The Asymmetrical Route Synchronization (ARS) algorithm is based on agents, and supports QoS routing, resource reservation, and admission control functions using ant-like agents [2, 6]. It works under two conditions: (i) every node in a network supports the weighted fair queuing algorithm and (ii) network users require two service class deliveries: datagram and real-time flow. The ARS efficiently allocates network resources according to user requests. Details of the algorithms and their performance comparison can be found in “Ant-Routing vs. Q-Routing in Telecommunication Networks” by Pacut et al., 2006 [10].

In this paper, a new algorithm is proposed by considering the QoS, the issue of load balancing and optimum routing for Wireless Networks. The algorithm is based on a specific self-organizing behaviour of ant colonies, the shortest path discovery, and on the related framework of ant colony optimization [7].

This paper is organised as follows. Section II, describes related routing algorithms based on ants colonies behaviour. Section III, describes the IANRA algorithm. Section IV, presents performance evaluation for IANRA,

<sup>1</sup> A. A. Moghanjoughi (✉), E-mail: ayyoub@ieee.org

followed by the simulation results in Section V and finally conclusion in Section VI.

## II. ROUTING ALGORITHMS BASED ON ACO BEHAVIOUR

One of particularly brilliant experiment was designed and run by Denebourg and colleagues [11]. They ran experiments varying the ratio,  $r = l_l / l_s$ , between the length of two branches of a double bridge, where  $l_l$  was the length of the longer branch and  $l_s$  the length of the shortest one. In their first experiment (shows in Figure 1.a), the bridge had two branches of equal length ( $l_s = l_l = 1 \Rightarrow r = 1$ ). At the beginning, ant's based agents were left free to move between the N (Nest) and F (Food) and the percentage of ants that chose one or the other of the two branches were observed over time. Then used the other bridge with two different length of branches which are depicted in Figure 1.b ( $l_s = 1$  and  $l_l = 2 \Rightarrow l_s \neq l_l$ ). In Figure 1.a the probability of path/branch selection was same depending to the pheromone congestion on each of paths. In second experiment (Figure 1.b), the length ratio between the two paths was set to  $r = 2$ , so that the longer path was twice as long as short one [11].

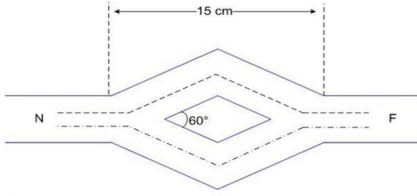


Figure 1.a. Branches with equal length ( $l_s = l_l = 1 \Rightarrow r = 1$ ).

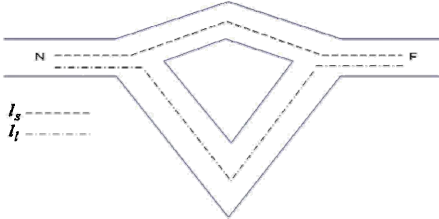


Figure 1.b. Branches with unequal length ( $l_s = 1$  and  $l_l = 2 \Rightarrow r = 2$ ).

If,  $\rho_{i\gamma}(t)$  is the probability of the ant arriving at decision spot  $i \in \{1, 2\}$  (see Figure 1.b) and selects paths  $\gamma \in \{s, l\}$ , also  $\phi_{i\gamma}(t)$  denote the total amount of pheromone on the path. For example, the probability of discovering the shorter path is given by Equation (1).

$$\rho_{is}(t) = \frac{(t_s + \phi_{is}(t))^\alpha}{(t_s + \phi_{is}(t))^\alpha + (t_s + \phi_{il}(t))^\alpha}, \quad \alpha = 2 \quad (1)$$

Where,  $\alpha$  is two and  $\rho_{il}(t) + \rho_{is}(t) = 1$ . Now suppose  $\psi$  be the number of ants cross the path per second in one direction at constant speeding of  $v$  cm/s that depositing one unit pheromone on the way. So, an ant will choose the short path traverse it  $t_s = l_s / v$  seconds. This time for longer path will be  $r \times t_s$  seconds, where  $r = l_l / l_s$ . To describe the evolution of the stochastic systems, the differential equations are as follows.

$$d\phi_{is} / dt = \psi \rho_{js}(t - t_s) + \psi \rho_{is}(t), \quad (i = 1, j = 2; i = 2, j = 1) \quad (2)$$

$$d\phi_{il} / dt = \psi \rho_{jl}(t - r \cdot t_s) + \psi \rho_{il}(t), \quad (i = 1, j = 2; i = 2, j = 1) \quad (3)$$

In the basic ACO algorithm the base idea for routing [8, 12] is the possession of routing information through the sampling of paths using small control packets, which are called ants. The ants are generated concurrently and independently at the nodes, with the task, to test a path, from Source (S) to Destination (D). Ants always sample complete paths, so that routing information can be updated in a pure Monte Carlo way, without relying on bootstrapping information from one node to the next [13]. The routing table for each destination contain a vector of real-valued entries. One entry indicates source to destination route for each known neighbour node. They are termed pheromone variables, and are continually updated according to path quality values calculated by the ants. The repeated and concurrent generation of path-sampling ants results, the availability of a bundle of paths, each with an estimated measure of quality, at each node. In turn, the ants use the routing tables to define the path they may sample.

The pheromone information is used for routing data packets, more or less in the same way as for routing ants. Packets are routed stochastically, giving higher probability to links with higher pheromone values. Usually, data for the same destination are spread over multiple paths (but with more packets travel over the best paths), resulting balanced load in the network. This mechanism is usually adopted to avoid low quality paths, while to make the ants more explorative, so that the good paths can be maintained and less good paths can be avoided. The path exploration is kept separate from the ongoing paths. Enough ants are sent to different destinations to make sure nodes have up-to-date information about the best paths and can automatically adapt their data load accordingly.

In computer networks, probability of using routing tables, in all nodes in networks for routing, can act as pheromone. This probability is calculated with ant's traffic or packages [4, 6]. Dorigo and Di Caro show a method, named Basic AntNet(B-AntNet) where the information that comes from Ant's appear in each node

like a routing table and a data structure (LTS) that involve information about local traffic and delay quantities[5, 6].

To accomplish this task, they are using ants to gather information about the traffic load in the network. Pheromone matrix is an isomorphic local data-base that helps router to decide where to forward data packets or not. It contains the information which specifies the next (neighbour) node to travel by a data packet to go to any possible destination in the network. Each routing table is organized as a set of all the possible destinations (all the nodes in the network) and the probabilities to reach these destinations through each of the neighbours of the node which is located in next hop.

The routing tables are organized like vector distance algorithms. The local model (LTS) shows structure of local traffic plays the role of a local adaptive model for the expected delay toward each possible destination. In AntNet, two different kinds of ants were used: Forward Ant (FA) and Backward Ant (BA) [4, 5, 6].

The quantity of  $\rho_{j,i}$  in the pheromone matrix (routing table) is calculated as follows: if only one ant arrive (from node  $i$ ) to node  $j$ , the probability of choosing that route will be increased base on Equation (4). In addition, probability of connective routes to other nodes will be decreased base on Equation (5), because the probability table must be normal. Finally, equation (6) is correct for all columns in the probability table in Fig 2 [4, 7].

$$\rho_{j,i} = \frac{(\rho_{j,i})' + \Delta\rho}{1 + \Delta\rho}, \quad (\rho_{j,i})' = \rho_{j,i} \text{ (OLD)} \quad (4)$$

$$\rho_{j',i} = \frac{(\rho_{j,i})'}{1 + \Delta\rho}, \quad j' \in \text{Neighbour} \Rightarrow j' \neq j \quad (5)$$

$$\sum_{j \in L_n} \rho_{j,i} = 1, \quad j \in [1, N], \quad L_n = \{\text{Neighbour}(n)\} \quad (6)$$

$$\Delta\rho = \eta + 0.005, \quad \eta = 0.08 / \lceil \xi \rceil \quad (7)$$

$\Delta\rho$  in Equation (7) shows the effect of ant movement at the entrance of probability table and it is created from delay reverse or ant age ( $\xi$ ), plus one fix number. If, in a route, ant age and delay are low, the probability of use of that route is high and then inclination of other Ant's to follow that route will be increased.

One of the most important problems that exist in this routing algorithms with ant's colonies is statistic route and not adaptive as per need in the network [14, 15, 16]. Statistic means that after Ants choose the  $\mathcal{R}_{OP}$  route as optimum route, the inclination of other ant's to choosing this route is usually increased. This causes traffic congestion due to increased traffic and at the same time the probability of use of other routes will be decreased. Choosing  $\mathcal{R}_{OP}$  route as the only optimized route in the network caused undesired result and problems like no optimization of  $\mathcal{R}_{OP}$ .

To solve this statistic nature route problem, Multiple Ant Colony Optimization (MACO) [14] use different ants with their special "pheromone" instead of using one kind of ant. In MACO, each kind of ant can find the separate optimum route and this increase route adaptation and solve the statistic route problem.

In addition, Bonabeaul has showed that success of ants when they act as gregarious to find the shortest route is proportional. If a group of ants choose non-optimum route,  $\mathcal{R}_{N-op}$ , by chance, the other ants will choose and follow the same  $\mathcal{R}_{N-op}$  at the end. The concentration of 'pheromone' will be increased and the non optimized route  $\mathcal{R}_{N-op}$  will be used as optimum route. Therefore, non-efficient routing can be created [15, 16].

In this paper, the proposed IANRA attempts to decrease the probability of choosing not optimized route instead of optimized one by more searches among possible routes. Also, instead of using one optimizing route, use of several optimized (or sub-optimized) routes can solve the congestion problems caused by a single optimizing route.

### III. THE PROPOSED INTELLIGENT ANTNET BASED ROUTING ALGORITHM

In this Algorithm, the used ants are able to produce new generation using *GA* concept. In the routing process when an ant from a particular family arrives to a junction node with  $m$  different connective routes, the ants regenerate  $m$  successive ants and then send them one per each connective route accordingly based on the same considered criterion. These new generations of ants inherit identifier of their family, the generation numbers and the routing information from their parents that includes route delay, bandwidth and identification of visited nodes. IANRA uses various kinds of ant for routing while one type of ant is applied in MACO. In addition, each kind of ant is used for one specific type of services for different QoS needs. To achieve a complete optimum routing table, several kinds of ants with different quality of services are used for each node.

The AntNet method converted to prepare QoS, required bandwidth and stage or Hop Count (*HC*) [3] by forcing the *FA* to use all available connective routes. Whenever a *FA* passes more than a certain number of stages,  $H_{max}$  it dies and the routing algorithm of AntNet is updated based on number of stage [16].

In IANRA, minimum bandwidth (*BW*) and maximum stage ( $H_{max}$ ) are considered as QoS parameters. As usual, here also, if *FA* passes the stage more than  $H_{max}$  it die and an ant produces  $m$  new ants when it arrives to a node with  $m$  connective routes in it to cover all possible routes. While ants move towards destination, like AntNet method, store the identification of visited nodes and



quantity of QoS routes in their stack. When an ant arrives to a new node, it first checks whether any co-family ant has already pass this node or not. If passed, it means an ant passes this node in a shorter route than the present ant, so the present route is not optimum and this ant will die there. Otherwise the ant regenerates like its parent and send new generation of ants towards the route that can prepare QoS conditions. This process continues until ants of that family arrive to destination. When the primary ant arrives to destination it finds out, first: whether the route that it has passed becomes the shortest route or not, second: the QoS condition. Therefore, the primary ant produces *BA* ant and transfer the information that has been gathered during route optimization period and then die there. Here *BA* has duty to use existing information stack transferred by *FA*, follow backward the same route created by *FA* and update the existing routing tables.

For example, in Figure 2, an ant in the *a* family at time *t* is going to search optimum route from *S* to *D*. First, this ant generates four new ants as there are total four routes from *S* to next neighbouring node. These four new ants are considered as 1st generation  $\delta I$  and will follow the routes *SA*, *SB*, *SC* and *SE*. In each stage, when an ant moves forward from one node to next, the generation number will be increased accordingly (depicted in Figure 2). But no ant will generate on *A* node for discovering *AD* route due to longest distance and weak signal in coverage area (i.e., Lack of QoS). Let  $\theta$  be a node number which are discovered by *FA*, and  $\tau$  denoted service time per node on supposed network, then the value of  $\xi$  for each of ant which is arrived to *D* from *S* is;

$$\xi = \sum_{k=1}^n \xi_{\delta k} + \sum_{l=1}^m \tau_{\theta} \quad (8)$$

Number of generation *s*,  $k = \{1, 2, 3, \dots, n\}$   
Number of nodes on  $\mathcal{R}_{op} \Rightarrow l = \{1, 2, 3, \dots, m\}$

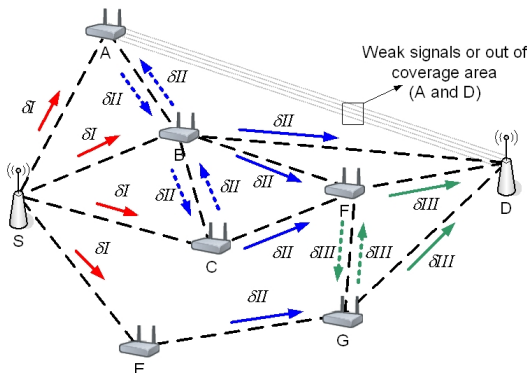


Figure 2. Distribution intelligent ant based agent's generations on sample network.

When ant  $\delta II$  arrives from node *F* to *G* and see that a co-family ant  $\delta II$  with smaller generation already passed

it, then the route, *SCFG* will not be optimized and ant  $\delta II$  dies there. Same thing happens for route *SEGF*.

At destination *D* three co-family ants  $\delta II$ ,  $\delta III$  have reached there from three different routes *SBD*, *SCFD* and *SEGD* respectively. It is considered that  $\delta II$  ant passes the shortest route compared the other Ants and finally the routing table will be updated with *SBD* as optimum route for *SD* connection. The similar process is used for all possible *S-D* pair for optimum and sub optimum (near-optimum or next possible optimum route) to balance the overall network load.

The second phase of IANRA algorithm is called route maintenance that is responsible for preservation of the paths during the communication. In IANRA algorithm no need to generate special packets for route maintenance, because while *FA* and *BA* construct pheromone tables for each node located on the path from source to destination, at the same time they are used for preservation of routes regularly. When *FA* arrives to a node between source and destination, the pheromone matrix is updated as in Biological System when, *S* relays a data packet to *D* via neighbours, it increases pheromone on the used path. It can be described as: suppose *FA* travels from node  $N_i$  to node  $N_d$  through neighbour nodes  $N_j$ , it increases the pheromone entry value  $(N_d, N_j, \rho)$ , by  $\Delta\rho$  according to equations (4) and (5). For each of next hops on  $\mathcal{R}_{OP}$ , the increased is  $\Delta\rho$  strength by arrived *FA*. The same procedure is followed by *BA* for backward paths.

The final stage of IANRA is to handle the link failure. This problem usually occurs due to movement of nodes in an optimized route (physical layer problem) and weak signals in coverage area. Since IANRA considers IEEE 802.11 both physical and the MAC layer, it's able to recognize a link failure through a missing acknowledgment message in the MAC layer. On the other hand in IANRA, each node maintains an updated view of its immediate neighbours at each instant. Hence it is easy to discover link failures as rapidly as possible, before they can lead to transmission errors and packet loss.

If a node receives *Link failure Notification* message for a settled path, it first removes this path from routing table by change pheromone value to zero (0) for failed link. Subsequently, the mentioned node starts to search selected new alternative path for data transmission. For discovering alternative path, a node updates its pheromone matrix. If there is another available route to destination which supports QoS, it will be used to send data packets instantly. Otherwise, the node sends *FA* to inform neighbours for updating its pheromone matrix and discovers a valid optimum route. In this process, if a *FA* can not travel a link to the next hop (due to link failure), it is backtracked to pervious hop to proceed further.

#### IV. IANRA PERFORMANCE EVALUATION

For performance evaluation of IANRA the OMNeT++ [17] simulation software is used, and then the obtained results are compared with AntNet and AODV. The considered performance evaluation parameters are average end-to-end delay per packet, packet delivery ratio and packet overhead. These 37 wireless nodes are placed randomly in an area of  $1400 \times 250 m^2$ . NED and C++ are applied for developing source code to make scenario of IANRA. Definitely, some headers and objects are created for these source codes. The arrival rates and the distribution of packets in each session follow a Poisson distribution. The data packet size is exponentially distributed. Size of  $BA$  is calculable by number of Visited Nodes ( $VNs$ ),  $FA$  and size of each  $FA$ . The  $BA$  calculation is as follows:

$$BA_S = NV + FA_S, NV = \delta + 1, FA_{Size} = \sum_{i=1}^m (FA_{Size})_{\delta_i} \quad (9)$$

To avoid increasing load in network, the number of generated ants in each node is two and the process of generation continues until the final generation. Data traffic flows contain 15 Constant Bit Rate (CBR) sources each transmitting 64 byte packet per second. Transmission time for each node follows stochastic process between 0 and 80 second after the start of the simulation. The bandwidth of each of all connective routes is 2.0 Mbits/s. IANRA designed buffer model for each of nodes that placed on network topology. Also the maximum speed in MANET scenario considered 20 m/s. In this section the considered buffers model at the node in IANRA algorithm are discussed. Each connected node has three kinds of buffers; (i) Input Buffer, (ii) Output buffer with high- and low-priority queues for every neighbour and (iii) Wireless Networks (WNs) buffer with one queue for each destination in the network. In order to make coverage of all aims for the proposed Algorithm, we implemented an extra buffer. A buffer called WNs buffer.

#### V. SIMULATION RESULTS

Experiment results show that the average end-to-end delay is decreased and the packet delivery ratio is increased in IANRA compared to two other algorithms AntNet and AODV Routing algorithm. Figure 3, shows that the average end-to-end packets delay is increased with increasing number of packets in the network. For example, when numbers of packets approach to 3300, the average end-to-end packet delays are around 0.17, 0.19 and 0.28 seconds for IANRA, B-AntNet and AODV respectively. This shows that for WMNs, the average end-to-end packet delay of IANRA algorithm is about 10.2 % and 7.21% lower than AODV and B-AntNet algorithms

respectively. Also, for Ad-hoc networks the average end-to-end packet delay is about 0.24 seconds which is lower than AODV and B-AntNet. Figure 4 shows the packet delivery ratio versus frequency of packet generation per second. It is seen that, with 20 packet/sec, IANRA gives higher packet delivery ratio about 92 % and 87% for WMNs and Ad-hoc networks respectively.

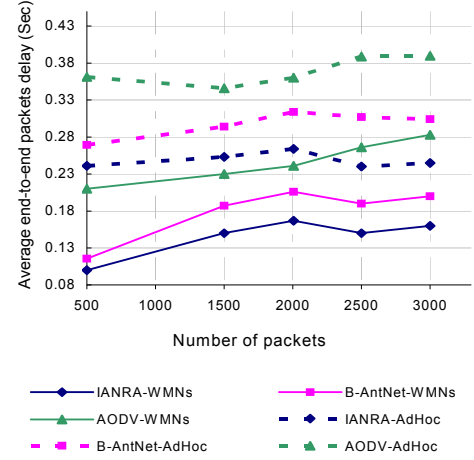


Figure 3. Average end-to-end delay with increasing number of data packets in the network.

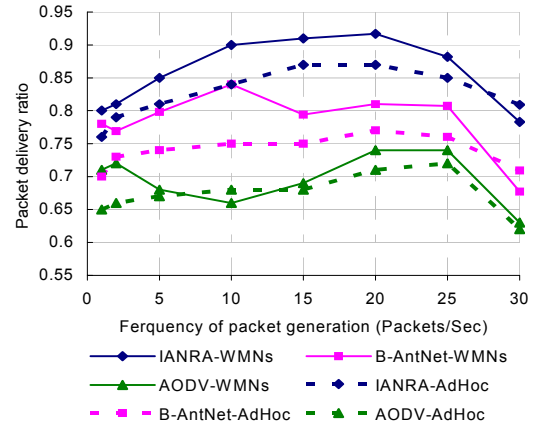


Figure 4. Packet delivery ratio with changing packet generation frequency.

In Figure 5 the average end-to-end packets delay versus number of network nodes with low data load shown. Low data load is considered as five packets per node per second approximately in the network.

Results show that, the average end-to-end packet delay is 10.75% and 5.65% lower than AODV and B-AntNet algorithms respectively. Also it reveals the related results of average end-to-end packet delay with high data load in the network. High data load is considered as 20 packets per node per second. The obtained result shows that, with 37 nodes, IANRA gives lower average end-to-end packet



delay about less than 0.12 seconds, where as it is for B-AntNet and AODV more than 0.13 and 0.14 seconds respectively. On the other hand, experiment result for packet overhead is depicted on Figure 6. Overhead results show that with increasing packet generation rate in the starting point, packet overhead is reduced. Also in this graph exponential behaviour of three algorithms are shown as well.

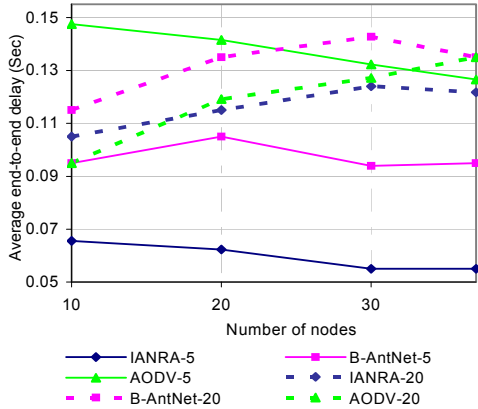


Figure 5. Average end-to-end packets delay with low (5 p/sec) and high (20 p/sec) data load on network.

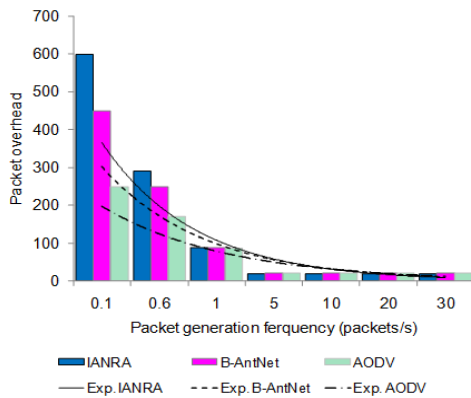


Figure 6. Packet overhead with changing packet generation frequency.

## VI. CONCLUSION

In this paper IANRA algorithm for load balancing with QoS has been introduced based on Ant Colonies Optimization. Three current problems in Ant colony Algorithms are (i) routing table freezes due to selecting same path, (ii) Statistic network and (iii) choosing a non-optimized route, have been eliminated by considering ant's and their successive generations, which shows the better load distribution among network nodes. Finally, IANRA is able to solve the routing, load-balancing and QoS problem in any wireless and mobile networks; including Ad-hoc, WSN etc.

## REFERENCES

- [1] Dorigo M, Stutzle T. Ant Colony Optimization. MIT Press: Cambridge, MA., (Year of publication: 2004 ISBN-13:978-0-262-04219-2).
- [2] L. Varzandeh et al., Load Balancing in QoS Based Computer Networking with Ant Net Algorithm, International Conference on Electrical Engineering ICEE2006 International Review on Computers and Software, pp. 94-98, 2006.
- [3] L. Yong, Z. Guangzhou, S. Fanjun, Adaptive ant based dynamic routing algorithm, proceeding of the 5th IEEE Conference on Intelligent Control and Automation, pp. 2694-2697, 2004.
- [4] Z. Subing, L. Zemin, A QOS routing algorithm based on Ant algorithm, IEEE International Conference on Communications, Vol. 5, No. 1, p.p.1587-1591, 2001.
- [5] Dorigo M., Di Caro G., Ant Net: A Mobile Agents Approach to Adaptive Routing Technical. Report, IRIDIA-Free Brussels University, Belgium, 1997.
- [6] K. M. Sim, W. H. Sun, Ant Colony Optimization for Routing and Load-Balancing: Survey and New Directions, IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, Vol. 33, No. 5, pp. 560-572, 2003.
- [7] Dorigo M, Di Caro G, Gambardella LM. Ant algorithms for discrete optimization. Artificial Life 1999; Vol. 5, No.2, pp.137–172. 1999.
- [8] Goss S, Aron S, Deneubourg JL, Pasteels JM. Self-organized shortcuts in the Argentine ant. Naturwissenschaften, Vol. 76, pp. 579–581, 1989.
- [9] G. Di Caro, F. Ducatelle, L. M. Gambardella, AntHocNet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks, EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS, Euro. Trans. Telecomms. AEIT2005, Vol. 16, pp. 443–455, 2005
- [10] Pacut, M. Gadomska, A. Igielski, Ant-Routing vs. Q-Routing in Telecommunication Networks, Proceedings of the 20-th ECMS Conference, 2006.
- [11] Dorigo, M.; Birattari, M.; Blum, C.; Gambardella, L.M.; Mondada, F.; Stützle, Th. Ant Colony Optimization and Swarm Intelligence. 4th International Workshop, ANTS 2004, Brussels, Belgium, September 5-8, 2004, Lecture Notes in Computer Science, Vol. 3172.
- [12] K. M., W. H. Sun, Multiple ant-colony optimizations for network routing, proceedings of the 1st Int. Symp. Cyber world, Tokyo, Japan, pp. 277–281, 2002.
- [13] R. Schoonderwoerd, O. Holland, J. Bruten, L. Rothkrantz, Ants for Load Balancing in Telecommunication Networks, Hewlett Packard Lab., Bristol, U.K., Technical Report HPL-96-35, 1996.
- [14] B. Baran, R. Sosa, A new approach for Ant Net routing, 9th International Conferences on Computer Communications Networks, Las Vegas, NV, 2000.
- [15] D. Caro, M. Dorigo, Ant Net: Distributed stigmergetic control for communications networks, Journal Artificial Intelligence Resource, vol. 9, pp. 317–365, 1998.
- [16] K. Oida, M. Sekido. An agent-based routing system for QOS quarantines. Proceedings of the IEEE International Conference on Systems, Vol. 3, 12-15, pp. 833-838, 1999.
- [17] OMNeT++ Discrete Event Simulation System, Version 3.3, 2006. <http://www.omnetpp.org>.

# An Artificial Ant Based Novel and Efficient Approach of Regular Geometric Shape Detection from Digital Image

Ayan Acharya, Kaushik Chattopadhyay, Deepyaman Maiti, and Amit Konar

Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata:  
700032, India

**Abstract-** The paper presents a novel and efficient method of regular geometric shape detection from gray scale images. Artificial ant based methods have not been used much in the field of image processing. This paper demonstrates how artificial ants can be used effectively to extract regular geometric shapes from images. We propose here Ant Regeneration and Recombination System (ARRS), an entirely new approach developed by ourselves. Our scheme of detection of shapes comprises of three steps. Firstly, MATLAB edge detection operator converts a gray scale image into a binary one. Ant Regeneration and Recombination System algorithm is then applied on this binary image to detect closed loops. Finally, these closed loops are tested for different geometric shapes like circle, ellipse, rectangle and square. The most important aspect of the scheme is it can detect both intersecting as well as non intersecting regular shapes from images consisting of different open and closed loop configurations. It is the incredible time and memory efficiency of the scheme that makes it useful in real time applications where decisions have to be taken within a very small time interval by observing an image.

**Index terms**— Edge Detection, Loop detection, Shape detection, Ant Regeneration and Recombination System.

## I. INTRODUCTION

Detection of regular geometric features like circles, ellipses, rectangles and polygons in digital images is an important exercise in image analysis and computer vision applications such as automatic inspection and assembly. Various methods for regular geometric shape detection have been researched till date. The Hough Transform (HT) technique, one of the most popular one, has been used extensively to extract analytic features, such as straight lines, circles, and ellipses ([2],[3],[5],[6]) because of its robustness against noise, clutter, object defect, shape distortion etc. A serious drawback of this technique is its huge time and space complexity. In HT technique, edge detected image is first mapped into a parameter space. For detection of circles, each edge point in the edge image is mapped to a region in the parameter space or accumulator as one count (or vote) for each point in the appropriate region. The parameter space for a circle is a 3-dimensional space where the center coordinates and the radii of

probable circle in the image are specified. Detection of ellipse requires a 4-dimensional or a 5-dimensional parameter space, depending on whether the ellipse has been rotated about its center. The time complexity grows exponentially with the number of parameters used to characterize the shape. For circles this gives rise to a complexity of  $O(n^3)$ . For ellipse, the complexity, in general, is of  $O(n^5)$ . Consequently, in many simple real time applications this method often turns out to be far from practical.

In this paper, we propose a completely new approach to detect regular shapes efficiently from images using Ant Regeneration and Recombination System (ARRS), an entirely novel algorithm. The ant colony system **ARRS** developed by ourselves is inspired by the ant system algorithm ([11],[12],[13],[14]) and its numerous variants but conceptually **ARRS** is widely different from these existing algorithms. Moreover, as far as our knowledge goes, we are not aware of any work related to the extraction of both intersecting and non-intersecting shapes from images efficiently, accurately and quickly. These points specify the main contribution of our paper.

The paper has been subdivided into eight sections. We introduce our scheme Ant Regeneration and Recombination System in section II. Sections III and IV deal with the implementation and pseudo code of the scheme. A modified approach of closed loop detection based on the graph explored by the scheme is presented in section V. Algorithms for detection of various geometric shapes are provided in section VI. Simulation results are provided in section VII. The paper ends at section VIII with conclusion and scope of future work.

## II. ANT REGENERATION AND RECOMBINATION SYSTEM (ARRS)

In an edge detected image, a string of edge pixels stand for the boundary of an object present in the image. When we are talking about intersecting shapes, it is clear that such boundaries will cross each other. We can say, a **graph** (in not too technical a sense) is formed. A string of edge pixels, where each pixel is connected to only two adjacent pixels, is called a **branch**. Where more than two branches meet, we call that pixel a **node**. A branch is completely described by the pixel co-ordinates of that branch, taken in order. A node is completely specified if we have its co-ordinates along with the information that,

to which other nodes it is connected and by which branches. This connectivity information is gathered in a matrix form which we call the **incidence matrix**.

Ants are basically one-pixel agents that explore the graph. Each ant has its own co-ordinates which indicate on which pixel it is, and those co-ordinates are continuously updated as the ant moves. Before making a move an ant finds the number of directions available to move, by counting the number of edge pixels around the pixel it is resting upon, excluding the pixel it has just come from (this pixel is not to be covered again, i.e. the ant cannot move back).

If an ant moves into a pixel on a branch, then it updates its own co-ordinates and stores the co-ordinates of the previous pixel in an array of its own called the **path[ ]** array. If an ant discovers a node, the co-ordinates of the node are recorded. The ant also realizes that the node it has come from and the newly discovered node are connected to each other by the branch it has traversed through. At this point, this ant becomes inactive (i.e. it stops travelling, its co-ordinates are never updated again) but reproduces new ants and place them on the starting pixels of each of the new branches originating from that particular node. The '**child**' ant now continues further exploration of the **graph**. We call this process as **regeneration** of ants. It can easily be perceived that this process allows complete exploration of the **graph** no matter how complex it is.

In fact, the process of searching starts with placing of an ant (we call it the 'mother' ant) at a randomly selected edge pixel. The mother ant moves up to a node (this is the first discovered node) and generates new ants which travel along branches originating from this node and create newer ants upon discovering new nodes. In this way a whole system of ants builds up. Thus far we observe three properties of the ants:

**Property I:** All ants move at the same speed, i.e. one pixel at one step.

**Property II:** Ants are generated at nodes.

**Property III:** Ant remembers the node it has originated from

The above listed properties of ants have an interesting consequence. Suppose, between two nodes A and B, there is a shorter path S and a longer path L. The ants are generated at A and move towards B. The one along S will reach B sooner and place a new ant on the pixel that connects L to B, oblivious of the fact that there is already one coming towards B along L. The new ant will move towards A along L. Thus the two ants on L move toward each other and after some time their co-ordinates will become the same. We then say that the two ants **recombine**. Whenever two ants recombine, we take advantage of that as follows: (i) Each of them has partially discovered the same branch. They recorded pixel co-ordinates of the branch from two sides in their **path[ ]** arrays. If we join the **path[ ]** arrays of these two ants back

to back, we get an array which has all the pixel co-ordinates of the branch, in order. (ii) Since each ant remembers from which node it had originated, when they meet, the connection between those two nodes is discovered. We use this information to build up the **incidence matrix** (how this is done is described in the next section).

After meeting each other, both the ants are made inactive. This keeps one ant from going through the same pixels that has already been visited by some other ant, which guarantees that each edge pixel is visited only once and by only one ant.

The ant can stop moving even when it finds a dead end. Thus we get another property:

**Property IV:** An ant must end up either reaching a dead end or a node or meeting another ant, and thereby becoming inactive.

### III. IMPLEMENTATION OF ARRS

We here explain the data structures modeled for our algorithm and provide the pseudo code to make the entire aspect of the algorithm comprehensible.

**Ants:** Each ant is represented as a structure having the following fields.

- i) **(x, y)** : its current co-ordinates ,
- ii) **(x<sub>n</sub>, y<sub>n</sub>)**: co-ordinates of the node it came from
- iii) **path[ ]**: this array stores the pixel co-ordinates of the branch it is moving along

Several ants are represented as blocks of a linked list. Each block corresponds to a single ant. As new ants are generated, new blocks are added to the list. When an ant becomes inactive, the block corresponding to that particular ant is deleted from the list.

**Branches:** Each branch is represented as a structure which has the following fields:

- i) **index**: if index is **n**, we call that branch the **n**th branch or branch **n**
- ii) **branch\_arr[ ]**: this array contains the co-ordinates of the pixels on the branch

Each branch is a block of a linked list. When a new branch is discovered, a block is added to the list. When an ant travels from one node to the other, it discovers a branch completely. Its **path[ ]** array becomes the **branch\_arr[ ]** of the newly discovered branch. When two ants move toward each other from the terminal nodes of a branch and eventually meet, each has a part of the branch stored in their **path[ ]** arrays. As explained earlier, we join these two **path[ ]** arrays back to back (to keep the order of the pixels right) and this becomes the **branch\_arr[ ]** of this newly discovered branch. There is one variable called **B**. This is initially set to zero, is incremented each time a branch is discovered and becomes the **index** of the newly discovered branch. After the ants finish their exploration, all the branches would have been discovered and the

variable  $B$  will be the total number of branches. Also  $B$  will be the highest branch index.

**Nodes:** Each node is also represented as a structure with the following fields:

i) **index:** if the index is  $n$ , we call it the  $n$ th node or node  $n$ , ii)  $(x, y)$ : co-ordinates of the node, iii) **connectivity** [ ].

The array connectivity [ ] has some special importance. If an ant coming from the  $i$ th node reaches the  $j$ th node or meets another ant coming from  $j$ th node, through the  $k$ th branch (i.e. branch with index  $k$ ), then we do the following operations: node  $p$ .connectivity[ $k$ ] = 1 for  $p=i, j$  : node  $p$ .connectivity[ $k$ ] = 0 otherwise.

Since branch indices are array positions of the connectivity[ ] arrays, after exploration is complete, these arrays will have length that corresponds to the maximum branch index, i.e.  $B$ . Let the total number of nodes be  $N$ . Now we stack the connectivity[ ] arrays of the nodes one below the other, starting with node 1 at the top, node 2 below that, then node 3 and so on. Then it is clear that the  $N \times B$  matrix so formed will be the **incidence matrix** of the **graph**. If somehow  $i=j$ , i.e. both the ants come from the same node then existence of a branch is detected which is itself a closed loop and consists the node the ants came from. In that case we join the path[ ] arrays of the ants to get the pixel co-ordinates of this 'loop-branch' and then test it. We do not include such branches which are connected to one node only in the incidence matrix for ease of computation.

#### IV. PSEUDO CODE FOR ARRS

First an ant, i.e. the mother ant is placed at any randomly selected edge pixel. As it moves along the edge pixels, if it comes back to the starting pixel without encountering any node, then it has discovered a closed loop. The path[ ] array of this ant stores the co-ordinates of the pixels forming this loop. This array is tested for specified shapes. But, if the mother ant reaches a node, it places new ants at the starting pixels that originate from this node, the node is recorded as the first node discovered, and this ant itself becomes inactive. The pseudo code describing the movement of the ants from this point onwards is given.

Let the total number of active ants be  $M$  and the number of branches discovered be  $B$ , at any time. Searching ends when  $M$  becomes zero, i.e. no active ants are left. The algorithm described above is presented below in a more compact form.

```

while $M > 0$
 $i = 1$;
 while $i \leq M$
 find d , the number of edge pixels around the pixel
 occupied by the i th ant ;
 if $d > 1$
 add a block to the list of nodes and record this
 pixel as a node;

```

```

 $B = B + 1$;
 add a block to the list of branches and make
 the path[] array of this ant the branch_arr[] of
 the new branch; the index of this branch is B ;
 if the ant was coming from node p and has
 reached node q
 then node p .connectivity[B] = 1;
 node q .connectivity[B] = 1;
 end
 delete the block of this i th ant from the list of
 ants; place $d-1$ new ants at the starting pixels of
 the $d-1$ new branches originating from this
 node by adding $d-1$ new blocks to the list of
 ants;
 $M = M - d + 2$;
 else if $d = 1$
 if the only pixel available to move into is
 occupied by some other ant let one ant be
 from node p and the other from node q
 then delete the block of the two ants from the
 list of ants;
 $M = M - 2$;
 if $p \neq q$
 then $B = B + 1$;
 add a new block to the list of branches;
 join the path[] arrays of the ants back to
 back and make this the branch_arr[] of
 the new branch ;
 the index of the branch is B ;
 node p .connectivity[B] = 1;
 node q .connectivity[B] = 1;
 connectivity[B] = 0 for all other nodes;
 else
 join the path[] arrays of the two ants
 and test for circle;
 end
 else update the co-ordinates of the i th ant
 according to the adjacent pixel it moves into;
 record the co-ordinates of this pixel in the
 path[] array of the i th ant;
 end
 end
 $i = i + 1$;
end
end

```

#### V. MODIFIED ALGORITHM FOR FINDING CLOSED LOOPS

Suppose a set of  $k$  branches are selected from the set of all  $B$  number of branches. Two points are to be noticed here:

I. Each branch is associated with the two nodes it connects. Let, for branch  $b_i$  the terminal nodes are  $n_{i_1}$  and  $n_{i_2}$ . If  $b_i$  is a part of a closed loop then the loop must pass

through  $n_{i_1}$  and  $n_{i_2}$ . Then it is clear that if the selected branches form a closed loop, the loop must pass through all the nodes that belong to the set of nodes  $N_k = \{(n_{i_1}, n_{i_2}) \cap (n_{j_1}, n_{j_2}) \cap \dots \cap (n_{k_1}, n_{k_2})\}$  i.e. the set of all nodes that these selected branches are associated to. On a closed loop, the number of branches forming the loop and the number of nodes lying on the loop must be equal, i.e. the cardinality of  $N_k$  must be equal to  $k$ .

**II.** Each node in the set  $N_k$  must be connected to exactly two of the branches in  $B_k$ .

The above mentioned conditions I & II together are necessary and sufficient for any selected set of branches to form a closed loop. However they do not suggest the order in which they form a loop.

Let the columns of the incidence matrix that correspond to the selected branches be augmented to form a new  $M \times k$  matrix, let us call it  $I_k$ .  $I_k$  has only 1 and 0 entries since it is formed from the incidence matrix. Since a branch can be connected between only two nodes, there are exactly two 1's in a column. Each row of  $I_k$  corresponds to a node and the number of 1's in a row gives the number of branches connected to the node. If branches in  $B_k$  form a loop then in  $I_k$  the rows corresponding to nodes belonging to  $N_k$  must have two non-zero entries each (condition II) and all other rows (i.e. rows not belonging to set  $N_k$ ) should have 0 entries, i.e. if a row has a 1 then it should have exactly two such 1's otherwise all its entries are 0's.

We start at a cell  $(r, c)$  of  $I_k$  which contains a 1 and where  $r$  is the row number and  $c$  is the column number. Now an entry in row  $r$  is 1, therefore there should be exactly another 1 in  $r$  (otherwise the branches in set  $B_k$  would not form a closed loop and then we reject  $B_k$  and look for another such set). We update the row number  $r$  and go to the other non-zero cell in  $c$ . This is equivalent to moving from one end of the branch to its other end. This new row has also a 1 (this cell), so we check for the other 1 (if that is not present then we reject  $B_k$ ) and update the column number  $c$  accordingly to move into that cell. Modifying  $c$  implies that we are on the other branch that is connected to the same node and we shall go to the other end of this branch (modifying  $r$ ) and then to the other branch connected to that end (modifying  $c$ ) and so on, each time checking for condition II for each node. A variable  $p$ , initially set to zero is maintained. Each time we modify  $r$ , we cover one branch (we move from one node to the other) and increment  $p$ . When  $p$  becomes  $k$ , i.e. all branches are covered, we check whether we have come back to the starting node (the row in we had started with); if yes, then condition I is satisfied (condition II is already satisfied because we came up to this point checking condition II for each node in  $N_k$ ) therefore a closed loop is formed, otherwise not. If we come back to the starting

node then we check whether  $p$  has become equal to  $k$ . If it happens, then branches in  $B_k$  form a closed loop, else not. We also record the column numbers in an array in order each time we modify  $c$ . Then it is obvious that this array will contain the branches of  $B_k$  in the order in which the form loop. So, this algorithm accomplishes two tasks at the same time, checks whether a set of branches forms a loop or not and finds their order if they do form a loop. Since non-zero entries mean only 1, if there are two non-zero entries in a row the sum of all entries in the row must be two.

Now, in a special case, where we are required to detect only circles from images, we can simplify calculations further by reducing some complexity. Now, only a single circle can be drawn using a branch if at all. This is because if two or more circles have a common branch then they have to intersect at more than two points (a branch by definition consists of three or more points) which is practically impossible. So, if a set of branches indeed form a circle, we can delete all these branches from the graph without affecting the possibility of finding other circles from the graph. Therefore, once it is found that some  $k$  number of branches form a closed loop which is a circle, we delete the blocks corresponding to those branches from the list of branches and reduce number of branches  $B$  by  $k$ . This reduces the time complexity by leaps and makes the overall process faster.

## VI. ALGORITHM FOR DETECTION OF VARIOUS SHAPES

**Algorithm for circle detection:** The algorithm used for circle detection from closed loop follows from the geometric property that a unique circle can be drawn through three non collinear points. The steps are described here: **1.** Choose three equidistant points on a closed loop randomly and find equation of the unique circle that passes through these three points. **2.** Compare distances between every other point in the loop and the centre of the circle with the radius of the circle. **3.** If deviation error  $\delta > \epsilon$  (threshold value, which is a function of the radius of the circle), the loop is not considered as a circle. **4.** Even if condition 3 is false, if more than one fourth of the points on the loop result in deviations greater than half of the threshold value, then discard the loop from being a circle. **5.** A close loop that has passed through all the above four conditions is considered to be a circle. The threshold value  $\epsilon$  is taken as a function of the radius of the probable circle as for a larger circle we can afford greater deviation. We have chosen  $\delta$  proportional to the radius. The proportionality constant, called Deviation Constant  $\lambda$ , is determined from experiments in such a way that more or less all visible circles fall within this range. It should be noted that due to discrete points in the graph each detected circle will produce some deviation even if it perfectly looks like a circle.



**Algorithm for ellipse detection:** Detection of ellipse is a multistage process which is described as follows: 1. Find  $x_{max}$ ,  $x_{min}$ ,  $y_{max}$ ,  $y_{min}$  from the entire array containing the pixels that form the closed loop. Now the centre of prospective ellipse would be given by  $x_c = (x_{max} + x_{min})/2$  and  $y_c = (y_{max} + y_{min})/2$ . 2. Find the point  $(x_1, y_1)$  farthest from the centre from one half of the loop and the other such point  $(x_2, y_2)$  from the other half. If  $|x_c - (x_1 + x_2)/2| < \epsilon_x$  and  $|y_c - (y_1 + y_2)/2| < \epsilon_y$  then the line joining these two points is the major axis otherwise the loop cannot be an ellipse and no further steps are needed. Let,  $a_1 = \sqrt{(x_c - x_1)^2 + (y_c - y_1)^2}$ ,  $a_2 = \sqrt{(x_c - x_2)^2 + (y_c - y_2)^2}$ . The semi major axis of the ellipse is taken as  $a = (a_1 + a_2)/2$ . And the inclination of the major axis is given by  $\alpha = \tan^{-1}((y_2 - y_1)/(x_2 - x_1))$ . 3. From one half of the loop find the point  $(x_3, y_3)$  which when joined to  $(x_c, y_c)$  makes an angle with the major axis closer to  $90^\circ$  and within as compared to other points. Find the other such point  $(x_4, y_4)$  from the other half of the loop. Now, if  $|x_c - (x_3 + x_4)/2| < \delta_x$  and also  $|y_c - (y_3 + y_4)/2| < \delta_y$ , then the line joining these two points forms the minor axis of the supposed ellipse else we discard the loop and make no further tests.

Let,  $b_1 = \sqrt{(x_c - x_3)^2 + (y_c - y_3)^2}$ ,  $b_2 = \sqrt{(x_c - x_4)^2 + (y_c - y_4)^2}$ . The semi minor axis of the ellipse is taken as  $b = (b_1 + b_2)/2$ . The eccentricity of the ellipse is  $e = \sqrt{1 - (b/a)^2}$ . The co-ordinates of the foci can be found to be  $(x_c + a \cos \alpha, y_c + a \sin \alpha)$  and  $(x_c - a \cos \alpha, y_c - a \sin \alpha)$ . For an ideal ellipse the sum of focal distances of each point on the ellipse is  $2a$ . Here we test the following condition for each pixel on the closed loop:  $|2a - \text{sum of focal distance}| < \Delta l$  where  $\Delta l$  had to be considered for the error due to digitization. If all the pixels of the loop satisfy the above condition then the loop is an ellipse with parameters  $x_c$ ,  $y_c$ ,  $\alpha$ ,  $a$  and  $b$ . The procedure is advantageous because it makes use of various parameters like  $\epsilon_x$ ,  $\epsilon_y$ ,  $\delta_x$ ,  $\delta_y$ ,  $\delta$ ,  $\Delta l$  which can be set to desired values as per requirement to suit different levels of accuracy and to allow different kinds of distortions in the shape of an ellipse in the image. This method also uses intermediate check-points where a loop can be discarded without any need for further computation.

**Algorithm for detection of rectangles and squares:** This algorithm can be presented as follows: 1. While moving from one pixel to the next note the gradient difference  $\frac{\Delta y}{\Delta x}$ .

Define  $\theta = \tan^{-1}[\frac{\Delta y}{\Delta x}]$ . If  $\theta$  is nearly equal to  $90^\circ$  degrees, mark any of the pixels as a possible corner of a rectangle or square. If not, we consider the two pixels as part of the same straight line. 2. Four such abrupt gradient differences indicate that the closed loop is either a rectangle or a square. A measurement of sides distinguishes further between a rectangle and a square.

## VII. SIMULATION RESULTS

We test our algorithm on four different sets of images. A comparative study of the system performance under gradually increasing complexity is provided in tabular form for the first set of images. Complexity increases from fig 1(a) to 1(d).

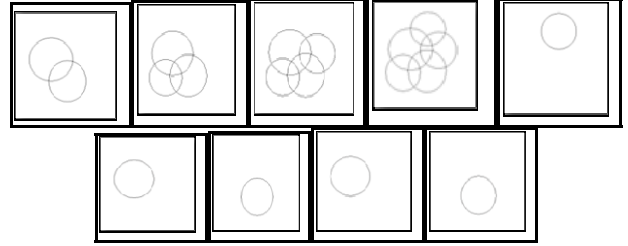


Fig : 1(a), 1(b), 1(c), 1(d) and 1(d.I) to 1(d.V)

Table 1: Data for Figures 1(a) to 1(d)

|                   | Fig 1(a)                                        | Fig 1(b)                       | Fig 1(c)                            | Fig 1(d)                            |
|-------------------|-------------------------------------------------|--------------------------------|-------------------------------------|-------------------------------------|
| edge pixels       | 684                                             | 968                            | 1312                                | 1611                                |
| (N,B)             | (2,4)                                           | (6,12)                         | (11,21)                             | (14,28)                             |
| Search Time(s)    | 0.2650                                          | 0.3750                         | 0.4840                              | 0.5630                              |
| Detection Time(s) | 0.0630                                          | 0.1250                         | 1.6750                              | 9.1870                              |
| Centers and radii | {{(107,132.5), 63.25};<br>{(156,194.5), 58.25}; | Column 1 +<br>{{(85,200), 51}} | Column 2<br>+{{(191,135.5), 54.25}} | Column 3<br>+{{(159, 73.5), 50.75}} |



Fig: 2(a) to (g)

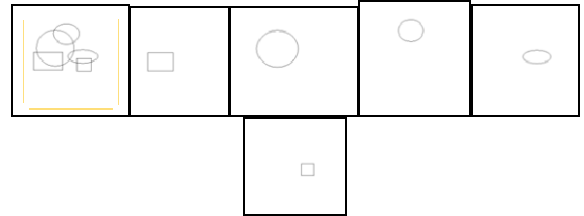


Fig: 3(a) to (f)

of parameters like computing time, memory usage, execution complexity and accuracy of system output in presence of noise.

#### REFERENCES

- [1] J. Yao, N. Kharm and P. Grogono, "A multipopulation genetic algorithm for robust and fast ellipse detection", Electrical and Computer Engineering Dept. and Computer Science Dept., Concordia University.
- [2] P.V.C Hough "Methods and means for recognizing complex patterns" US patent 3069654,1962.
- [3] R. O. Duda and P. E Hart, "Use of Hough transformation to detect lines and curves in pictures", Comm. Assoc. Comput, 15(1), pp. 11–15, 1972.
- [4] C.Kimme, D.H ballard and J.Klansky "Finding circles by an array of accumulators" Comm. ACM vol. 18 pp. 120-122, 1975.
- [5] R.K.K. Yip, P.K.S. Tam and D.N.K. Leung, "Modification of the Hough transform for circles an ellipses detection using a 2-dimensional Array", Pattern Recognition, 25, pp. 1007-1022, 1992.
- [6] A.S. Aguado, M.E. Montiel and M.S. Nixon, "Ellipse Detection via gradient direction in the Hough transform", Proceedings of IEE International Conference on Image Processing and its Applications, pp 375-378, July,1995
- [7] L. Costa and R. Cesar Junior, "Shape Analysis and Classification: Theory and Practice", CRC Press, 2000.
- [8] L.A.F. Fernandes, M. Oliveira, "Real-time line detection through an improved Hough transform voting scheme", Pattern Recognition, Vol. 41, pp. 299-314, 2008.
- [9] M. Blanco, M. G. Penedo, N. Barreira, M. Penas and M. J. Carreira, "Localization and Extraction of the Optic Disc Using the Fuzzy Circular Hough Transform", ICAISC 2006, pp. 712-721.
- [10] L. Iverson, S. Zucker, "Logical/Linear Operators for Image Curves", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.17, pp. 982-996, October 1995.
- [11] M. Dorigo, G. Di Caro and L. M. Gambardella, "Ant algorithms for discrete optimization", Artificial Life, vol. 5, no. 2, pp. 137-172, 1999.
- [12] M. Dorigo and C. Blum. 2005 "Ant colony optimization theory: A survey" Theoretical Computer Science 344, pp. 243 – 278.
- [13] M. Dorigo, V. Maniezzo and A. Colomi. "The Ant System: Optimization by a colony of cooperating agents" IEEE Transactions on SMC–Part B, Vol.26, No.1, pp.1-13,1996
- [14] C. Blum and M. Dorigo. "Search bias in ant colony: On the role of competition balanced systems" IEEE Transactions on Evolutionary Computation, vol. 9, no.2, pp. 159-174, 2005.

**Table 2: Data for Images 2(a) and 3(a)**

|                                     | Figure 2(a)                     | Figure 3(a)        |
|-------------------------------------|---------------------------------|--------------------|
| size                                | 270*366                         | 300*300            |
| edges                               | 1389                            | 1258               |
| (N,B)                               | (15,27)                         | (8,16)             |
| Exploration time(s)                 | 0.3280                          | 0.2050             |
| Detection time(s)                   | 3.4220                          | 2.143              |
| Results                             | (70.5,126):42.8                 | (100,110): 60      |
| (centre_x,centre_y):r               | (155,221):38.5                  | (135,61): 41       |
| (centre_x,centre_y):a,b,e, $\alpha$ | (272,126):40                    | (187,133):48,22,0. |
| Vertex_x:vertex_y                   | (203,128)::56, 26,0.88,0.02     | 88,-0.07           |
|                                     | (129,125)::53, 28,0.85,0.03     | {(216,140)(216,181 |
|                                     | (161.5,179.5)::54.5,24.5,0.9,88 | )(171,181)         |
|                                     | .28                             | (171,140)}         |
|                                     |                                 | {(122,120)(122,177 |
|                                     |                                 | )(36,177)(36,120)} |

Time efficiency of our scheme is obvious from above results. But the scheme is memory efficient too. Memory is used to store (i) the two dimensional matrix corresponding to the intensity information of the image pixels (ii) the co-ordinates of the pixel corresponding to each discovered branch (iii) the incidence matrix storing the connectivity information..

#### VIII. CONCLUSIONS AND SCOPE OF FUTURE WORK

The algorithm and ant system **ARRS** are entirely novel concept and have been proved to be efficient and applicable in real life situation. But the time complexity and memory requirement of our scheme have not been compared with other existing methods such as Hough Transform or GA based method ([1]). We are trying to compare our algorithm with those mentioned on the basis

# Codebook Design Method for Speaker Identification based on Genetic Algorithm

Md. Rabiul Islam, Md. Fayzur Rahman\*, Md. Ruhul Amin\* and Md. Shahid Uz Zaman

Department of Computer Science & Engineering

\*Department of Electrical & Electronic Engineering

Rajshahi University of Engineering & Technology (RUET), Rajshahi-6204, Bangladesh

E-mail: rabiul\_cse@yahoo.com, mfrahman3@yahoo.com, aminr\_bd@yahoo.com and szaman22@yahoo.com

**Abstract** — In this paper, a novel method of designing a codebook for speaker identification purpose utilizing Genetic Algorithm has been proposed. Speech features have been extracted using standard speech parameterization method such as LPC, LPCC, RCC, MFCC,  $\Delta$ MFCC and  $\Delta\Delta$ MFCC. For each of these techniques, the performance of the proposed system has been compared. In this codebook design method, Genetic Algorithm has the capability of getting global optimal result and hence improves the quality of the codebook. Experimental results show the superiority (i.e. 92% accuracy has been achieved) of this proposed codebook design method.

**Index Terms** — Automatic speaker identification, codebook design, genetic algorithm, speech pre-processing, speech parameterization.

## I. INTRODUCTION

Person recognition can be based on both the separate or combined use of several biometric features (voice, face, fingerprints, iris, etc...). A usual approach, to cope with the speaker recognition problem, involves in classifying acoustic parameters derived from the input speech signal by short time spectral analysis. These parameters contain both phonetic information, related to the uttered text, and individual information, related to the speaker [1]. According to the application area, speaker recognition systems can be divided into speaker identification systems and speaker verification systems [2].

Speaker Identification is the task of finding the identity of an unknown speaker among a stored database of speakers. To implement the speaker identification system in real time environment, codebook design is essential. The LBG algorithm is most popular to design the codebook due to its simplicity. But the limitations of the LBG algorithm are the local optimal problem and its low speed. It is slow because for each iteration, determination of each cluster requires that each input vector be compared with all the codewords in the codebook. This paper deals the efficient approach for implementing the codebook design method for real time close-set text-dependent speaker identification system.

## II. SYSTEM OVERVIEW

The proposed codebook designed method can be divided into two operations. One is the encoder and another is the decoder. The encoder takes the input speech utterance and outputs the index of the codeword with considering the minimum distortion. To find out the minimum distortion, different types of genetic algorithm operations have been used. In decoding phase, when the decoder receives the index then it translates the index to its associate speaker utterance. Fig. 1 show the block diagram of this proposed codebook designed method.

## III. SPEECH SIGNAL PREPROCESSING

### A. Speech start-end points detection and silence parts removal

To capture the speech signal (Fig. 2) high quality microphone has been used according to the parameter set (Sampling frequency 8000 Hz, resolution 16-bits/sample and recorded file format = \*.wav). After taking the speech signal, start and end point detection algorithm [3] has been used. Speech endpoints detection is used to detect the presence of speech and to remove the silences in background noises [4]. Fig. 3 shows the signal status after start and end point detection algorithm. Short-time energy is implemented to remove the silence parts by using the following equation.

$$E_i = \sum_{t=n_i}^{n_i+N-1} S^2(t) \quad (1)$$

Where  $s(t)$  is the data sample,  $i$  is the frame number,  $E_i$  is the frame energy,  $N$  corresponds to the frame length and  $n_i$  is the index of first data sample in the frame. This results show in Fig. 4.



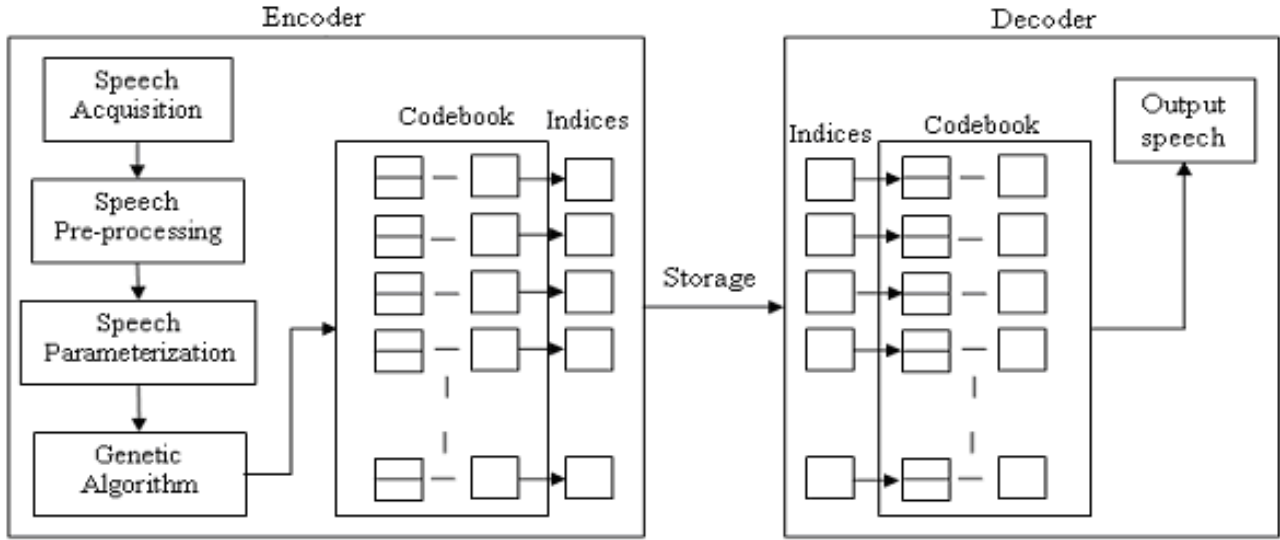


Fig. 1. Paradigm of the proposed codebook (both encoding and decoding) design method for automatic speaker identification system.

### B. Pre-emphasizing Filtering

The Pre-emphasizing filter has been used here as a noise reduction system to increase the amplitude of the input signal at frequencies where signal-to-noise ratio (SNR) is low. Pre-emphasis has been used to balance the spectrum of voiced sounds that have a steep roll-off in the high frequency region [5]. Fig. 5 displays the output after applying the pre-emphasizing filtering technique by using the following equation.

$$H(z) = (1 - az^{-1}) \quad (2)$$

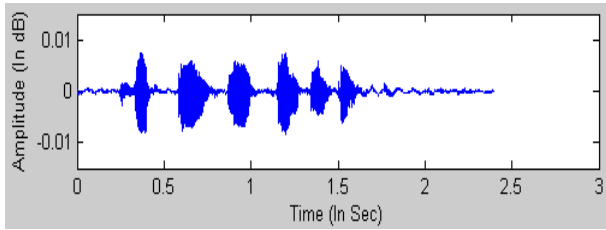


Fig. 2. Sample form of recorded speech signal.

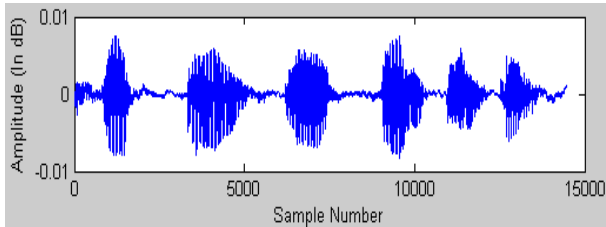


Fig. 3. Speech after start and end points detection algorithm.

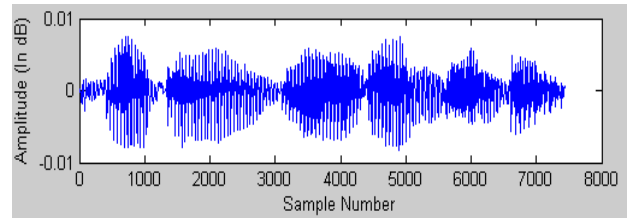


Fig. 4. Removing the silence parts from the speech.

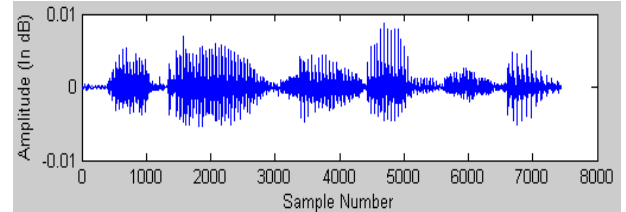


Fig. 5. Speech after Pre-emphasis filtering.

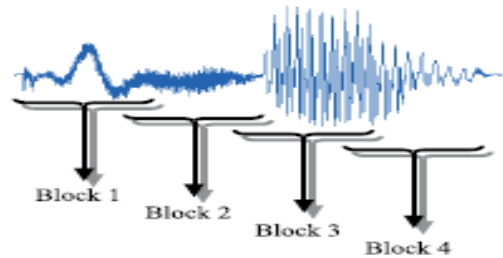


Fig. 6. Processes of segmented speech.

### C. Speech Segmentation Or Frame Blocking

After pre-emphasizing, speech signal has been blocked into frames of  $N$  samples, with adjacent frames being separated by  $M$  ( $M < N$ ). The purpose of the overlapping analysis is that each speech sound of the input sequence would be approximately centered at some frame [6], [7]. 10-30 milliseconds frame length and around 25% to 75% frame overlap has been considered for frame blocking. Fig. 6 shows the process of frame blocking.

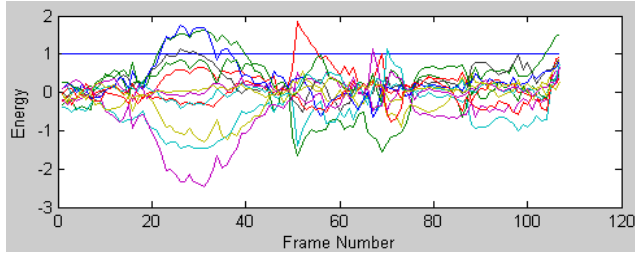
### D. Windowing

Among various types of windowing techniques (i.e. hamming window, blackman window, hanning window etc.) hamming window has been used. The purpose of using windowing is to reduce the effect of the spectral artifacts that results from the framing process [8]. The hamming window has been implemented by the equation:

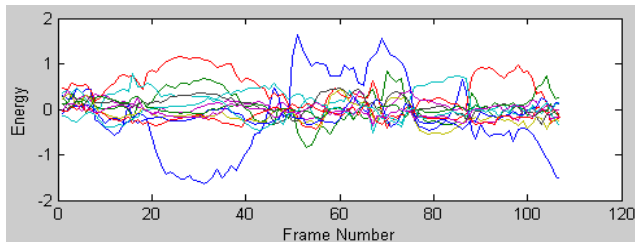
$$w[k+1] = 0.54 - 0.46 \cos(2\pi \frac{k}{n-1}), k=0, 1, \dots, n-1 \quad (3)$$

## IV. SPEECH PARAMETERIZATION

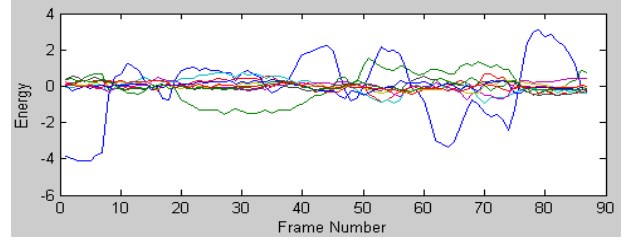
To extract the features from the speech some effective and standard speech parameterization methods such as RCC, MFCC, MFCC, MFCC, LPC, LPCC have been applied [9], [10] because the quality of the speaker modeling and pattern matching strongly depends on the quality of the feature extraction methods. Fig. 7 shows the features after applying different types of feature extraction methods.



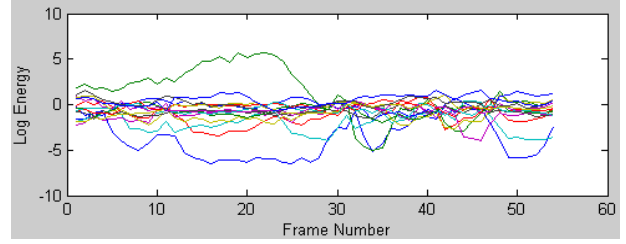
(a) LPC features of 12th order.



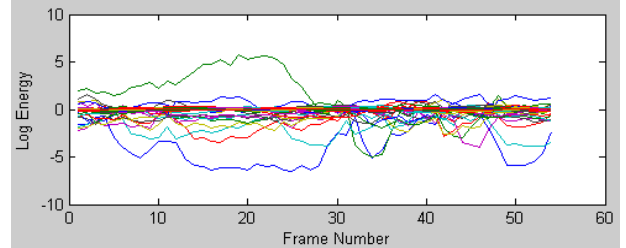
(b) LPCC features of 12th order.



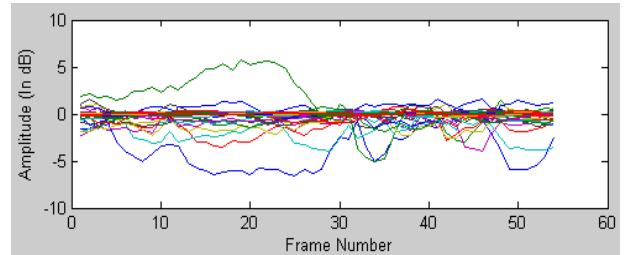
(c) RCC features with 15 coefficients.



(d) MFCC features with 15 coefficients.



(e) MFCC features with 15 coefficients.



(f) MFCC features with 15 coefficients.

Fig. 7. Feature extraction form of speech.

## V. GENERATE SEARCH ENGINE ON GENETIC ALGORITHM

Genetic Algorithm has been applied in two ways for the encoding and decoding purposes. On encoding every speaker utterance is compared with an environmental noise utterance and made some groups. In each group one utterance is selected which is defined as the codeword of that group. As a result on encoding, some groups have been

defined and one speaker utterance will lead one group. On decoding side, when unknown speaker utterance will come then it is matched with a leading utterance. The unknown utterance will then find out within that group.

In GA processing selection, crossover and mutation operators have been used here. The fitness function is expressed as follows:

$$fitness = (unknown\ speech\ features \times each\ stored\ speech\ features)$$

In the recognition phase, for each unknown group and speaker within the group to be recognized, the processing shown in Fig. 8 has been carried out.

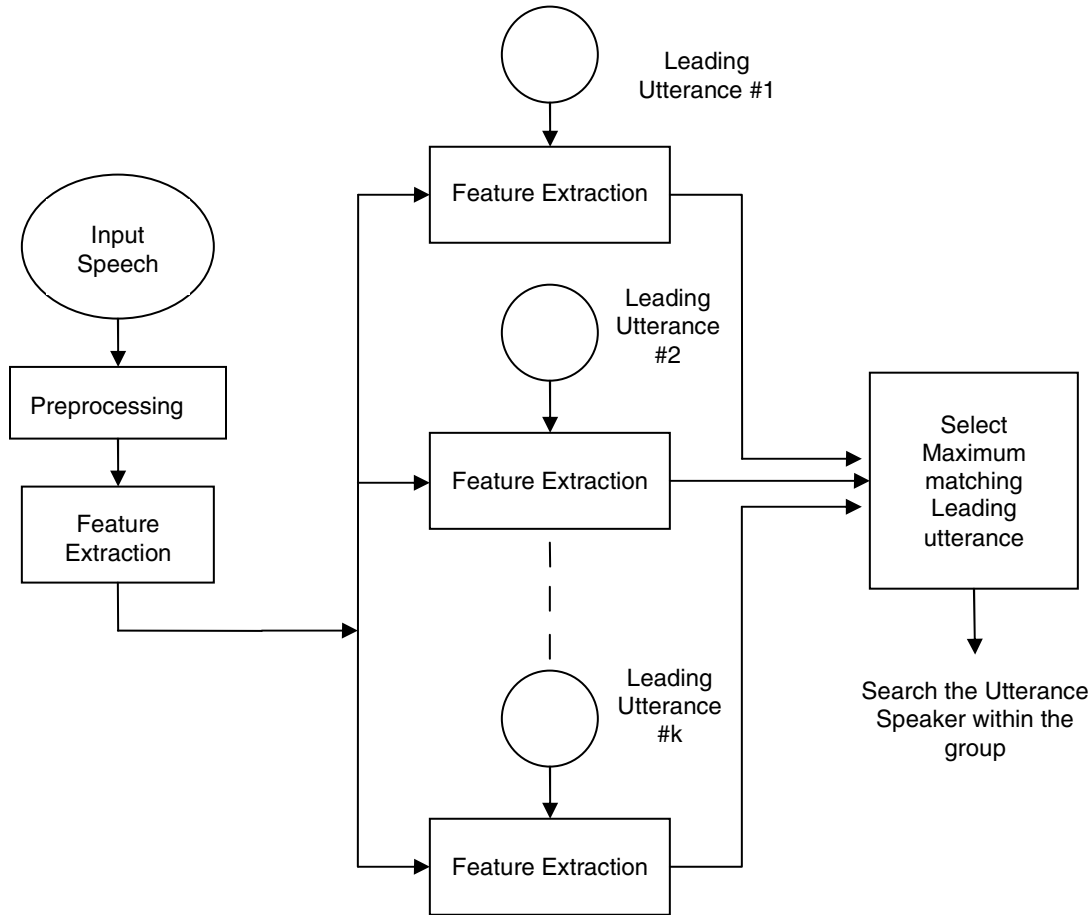


Fig. 8. Recognition model on Genetic Algorithm for the automatic speaker identification system.

## VI. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

### A. Experiments on the Crossover Rate

The identification rate has been measured according to the various crossover rates. Fig. 9 shows the comparison among results of different crossover rates. It is shown that the highest identification rate was achieved at crossover rate 5.

### B. Experiments on the No. of Generations

The number of generations has also been varied to measure the best performance of this codebook design

method. According to the number of generation 5, 10 and 20 (with crossover rate 5), a comparative identification rate was found which is shown in Fig. 10. When the comparison is continued up to 5th generation, highest speaker identification rate is achieved.

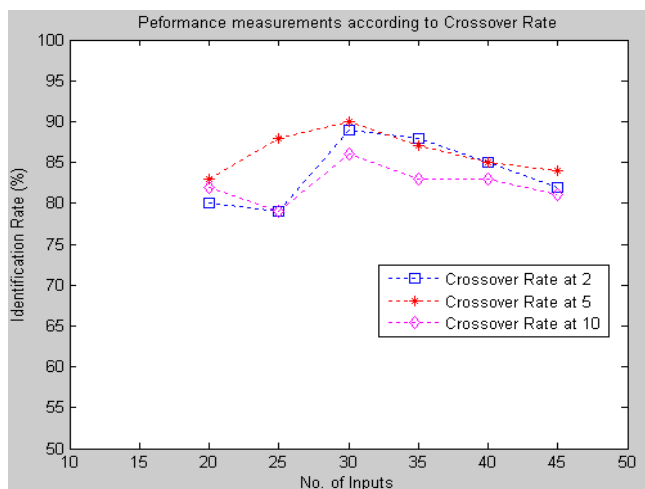


Fig. 9. Performance comparison among different crossover rate.

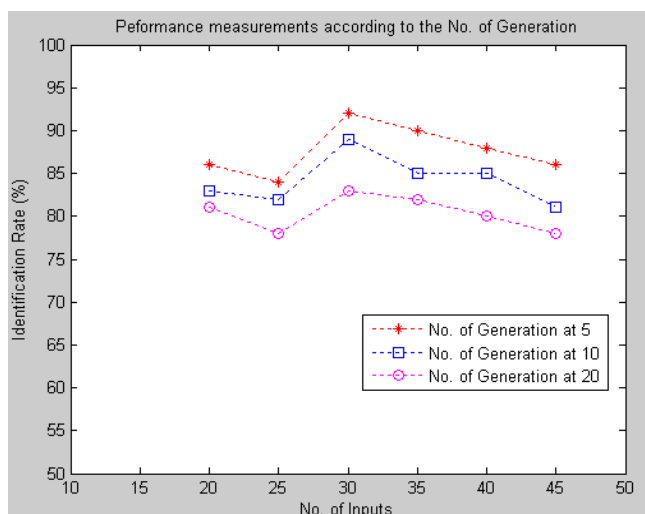


Fig 10. Performance comparison among various numbers of generations.

## VII. CONCLUSION AND OBSERVATION

The experimental results reveal that the performance of the proposed codebook design method yields about 92% identification rate that is seemingly higher than the previous technique that utilized LBG clustering method. However, a benchmark comparison is needed to establish the superiority of this proposed method and which is underway. In the speaker identification technique, noise is a common factor that influences the performance of this technique significantly. In this sense, the conventional techniques are not very efficient and are not flawless. So, GA based codebook design method is capable to protect the system from noise distortion. The performance of this system may be improved by using some noise removing techniques and some advanced signal pre-processing methods.

## REFERENCES

- [1] Online documentation from <http://ditelo.itc.it/people/falavi/IdVe.html>
- [2] Evgeny Karpov, "Real-Time Speaker Identification", *Master's Thesis*, 2003, Department of Computer Science, University of Joensuu.
- [3] Kaushik Roy and M. Ganger Ali, "Implementation of Speech Recognition System using Artificial Neural Network", *Third International Conference on Electrical, Electronics and Computer Engineering*, 2003, Dhaka, Bangladesh.
- [4] J. Picone, "Signal modeling techniques in speech recognition", *Proceeding of the IEEE*, 81,9 (1993), pp. 561-580.
- [5] J. Harrington and S. Cassidy, *Techniques in Speech Acoustics*, Kluwer Academic Publishers, Dordrecht, 1999.
- [6] Claudio Beccchetti and Lucio Prina Ricotti, *Speech Recognition Theory and C++ Implementation*, John Wiley & Sons. Ltd., pp.124-136.
- [7] L.P. Cordella, P.Foggia, C. Sansone, and M. Vento, "A Real-Time Text-Independent Speaker Identification System", *Proceeding of the 12th International Conference on Image Analysis and Processing*, ICIAP'2003.
- [8] J. Proakis and D. Manolakis, *Digital Signal Processing, Principles, Algorithms and Applications*, second ed. Macmillan Publishing Company, New York, 1992.
- [9] Mike Lincoln, "Characterization of Speakers for Improved Automatic Speech Recognition", *A thesis submitted for the Degree of Doctor of Philosophy*, School of Information Systems, University of East Anglia, Norwich, December 13, 1999.
- [10] Dan Bakmand-Mikalski, Anders Havnso Rasmussen and Niels-Ole Christensen, "Speaker identification", *Master thesis*, DTU Oktober 2007.

# Emotion-Driven Learning Agent for Setting Rich Presence in Mobile Telephony

Sumanta Saha\* and Rumana Quazi

\*The Royal Institute of Technology (KTH), Sweden and University of Dhaka, Bangladesh  
sumanta@kth.se, rquazi@hotmail.com

**Abstract**—Presence or personal status information is going to be an integral part of human life in the near future. With the possibility of personalizing user preferences in a fine grained way, mobile presence will appeal to most users. Among all the advantages, one of the most annoying problems is to set the presence status manually each time. This paper discusses the development of an intelligent agent based presence client that will learn and make decisions on behalf of the user about his or her presence status. The decision is emotion driven and the learning depends on real world experience. The proposed system utilizes a neural network (NN) based emotion-driven agent to learn user preferences. As a NN learning algorithm, two approaches based on Differential Evolution and Reinforcement have been proposed, of which either one can be used. Rich presence status is set using a scripting language named Call Processing Language; and SIP is used for publishing the presence to others.

**Index Terms** — Presence; neural network; ANN; CPL; context aware; learning agent; SIP; DE; reinforcement

## I. INTRODUCTION

For a long time emotion has been kept out of the computing world. And then slowly, but surely it started to gain attention to the computing field. Scientists started to build systems with human like behavior that can respond to human emotion, and can modify their own behavior accordingly. Unfortunately the intrusion of emotion into computing has largely been in the fixed systems, rather than in mobile computing. Knowing that mobile computing is going to be a major part of a human's life in the coming days, it is hard to keep human feelings out of it.

In the present situation, mobile devices cannot understand how irritated the user gets when he or she receives a call while in a meeting or while driving on a busy road. The user may prefer that the call be diverted to voice mail, rather than ringing in the user's jacket pocket. With mobile devices, the best one can do is to set his own preferences explicitly as he or she changes context. Mobile services have been developed which are context aware and enable the user to set preferences according to their context, such as, divert my calls while I am at the office or keep my mobile device in silent mode while I am in meeting room. But still, these

preferences have to be set explicitly by the user prior to the event.

If mobile devices can be smart enough to sense the emotion of human users and act accordingly, thus irritation and many embarrassing situations can be avoided permanently. We have performed a survey in Stockholm, Sweden and received a very healthy response in favor of an intelligent presence agent. In this paper, we have focused on making the mobile device a learning agent that will learn human feelings and try to mimic the act of a human in a certain situation. To determine the context the user is in, location awareness has been used. A neural network (NN) based system has been proposed that will learn in an emotion based reinforced environment. The network will make a decision depending on two inputs: the location of the device and the decision that the user took previously in that location. To make the decision more realistic and avoid false positives, the agent will memorize a decision only when it has a specific frequency more than a threshold. The learnt presence statuses can be transcribed to Call Processing Language (CPL) script by the agent and published to subscribers with Session Initiation Protocol (SIP).

The rest of the paper is organized into a discussion about related works and a motivation behind our work in section two and three. An introduction to presence, location awareness, and the learning agent have been presented in sections four, five, and six respectively. Section seven presents the proposed algorithm and its analysis and section eight compares two approaches of performing the solution. The paper ends following a discussion of possible future works in section nine.

## II. RELATED WORK

With the introduction of the highly popular Instant Messenger (IM) ICQ (I Seek You), people understand the power of presence information. Since then, there has been much work on presence to make it configurable and customizable. In [1], the authors propose a personalized service where a user can configure his own preferences about presence using a simple scripting language called "Call Processing Language" (CPL). In some papers, such as [2], presence information has been used as a means of efficiently manage other tasks, for example, session setup of Push-to-

Talk over Cellular (PoC). Little work has really been done as of today to build intelligence into the presence system, to make it a learning agent which responds to human emotion which acts thereby. To learn more about learning agents and context aware preference learning please refer to [3] and [4].

### III. PROBLEM SPECIFICATION AND MOTIVATION

#### A. Problem specification

According to Ashworth [5], social interaction among humans depends on the functions of perception, motivation, and behavior. Humans perceive their environment, feel the context, and behave with other fellow human beings according to their perception and emotion. When it comes to mobile communication, much depends on the context of the receiver. The attitude towards the incoming call varies according to the environment. Based on our day to day experience, we can infer that the outcome of human emotion mostly depends on context and location in the case of mobile communication.

A learning agent or intelligent agent [6] is a software or hardware agent which observes the environment and acts accordingly to be as close as possible to the performance goal set by the user. In our paper, we propose a context-aware intelligent agent based system to enable rich presence information on mobile devices. The system utilizes the context of the mobile device and the behavior of its owner in that particular context and tries to correlate the context with the user's device setting. If we specify the decision taken by the agent as  $d$  and the reward factor as  $R$ , then  $R$  can be modeled as a function of the user's satisfaction with  $d$ , where:

$$R = f(d) \quad (1)$$

The human contributes by modifying the reward factor  $R$ , which he or she does by changing the mode of presence set by the agent. Based on the discussions, a simplified version of human-agent interaction is shown in figure 1.

#### B. Motivation

Presently, there are many services such as instant messaging applications which provide us with presence information but most of them are very limited and have to be set by the user. However, when it comes to mobile devices, presence information brings a lot of new possibilities. Inclusion of CPL ensures that the user has very fine grained control over the presence information and the way he wants to publish it. Still the problem remains, often the user forgets to set their status or presence information causing embarrassing situations. It could be avoided altogether, if the mobile could decide itself about its status.

While [1] describes how to extend presence information by extending CPL to provide rich information, our proposed

presence system is not limited to user input, but rather introduces an intelligent agent who acts on behalf of the user.

### IV. PRESENCE

*Presence* is a popular way to convey to your friends or contacts your willingness to communicate and the means of communication. This has been widely used in Instant Messengers (IM) for quite a long time. Most popular "Presence" statuses have been "Busy", "Away", or "Stepped out". The application of presence is not only limited to IM but can be very successfully applied in cellular communications. Knowing where a friend will eliminate the requirement of calling many times when the sole reason of calling is to know where he or she is. Additionally, it will help to determine the best way to communicate with a friend if necessary. When the reason to call is just for social interaction, it is better not to call if the person is busy or her mood is sad.

'Presence' has been called "The best thing that ever happened to voice" by Jonathan Rosenberg [7], by one of the authors of SIP RFC and inventor of SIP SIMPLE protocol. He declared the extinction of caller ID; presence will be the way to convey who I am, where I am, and what kind of communication I am willing to receive.

'Presence' could be much more diverse in cellular world than it is in IM world. Probable scenarios include:

- You want to call one of your friends, but find that he is now driving in downtown, and as it was a social call, you decide to leave a voice message instead.
- When you are in a meeting the calls would be automatically redirected to voice mail.
- You expressed your interest in real estate before and when you are wondering in near an available property a descriptive message is automatically delivered.
- The location of your best friends nearby can be detected and the phone can inform you with an integrated map with directions.

The Internet Engineering Task Force (IETF) RFC 2778 [8]

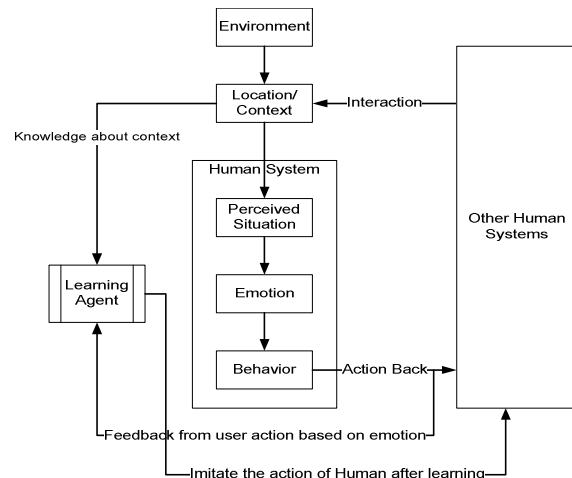


Figure 1: Human-Agent Interaction

defined a model to describe presence information. A SIP implementation of this RFC has been defined in [9]. In [9], the presence information of a user is published as an event to the interested parties defined as *watchers*. While several technologies can be used to enable a presence service, SIP is well suited for this purpose because of its simplicity, focus on the Internet, compact format, and simplicity.

#### A. Session Initiation Protocol (SIP)

In the contemporary world, two signaling protocols are dominant for session setup and teardown, namely: IETF's SIP from and ITU-T's H.323 [10]. Although both are quite successful and have similar features, SIP is the predominant protocol today due to its simplicity. In our paper, we have used SIP for signaling and session setup, although our concept applies equally to both. SIP is an application layer protocol for establishing, maintaining, and terminating multimedia sessions or calls. Two SIP extension messages for presence have been defined by IETF, namely: SIP SUBSCRIBE and SIP NOTIFY. These messages are defined and discussed in [9, 11]. Later an event state publication method SIP PUBLISH [12] was proposed. This message can be used for presence event publication in a presence system.

With the advent of many customized services, it has become critical to control and manage them. The Call Processing Language (CPL) is a solution built on top of SIP to let the end users specify and control services.

#### B. Call Processing Language (CPL)

CPL [13] is a simple markup language to specify and control Internet telephony services. It is based on XML syntax and made to be easily used by non-professional users to control their services. According to [13], "It is meant to be simple, extensible, easily edited by graphical clients, and independent of operating system or signaling protocol. It is suitable for running on a server where users may not be allowed to execute arbitrary programs, as it has no variables, loops, or ability to run external programs." For examples of CPL scripts please refer to [1].

A CPL script is usually associated with a user and only activated when a SIP INVITE message for that particular user arrives. So, the scripts can be on the user device as well as in a proxy server (freeing the mobile device from the load of processing) while providing the necessary privacy and personalization.

### V. LOCATION AWARENESS

Location awareness enables a device to be geographically located, while presence allows a user to know the status of another user. These two services act as a complement to each other. The application of location-awareness along with presence is so promising that people in telecommunication industry (as well as statistical research organizations) are

predicting that location aware services will increase exponentially every year. One reason behind such high expectations is a 1996 mandate requiring that by December 2005 all mobile carriers in the US be able to locate a subscriber making an emergency 911 call to within an accuracy of 50 to 100 meters 60% of the time.

A mix of technologies is used today for location awareness. The most accurate one is Global Positioning System (GPS) which can pinpoint a device with the highest accuracy and another is Cell triangulation [14]. In triangulation, using the signal strength of the cellular towers a triangle is formed and thus the location of a device is discovered. Another technology is ultra wideband [15], which is even more accurate than GPS but is still under active research.

### VI. LEARNING AGENT

A learning agent is a programmable software or hardware agent that helps human beings by acting on their behalf. The intelligence in a learning agent can be implemented in various ways. For example: an Artificial Neural Network (ANN), Bayesian network, Markov random fields, and so on. The learning process can be supervised, unsupervised, reinforced, semi-supervised, transduction, and learning to learn.

In our proposal we have used ANN with both Reinforced learning and Unsupervised learning with a Learning to learn technique. ANN uses adjustable weights attached to each input of the layers to vary the influence they have on the output [16]. A model of the neural network used in this paper is presented in figure 2.

#### A. Reinforced Learning

In many situations, while learning, it is hard to find any feedback. There may not be any benchmark set of data available to compare the output with and it becomes pretty impossible to evaluate whether the agent is providing the proper output. Sometimes, there is only a little information available after a long sequence of input to indicate whether the output is right or wrong. Reinforced learning is

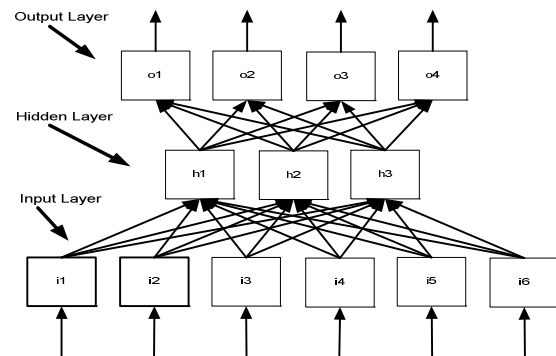


Figure 2: Artificial Neural Network

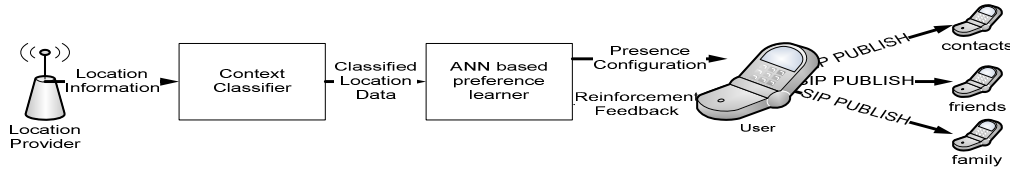


Figure 3: High level architecture of the system

developed to deal with these kinds of situations. It is a kind of supervised learning where the only feedback available is from the environment. Moreover, the information is evaluated rather than instructive. So, reinforced learning is often called “Learning with a critic” rather than “Learning with a teacher”.

### B. Unsupervised learning with Learning to Learn

In unsupervised learning there is no teacher at all. The ideal outcome is completely unknown. The agent has to evolve over the course of time to learn how to react to the inputs to provide the best output which fits the environment. It adjusts through direct confrontation with new experiences. With Learning to learn technique the bias of its weights are also changed automatically.

## VII. METHOD

In this section, we describe the architecture of the proposed presence system while detailing the context identification and the strategy for learning user preferences based on location and emotion. The total system can be divided into high level blocks where each level does its work based upon input from the previous level. The architecture of system looks like figure 3. The context identification part of the architecture helps the ANN by providing a scratchy model output set.

### A. Context identification

The context identification method of our system works on location data available from any location-provider services in the phone. Day by day, services like GPS are getting popular and so our system presumes that such a system is available either in users’ devices or in the cellular infrastructure. The system allocates a node represented by a binary number to each location that is relevant for the system. To evaluate relevancy, it takes the following approach:

1. Detect the location information when the user changes the status of her device
2. Store the information in a temporary buffer and initialize a counter with respect to the (location, status) combination
3. The next time the user comes to this location:
  - a. Check the status information she set
  - b. If it is same as the previous status then increase the counter  
else if the status set is different than that in buffer do

- i. start multiplicative decrease of the counter value
  - c. else if the user sets no status then do nothing
- If the value of the counter reaches a threshold then in a separate process goto step 4.
4. Make the location value permanent and feed it as input to the Neural Network
5. Store the value of status as a model outcome of user preference to use with the Neural Network

The location data is fed to the learning process if and only if the system finds that the user is repetitively setting the same status for that location.

### B. Learning process

The learning process of the agent is done online. Whenever a location node is added to the systems input, the system starts learning the users’ decision in that location. The learning process is done by a Neural Network with one hidden layer. The process is similar to an environmental learning process such as face detection, which does not depend on any explicit equation rather on environmental feedback on a particular decision.

### C. Construction of the network

The Artificial Neural Network (ANN) has a very simple design of an input layer, one hidden layer, and one output layer. The design need not to be complex because the decision to be taken does not depend on many variables. Rather the agent only learns what is trivial for the user to save the effort and disturbance from the users’ part. The input layer takes the location information as input from the previous block of the system, the context identification block. The location is encoded as a 6 bit binary value in our system based on the assumption that a single user will not go to more than  $2^6$  places regularly. It is possible to change this number of bits if necessary. The hidden layer consists of 3 neurons and the output layer consists of 4 neurons. The decision about the output layer is taken under the assumption that a user normally does not have more than  $2^4$  presence statuses. As before, these values are changeable if necessary.

### D. Learning algorithm

We designed the solution with two different learning algorithms where both are equally capable of solving the problem having their own advantages and disadvantages. Both algorithms are specially suited to this kind of situation



where the correct output is not known beforehand, but in the course of learning, environmental feedback is available.

**Differential Evolution:** Differential evolution [17] is an effective, efficient, and robust optimization technique based on an evolutionary strategy. It can be considered as a stochastic search algorithm to determine the best solution in the complex solution space. DE exploits a population of potential solutions, individuals, to probe the search space. New set of solutions are generated by combination and evolution of randomly chosen solutions. To be more specific, for each individual solution  $X_i$  of current generation, where  $i = 1, 2, \dots, N$  and  $N$ =number of random initial solution, a new individual solution  $Y_i$  is generated according to the following equation [18]:

$$Y_i = X_{i1} + F(X_{i2} - X_{i3}) \quad (2)$$

where,  $X_{i1}$ ,  $X_{i2}$ , and  $X_{i3}$  are randomly chosen individual from the current generation to produce offspring  $Y_i$  and  $X_{i1} \neq X_{i2} \neq X_{i3}$ .  $F$  is called the scaling factor or amplification factor. This operation is similar to what is commonly known as mutation in evolutionary biology. The final offspring  $X_{i+1}$  of the next generation are produced from mutation of  $X_{i1}$  and  $Y_i$ . The parameter which regulates how many consecutive genes of the mutated individual on average are copied to the offspring is called crossover rate CR. Finally the offspring  $X_{i+1}$  is evaluated and replaces its parents  $X_{i1}$  in next generation if and only if its fitness is better than that of its parent. This is called the selection process.

The fitness function is the central part of the algorithm where the new generation is evaluated against the parent for fitness. If it is better suited to the ultimate performance goal, then the new breed replaces the parent otherwise the parent is kept. In our algorithm, the fitness function depends on the proximity to the user preferred value of presence status in response to the input location value. A model result is taken from the database as created at the context identification step of the system. If the model result vector is  $Y_i$ , the new generation of weight is  $W_i$  and the input vector is  $X_i$ ; then the fitness function becomes:

$$Y_i \approx \text{ANN}_{W_i}(X_i), \text{ with threshold } t \quad (3)$$

where ANN is the neural network. The less the difference between the model  $Y$  and the derived  $Y$  the more fit is the generation. The approach followed for the ANN in our system in this case is as follows:

1. Take location input from the context identifier
2. Initialize the parameters( $F$ , CR, Population Size) of DE
3. Randomly Initialize the first (initial) population
4. for  $i=1$  to PopSize do
  - a. for  $j=0$  to Number of Input Nodes do Create random candidate weights  $w[i][j]$
  - b. Calculate fitness value of the individual using ANN fitness function

5. While iteration < max iteration (or termination criteria not satisfied)
  - for Each Individual of Current Population
    - i. Select three random and mutually exclusive individuals
    - ii. Create new breed with DE operators and Crossover from the original three random parents
    - iii. Calculate fitness value of the newly generated offspring
    - iv. If (offspring fitness value > original parent fitness value) then promote the offspring to next generation
    - Else promote the original parent to next generation
  - v. End If
6. Report the best fitness value and the offspring as a weighted combination for the input nodes of the ANN.
7. Adjust the weights of ANN according to the output of DE.
8. In case of the user contradictions with any of the decisions made by ANN, multiplicative decrease of the fitness score of that offspring and start the learning process again.

**Reinforced Learning:** In reinforced learning, the weight matrix  $W$  has the size of  $n$  by  $m$  where  $n$  is the number of possible inputs and  $m$  is the number of possible outputs. The value at  $w_{ji}$  denotes the possibility of taking action  $i$  while in state  $j$ . If  $s$  is a new state reached by taking action  $i$  at state  $j$ , we can compute the reward of this move as:

$$R^k = R(i|j) = R_s \quad (4)$$

Let  $l$  be the change of reward score between two moves:

$$l^k = R^k - R^{k-1} = R_s - R_j \quad (5)$$

Equation (5) represents the human attitude towards the decision made. Depending upon that, the rule of updating the weights can be as simple as:

if  $l_k \geq 0$ :

$$w_{ji}^{k+1} = w_{ji}^k + a(1 - w_{ji}^k),$$

$$w_{jh}^{k+1} = (1 - a)w_{jh}^k \quad \text{for } h \neq i \quad (6)$$

if  $l_k < 0$ :

$$w_{ji}^{k+1} = (1 - a)w_{ji}^k,$$

$$w_{jh}^{k+1} = \frac{a}{m-1} + (1 - a)w_{jh}^k \quad \text{for } h \neq i \quad (7)$$

Where  $a$  is the learning rate [19].

The rules encourage the decisions that have better score. If the decision is supported by a better score, such as  $l^k \geq 0$ , it is rewarded by (6). Otherwise it will be penalized by (7). In this case of mobile presence decision making, no feedback from user indicates that the user is happy with the decision made by the agent. So if there is no change in the reward score, we can safely assume that the decision was correct and it can be rewarded. But if the user changes the decision, then the decision made by the agent will be penalized. A more detailed mathematical analysis of reinforced learning can be found at [20].

### E. Realization of Presence status change

In the system, the user is able to create his own presence statuses using a simple web-based client where he can specify every single detail of his preferences. The statuses are translated to a CPL script and stored as output nodes in the system. According to the output from the ANN, a particular node is selected and a SIP PUBLISH method is used to publish the presence status to the subscribers. CPL also sets the status of various mobile device properties such as vibration, ringer, etc.

## VIII. COMPARISON

The main advantage of using differential evolution as a weight adjusting algorithm is that it is unsupervised and it learns the learning rate itself by driving the learning rate through the evolution process. We do not have to provide anything other than the initial fitness function. According to the user's emotion, which is realized by changing an agent decision, the fitness value is adjusted. When using differential evolution, we do not have any problem similar to the typical back propagation where we can be trapped in local minima, and which scales exponentially with the complexity of the problem. On the other hand, reinforcement algorithms such as the one stated above has the property of being extremely fast to learn and adapt. It quickly finds the slope to the reward function and modifies itself by propagating the findings towards the other layers of neuron. The only limitation is that sometimes reinforcement algorithms are trapped in local minima. A good evaluation of differential algorithm and the back propagation approach is given at [21].

## IX. CONCLUSION

An approach to make the mobile devices intelligent for making decisions about the user's presence status in accordance to their location and previous decisions is presented in this paper. Presence status will be an inevitable part of the mobile technology in the near future where everyone can learn about the current status of their contact to make knowledgeable decision about contacting them. Life of most of the people in the world is limited to a certain boundary and in most cases they seem to be occupied in the same kind of work at a certain place. This makes an intelligent agent perfectly capable of predicting or deciding what their status should be if they are in a particular place depending upon what the user did before for the same position. With the plethora of location data providing systems, this approach of human-computer interaction seems to be inevitable. Two algorithms for decision making using artificially intelligent agent along with a market research about the approach have been presented in this paper.

## REFERENCE

- [1] D. Jiang, R. Liscano, L. Logrippo, "Personalization of internet telephony services for presence with SIP and extended CPL", *Computer Communications*, vol. 29, issue 18, pp. 3766-3779, Nov 2006.
- [2] L. Wang, D. Gu, "A study on session setup for group communication in Push-to-Talk over Cellular using rich presence", *ISCIT'07*, pp. 415-420, Oct 2007.
- [3] A. Krause, A. Smailagic, D. P. Siewiorek, "Context-aware mobile computing: Learning context-dependant personal preferences from a wearable sensor array", *IEEE Transaction on Mobile Computing*, vol. 5, no. 2, pp. 113-127, Feb 2006.
- [4] C. Ono, M. Kurokawa, Y. Motomura, H. Asoh, "A Context-Aware mobile preference model using a Bayesian network for recommendation and promotion", *UM2007*, pp. 247-257, 2007.
- [5] P. D. Ashworth, "Social Interaction and Consciousness", *John Wiley & Sons*, 1979.
- [6] Intelligent Agents, In [http://en.wikipedia.org/wiki/Intelligent\\_agent](http://en.wikipedia.org/wiki/Intelligent_agent), last accessed: April, 2008.
- [7] J. Rosenberg, "Presence: The best thing that ever happened to voice", <http://www.callcentermagazine.com/article/CTM20001023S0001>, last accessed: May 2008.
- [8] M. Day, J. Rosenberg, H. Sugano, "A model for Presence and instant Messaging", *RFC 2778*, <http://www.faqs.org/rfcs/rfc2778.html>, last accessed: May 2008.
- [9] J. Rosenberg, "A Presence Event Package for the Session Initiation Protocol", *RFC 3856*, <http://www.faqs.org/rfcs/rfc3856.html>, last accessed: May, 2008.
- [10] I. Dalgic, H. Fang, "Comparison of H.323 and SIP for IP telephony signalling", *Proceedings of Photonics*, 1999.
- [11] A. B. Roach, "Session Initiation Protocol (SIP)-Specific Event Notification", *RFC 3265*, <http://www.faqs.org/rfcs/rfc3265.html>, last accessed: May 2008.
- [12] A. Niemi, "Session Initiation Protocol (SIP) extension for Event State Publication", *RFC 3903*, <http://www.faqs.org/rfcs/rfc3903.html>, last accessed: May, 2008.
- [13] J. Lennox, H. Schulzrinne, "Call Processing Language (CPL): A Language for User Control of Internet Telephony Services", *RFC 3880*, <http://www.faqs.org/rfcs/rfc3880.html>, last accessed: May 2008.
- [14] Alabama chapter of NENA, "Triangulation Method", [http://www.al911.org/wireless/triangulation\\_location.htm](http://www.al911.org/wireless/triangulation_location.htm), last accessed: May 2008.
- [15] Intel Corp. "Ultra-wideband (UWB) Technology", <http://www.intel.com/technology/ultrawideband/downloads/Ultra-Wideband.pdf>, last accessed: May, 2008.
- [16] C. Stergiou, D. Siganos, "Neural networks", [http://www.doc.ic.ac.uk/~nd/surprise\\_96/journal/vol4/cs11/report.html](http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html), last accessed, May 2008.
- [17] R. Storn, K.V. Price, "Differential Evolution – a simple and efficient heuristic for global optimization over continuous space". *Journal of Global Optimization*, vol. 11(4), pp. 341-359, Dec 1997.
- [18] R. Storn, J. A. Lampinen, "Differential Evolution: A Practical Approach to Global Optimization" *Springer*, 2007.
- [19] M. Zeidenberg, "Neural Networks in Artificial Intelligence", *Chichester: Ellis Horwood Limited*, 1990.
- [20] Y. Wang, D. Lee, "Reinforcement Learning for a Human-Following Robot", *ROMAN'06*, pp. 309-314, sep 2006.
- [21] J. Ilonen, J. Kamarainen, J. Lampinen, "Differential Evolution Training Algorithm for Feed-Forward Neural Networks", *Neural Processing Letters*, Springer, vol. 17, no. 1, pp. 93-105, 2004.

# An HPSG Analysis of Arabic Passive

Md. Shariful Islam Bhuyan and Reaz Ahmed

Department of Computer Science and Engineering,  
Bangladesh University of Engineering and Technology,  
Dhaka, Bangladesh

**Abstract**—In spite of being a successful syntactic theory in many respects, Head-driven Phrase Structure Grammar (HPSG) has inadequate coverage for morphological constructions, especially for nonconcatenative morphology, which is prominent in the Semitic languages such as Arabic, Hebrew etc. Among various syntactic and semantic phenomenon, passive constructions draws attention of many researchers in theoretical linguistics due to their diversity. Arabic exhibits lexical passives with nonconcatenative morphology. In this paper, we extend the HPSG framework to support the nonconcatenative construction of Arabic passives along with the necessary lexical rules and type hierarchy.

**Index Terms**—Passive Construction, Arabic Morphology, Nonconcatenative Morphology, Head-driven Phrase Structure Grammar, Constraint-based Grammar

## I. INTRODUCTION

Broad-coverage precision grammar [1], [2] and computational lexicon development for deep linguistic processing is a research-intensive area with several potential applications [4]. Amidst the vast literature on formal linguistic theory [5], Head-driven Phrase Structure Grammar (HPSG) [6] has a unique position since it combines the best features of the contemporary approaches as well as establishes an integrated framework for cross-layer representation comprising phonology, morphology, syntax, semantics, pragmatics and discourse. Although, HPSG successfully describes numerous syntactic and semantic phenomena, it lacks rigorous analyses for morphological phenomena, especially for non-concatenative morphology [7]. Nonconcatenative morphology illustrates an interesting paradigm of morphological operations, which is prominent in the Semitic languages such as Arabic, Hebrew etc [9].

From the inception of mathematical formalism in theoretical linguistics, passive construction draws attention of many researchers. HPSG is used in successful analysis of passive for several languages [8], [9]. However, there is no HPSG analysis for Arabic passives. Among the living languages, Arabic demonstrates the best instance of nonconcatenative morphology. Arabic verb system exhibit both concatenative and nonconcatenative morphology, capable of lexically expressing diverse syntactic and semantic phenomena including passive. In this paper, we extend the HPSG framework to support Arabic passive along with its nonconcatenative morphology.

## II. PASSIVES IN ARABIC

Arabic word formation is an excellent example of root-pattern morphology [8], [9]. A combination of root letters are plugged in a variety of morphological pattern with priory fixed letters and particular vowel melody that gives rise to corresponding syntactic and semantic phenomena.

An Arabic word can encode a complete sentence. For example, *sayaktubuhu* – *He will write it*. We can break the word in the following component.

- 1) Prefix: *sa* – the particle indicating future.
- 2) Suffix: *hu* – the object pronoun attached as a clitic
- 3) Root: *k, t, b* – the root letters bearing the concept of writing
- 4) Measure: *ya\_\_u\_\_u* – bearing the syntactic and semantic information of the event

It may be possible to concatenate multiple prefixes and suffixes. However, there must be a single measure and single set of root letters, where the measure packages syntactic and semantic features and root supplies the core concept. Here the measure indicates that – the actor is in 3rd person, singular number, masculine gender; the verb is in indicative case, active voice and derived form I; it also indicates that the event has not yet been completed. If we plug in another set of root letters, for example, *n, S, r* – which bears the concept of helping, we get *sayanSuruhu* – *He will help him*.

Arabic passives are constructed using measures. There are dedicated measures for expressing passives. For example, the passive of *yaktubu* is *yuktabu*. The passive measure is *yu\_\_a\_\_u*. Several syntactic and semantic information has been modified in this measure. First, the straight forward change is stem alternation. The measure *ya\_\_u\_\_u* has been changed to *yu\_\_a\_\_u*. Next, its subcategorization feature has been changed. In general, only the transitive verbs can exhibit passive sense. In active form, these verbs expects syntactic object(s). In passives, these objects are promoted to the position of syntactic subject. Subject in active forms are realized by obligatory prepositional phrase in languages like English. However, Arabic passives do not expect this subject in active forms to be its syntactic argument. Rather, they remain unexpressed. The object in the active form becomes the subject in passive. Another important concept is hidden

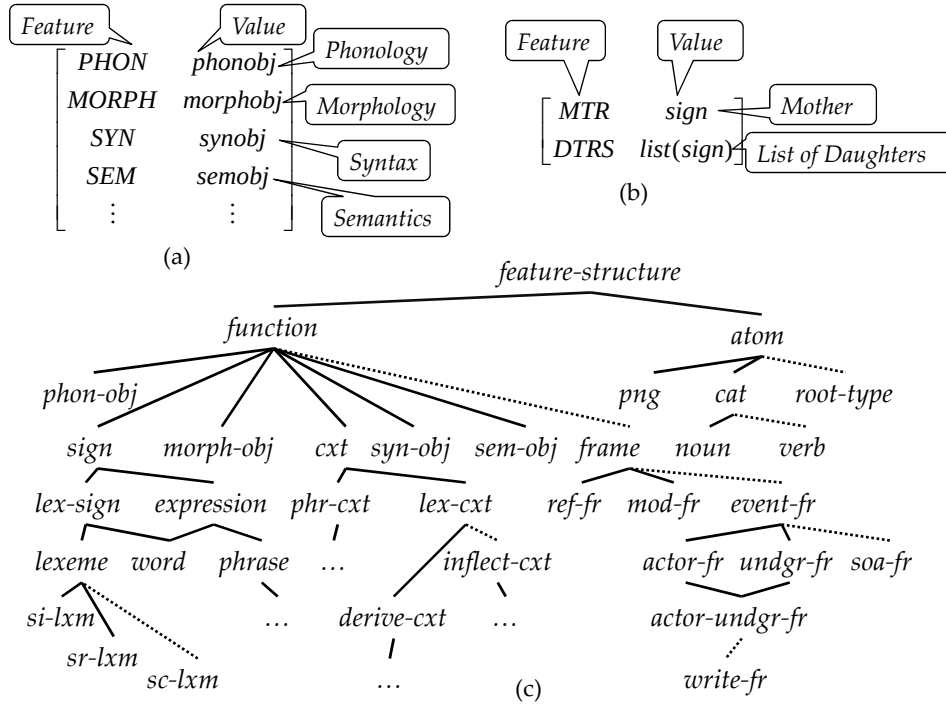


Fig. 1. (a) An HPSG sign (b) An HPSG construction (c) A standard HPSG type hierarchy

pronoun. The semantic actor of Arabic verbs may not be syntactically realized. In that case, a hidden pronoun is introduced which plays the role of semantic actor. In passive, this hidden pronoun plays the role of semantic undergoer. In the next sections, we present the necessary formalisms to capture Arabic passives under HPSG framework.

### III. AN HPSG PRIMER

Natural languages generally consist of two components. First, the utterances that can be used by human. Second, the linguistic rules that license those utterances. For example, in English, "He writes books", "writes books", "writes" - all are valid utterances. However, "Writes he books", "writes he", "rwrite" are not valid, since the rules do not license them. HPSG is a mathematical theory for natural languages that formally captures these two core linguistic components. Utterances are modeled using a mathematical object *Sign* (a formal representation of words, phrases as well as sentences) and rules are captured using another mathematical object *Construct* (a formal representation of grammar rules or schema which are used to license signs). Both sign and construct are described using feature structure - a collection of features of corresponding linguistic objects along with their values. These features and their values constitute a very detailed type hierarchy (see figure 1(c)).

An utterance can have linguistic feature spanning multiple layers, e.g. phonological morphological, syntactic, semantic, pragmatic and others. To capture these features, the description of a typical HPSG sign looks like figure 1(a). To capture grammatical rules, the feature structure of a construct has a *mother* (MTR) feature and a *daughters* (DTRS) feature. The value of the MTR is a sign and the value of the DTRS is a nonempty list of signs. The description of a typical HPSG construct looks like figure 1(b). The licensing of signs follows *The Sign Principle* which states that "Every sign must be lexically or constructionally licensed, where, lexically licensed only if it satisfies some lexical entry, and constructionally licensed only if it is the mother of some construct" [6].

From the type hierarchy of figure 1(c), we can see that there are two types of feature structure. Functions are those feature structures which are described using an attribute value matrix (AVM). They map features to feature structures. Atoms are atomic types that can be used as the value of features. Notable functions are sign, cxt (construction), lexeme, phrase and others. To model a linguistic phenomenon we first need to identify the involved signs with their hierarchy. Next, we need to design functional feature structures for them with linguistically motivated features. Then we also need to define the necessary constructs as well as the atomic type hierarchy. In the next section, we

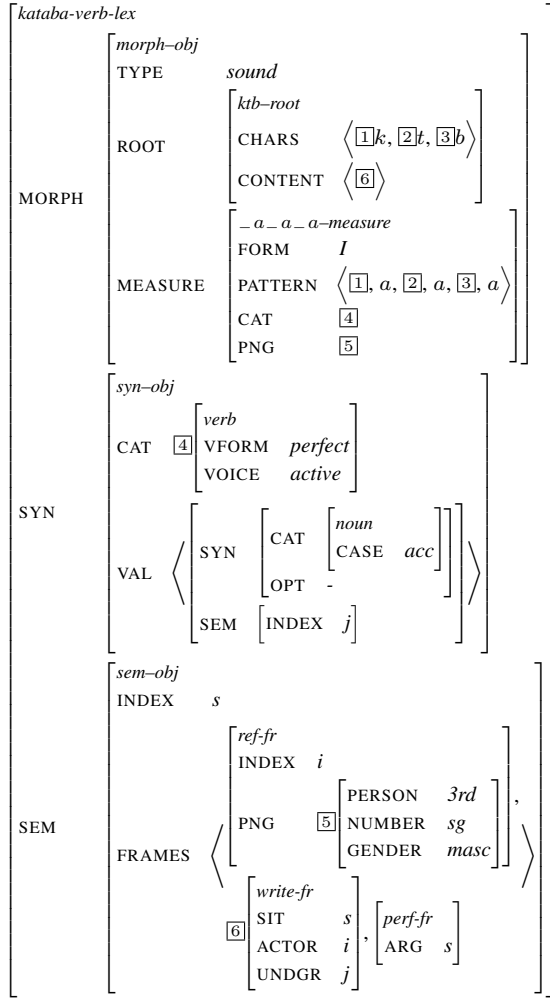


Fig. 2. HPSG Sign for *kataba*

build these ingredients for Arabic passives.

#### IV. HPSG ANALYSIS OF ARABIC PASSIVE

Here we give the attribute value matrix (AVM) for an Arabic verb *kataba* – *He wrote* in active form and its corresponding passive form *kutiba* – *It was written*, using our analysis.

In the figure 2, we have three features associated with morphology. First, the feature TYPE, which denotes the associated root class. Arabic roots are classified into several root class according to their derivational and inflectional paradigm. This feature affects both root and measure. So, it has been taken out to a first level morphological feature. In this case, its value is *sound*. Next, the feature ROOT, which has a list of root letters as well as the CONTENT feature, which gives the semantic contribution made by root letters. In this case, its value is structure-shared with the *write-fr* in the FRAMES feature. Next, the feature MEASURE, which contains the morphological, syntactic and semantic information contributed by measure. First,

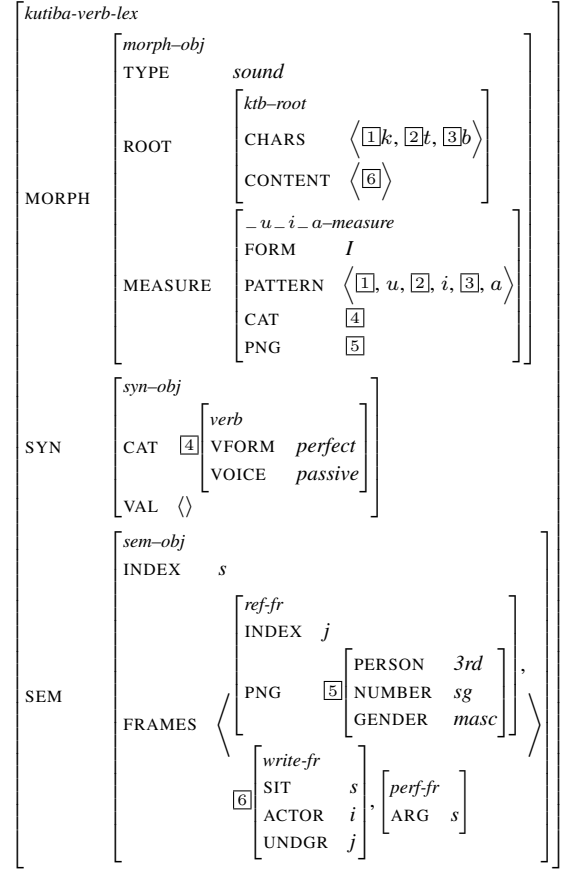


Fig. 3. HPSG Sign for *kutiba*

the feature FORM, which denotes the semantic paradigm. *kataba* - is a form-I derivative. Next, the feature PATTERN which captures the stem along with the root letters using structure sharing. Then, the feature CAT, which contains the syntactic category for this measure. Its value is structure-shared with the syntactic feature CAT. Finally, the feature PNG, which captures the PERSON, NUMBER and GENDER information of our semantic actor in the case where it is not syntactically realized.

We present the syntactic and semantic information using the SYN and SEM feature, for the word - *kataba*. First, the CAT features identifies the syntactic category of - *kataba*. It contains the VFORM and VOICE feature of Arabic, which governs the derivational paradigm of verb lexeme. In this case their values are *perfect* and *active* respectively. Next, the VAL feature, which captures the subcategorization of verbs. VAL is a list of signs, which are required by the syntactic head. In this case, the verb - *kataba*, requires an object. The verb - *write* is a transitive verb that takes an object. We should note that the hidden pronoun. *he* is encoded by the inflectional morphology, when no explicit subject is used. The semantic actor is not realized syntactically. So, the verb only subcategorizes for

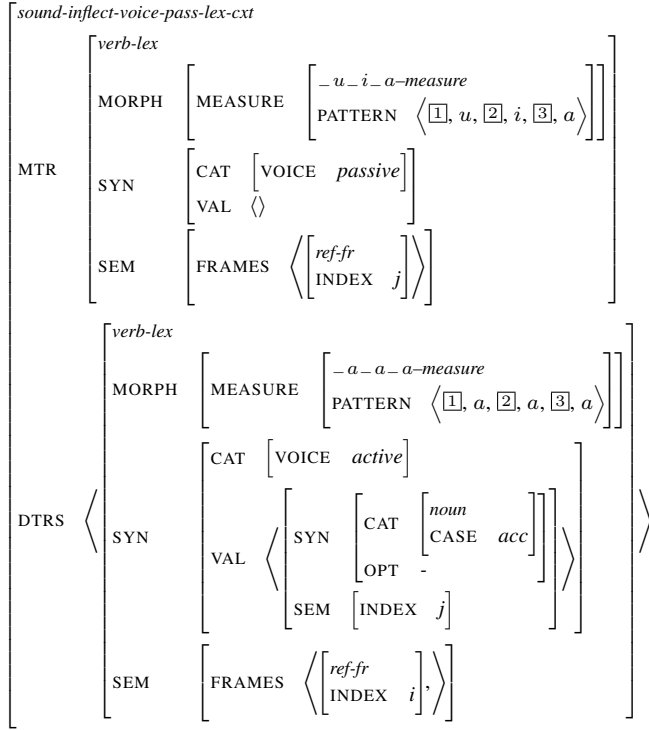


Fig. 4. Lexical rule for passive construction

syntactic object. We can also see the constraints imposed over the object. In this version, its syntactic head should be a noun phrase with the value of its CASE feature set to *accusative*. The negative value of the OPT feature indicates that this object is not optional, rather required to be syntactically correct.

Next, we need to consider some semantic features. Here, we use a type feature version of predicate logic to capture semantics of natural language. First, we consider the INDEX feature, which is a reference to a discourse entity. Then, the PNG feature, which capture the semantics of PERSON, NUMBER and GENDER. Next, the FRAMES feature, which serves as a bag for elementary predicates to describe the situation at hand. For example, in the case of *kataba*, the event of writing is expressed. The event is completed in the past and there is a discourse referent to the actor. To capture the core event, write-predicate is introduced. To capture the temporal constraint, we use the perfect-predicate. Finally, to express the actor of the event, the hidden pronoun, we introduce a discourse referent with corresponding PNG feature.

Predicates have their respective arguments. write-predicate has a situation hook, expressed by the feature SIT. There are two semantic role associated with this predicate. First, we consider the role of writer, who plays a doer role, expressed by the feature ACTOR. Second, we consider the role of written, who plays an undergoer role, expressed by the feature UNDGR. The perfect-predicate

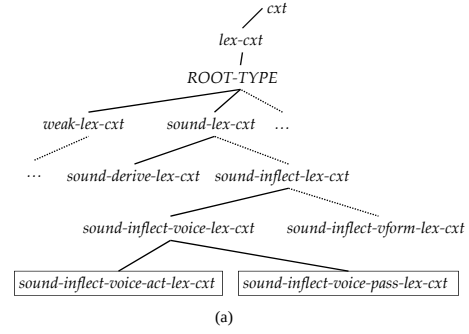


Fig. 5. Type hierarchy for lexical construction

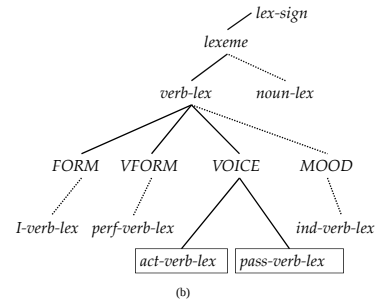


Fig. 6. Lexical type hierarchy

takes a situation hook as an argument, which is expressed as the feature ARG. We use the technique of co-indexing for sharing semantic objects. The discourse referent predicate is actually the actor of the write-predicate. To denote this constraint, the INDEX value of hidden pronoun and the ACTOR value of the write-predicate are co-indexed, both are given the value *i*. This is an example of reference co-indexing. We also use event co-indexing. The event hook SIT of write-predicate, situation hook of the entire scenario and argument ARG of the perfect-predicate, all are co-indexed and expressed using the value *s*. Another important issue of HPSG representation is the syntax-semantics interface. In this example, this is done by co-indexing the INDEX value of the syntactic object and the UNDGR value of the write-predicate with a value *j*. This indicates that the syntactic object is our semantic undergoer whereas from our previous discussion we can note that the semantic actor is not syntactically realized.

In the figure 3, we show the HPSG representation of the passive - *kutiba*. We identify the associated changes for this conversion. The PATTERN feature is changed to capture the derivational morphological operation. Next change can be found obviously in the feature VOICE, changing

its value to passive. Unlike English, which can have a prepositional complement in passives, Arabic passives do not subcategorize for a subject or any other argument. For this reason, the VAL list is empty. Moreover, the discourse referent in the feature FRAMES is now co-indexed with the UNDGR feature of the write-predicate, expressed by the value *j*. Semantic actor now completely unknown by not having any syntactic or semantic reference, which is a distinctive property of Arabic passive. These changes can be succinctly captured by the lexical construct in figure 4. We also give the type hierarchy for active and passive verb lexeme and lexical construction for passive. They are given in the figure 5 and figure 6.

## V. CONCLUSION

In this paper, we give the proposal how to capture Arabic passives along with their nonconcatenative morphology within the framework of HPSG. There are lots of syntactic and semantic phenomena that needs to be captured for the development of a full-fledged computational grammar for Arabic. Results will be immensely helpful for the construction of resource grammar for languages with rich nonconcatenative morphology.

## REFERENCES

- [1] A. Copestake, and D. Flickinger, "An open-source grammar development environment and broad-coverage English grammar using HPSG," Second conference on Language Resources and Evaluation, 2000.
- [2] M. Marimon, N. Bel, S. Espeja, and N. Seghezzi, "The Spanish Resource Grammar: pre-processing strategy and lexical acquisition," ACL Workshop on Deep Linguistic Processing, 2007.
- [3] J. Tseng, "English Prepositional Passive Constructions," HPSG Conference, pp. 271-286, 2007.
- [4] F. Bond, S. Oepen, M. Siegel, A. Copestake, and D. Flickinger, "Open source machine translation with DELPH-IN," Open-Source Machine Translation Workshop at the 10th Machine Translation Summit, pp. 15-22, 2005.
- [5] P. Sells, Lectures on Contemporary Syntactic Theories, Stanford: CSLI Publications, 1985.
- [6] I. A. Sag, and T. Wasow, Syntactic Theory: A Formal Introduction, Stanford: CSLI Publications, 1999.
- [7] S. Bird, and E. Klein, "Phonological Analysis in Typed Feature Systems", Computational Linguistics, vol. 20, pp. 55-90, 1994.
- [8] K. R. Beesley, "Finite-State Morphological Analysis and Generation of Arabic at Xerox Research: Status and Plans in 2001", ACL Workshop on Arabic Language Processing: Status and Prospects, pp. 1-8, 2001.
- [9] S. Z. Riehemann, A Constructional Approach to Idioms and Word Formation, PhD Dissertation, Stanford University, 2001.
- [10] S. Song, and J. Choe, "Type Hierarchies for Passive Forms in Korean," HPSG Conference, pp. 250-270, 2007.

# WalkSAT Approach in Solving The Staff Transfer Problem

Sriyankar Acharyya

Department of Computer Science & Engineering, West Bengal University of Technology, Kolkata,  
West Bengal, email: srikalpa8@yahoo.co.in, India

**Abstract** — Staff Transfer is an important issue in Human Resource Management. It is concerned with the assignment of transfer postings to employees in large organizations. A large organization has offices and worksites at multiple locations and it is customary to transfer a subset of employees at periodic intervals. The Staff Transfer Problem (STP) can be viewed as a Constraint Satisfaction Problem(CSP). The deterministic methods were proved to be inferior to local search methods in solving this problem. In earlier investigation GSAT(L) easily outperformed the Simulated Annealing (SA) which was the then best among all approaches. In this work computer experiments indicate that WalkSAT(L,p) easily outperforms GSAT(L) in most of the instances.

**Index Terms** — Constraint Satisfaction, Local search, Satisfiability, Tabu Search, Staff Transfer

## I. INTRODUCTION

A transfer is a lateral movement of an employee in an organization not involving a change in rank. In many large organizations that maintain offices and work-sites at multiple locations, it is customary to transfer a subset of employees at periodic intervals from one office or work site to another. Examples of such organizations are the armed forces, government departments, commercial banks, and construction firms. Transfers play a major role in human resource flow in organizations, and the satisfactory assignment of transfer postings to employees is an important issue in Human Resource Management [4,5].

The *Staff Transfer Problem* (STP) [1] is described and formulated as a Constraint Satisfaction Problem(CSP). The variables take values that correspond to transfer postings of employees. A typical constraint enforces the requirement that the total number of employees assigned to a transfer posting should not exceed the specified number of vacancies. The objective is to find transfer postings for a given subset of employees that satisfy all the constraints. Fig. 1 represents a diagram showing the basic structure of the problem

The constraints here are not binary but involve a number of variables. The constraints are of two types, *strict* and *desirable*. A strict constraint must be satisfied; for example, an employee can only be transferred to a location where there is a vacancy for a job the employee can perform. A solution that satisfies all strict constraints is a feasible solution. A desirable constraint should be

satisfied if possible. For example, it is desirable that senior employees get priority over their juniors in the assignment of transfer postings. The quality of a feasible solution is determined by the degree to which desirable constraints are satisfied. The objective is to find feasible solutions of high quality.

The transfer policy that is adopted in an organization must be both fair and systematic to be acceptable to employees. When the transfers involve a large number of employees, it is typically quite difficult to find suitable alternative positions for everyone. Some preliminary results are reported in [2]. More detailed results comparing some CSP techniques and showing the superiority of Simulated Annealing over other methods are reported in [3]. But when the Staff Transfer Problem(STP) was converted to Satisfiability (SAT) and GSAT with tabu list was applied, the Simulated Annealing was outperformed by GAST(L) significantly[13]. In this paper we have shown that WalkSAT(L,p) approach is significantly better than GSAT(L) in performance.

In this paper the sections are arranged as follows. Section 2 describes the Staff Transfer Problem (STP) in detail. Section3 presents solution methods applied to this problem. Experimental results are provided in Section 4. Section 5 summarizes the paper and discusses some unsolved issues.

## II. PROBLEM DESCRIPTION

There is a set of  $E$  employees who are to be transferred. The organization has other employees who are not being transferred. At any office or work site, the number of vacancies is the difference between the number of sanctioned positions and the number of employees who are not being transferred. An employee can perform only certain types of jobs. An employee has a seniority level, which is determined by factors such as the level of the current position in the organizational hierarchy, the date of joining the current position, and the date of birth. There is a set  $LOC$  of locations participating in this process. A location corresponds to an office or a work site. There is a set  $J$  of job categories. Typical job categories are electrician, typist and manager. For each pair  $(loc, j)$ ,  $loc$  in  $LOC$  and  $j$  in  $J$ , the number of vacancies  $vacancy(loc, j)$  is known. This number includes the chain vacancies that



arise as a result of the E transfers. It can be zero, indicating that job type j does not exist at location loc or that there is currently no vacancy for job type j at location loc. Fig. 1 shows a schematic representation of the Staff Transfer Problem.

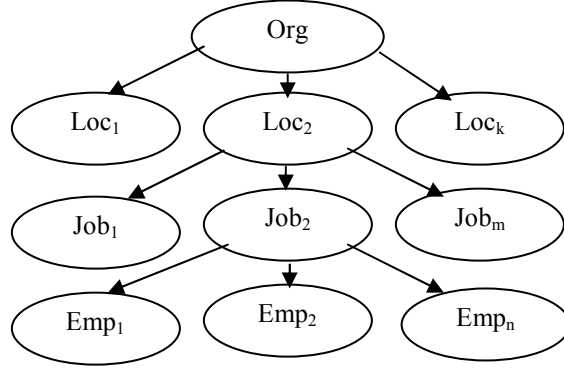


Fig.1. A schematic diagram of the Staff Transfer Problem where no of locations = k, no of job types in location 2 = m, no of employees in bucket (loc<sub>2</sub>, job<sub>2</sub>) = n

An employee who is due for transfer submits options for  $T \geq 2$  destinations, each represented by a (loc,j) pair. The job types in the transfer options correspond to jobs that the employee can perform. The parameter T has a typical value of 3. The options are arranged by preference, the first option having the highest preference for the employee, the second the next highest, and so on.

For convenience, we assume that the employees to be transferred are numbered 1 through E in decreasing order of (rank, seniority). Thus employee 1 has the highest rank and is seniormost, and employee E has the lowest rank and is juniormost. Let *vac* be the total number of vacancies. Then  $vac = E \cdot (1 + extra)$ , where *extra* gives the total number of positions lying vacant prior to the transfers, expressed as a fraction of E. It is possible that *extra* = 0, in which case the chain vacancies are the only vacancies. To avoid problems in implementation arising out of the chain vacancies, we assume that all E transfers take effect at the same instant of time. Let us call each (loc,j) pair a *bucket*. The buckets are numbered 1 through B. The number of vacancies in bucket  $k = (loc,j)$  is  $M_k = vacancy(loc,j)$ .

As already mentioned, constraints are either strict or desirable. There are three strict constraints:

SC1: Each of the N employees must be assigned a new posting. A new posting is a (loc,j) pair that differs from the original (loc,j) pair in at least one position.

SC2: An employee can only be transferred to one of the T options specified by the employee.

SC3: For each (loc,j) pair, the total number of persons transferred to (loc,j) must not exceed vacancy(loc,j).

The problem can be formulated mathematically as follows. An ExTxB matrix  $A = [a_{ijk}]$  is given, where  $a_{ijk} = 1$  if transfer option j of employee i is bucket k, and 0 otherwise. We have to determine the ExT solution matrix  $X = [x_{ij}]$ , where  $x_{ij} = 1$  if employee i is assigned transfer option j, and 0 otherwise. An employee is assigned exactly one new transfer option. The solution is *feasible* if the number of employees transferred to any bucket k does not exceed the number of vacancies  $M_k$ . This condition can be expressed as follows:

$$\sum_i x_{ij} a_{ijk} \leq M_k \text{ for } 1 \leq k \leq B$$

where the summations are over  $1 \leq i \leq E$ ,  $1 \leq j \leq T$ . A feasible solution satisfies all the three strict constraints mentioned above.

There are two desirable constraints:

DC1: Among the T options, an employee should be assigned an option of as high a preference as possible.

DC2: When two employees contend for the same transfer posting, the employee who is senior should get priority over the employee who is junior.

DC1 can be expressed by the condition

$$\sum_i x_{ij} \geq \sum_i x_{ij+1} \text{ for } 1 \leq j < T,$$

where the summation is over  $1 \leq i \leq E$ . This says that more employees should be assigned an option of preference j than an option of preference j+1.

DC2 is satisfied if *seniority violations* can be avoided. Such a violation occurs if there exist employees  $k_1$  and  $k_2$ , where  $k_1 < k_2$  and integers r, s and t,  $1 \leq r, s < t \leq T$ , such that:

- i) the junior employee  $k_2$  is assigned a transfer option (loc,j) that has preference r for  $k_2$ ;
- ii) the senior employee  $k_1$  is assigned an option of preference  $t > r$ ;
- iii)  $k_1$  has specified the same (loc,j) pair assigned to  $k_2$  as an option with preference  $s < t$ .

The condition for no seniority violation, of an employee i assigned option j, by a junior employee q assigned option r is

$$\sum_i \sum_j \sum_k \sum_l \sum_p \sum_q x_{ij} a_{ipk} x_{qr} a_{qrk} = 0$$

where the summations are over  $1 \leq i < q \leq E$ ,  $1 \leq p, r < j \leq T$ , and  $1 \leq k \leq B$ . In explanation, we note that: i) the juniormost employee cannot suffer a seniority violation; ii) it must be the case that  $j > 1$ , for if employee i gets the first transfer option then no seniority violation can occur; and, iii) q must get an option  $r < j$  to violate the seniority of i. Stated in this form, the Staff Transfer Problem is just a CSP and not an optimization problem, since there is no objective function to optimize.

We can view the Staff Transfer Problem, with seniority violations ignored, as a problem in which a complete matching of maximum weight must be found in a weighted bipartite graph [7]. Let an instance of the Staff Transfer Problem be given in which all bucket sizes are

unity. We construct a bipartite graph  $G = (V_1 \cup V_2, E')$ , where  $V_1$  is the set of vertices on the north side,  $V_2$  the set of vertices on the south side, and  $E'$  is the set of edges.  $V_1$  and  $V_2$  are assumed to be disjoint. In our case,  $V_1$  represents the set of employees and  $V_2$  the set of buckets. An edge from a vertex  $v_1$  of  $V_1$  to a vertex  $v_2$  of  $V_2$  indicates that  $v_2$  is a transfer option of employee  $v_1$ . We assign to each edge a positive integer weight that depends on the employee, the bucket, and the preference of the transfer option. This weight reflects the value of this particular option to the organization and to the employee. A transfer option that is a first preference has a higher weight than one that is a second preference, and so on. Since the buckets have unit size, our objective here is to find a matching that is complete for  $V_1$  and maximizes the sum of weights of the selected edges. This problem can be solved in time that is polynomial in the size of the input. Now, suppose there is a bucket that has a size  $k > 1$ . In  $V_2$  we split this big bucket into  $k$  smaller buckets each of unit size. If the big bucket happens to be a transfer option with weight  $w$  for employee  $v_1$ , we join  $v_1$  to each of the  $k$  smaller buckets with edges of weight  $w$ . As a result we are again left with a complete matching problem. Since bucket sizes are constant integers supplied as input, the solution can still be obtained in time that is polynomial in the size of the input.

When seniority violations must be resolved, it becomes much harder to assign transfer options to employees in a satisfactory manner. It is not yet known whether the Staff Transfer Problem becomes NP-complete in this case.

The Staff Transfer Problem models an idealized situation. In practice, an organization might want to impose additional conditions on the transfer options of employees or change some of the assumptions. For example, the number  $T$  of transfer options need not be the same for all employees. The organization could decide how many transfer options an employee would be entitled to submit. Assuming an employee is allowed to submit options freely, it is advantageous to submit as few options as possible, since this means the employee is more likely to get a desired posting. If the organization wants to impose some degree of control over the transfers, it could itself supply one of the transfer options, perhaps the one of highest preference, and the employee could be requested to submit the remaining options. There are many other possibilities.

### III. SOLUTION METHODS

A satisfactory analytical solution procedure for the problem, even in its idealized form, has not yet been found. We are thus forced to use heuristic methods. In our earlier work[2,3,13] it was shown that GSAT(L)

outperformed other heuristic and deterministic methods. Here, we have tried to solve randomly generated problem instances using WalkSAT(L,p). The implementations given below consider only the basic version of the problem consisting of: i) the strict constraints and ii) desirable constraint DC2.

Here, the STP is viewed as a SAT. One way to express the Staff Transfer Problem in terms of logical satisfiability is as follows. Let the  $T$  transfer options of employee  $k$  be  $(loc_{1,j_1}), (loc_{2,j_2}), \dots, (loc_{T,j_T})$ . Then employee  $k$  must be assigned a transfer posting at one of these  $T$  destinations. This can be mathematically expressed by the logical proposition

$$X(k, loc_{1,j_1}) \cup X(k, loc_{2,j_2}) \cup \dots \cup X(k, loc_{T,j_T})$$

where  $X(k, loc_{i,j_i})$ ,  $1 \leq i \leq T$ , are logical variables that correspond to positive literals. The literal  $X(k, loc_{i,j_i})$  is *true* if  $k$  is assigned bucket  $(loc_{i,j_i})$ , and is *false* otherwise. There are as many such propositions as there are employees. We refer to this set of clauses as the first set.

Another set of clauses arise out of the size restrictions on the  $(loc_{i,j_i})$  buckets. Suppose  $m'$  employees  $k_1, k_2, \dots, k_{m'}$ , have all given bucket  $(loc_{i,j_i})$  as the transfer option. If  $m' > M = \text{vacancy}(loc_{i,j_i})$ , bucket  $(loc_{i,j_i})$  could become overfull. To prevent the bucket from becoming overfull, for *every* subset  $\{k'_1, \dots, k'_M, k'_{M+1}\}$  of size  $M+1$  of the set  $\{k_1, k_2, \dots, k_{m'}\}$ , we generate a clause of the form

$$\sim X(k'_1, loc_{i,j_i}) \cup \sim X(k'_2, loc_{i,j_i}) \cup \dots \cup \sim X(k'_{M+1}, loc_{i,j_i})$$

We refer to this set of clauses as the second set. The number of such clauses for this particular bucket is  $C(m', M+1)$ , the number of combinations of  $m'$  objects taken  $(M+1)$  at a time.

Seniority violations can be incorporated in our scheme as follows. Suppose a seniority violation has occurred between employees  $k_1$  and  $k_2$ , where  $k_1 < k_2$ . Then, as stated before, there exist integers  $r, s$  and  $t$ ,  $1 \leq s, r < t \leq T$ , such that: i) the junior employee  $k_2$  is assigned a transfer option  $(loc_{i,j_i})$  which has preference  $r$  for  $k_2$ ; ii) the senior employee  $k_1$  has specified the same  $(loc_{i,j_i})$  pair as an option with preference  $s$ ; iii)  $k_1$  is assigned an option of preference  $t > r$ . We can express this constraint as a clause

$$\sim X(k_1, loc_{i,j_i}) \cup \sim X(k_2, loc_{2,j_2})$$

where option  $t$  corresponds to  $(loc_{i,j_i})$  and option  $r$  corresponds to  $(loc_{2,j_2})$ . To generate all such clause, the transfer options of all the  $E$  employees must be examined and potential seniority violations identified. We refer to this set of clauses as the third set. Many such clauses would be generated in a typical problem instance.

The satisfaction of the first set of clauses ensures that each employee is assigned a new posting. This means that strict constraints SC1 and SC2 are satisfied. The satisfaction of the second set of clauses ensures that there is no overfull bucket, so the strict constraint SC3 is satisfied. The satisfaction of the third set of clauses

ensures that there are no seniority violations, so desirable constraint DC2 is satisfied. No effort was made to satisfy desirable constraint DC1, because this constraint is difficult to state in terms of clauses.

#### A. GSAT(L)

The greedy local search procedure GSAT assigns truth values to variables in an effort to satisfy all the clauses in a given set [11,12]. GSAT looks for the variable with the property that its truth-value when flipped causes the *largest net decrease* in the number of unsatisfied clauses. It flips the truth-value of this variable, and again looks for such a variable. Ties are resolved arbitrarily. This is repeated until a satisfying assignment is found. If no satisfying assignment is found within a specified number of flips (*maxflips*), the procedure is restarted with a new initial truth assignment. This cycle is repeated a specified number of times (*maxtries*).

The performance of GSAT can be significantly improved in another way, namely by incorporating a *tabu search* strategy to ensure that the same variable is not flipped again and again [10]. The tabu list is initially empty and is implemented as a FIFO queue. A variable that has just been flipped is inserted into the list. The variable to be flipped next is selected randomly [6] from among those variables *not* in the tabu list that cause the largest decrease in the number of unsatisfied clauses. As new variables get flipped and enter the tabu list, older entries fall out of the list at the other end. Thus some variables are prevented from being flipped for a limited period of time, determined by the length of the tabu list. In some applications this length plays a critical role in the performance of GSAT. We represent GSAT with a tabu list of length L as GSAT(L). In Procedure GSAT(L), the values of *maxflips* and *maxtries* are selected by trial and error so that good solutions are obtained in reasonable time. GSAT outputs feasible solutions that are *completely free* of seniority violations. Thus the solutions obtained by GSAT(L) are superior to those obtained by SA.

#### B. WalkSAT(L,p)

The algorithm with a tabu list of length L and random walk probability p is denoted by WalkSAT(L,p). The WalkSAT algorithm [11] also starts from a randomly chosen variable assignment but the variable selection method differs from the ones used in GSAT. Here the variable is always selected from a randomly chosen unsatisfied clause. If in the chosen clause, variables can be flipped without unsatisfying other clauses, one of these is randomly chosen. Otherwise, with a given probability p a variable is randomly chosen from the clause, and with probability 1-p a variable is picked that minimizes the number of clauses that are currently satisfied but would

become unsatisfied when the variable is flipped. A tabu list can also be used with WalkSAT

```
Procedure WalkSAT(L,p){
 for (try = 1; try ≤ maxtries; try++) {
 X = TA; /* TA is random truth assignment */
 for (flip = 1; flip ≤ maxflips; flip++) {
 if (X satisfies all clauses) return X;
 }
 else {
 randomly select an unsatisfied clause; determine
 the set of variables in the selected clause, not
 present in the tabu list, flipping any of which
 does not unsatisfy a satisfied clause;
 if (the set is non-empty)
 arbitrarily select a variable from this set;
 else {
 with probability p, pick a variable randomly
 from the selected clause; with probability 1-p,
 pick a variable from the selected clause, not in
 the tabu list, flipping which unsatisfies the
 minimum number of satisfied clauses;
 }
 }
 flip the selected variable and insert into the list;
 }
 announce failure;
}
```

## IV. EXPERIMENTAL OBSERVATIONS

We now summarize our experimental observations on the Staff Transfer Problem. The methods were programmed in C in Linux Environment, and run on Pentium-IV, 1.8 GHz, 1024MB RAM Machine. Identical problem instances were run for the two methods. We have not run other methods because they have been proved to be inferior as compared to the Simulated Annealing [3]. We wanted to create random instances of the STP that were realistic and indicative of real life situations. So we kept the number E of employees to be transferred between 500 and 2000. Specialists in Human Resource Management informed us that more than 2000 employees are rarely transferred by any organization in one lot. For all runs we took T = 3. Initially, the E employees to be transferred were randomly assigned to buckets. This information was used for computing chain vacancies. The value of parameter *extra* was chosen to lie between 0% and 40% of E; it was felt that a higher vacancy rate would be unrealistic. The parameter *extra* is calculated as  $extra = MPOST * loc * j * 0.5 * 100 / E$ , where MPOST is an integer for generating extra vacancies randomly. Additional vacant positions were created for each value of *extra* and randomly assigned to buckets. For each employee, T

transfer options were also randomly created. Transfer options to buckets having no vacancies were not permitted. No restrictions were imposed on which jobs an employee could perform, it being felt that such restrictions were unlikely to affect the runtime. 100 problems were generated for each set. We determined the number of problems solved in a set and the average runtime in seconds. We also computed, per problem instance, the average numbers of employees who were assigned their first, second and third transfer options. The averages were taken over solved instances. The methods were compared on the basis of three criteria: i) the number of problems solved in each set of 100; ii) the runtime averaged over solved problems, iii) the quality of the solution obtained, quality being determined by the extent to which desirable constraints were satisfied.

Our observations on the experiments are as follows:

The Staff Transfer Problem is a difficult problem. The run times are high when E is large. In the earlier implementations [2,3,13], the GSAT(L) was the best method. But, in this paper, we have shown that WalkSAT(L,p) has outperformed the GSAT (L) significantly. In Table 1, the comparative performance of WalkSAT(L,p) and GSAT(L) is shown. We see that in most of the cases WalkSAT(L,p) performs better in solving the more number of instances as well as in runtime. Of course, not every randomly generated problem instance has a feasible solution. When *extra* = 0, it is quite possible that the problems may not have feasible solutions. From our experimental results it appears that this is very likely the case. The problem instances were all generated randomly. If the transfer scheme described in this paper is used in real life, the transfer options submitted by employees could exhibit bias in favor of certain postings and against some other postings. This

TABLE I  
COMPARISON OF THE PERFORMANCE BETWEEN GSAT AND WALKSAT

| E    | loc,j | T | Extra (%) | Method  | Prob | Solved | 1 <sup>st</sup> opt | 2 <sup>nd</sup> opt | 3 <sup>rd</sup> opt | Time (s) |
|------|-------|---|-----------|---------|------|--------|---------------------|---------------------|---------------------|----------|
| 500  | 15,10 | 3 | 15        | GSAT    | 100  | 69     | 209                 | 172                 | 119                 | 12.12    |
|      |       |   |           | WalkSAT |      | 74     | 192                 | 175                 | 133                 | 9.19     |
| 600  | 20,10 | 3 | 16.6      | GSAT    | 100  | 55     | 243                 | 206                 | 151                 | 16.02    |
|      |       |   |           | WalkSAT |      | 77     | 234                 | 210                 | 156                 | 12.40    |
| 700  | 20,10 | 3 | 13.7      | GSAT    | 100  | 70     | 289                 | 242                 | 169                 | 24.53    |
|      |       |   |           | WalkSAT |      | 77     | 273                 | 242                 | 185                 | 22.75    |
| 750  | 25,10 | 3 | 16.6      | GSAT    | 100  | 52     | 303                 | 261                 | 186                 | 12.65    |
|      |       |   |           | WalkSAT |      | 69     | 291                 | 260                 | 199                 | 21.84    |
| 900  | 30,10 | 3 | 16.6      | GSAT    | 100  | 54     | 362                 | 315                 | 223                 | 42.03    |
|      |       |   |           | WalkSAT |      | 67     | 345                 | 316                 | 240                 | 25.61    |
| 1000 | 30,10 | 3 | 15        | GSAT    | 100  | 47     | 414                 | 342                 | 244                 | 38.89    |
|      |       |   |           | WalkSAT |      | 67     | 390                 | 348                 | 262                 | 44.82    |
| 1200 | 40,10 | 3 | 16.6      | GSAT    | 100  | 22     | 487                 | 413                 | 300                 | 58.72    |
|      |       |   |           | WalkSAT |      | 49     | 469                 | 418                 | 313                 | 50.25    |
| 1250 | 40,10 | 3 | 16        | GSAT    | 100  | 23     | 520                 | 427                 | 303                 | 112.97   |
|      |       |   |           | WalkSAT |      | 47     | 487                 | 434                 | 329                 | 133.28   |
| 1500 | 40,10 | 3 | 13.3      | GSAT    | 100  | 19     | 624                 | 507                 | 369                 | 298      |
|      |       |   |           | WalkSAT |      | 32     | 590                 | 517                 | 393                 | 123.11   |

would make the data less random, and the problems could become harder to solve using randomized CSP techniques. One way to resolve this difficulty might be the following. The organization could generate an extra transfer option for each employee, and this could be the option of least preference. If this last option is allocated properly among employees, an acceptable solution would always be found.

## V. CONCLUSION

In this paper, the Staff Transfer Problem is viewed as a CSP. It is converted to a SAT and the experiments indicate that WalkSAT(L,p) has solved this problem very efficiently. In most of the instances it has easily outperformed GSAT(L) which was the best among earlier implementations. But SAT techniques have a limitation with the size of the problem. The number of clauses increases with the increase in size of the problem. In the present machine with 1 GB RAM we could not run instances having more than 1500 employees.

Some interesting issues remain open. The first concerns the formulation of an analytic solution procedure for the STP. If the problem can be suitably expressed in the language of Mathematical Programming, we can compare the runtime and the quality of solution obtained by an analytical procedure with that obtained by WalkSAT(L,p). The second relates to the NP-completeness of the Staff Transfer Problem when seniority violations are taken into account. The difficulty level of the problem suggests that it is NP-complete, but this has not yet been proved formally.

## REFERENCES

- [1] S Acharyya, "The Satisfiability Problem: A Constraint Satisfaction Approach", Ph D Thesis, Computer Science & Engg., Calcutta University, 2001
- [2] S Acharyya, A Bagchi, "Staff Transfers in a Large Organization: A Constraint Satisfaction Approach, *Proc KBCS-98*, International Conference on Knowledge-Based Computer Systems, Mumbai, India, pp 51-63, 1998
- [3] S Acharyya, A Bagchi, "Constraint Satisfaction Methods for Solving the Staff Transfer Problem", *OPSEARCH*, Vol 42, No 3, September 2005, pp 179-198, 2005
- [4] M Beer *et al*, *Human Resource Management*, A General Manager's Perspective, Text & Cases, The Free Press, 1985
- [5] G Dessler, *Human Resource Management*(17<sup>th</sup> Ed), Prentice Hall India, 1997
- [6] A. S. Fukunaga, "Variable Selection Heuristics in Local Search for SAT", *Proc AAAI-97*, pp 275-280, 1997
- [7] C H Papadimitriou & K Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice Hall, 1982
- [8] D. S. Johnson, C. R. Aragon, L. A. McGeoch & C. Schevon, "Optimization by Simulated Annealing: An Experimental Evaluation, Part II, Graph Colouring and Number Partitioning", *Operations Research*, Vol 39, pp 378-406, 1991
- [9] C R Reeves, *Modern Heuristic Techniques for Combinatorial Problems*, Orient Longman, 1993
- [10] B. Mazure, L. Sais & E. Gregoire, "Tabu Search For SAT", *Proc AAAI-97*, pp 281-285, 1997
- [11] B. Selman, H. Kautz H & B. Cohen, "Noise Strategies for Improving Local Search", *Proc AAAI-94*, pp 337-343, 1994
- [12] B. Selman, H. J. Levesque & D. J. Mitchell, "A New Method for Solving Hard Satisfiability Problems", *Proc AAAI-92*, pp 440-446, 1992
- [13] S. Acharyya, A. Bagchi, "A SAT Approach for solving the Staff Transfer Problem", *Proc IMECS-08*, Vol. 1, pp 64-68, 2008.

# Simulation and Bit Error Rate Performance Analysis of 4G OFDM Systems

Md. Sadek Ali, Md. Mahbubur Rahman, Md. Delowar Hossain, Md. Shariful Islam, Md. Moinul Islam

Communication Research Laboratory, Dept. of Information & Communication Engineering, Islamic University, Kushtia-7003, Bangladesh. E-mail: (sadek\_ice, drmahbub\_07, shohelictiu2005, s\_ice\_iu, mmislam\_ice@yahoo.com)

**Abstract** — In this paper a simulation based system developed for analysis and Bit Error Rate (BER) performance measurement of Orthogonal Frequency Division Multiplexing (OFDM) systems. This paper involves the basics of OFDM system. In this paper bit error rate is calculated and measured with respect to Signal to Noise Ratio (SNR), Doppler Effect, and guard interval. BPSK, QPSK and 16PSK are used as modulation techniques. Additive White Gaussian Noise (AWGN) is used as a communication channel. The effect of SNR, Doppler Effect, and guard intervals on OFDM signals improves the system performance.

**Index Terms** — AWGN, Bit Error Rate (BER), Doppler Effect, Guard interval, OFDM, Signal-to-Noise Ratio (SNR).

## I. INTRODUCTION

The goal for the next generation of mobile communications system is to seamlessly integrate a wide variety of communication services such as high speed data, video and multimedia traffic as well as voice signals. The technology needed to tackle the challenges to make these services available is popularly known as the Fourth Generation (4G) OFDM System. Although the principles of Orthogonal Frequency Division Multiplexing (OFDM) modulation [1] [2] have been in existence since 1960, in the last few years OFDM modulation is emerged as a key modulation technique of commercial high speed communication systems. The principal reason of this increasing interest is due to its capability to provide high-speed data rate transmissions with low complexity and to counteract the intersymbol interference (ISI) introduced by dispersive channels. For this reason OFDM modulation has been adopted by several digital wireline and wireless communication standards, such as the European digital audio and video broadcasting standards, as well as local area networks.

The rest of the paper is organized as follows. In section II shows the literature review. Section III describes the mathematical tools and techniques for OFDM systems. Section IV describes OFDM system design. Simulation results are described in section V. Finally, section VI concludes the paper.

The main idea behind OFDM is the so-called Multi-Carrier Modulation (MCM) transmission technique. MCM is the principle of transmitting data by dividing the input bit stream into several parallel bit streams, each of them having a much lower bit rate, and by using these sub-streams to modulate several carriers. The first systems using MCM were military high frequency radio links in the late 1950s and early 1960s, like Kineplex, Andeft and Kathryn. OFDM is a special form of MCM with densely spaced sub-carriers and overlapping spectra, whose main idea was patented by Chang, from the Bell Labs, in 1966. OFDM abandoned the use of steep bandpass filters that completely separated the spectrum of individual sub-carriers, as it was common practice in older analogue FDMA systems. Instead, OFDM time domain waveforms are chosen such that mutual orthogonality is ensured even though carrier spectra may overlap. Orthogonality is achieved by performing a Fourier Transform (or equivalently a Fast Fourier Transform) on the input stream.

Implementation aspects have delayed the introduction of OFDM into real systems. For instance, the complexity of a real-time FFT requires powerful digital signal processors, which were not available at the time when OFDM was invented. Furthermore, in order to keep the orthogonality between the different subcarriers, highly stable oscillators are required in the transmitter and receiver, as well as linear power amplifiers. Now a days many of the implementation problems appear solvable and OFDM has gained a big interest since the beginning of the 1990s [3]. This interest is mainly motivated by the advantages OFDM offers when transmitting through a fading channel.

## II. MATHEMATICAL TOOLS AND TECHNIQUES FOR OFDM SYSTEMS

OFDM is a special case of multicarrier transmission, where a single data stream is transmitted over a number of lower rate subcarriers. OFDM is a combination of modulation and multiplexing. Multiplexing generally refers to independent signals, those produced by different sources. In OFDM, multiplexing is applied to independent signals but these independent signals are a sub-set of the one main signal. The signal itself is first split into independent

channels, modulated by data and then re-multiplexed to create the OFDM carrier.

In a 1971 paper [4], Weinstein and Ebert suggested that the multicarrier modulated signal is effectively the Fourier transform of the serial data stream and the bank of coherent demodulators the inverse Fourier transform. In this paper the following is the mathematical derivation used to show the validity of the argument:

Consider a data sequence  $(d_0, d_1, \dots, d_{N-1})$ , where each  $d_n$  is a complex number  $d_n = a_n + jb_n$ . The result of a DFT operation on the vector  $\{2d_n\}_{n=0}^{N-1}$  is then the vector  $S = (S_0, S_1, \dots, S_{N-1})$  of  $N$  complex numbers, with

$$S_m = \sum_{n=0}^{N-1} 2d_n e^{-j(2\pi mn/N)} = 2 \sum_{n=0}^{N-1} d_n e^{-j2\pi f_n t_m}, m = 0, 1, \dots, N-1 \quad (1)$$

Where

$$f_n \triangleq \frac{n}{N\Delta t} \quad (2)$$

$$t_m \triangleq m\Delta t \quad (3)$$

and  $\Delta t$  is an arbitrarily chosen interval. The real part of  $S$  has components:

$$Y_m = 2 \sum_{n=0}^{N-1} (a_n \cos 2\pi f_n t_m + b_n \sin 2\pi f_n t_m), m = 0, 1, \dots, N-1 \quad (4)$$

If these components are applied to a low pass filter at time intervals  $\Delta t$ , a signal is obtained that closely approximates the frequency division multiplexed signal as

$$y(t) = 2 \sum_{n=0}^{N-1} (a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t), 0 \leq t \leq N\Delta t \quad (5)$$

In order to recover the modulated data, a DFT with twice the sampling rate is employed. This is necessary since only the real part of the modulated signal is transmitted. Therefore, the DFT operates on the  $2N$  samples

$$Y_k = y\left(k \frac{\Delta t}{2}\right) = 2 \sum_{n=0}^{N-1} \left( a_n \cos \frac{2\pi nk}{2N} + b_n \sin \frac{2\pi nk}{2N} \right), k = 0, 1, \dots, 2N-1 \quad (6)$$

The result of the DFT operation is then

$$z_l = \frac{1}{2N} \sum_{k=0}^{2N-1} Y_k e^{-j(2\pi lk/2N)} = \begin{cases} 2a_0, & l=0 \\ a_l - jb_l, & l=1, 2, \dots, N-1 \\ \text{irrelevant}, & l > N-1 \end{cases} \quad (7)$$

The original data  $a_l$  and  $b_l$  can then be extracted as the real and imaginary parts of  $z_l$  (except at  $l=0$ ) [5]. Since the sinusoidal components of the parallel input are time limited,

they have a  $\left[\frac{\sin f}{f}\right]^2$  - shaped power spectrum. This special shape ensures that as long as the components are samples at the right instance, the neighboring components have zero contribution. This orthogonal nature of the OFDM symbols helps prevent ICI.

In this paper, we also proposed some simple equalization schemes and studied the effect of linear channel distortions.

### III. OFDM SYSTEM DESIGN

For performing the simulations, the chain shown in Figure 1 was developed under MATLAB environment. This paper is mainly dependent on this system model. This model has implemented the basic characteristics of OFDM systems. MATLAB environment has assisted to implement that purposes.

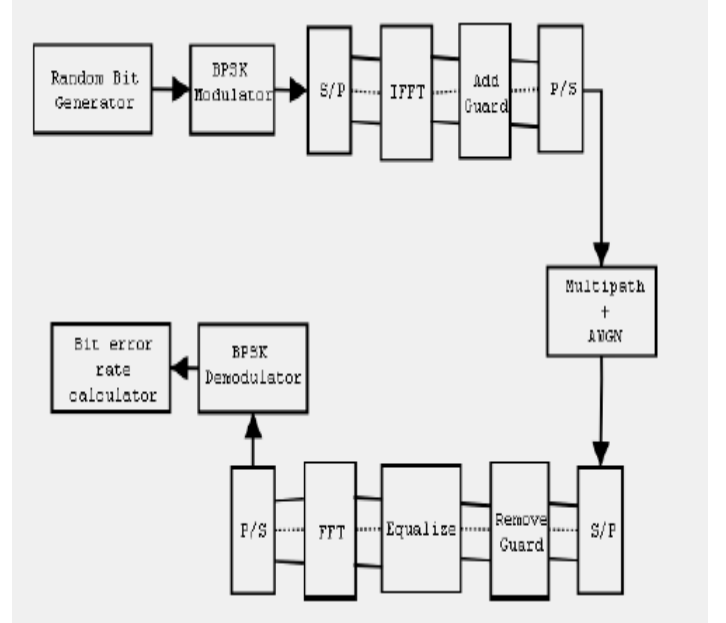


Fig. 1. Block diagram of the system implemented.

#### A. Algorithm

The following is a general overview on the operation of the system.

1. Generate a binary message with a length divisible by the number of subcarriers
2. Modulate the signal using BPSK. This is done quite easily by the following operation  $\text{out} = \text{msg} \cdot 2 - 1$
3. Perform Serial-to-Parallel (S/P) conversion using the MATLAB function `reshapes`
4. Perform the Inverse FFT

5. Add a cyclic extension to each symbol. The length of the extension to add in us is defined by the input parameter guard length
6. Perform Parallel-to-Serial (P/S) conversion to make the signal ready for transmission
7. Pass it through the Rayleigh fading multipath channel with Additive White Gaussian Noise. An estimate of the channel at a rate defined by estRate is also performed simultaneously
8. Perform S/P conversion
9. Remove the guard cyclic prefix
10. Using the channel estimate obtained, perform equalization of the received signal
11. Perform the FFT to recover the signal
12. Perform P/S conversion
13. Demodulate the BPSK symbols using a decision boundary of 0: demod = double (real (RXmsg) >= 0)
14. Calculate the bit error rate of the system

#### IV. SIMULATION RESULTS

The effect of SNR, Doppler Effect, and guard intervals on OFDM signals are simulated in this paper. The simulation results are given below:

##### A. Simulation Results of BER Performance vs. SNR

At first we generate a random signal using a random bit generator. The signal is modulated using different modulation techniques such as BPSK, QPSK and 16PSK. Then the signal is passed away through AWGN channel. The signal is demodulated and checked the errors. The simulation is dependent on signal to noise ratio (SNR). Here SNR used as average symbol-energy-to-noise ratio. Different values of BER are obtained from the simulation graph with respect to different values of SNR. The simulations are performed using sub carriers 100 and bit rate 100bps.

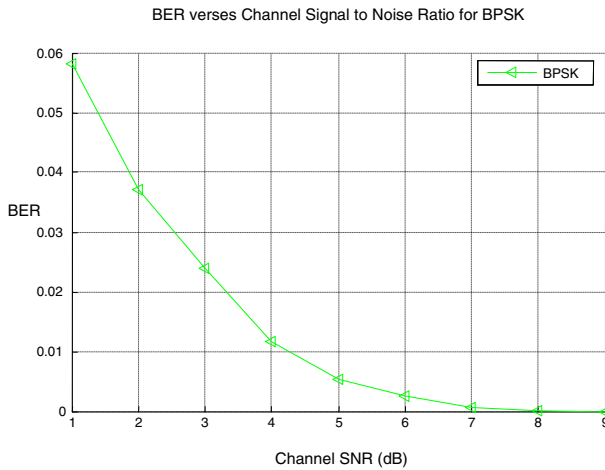


Fig. 2. BER vs. Channel SNR for BPSK.

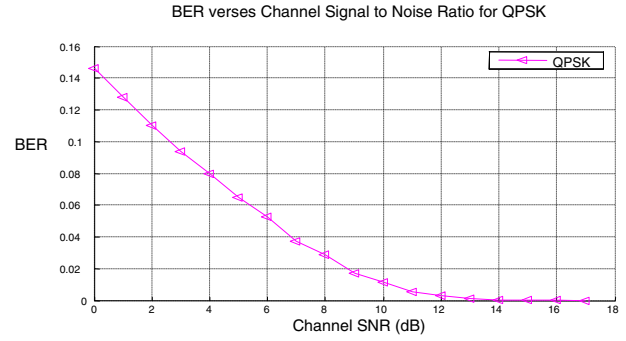


Fig. 3. BER vs. Channel SNR for QPSK.

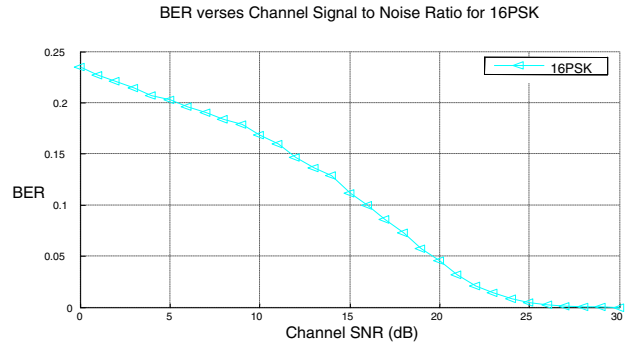


Fig. 4. BER vs. Channel SNR for 16PSK.

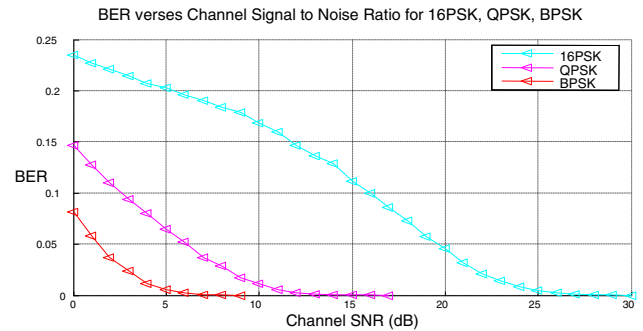


Fig. 5. BER vs. Channel SNR for 16PSK, QPSK and BPSK.

From the simulation results we can show that the transmission can tolerate a SNR > 28 dB using 16PSK. QPSK transmission can tolerate a SNR of > 15-17 dB. BPSK transmission can tolerate a SNR of > 7-10 dB. Among these modulation techniques 16 PSK can tolerate highly than BPSK and QPSK. Bit error performance can be calculated and measured approximately using 16PSK modulation techniques.

##### B. Simulation Results of Guard Length on the BER Performance

In order to analysis the effect of a cyclic prefix, the maximum expected channel delay spread was fixed at 5μs and the length of the guard prefix was varied as 1μs, 4μs, and 20μs. The multipath simulation was carried out with 2



different paths and the maximum Doppler spread used was fixed at 1 kHz. A plot of the relative performance of the different lengths is shown in following figures.

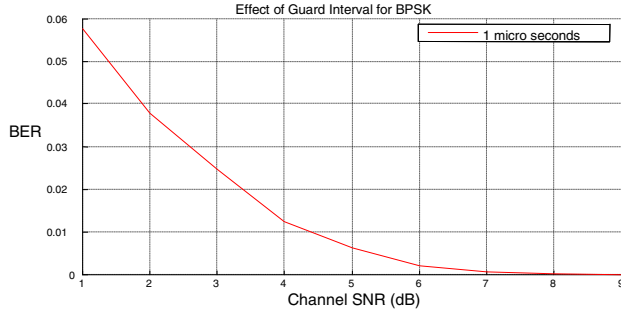


Fig. 6. Effect of guard length on the BER performance for 1 micro seconds for BPSK.

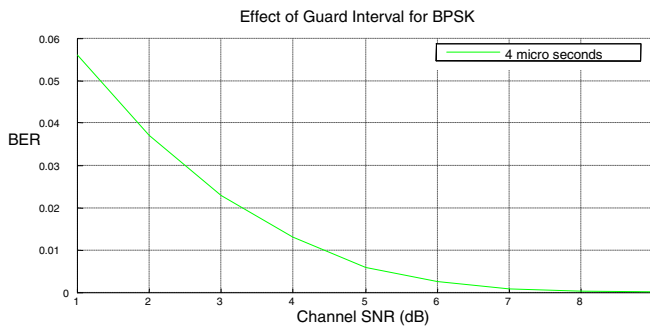


Fig. 7. Effect of guard length on the BER performance for 4 micro seconds for BPSK.

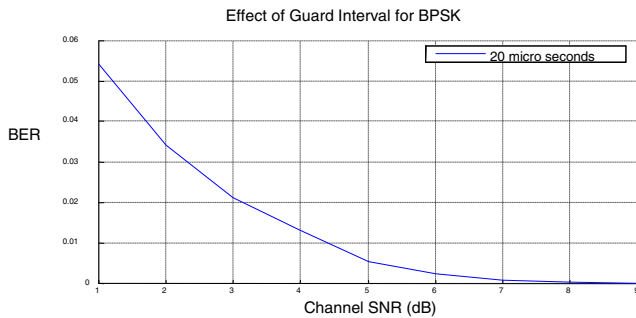


Fig. 8. Effect of guard length on the BER performance for 20 micro seconds for BPSK.

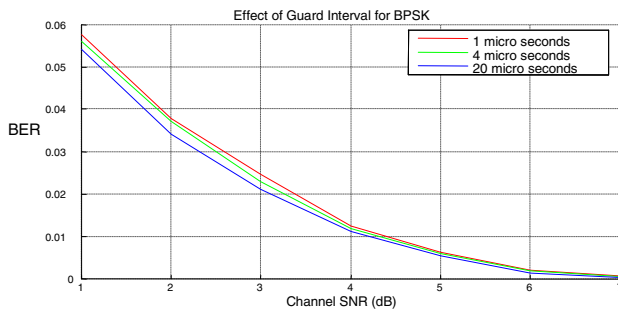


Fig. 9. Effect of guard length on the BER performance for 1, 4 and 20 micro seconds for BPSK.

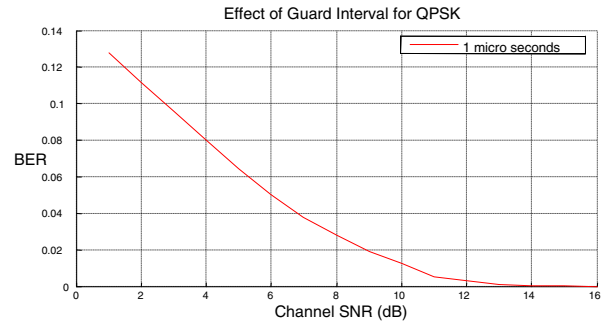


Fig. 10. Effect of guard length on the BER performance for 1 micro seconds for QPSK.

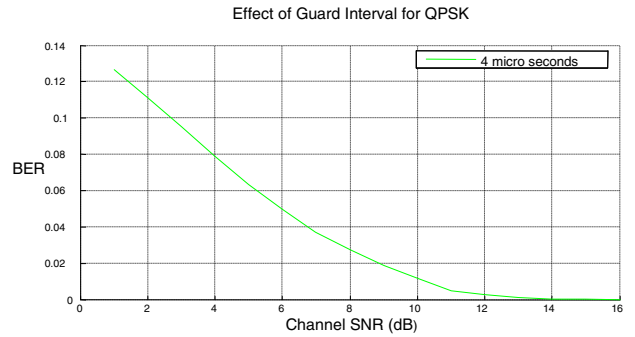


Fig. 11. Effect of guard length on the BER performance for 4 micro seconds for QPSK.

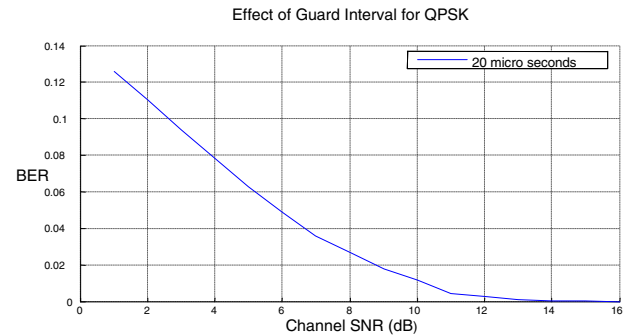


Fig. 12. Effect of guard length on the BER performance for 20 micro seconds for QPSK.

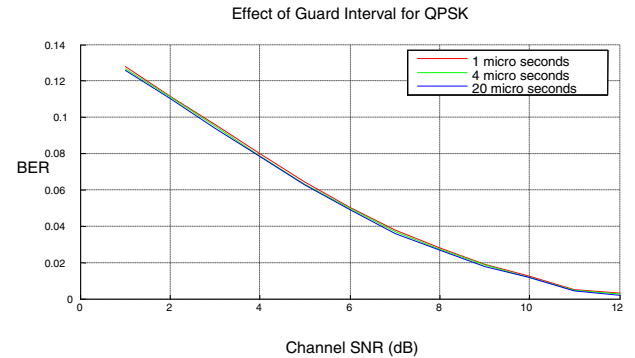


Fig. 13. Effect of guard length on the BER performance for 1, 4 and 20 micro seconds for QPSK.

As expected when the length of the cyclic guard prefix is less than the expected  $5\mu\text{s}$  of the channel, the OFDM symbols undergo ISI and the bit error rate of the system goes up. The trend in the plot is also clear; as long as the guard length is greater than or equal to  $5\mu\text{s}$ , the bit error rate is low. When the guard is made a little shorter than needed, i.e.  $4\mu\text{s}$ , the bit error rate increases. When it is further reduced to just  $1\mu\text{s}$ , there is severe ISI and hence the bit error rate is much higher.

Finally, it can be seen that for both BPSK and QPSK modulation techniques, if guard length is reducing the bit error rate is gradually increasing and vice-versa.

### C. Simulation Results of Maximum Doppler Effect on the BER Performance

In this case of simulation we consider various channel estimation rates such as  $3^{\text{rd}}$ ,  $5^{\text{th}}$  and  $10^{\text{th}}$ . Here 2 kHz is the maximum Doppler spread. Once again the channel is simulated considering only two different paths to the receiver. A cyclic prefix of  $5\mu\text{s}$  is used to keep ISI from affecting the results. The plots obtained are shown in the following figures.

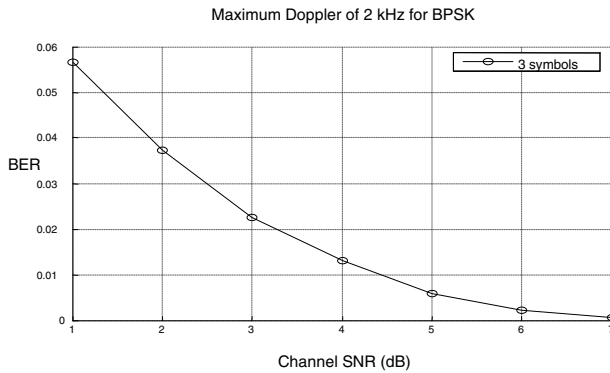


Fig. 14. Effect of estimation rates on the BER performance for 3 symbols.

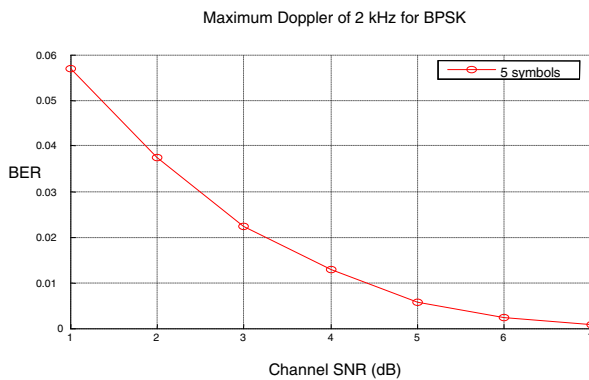


Fig. 15. Effect of estimation rates on the BER performance for 5 symbols Using BPSK

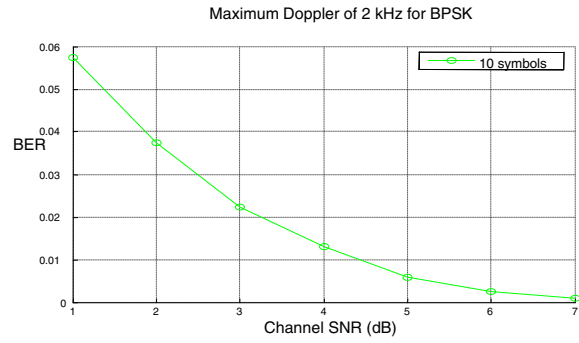


Fig. 16. Effect of estimation rates on the BER performance for 10 symbols Using BPSK.

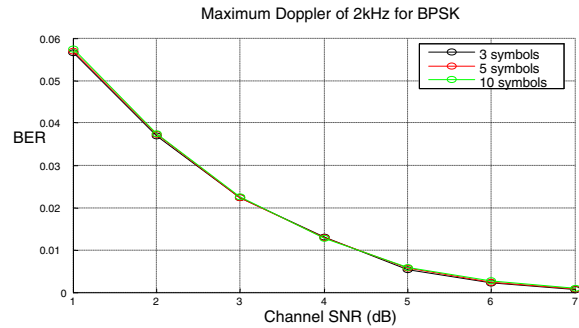


Fig. 17. Effect of estimation rates on the BER performance for 3, 5, and 10 symbols Using BPSK.

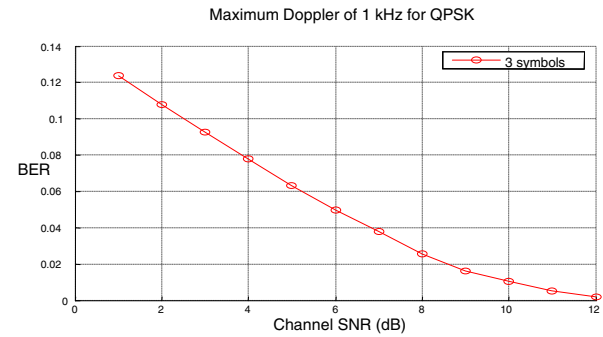


Fig. 18. Effect of estimation rates on the BER performance for 3 symbols using QPSK.

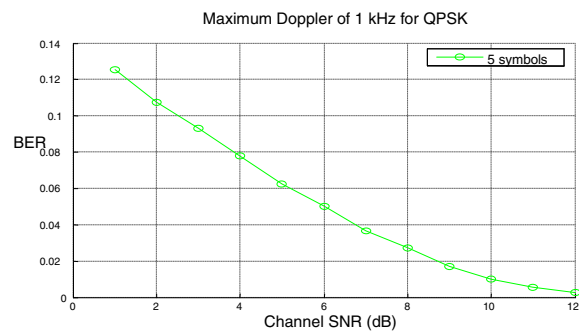


Fig. 19. Effect of estimation rates on the BER performance for 5 symbols using QPSK.

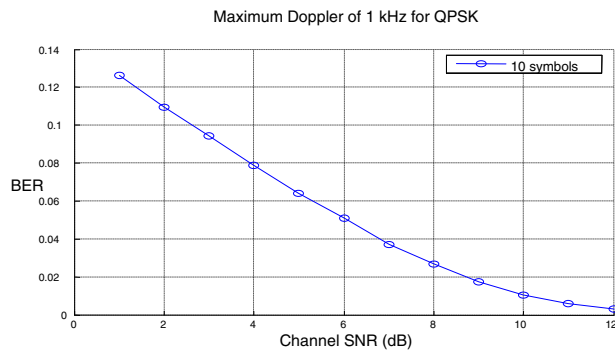


Fig. 20. Effect of estimation rates on the BER performance for 10 symbols using QPSK.

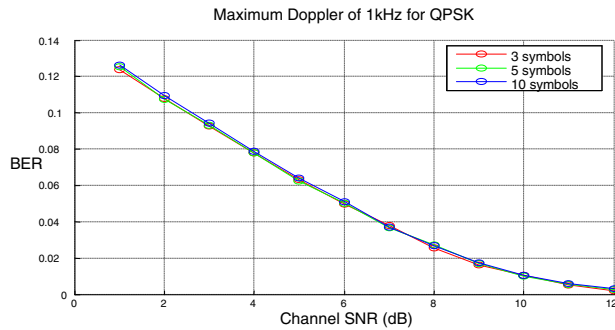


Fig. 21. Effect of estimation rates on the BER performance for 3,5,10 symbols using QPSK.

Here we have used two modulation techniques such as BPSK and QPSK. From this particular simulation, it is possible to determine a requirement on the minimum estimation rate for a given maximum Doppler spread. Comparing the three symbol plots, it can be seen that if symbols are increasing, bit error rate is gradually increasing and vice-versa. It is clear for both modulation techniques.

## V. CONCLUSION

Inter Symbol Interference (ISI) and Inter Carrier Interference (ICI) are dominant problems in OFDM systems. These problems are solved with the help of guard period/cyclic prefix. In the case of BER vs. SNR when we are increasing signal to noise ratio (SNR), bit error rate (BER) is decreasing in the same way. At that time bit error rate is zero. We use three modulation techniques such as BPSK, QPSK and 16PSK. BER depends on sub-carriers and symbol time. The minimum signal to noise ratio (SNR) required for BPSK is 7dB, 12dB for QPSK and 26dB for 16PSK. In case of Guard length on the BER performance when we are increasing guard length, bit error rate is reducing in the same way. At that time bit error rate is zero. In case of Doppler Effect on the BER performance when we are increasing number of symbols, bit error rate is increasing in the same way.

Finally, from the analysis of the system simulations, it is concluded that a cyclic prefix at least as long as the maximum multipath delay spread is required for complete protection against ISI effects. Also during low velocity and hence Doppler spread conditions it is possible to have lower channel estimation rate for improved data throughput efficiency.

## ACKNOWLEDGEMENT

I am most grateful to my all assistant. Special thanks should go to IEEE Bangladesh section and organizing committee of ICCIT-2008 for sending necessary information.

## REFERENCES

- [1] R. van Nee and R. Prasad, "OFDM for wireless multimedia communications," *Boston, MA: Artech House*, 2000.
- [2] A. R. S. Bahai and B. R. Saltzberg, "Multi-carrier digital communications. theory and applications of OFDM," *Kluwer Academic Publishers*, 1999.
- [3] John A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come", *IEEE Communications Magazine*, pp. 5 - 14, May 1990.
- [4] S. B. Weinstein, Paul M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform", *IEEE Transactions on Communication Technology*, Vol. COM-19, No.5, October 1971, pp.628 – 634.
- [5] J. Armstrong. "New OFDM peak-to-average power reduction scheme". In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 1, pages 756–760, May 2001.

# Lightpath Reconfigurations in IP-over-CWDM Networks with Stackable ROADMs

MD. Nooruzzaman, Yuichi Harada, Michio Hashimoto, Osanori Koyama, and Yutaka Katsuyama

Graduate School of Engineering, Osaka Prefecture University,  
Gakuen-cho 1-1, Naka-ku, Sakai, Osaka, 599-8531 Japan

**Abstract** — Stackable ROADMs (S-ROADMs) have been proposed for use in regional IP-over-CWDM networks. The S-ROADM can be constructed by connecting modules with different wavelengths required in the node. The experimental results clarified that the S-ROADM could mux and demux the wavelengths successfully, and gave no limit to the passing-through wavelengths, making the network be wavelength transparent. Contrary to the S-ROADM, the existing fixed ROADMs were not wavelength transparent. The bidirectional S-ROADMs have also been proposed, and lightpaths could be reconfigured successfully in an IP-over-CWDM network with the bidirectional S-ROADMs.

**Index Terms** – CWDM, Stackable ROADM, Wavelength Transparency, Lightpath Reconfiguration

## I. INTRODUCTION

Optical networks have been introduced commonly to satisfy the rapid increase of traffic demands. Optical networks are also expected to provide effective broadband communication in optical LAN (Local Area Network) for business use, campus networks, and factory networks to meet different application requirements. The CWDM (Coarse Wavelength-Division Multiplexing) technologies are effective for such applications, because no wavelength control is necessary and the devices are low cost. Such networks have to be operated and maintained by the network administrators. One of the important tasks by the administrators is lightpath reconfigurations [1,2] to avoid traffic congestion, depending on the demand changes.

For the reconfiguration purpose, reconfigurable optical add/drop Multiplexers (ROADMs) were proposed [1] and investigated. Lightpaths are reconfigured by ROADMs consisting of WDM devices and wavelength-tunable transceivers (SWs) [1]. The ROADMs with SWs are simple and made of low-cost devices, leading to a promising candidate for such applications as optical LAN and campus networks. We investigated the ROADMs with SWs, and it was clarified that the ROADMs could reconfigure lightpaths successfully in IP (Internet Protocol) networks. However, the ROADMs have some problems: i.e. the ROADMs are not wavelength transparent, restricting network scalability and flexibility. We propose stackable ROADMs to solve the problems, which are described in this paper together with the performance evaluated by experiments.

## II. PROPOSAL OF STACKABLE ROADM

Figure 1 shows a basic structure of the ROADM developed to reconfigure 4 wavelengths  $\lambda_1 - \lambda_4$  in our experiment. The ROADM could reconfigure lightpaths successfully in CWDM networks [4]. The ROADMs were controlled by receiving control packets via the network interface. However, the ROADM had some problems. One of the problems is that the coupler/splitter should mux/demux more wavelengths than the 4 wavelengths  $\lambda_1 - \lambda_4$  of the ROADM, if other wavelengths are used in the network: e.g. the coupler/splitter should mux/demux 8 wavelengths, if wavelengths  $\lambda_5 - \lambda_8$  are used in other ROADMs. The 4 wavelengths  $\lambda_1 - \lambda_4$  can be reconfigured in this ROADM, and wavelengths  $\lambda_5 - \lambda_8$  pass through this ROADM. The second problem is the limited wavelengths in the network: i.e. the 4 wavelengths  $\lambda_1 - \lambda_4$  can be reconfigured by the ROADM, and other 4 wavelengths can be used in other ROADMs. But no other wavelengths can be used, if one more wavelength, e.g.  $\lambda_9$  is required in the network. The coupler/splitter must be replaced. Therefore, we call this type of ROADM fixed ROADM (F-ROADM). To distinguish the type of the wavelengths, we call the 8 wavelengths  $\lambda_1 - \lambda_8$  working wavelengths (W- $\lambda$ s) in the network, and call such 4 wavelengths as  $\lambda_1 - \lambda_4$  reconfigurable wavelengths (R- $\lambda$ s) of the ROADM in Fig.1. And we call such 4 wavelengths as  $\lambda_5 - \lambda_8$  passing-through wavelengths (PT- $\lambda$ s) of the ROADM. It is possible to have some wavelengths of the PT- $\lambda$ s to be R- $\lambda$ s in the ROADM, if the ROADM has reserve SWs, and

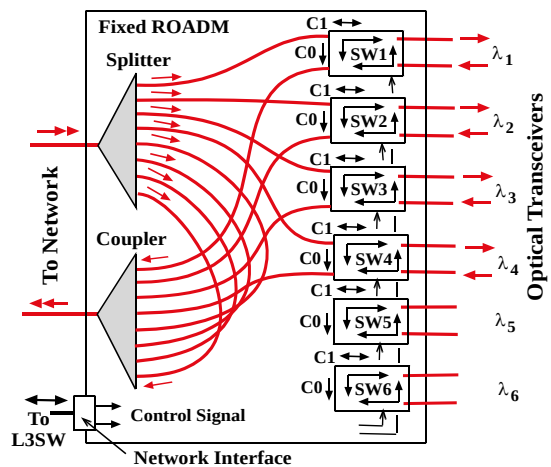


Fig.1 Structure of F-ROADM

the fibers connecting the coupler and splitter are re-connected so as to connect the coupler or splitter and the SWs. Two SWs are reserved in Fig.1, and therefore, 2 of the PT- $\lambda$ s, e.g.  $\lambda_5$  and  $\lambda_6$  can be used as R- $\lambda$ s in this F-ROADM. We call such wavelengths as  $\lambda_5$  and  $\lambda_6$  reserve wavelengths (RV- $\lambda$ s). As a result, the W- $\lambda$ s are  $\lambda_1 - \lambda_8$  in the network, where the F-ROADM shown in Fig.1 is used, but no other wavelength, e.g.  $\lambda_9$  can be used in the network. Thus, the F-ROADMs restrict the wavelength transparency of the network, giving the limitation to the network scalability. The R- $\lambda$ s of the F-ROADM are  $\lambda_1 - \lambda_6$ , including 2 RV- $\lambda$ s, but no other wavelengths can be reconfigured by this F-ROADM. Thus, the F-ROADMs restrict the lightpath reconfigurability in the node.

To remove all the problems of the F-ROADMs, we propose a stackable ROADM (S-ROADM), shown in Fig.2. The S-ROADM can be constructed by connecting stackable modules with different wavelengths. The stackable module consists of an OADM connected by a 2x2 optical SW. The SW has 2 states C1 and C0, and works in the same way as in the F-ROADM. When muxed lightwaves are input to a common input (Cin) port of the S-ROADM, only  $\lambda_1$  passes the OADM1 in the first module, and other wavelengths are reflected. When SW1 state is C1, the passed  $\lambda_1$  is dropped, and received by a  $\lambda_1$  optical transceiver ( $\lambda_1$ -OTR). The  $\lambda_1$  lightwave from the  $\lambda_1$ -OTR is added by SW1, muxed by OADM1, sent to the following modules, reflected by the modules, and sent to the network. When SW1 state is C0 in the first module,  $\lambda_1$  is sent back to the network: i.e.  $\lambda_1$  passes through this S-ROADM. Thus, one module has the reconfigurable function for the specified wavelength. The reflected lightwaves other than  $\lambda_1$  at the first module are input to the second module, and  $\lambda_2$  is added/dropped (A/D) or passes through (PT) the module, depending on the state of SW2.

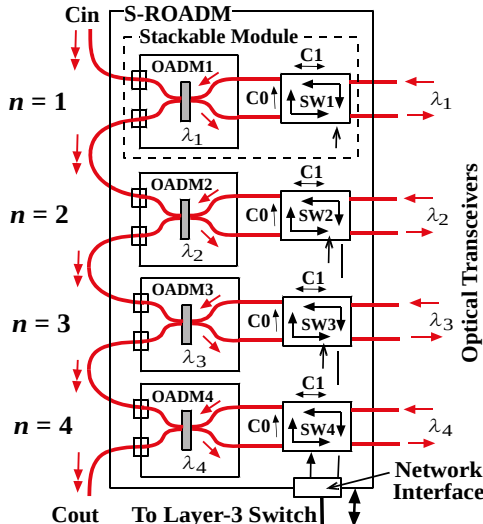


Fig.2 Proposed S-ROADM

Thus, each wavelength can be set to be A/D or PT at the module independently, and gives no influence on the other wavelengths. Therefore, S-ROADM can be constructed in a stackable way by connecting the modules with as many different wavelengths as required.

### III. PERFORMANCE OF STACKABLE ROADM

The loss properties of S-ROADM and F-ROADM were investigated and compared. Firstly, loss estimation equations were derived as follows. Three kinds of loss affect the transmission between nodes: i.e. the add, drop and PT losses denoted by  $L_{ad}$ ,  $L_{dp}$  and  $L_{PT}$ , respectively. The losses depend on the position of the module. When a module is located in the  $n$ -th position from the Cin port in the S-ROADM, as shown in Fig.2, the losses are given by

$$L_{ad}(n) = (N-n)L_r + L_p + L_{SW} \quad (1-1)$$

$$L_{dp}(n) = (n-1)L_r + L_p + L_{SW} \quad (1-2)$$

$$L_{PT}(N) = NL_r, \quad (1-3)$$

where  $L_r$  and  $L_p$  are the reflection and passing losses at/through the OADM, respectively, and  $N$  is the total module number connected in the S-ROADM. The loss of the SW is denoted by  $L_{SW}$ , assuming that the losses are the same value for the C1 and C0 states. In the equations the losses are assumed to be independent of wavelength. The losses of the F-ROADM can be estimated by the similar equations, because the coupler and the splitter have optical bandpass filters which can mux or demux one wavelength at each filter in a similar way as the connected OADM in the S-ROADM. Therefore, the estimation equations for the F-ROADM are given by

$$L_{ad}'(n) = (W-n)L_r' + L_p' + L_{SW} \quad (2-1)$$

$$L_{dp}'(n) = (n-1)L_r' + L_p' + L_{SW} \quad (2-2)$$

$$L_{PT}'(W) = (W-1)L_r' + 2L_p' \quad (2-3)$$

Here,  $W$  denotes the W- $\lambda$  number in the network, and primes denote the values of the coupler/splitter.

The values of  $L_r$  and  $L_p$  were measured for 4 sets of the OADMs used in the S-ROADM, and the values were obtained by averaging the 4 values at each  $\lambda$  of the 8 wavelengths from 1470 nm to 1610 nm. The wavelength dependence of the averaged values is found to be small. Therefore,  $L_r = 0.451\text{dB}$ , and  $L_p = 0.38\text{dB}$  were obtained by averaging all the values. The loss  $L_{SW} = 0.684\text{dB}$  was obtained by averaging measured losses of the SWs. In the same way,  $L_r' = 0.202\text{dB}$ , and  $L_p' = 0.38\text{dB}$  were obtained for the 3 sets of coupler and splitter in the F-ROADM.

Figure 3 shows optical loss properties of the S- and F-ROADMs, when the numbers of R- $\lambda$ s and W- $\lambda$ s are 4 and 8, respectively. The 4 R- $\lambda$ s are the standard values from  $\lambda_1=1470\text{nm}$  to  $\lambda_4=1530\text{nm}$ . The number  $m$  in Fig.3 denotes  $\lambda_m$ . The S-ROADM was constructed by connecting the 4 modules as shown in Fig.2, and the add and drop losses were measured. The results are plotted in



Fig.3. The calculated losses  $L_{ad}(m)$  and  $L_{dp}(m)$  by Eqs.(1-1) and (1-2) are also shown in Fig.3. In the same way, the F-ROADM was constructed as shown in Fig.1, and the add and drop losses were measured. The measured values are plotted in Fig.3, together with the calculated lines by Eqs.(2-1) and (2-2). It is found that the measured and calculated values are consistent with each other for S- and F-ROADMs, respectively, and the maximum losses are comparable between S- and F-ROADMs. The add losses of F-ROADM are higher than those of S-ROADM, whereas the drop losses of F-ROADM are lower than those of S-ROADM. The coupler/splitter of the F-ROADM must mux/demux 8 wavelengths, because the number of  $W\text{-}\lambda$ s is required to be 8. Therefore,  $W=8$  in Eq.(2-1) gives higher add loss. On the other hand, the angle of inclination for S-ROADM loss lines is steeper than that of F-ROADM loss lines, due to the large reflection loss  $L_r$  of OADM in S-ROADM. This is because the reflection loss  $L_r$  of S-ROADM contains the reflection loss at the OADM and the connection loss between modules. The connection loss by connectors in the S-ROADM is higher than the loss connecting adjacent bandpass filters in the waveguide structure of the coupler/splitter for the F-ROADM. As a result, the maximum losses are comparable between S- and F-ROADMs in the present case, but the maximum loss of S-ROADM shall be higher than that of F-ROADM, if numbers of  $R\text{-}\lambda$ s and  $W\text{-}\lambda$ s are 8.

As the next experiment, 2 sets of the S-ROADM were constructed by connecting 4 modules with wavelengths from  $\lambda_2=1490$  to  $\lambda_5=1550$ nm, respectively. The states of the 4 SWs were set to be C1, and the optical mux properties were evaluated. Four OTRs with wavelengths  $\lambda_2 - \lambda_5$  were connected to the add ports, respectively, and the wavelength dependence of the output optical power was measured. The result is shown in Fig.4. It is seen clearly that the 4 wavelengths could be multiplexed by the S-ROADM. The optical powers increase with increasing the wavelength number, consistent with the loss decreasing property of the S-ROADM, as expected by Eq.(1-1).

The passing-through (PT) properties of the S-ROADM were evaluated. Figure 5 (a) shows the measured loss between Cin and Cout ports of the first S-ROADM, when all the SW states were set to be C1. It is found that the losses at the 4  $R\text{-}\lambda$ s were around 22 dB, showing that the 4 wavelengths were dropped. It is also found that smaller PT loss ranges were obtained at wavelengths below  $\lambda_1=1470$ nm, and over  $\lambda_6=1570$ nm. The PT losses were almost constant, being about 1.9 – 2.3 dB. The values agree with the estimated value  $L_{PT}(4) = 4L_r = 1.8\text{dB}$ , given by Eq.(1-3). Thus, the S-ROADM gives no restrictions to the transmissions of the lightwaves with other wavelengths

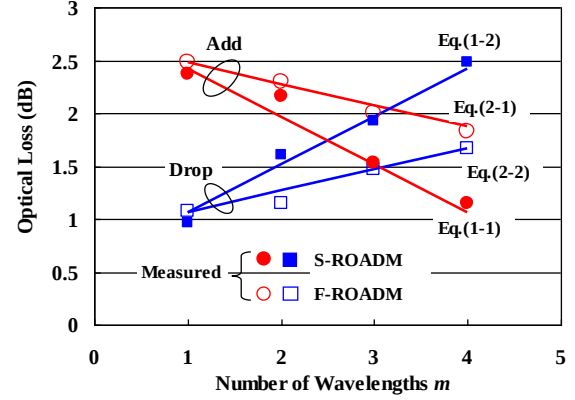


Fig.3 Optical loss of S-ROADM and F-ROADM

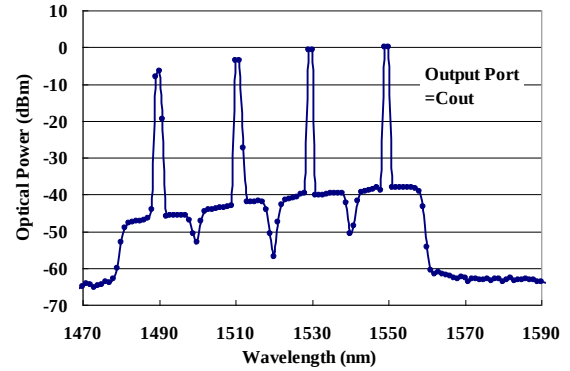
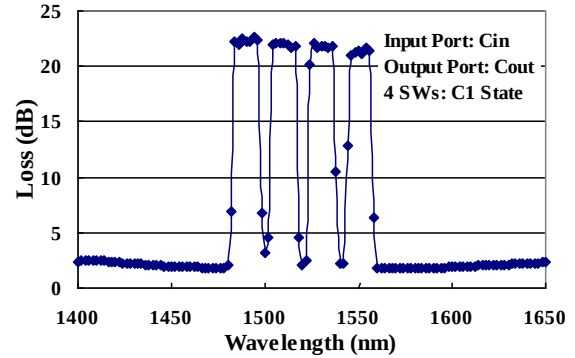
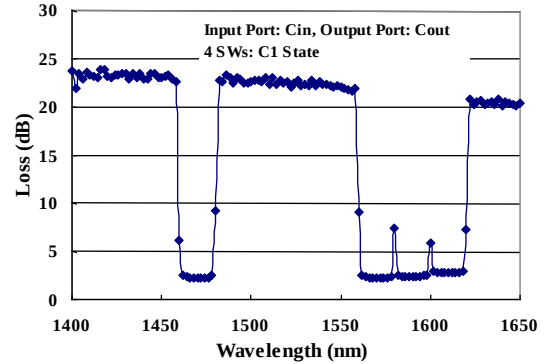


Fig.4 Measured multiplexing properties of S-ROADM



(a) S-ROADM



(b) F-ROADM

Fig.5 Measured passing-through losses

than the 4 R- $\lambda$ s of the S-ROADM, showing that PT transmissions are  $\lambda$  transparent.

The same PT properties were measured for the F-ROADM with 8 W- $\lambda$ s from  $\lambda_1$  to  $\lambda_8$ . The 4 SWs and the coupler/splitter were connected such that 4 wavelengths from  $\lambda_2=1490$  to  $\lambda_5=1550\text{nm}$  were R- $\lambda$ s, and other 4 wavelengths were PT- $\lambda$ s. The PT losses were measured for the F-ROADM with the 4 SWs set to be C1 state. The measured losses are shown in Fig.5 (b). It is found that the losses at wavelengths from  $\lambda_2$  to  $\lambda_5$  were around 22 dB, showing that the 4 wavelengths were dropped. And the PT losses at the 4 PT- $\lambda$ s were as small as about 2.3dB, which agrees with a value of  $L_{PT}(8)=2.17\text{dB}$  calculated by Eq.(2-3). Fig.5 (b) also shows that the PT losses are higher than 20dB at smaller wavelengths than  $\lambda_1$ . This property of the F-ROADM restricts the transmissions of other wavelengths than the W- $\lambda$ s: i.e. the F-ROADM is not  $\lambda$  transparent at other wavelengths than W- $\lambda$ s.

#### IV. PROPOSAL OF BIDIRECTIONAL S-ROADM

The packet transmissions may be uni- or bi-directional in the network with the S-ROADMs. We propose bidirectional S-ROADM (BS-ROADM) structure shown in Fig.6 so as to cover possible lightpath reconfigurations in more general way. The BS-ROADM has 2 sets of input and output ports, and the sets are used for clockwise (CW) and counter clockwise (CCW) transmissions, respectively. Two OTRs per one wavelength can be connected to the BS-ROADM, and are used for CW and CCW transmissions, respectively. In the BS-ROADM, 3 types of configuration are proposed, depending on the SW number and the fiber connections. Each wavelength can be given one of the types independently in the BS-ROADM.

Three SWs are used in Type-1 configuration: the connections for  $\lambda_1$  are the example for Type 1 shown in Fig.6. The CW  $\lambda_1$  lightwave from the left node is demuxed by OADM11 and the direction is determined by C1 or C0 state of SW11. When SW11 state is C1, the demuxed lightwave is dropped and received by a  $\lambda_1$ -OTR1. The lightwave from the  $\lambda_1$ -OTR1 is added, and is sent to the CCW or CW direction, according to C1 or C0 state of SW13, respectively. On the other hand, the lightwave demuxed by OADM11 passes through SW11, when SW11 state is C0, and is sent to the CCW or CW direction, according to SW13 state. The CCW  $\lambda_1$  lightwave from the right node is demuxed by OADM12 and specified to be dropped or PT by SW12. When the lightwave passes through the module, it is sent to the CCW or CW direction by SW13 in the same way as for the CW  $\lambda_1$  lightwave.

Table 1 summarizes possible transmissions by Type-1 configuration for one wavelength, showing the lightpath connections among OTRs installed in the left and right

nodes and the 2 OTRs in this node where the BS-ROADM is working. Since there are 3 SWs in Type-1, 8 states are possible. When #1 state, i.e. the states of all the 3 SWs are C1, CW lightwave from the left node is received by OTR 1 in this node, and the lightwave from OTR1 is sent back to the left node. This is the point-to-point (P2P) transmission, connecting this and the left nodes. The CCW lightwave gives also P2P transmission, connecting this and the right nodes. When #2 state (SW1 and SW2  $\rightarrow$  C1, and SW3  $\rightarrow$  C0), CW lightwave is received in this node, and the lightwave from this node is sent to the right node. This is the one-by-one (O-O) transmission [5], connecting nodes one-by-one with one wavelength. The CCW lightwave also gives O-O transmission in the #2 state. When #3 state (SW1 and SW3  $\rightarrow$  C1, and SW2  $\rightarrow$  C0), CW lightwave gives P2P transmission, because the 2 SW states are the same as #1. The CCW lightwave passes through SW1 and is sent back to the right node by SW3. This is a loop back transmission, which cannot be used for signal transmission, but may be used for the hardware test consisting of the loop back route, i.e. fibers and the optical components in the ROADMs. The OTR 2 is waiting for use in the #3 state. When #5 state (SW1 and SW3  $\rightarrow$  C0 and SW2  $\rightarrow$  C1), the CW lightwave passes through the 2 SWs, and is sent to the right node i.e. the lightwave passes through this node.

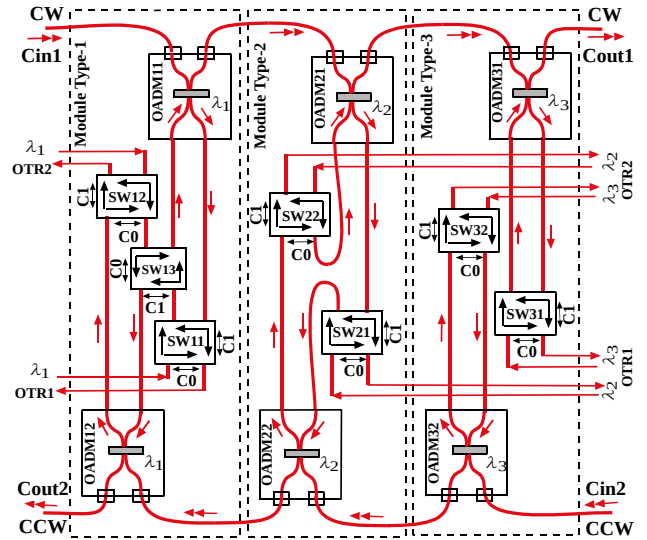


Fig.6 Bidirectional S-ROADM

Table 1 Possible transmissions by Type-1

| # | SW |    |    | Optical Transceiver |         |         |
|---|----|----|----|---------------------|---------|---------|
|   | 1  | 2  | 3  | Left Node           | OTR 1   | OTR 2   |
| 1 | C1 | C1 | C1 | P2P                 |         | P2P     |
| 2 | C1 | C1 | C0 | O-O/CW              |         | O-O/CCW |
| 3 | C1 | C0 | C1 | P2P                 |         | Waiting |
| 4 | C0 | C1 | C1 | Loop Back           | Waiting | P2P     |
| 5 | C0 | C1 | C0 | PT/CW               | Waiting | O-O/CCW |
| 6 | C1 | C0 | C0 | O-O/CW              |         | Waiting |
| 7 | C0 | C0 | C1 | Loop Back           | Waiting | Waiting |
| 8 | C0 | C0 | C0 | PT/CW               | Waiting | Waiting |

Therefore, OTR 1 in this node is not used, waiting for use. The other states i.e. #4 and #6-8 states have the same combinations of the 3 SWs, and therefore, the lightpath connections can be given by Table 1. It is possible to make such tables for Type-2 and Type-3, respectively.

To examine the lightpath reconfiguration performance, an experimental network was designed so as to connect 5 nodes with bi-directional dual fibers. The network configuration is shown in Fig.7. Three sets of the BS-ROADMs were prepared, and installed in Nodes 1, 2 and 3. The BS-ROADM configurations are shown by symbols schematically in Fig.7: e.g. when  $\lambda_1$  is configured by Type-1, it is denoted as  $\lambda_1/T1$  symbolically, and no precise SW configurations are shown. Each BS-ROADM was designed such that it has three Type-1 configurations. In Nodes 4 and 5, bidirectional fixed ROADMs (BF-ROADM) with 2 sets of coupler and splitter were installed. The SW configurations were adopted Type-1 for different 3 wavelengths. As a result, each node has 3 sets of OTRs with different wavelengths. By this configuration, the 5 nodes with BS-ROADM or BF-ROADM are connected by more lightpaths than the full-mesh connections. The 3 wavelengths  $\lambda_1 - \lambda_4$  of 1530, 1550, 1570 and 1590 nm were adopted, respectively. To control the BS-ROADMs, a control system (CS) was connected to the L3SW in Node 1. A LAN analyzer (LA) was also connected to the L3SW in Node 1 to monitor packets, when lightpaths are reconfigured. A personal computer (PC) is connected in each node as shown in Fig.7 to send and receive packets between nodes.

Firstly, the lightpath reconfiguration performance was examined by adding 6 lightpaths to the star connections shown in Fig.8 (a) to obtain the full-mesh connections shown in Fig.8 (b). While ICMP (Internet Control Message Protocol) echo requests were sent continuously from PC1 to PC2, PC3, PC4 and PC5, the CS sent control signals to the BS-ROADMs and BF-ROADM so as to change the SW states. In the reconfiguration processes, the packets were monitored by the LAN analyzer (LA), and the monitored results are shown in Fig.9. The observed packets were classified, depending on the packet types, and were plotted at the observed time in Fig.9. The echo request and reply packets were observed between Node 1 and other 4 nodes, respectively, via Node 5, before the control signals were observed. It is found that the ICMP echo request and reply packets via Node 5 disappeared in 7.62 s, after the TCP packets for the control were observed. Just after this, it was observed that the echo request and reply packets were sent by the direct lightpaths between Node 1 and other 4 nodes, respectively, as shown in Fig.9. This is because the 6 lightpaths were added by the BS-ROADM and BF-ROADM, OSPF (Open Shortest Path

First) found the new direct connections, and OSPF re-calculate the routing table dynamically. After this, the packets were transmitted by the new direct paths, leading to the observed packet changes. It is understood that it took 7.62 s to reconfigure the lightpaths, including the routing table re-calculation time by OSPF. The hello and dead intervals were 1s and 4s, respectively.

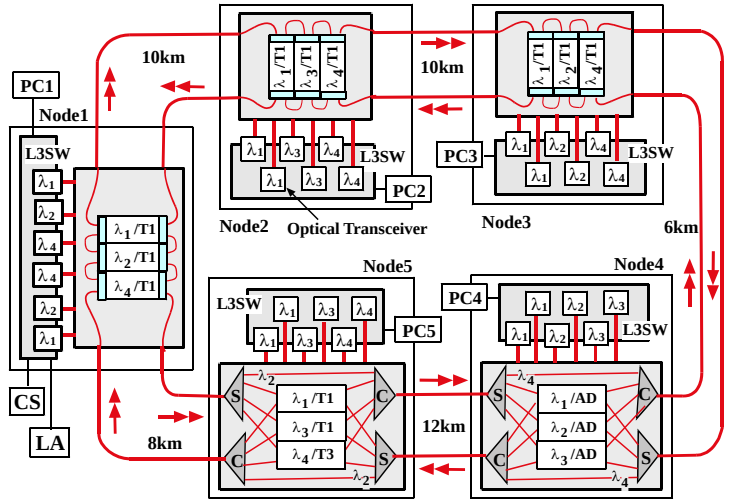


Fig.7 Experimental network configuration

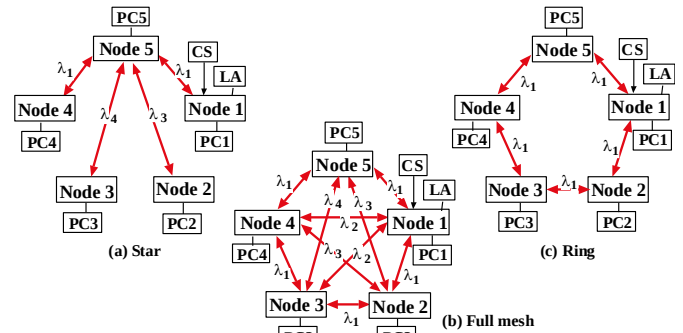


Fig.8 Lightpath connections made by BS-ROADMs in the experimental network

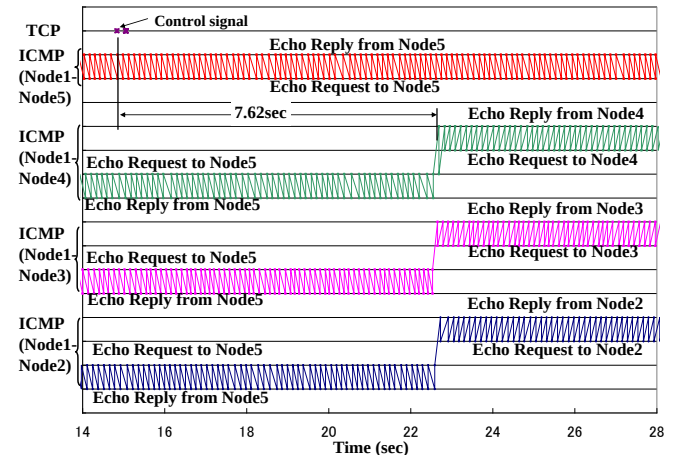


Fig.9 Monitored packets in lightpath reconfiguration from the star to full-mesh connections



As the next experiment, the lightpaths were reconfigured from the full-mesh to a ring connections, as shown in Figs.8 (b) and (c). To evaluate the performance, ICMP echo requests were sent continuously from PC1 to other 4 PCs in the same way as before in the full-mesh network (b), and the packets were monitored. The results are shown in Fig.10. The request and reply packets were observed by the direct paths between Node 1 and other 4 nodes, before the reconfiguration. After control signals were detected, the request and reply packets disappeared between Node 1 and Nodes 3 and 4, respectively, and the packets reappeared in 5.8s between Nodes 1 and 3 via Node 2, and between Nodes 1 and 4 via Node 5. This is because OSPF detected that the 2 direct connections between Node 1 and Nodes 3 and 4 disappeared, tried to find other routing paths, and found the possible route from Nodes 1 to 3 via Node 2, and from Nodes 1 to 4 via Node 5, taking 5.8s to complete the processes on the packet detection base. It is found that no changes were observed for the ICMP packet transmission between Node 1 and Nodes 2 and 5, respectively. This is because no lightpaths were changed between the nodes, proving that the reconfigurations for Node 1 and Nodes 3 and 4 gave no influence on the packet transmissions by other lightpaths.

## V. CONCLUSION

A S-ROADM structure has been proposed for use in campus networks, large-scale LANs and other local networks. The wavelength types were classified as W- $\lambda$ s, R- $\lambda$ s, PT- $\lambda$ s and RV- $\lambda$ s so as to determine what wavelength ranges are  $\lambda$  transparent and what wavelengths can be used for reconfiguration between particular nodes, when applying ROADMs to such networks. The S-ROADMs and F-ROADMs were manufactured for the case that the numbers of W- $\lambda$ s and R- $\lambda$ s are 8 and 4,

respectively, and the performance was evaluated. The equations to estimate the add, drop and PT losses were derived and the 3 types of losses were measured for the S- and F-ROADMs. The measured and calculated results are consistent with each other, showing that the maximum losses were identical for the 2 types of ROADMs. It has been clarified that the S-ROADM could mux the 4 R- $\lambda$ s successfully. It has been shown that the losses at the other wavelengths than the R- $\lambda$ s were as small as 1.9 – 2.3dB for the S-ROADM, clarifying that the S-ROADM provides  $\lambda$  transparent networks. Contrary to the S-ROADM, the losses at other wavelengths than the W- $\lambda$ s were as large as over 20 dB for the F-ROADM, showing that the network by F-ROADMs is not  $\lambda$  transparent. As a result, the S-ROADMs provide more flexible CWDM networks in terms of scalability and reconfigurability.

Three types of SW connection configurations have been proposed for bidirectional S-ROADM structures, enabling us to reconfigure lightpaths such that P2P and O-O links can be made and/or changed between/among nodes. Thus, the BS-ROADMs provide lightpath reconfiguration functions in more general way. To examine the BS-ROADM performance, 3 BS-ROADMs and 2 BF-ROADMs were manufactured, and an experimental network was constructed by connecting the ROADMs. A control system could send control signals to the ROADMs in the network, and it was found that the lightpaths could be reconfigured successfully from star to full-mesh lightpath connections, and full-mesh to ring lightpath connections. The SB-ROADMs together with the reconfiguration functions of the control system provide network administrators with easy management for the network in a skill-less and centralized way.

## ACKNOWLEDGEMENT

This research was supported in part by SCOPE (Strategic Information and Communications R&D Promotion Programme). The authors appreciate greatly for the supports.

## REFERENCES

- [1] B.Mukherjee, *IEEE J. Selected Areas in Communications*, vol.18, pp.1810-1824, Oct., 2000.
- [2] Hongyue Zhu, and B.Mukherjee, *IEEE J. of Lightwave Tech.*, vol.23, pp.2893-2901, 2005
- [3] M.Herzog, M.Maier, and M.Reisslein, *IEEE Comm. Surveys*, vol.6, No.2, pp.2-20, 2004.
- [4] Y.Katsuyama, M.Hashimoto, K.Nishikawa, A.Ueno, MD. Nooruzzaman, and O.Koyama, *The 32nd Annual IEEE Conference on Local Computer Networks*, pp.63-69, Dublin, Ireland, Oct., 2007.
- [5] S.Fujimoto, O.Koyama and Y.Katsuyama, *IEEJ Trans EIS*, 125, No.3, pp.534-535, 2005.

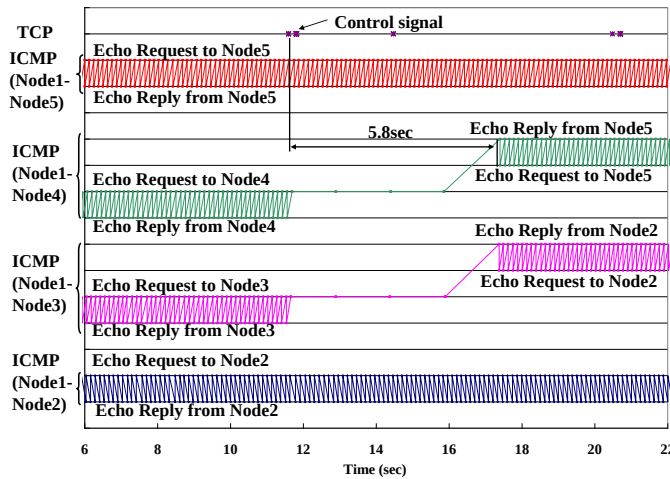


Fig.10 Monitored packets in lightpath reconfiguration from the full-mesh to ring connections

# Evaluation of TCP Performance over Mobile IP Wired-cum-Wireless Networks

Md. Asif Nashiry<sup>1</sup>, Md. Abdullah-Al-Mahmud<sup>1</sup>, Md. Mahbubur Rahman<sup>1</sup>, Md. Nashid Anjum<sup>2</sup>

<sup>1</sup>Department of Information & Communication Engineering, Islamic University, Kushtia, Bangladesh

<sup>2</sup>Department of Electronics & Communication Engineering, KUET, Khulna, Bangladesh.

**Abstract** — Mobility of user devices connecting to the Internet is of major interest in today's research in networking. For users as well as for applications, network mobility should be transparent. Reliable transport protocols like TCP has served well the wired Internet where the packet losses are mainly due to congestion, but is not ready for Mobile IP wired-cum-wireless environments where the significant packet losses are due to bit errors and handoffs. In this paper, we have investigated the performance of TCP among the various TCP variants. We have observed both TCP senders (Newreno & Vegas) and TCP receivers (Base & Delayed-Ack). Using ns-2, we have evaluated the TCP throughput and packet delay over a single TCP connection. The simulation results suggest that a particular combination (one TCP sender with one TCP receiver) of TCP shows the best result in such Mobile IP networks.

**Index Terms** — Wired-cum-Wireless Networks, TCP, Mobility, Mobile IP, Handoff.

## I. INTRODUCTION

TCP was basically developed assuming that it would run on wired networks. Wired networks usually have less bit error rates and hence less packet loss. Also, since mobility was not considered, there are no packet losses and delays caused by mobility. Hence, TCP was designed to assume that any segment loss is caused by the network congestion and on seeing any segment loss; TCP invokes its congestion control measures.

In various circumstances wired and wireless networks are connected together to take the advantages of both. Usually, the wireless links have high bit error rates. Also, temporary disconnections occur because of the factors like channel fading, and handoffs when a mobile node is in motion. The handoff period depends on both the link level handoff protocol being used as well as the IP level handoff protocol being used, if any. The standard TCP implementations assume the cause of any packet loss to be the network congestion. Then, they reduce the congestion size to a minimum. Also, TCP invokes the slow-start mechanism. If the network links are slow, it takes long time to grow the congestion window. Thus, reduction in the size of the congestion window reduces the transmission rate and hence degrades the performance. Also, TCP undergoes a binary exponential backoff, causing long pauses of communication. Because of this, some time may be unused even after the mobile node recovers from the temporary disconnection. It has been observed that even a single wireless link can reduce the TCP performance considerably. In theory we can use existing transport protocols like TCP on a wireless host to

communicate with a fixed network. Though this keeps the transport layer transparent to mobility, disconnection and other features of wireless and mobile hosts reduce the performance.

Thus TCP proved not so efficient in Mobile IP wired-cum-wireless environments and suffers from performance degradation introduced by the conditions that exist in those environments. Hence, an optimized reliable transport protocol is very important for the development of the wireless Internet.

This paper provides a realistic comparison of the performances of the various TCP variants over Mobile IP wired-cum-wireless network. The TCP variants include both TCP senders (TCP Newreno and TCP Vegas) and TCP receivers (Base TCP Sink and Delayed-Ack TCP Sink).

The reminder of this paper is organized as follows: Section II depicts the related works; mobile IP and TCP variants are described in Section III. We define the simulation model in Section IV and present the results and discussions in Section V. The conclusion and future work follow in Section VI.

## II. RELATED WORKS

TCP/IP is the standard networking protocol stack for Internet. So it is expected to be deployed over wireless networks to allow seamless integration with the Internet. However, earlier research suggested that TCP performs poorly over cellular (single-hop) wireless networks [1]. A good amount of research has been done in the past to improve the performance of TCP in the conditions that prevail in mobile and wireless environments.

Split connection approach is one of them. It separates TCP connection, one for the wireless network and the other for the wired network between a sender and a receiver. The motivation behind the use of separate connections is the fact that the characteristics of a wireless network are different from the characteristics of a wired network, hence a separate connection that is optimized to the kind of network it operates on, can be used. I-TCP [2], ELN [3] can be classified in this group.

Enhanced link layer approach is another solution. Since the basic factor that causes degradation in wireless environments is the nature of the transmission medium, an improvement in the link technology that operates over the physical layer can improve the overall performance. This was the idea behind enhanced link layer solutions.

Mechanisms like ARQ, FEC [4] were suggested to improve the performance at the lower layers which, in turn, can reduce the chances of TCP getting timed out and does not invoke its congestion control measures.

Most of the above-mentioned solutions require special support from the network infrastructure in some form or the other. Some other approaches were suggested that do not require any special support from the intermediate infrastructure and changes are confined to the end nodes. Path Prediction [5] and ACK Regulator [6] come under this category. These solutions are not bound by the problem of encrypted traffic, and they do not assume any special support from the network infrastructure.

As suggested earlier, most related work focuses on to improve the TCP performance in wireless network. Also most of the earlier solutions (except ELN) try to optimize the mobile host acting as a TCP receiver but do not consider the case of a mobile TCP sender, which can also be a common case in the future. Also a few of the earlier work measured TCP performance with combining both the TCP senders and the TCP receivers.

This paper provides a realistic comparison of performance of the various TCP variants over Mobile IP wired-cum-wireless network. The TCP variants include both TCP senders (TCP Newreno and TCP Vegas) and TCP receivers (Base TCP Sink and Delayed-Ack TCP Sink).

### III. MOBILE IP AND TCP VARIANTS

In order to achieve the mobility function, the Internet Protocol (IP) has extended to become the Mobile Internet Protocol (Mobile IP or MIP). Mobile IP provides hosts with the ability to change their point of attachment to the network without compromising their ability to communicate. The mobility support provided by Mobile IP is transparent to the other protocol layers so as not to affect the operation of applications which do not have the mobile capability.

Among various IP mobility proposals, Mobile IPv4 [7, 8] is the oldest and probably the most widely known mobility management proposal with IP. MIPv4 introduces three new entities required to support the protocol: the Home Agent (HA), the Foreign Agent (FA) and the Mobile Host (MH). Home and foreign agents are introduced for mobility management. Each time a mobile host connects to a network at a new location, it will obtain a temporary address, called Care-of Address (COA) from a foreign agent in the local network. Then the mobile host must inform its home agent of the new address by a registration procedure, which begins when the mobile host, possibly with the assistance of the foreign agent, sends a registration request with the COA. When the home agent receives this request, it may typically add the necessary information to its routing table, approve the request, and send a registration reply back to the mobile host.

A basic function of TCP is to provide reliable communication over an unreliable network layer. TCP ensures that the data is delivered from the sending process to the receiving process correctly and in order. TCP controls the traffic flow and network congestion by maintaining window (buffer to store data packets); and various TCP variants are classified according to the mechanism for maintaining this window.

#### A. TCP New-Reno

A modification to Reno TCP called New-Reno TCP [9]. Since most TCP sessions last only for a short period of time, the initial slow start period is significant for the overall performance. A method is proposed to estimate an optimum slow start threshold (ssthresh) value by calculating the byte equivalent of bandwidth delay product of the network when a new connection is made. TCP New-Reno also deals with multiple packet losses from a single window. If two or more segments have been lost from the transmitted data (window), the Fast Retransmission and Fast Recovery algorithms will not be able to recover the losses without waiting for retransmission time out. New-Reno overcomes this problem by introducing the concept of a Fast Retransmission Phase, which starts on detection of a packet loss (receiving 3 duplicate ACKs) and ends when the receiver acknowledges reception of all data transmitted at the start of the Fast Retransmission phase.

The transmitter assumes reception of a partial ACK during the Fast Retransmission phase as an indication that another packet has been lost within the window and retransmits that packet immediately to prevent expiry of the retransmission timer. New Reno sets the congestion window (cwnd) to one segment on reception of 3 duplicate ACKs (i.e. when entering the Fast Retransmission Phase) and unacknowledged data are retransmitted using the Slow Start algorithm. The transmitter is also allowed to transmit a new data packet on receiving 2 duplicate ACKs. While the transmitter is in the Fast Retransmission Phase, it continues to retransmit packets using Slow Start until all packets have been recovered (without starting a new retransmission phase for partial ACKs). Although this modification may cause unnecessary retransmissions, it reduces transmitter timeouts and efficiently recovers multiple packet loss using partial ACKs.

#### B. TCP Vegas

TCP Vegas [10] does not involve any changes to the TCP specification; it is merely an alternative implementation that is able to interoperate with any other valid implementation of TCP. In fact, all the changes are confined to the sending side. There are some techniques that Vegas employs to increase throughput and decrease losses.

It calculates the RTT more accurately. Using an accurate RTT estimate serves two purposes. First, it leads to a more accurate timeout calculation. Second, it is used

in conjunction with the mechanism described next to decide to retransmit a dropped segment in a more timely fashion. It includes the new mechanism for deciding to retransmit Vegas treats the receipt of certain ACKs as a trigger to check if a timeout should happen. Moreover, Vegas only decreases the congestion window if the retransmitted segment was previously sent after the last decrease. Any losses that happened before the last window decrease do not imply that the network is congested for the current congestion window size, and therefore do not imply that it should be decreased again. Another most important feature of TCP Vegas is that the detection of congestion. Vegas calculates the expected throughput by:

$$Expected = WindowSize / BaseRTT$$

where *WindowSize* is the size of the current congestion window and *BaseRTT* is the RTT of a segment when the connection is not congested. It then calculates the current *Actual* throughput of the network. Then, Vegas compares *Actual* to *Expected*, and adjusts the window accordingly.

### C. Base TCP Sink

The base TCP sink object is responsible for returning ACKs to a peer TCP source object. It generates one ACK per packet received. The size of the ACKs may be configured. The creation and configuration of the TCP sink object is generally performed automatically by a library call.

### D. Delayed-Ack TCP Sink

A Delayed-Ack sink object is available for simulating a TCP receiver that ACKs less than once per packet received. This object contains a bound variable, which gives the number of seconds to wait between ACKs. The Delayed-Ack sink implements an aggressive ACK policy whereby only ACKs for in-order packets are delayed. Out-of-order packets cause immediate ACK generation. It is possible to configure the bound variable (interval) of the Delayed-Ack TCP Sink.

## IV. SIMULATION MODEL AND DESIGN

To design the Mobile IP wired-cum-wireless network topology we have considered various simulation parameters, which are supported by the network simulator-2 (NS-2) [11]. We have done simulation for two different topologies and call them topology 'A' and topology 'B' respectively.

There is a roaming mobile node called Mobile Host (MH) that moves between its home agent and foreign agent in a certain velocity. Data will be exchanged between the roaming mobile node (MH) and the wired node W(0) via the two base-station nodes (HA or FA). Base station nodes are like gateways between wireless and wired domains and allow packets to be exchanged between the two different types of nodes. The initial positions, (x,y), of the nodes MH, HA and FA are

(50,100), (200,300) and (500,300) respectively in the 700mX800m topology. In this network, the mobile node MH starts moving from position (50,100) at 10 sec towards destination (650,100) at fixed speeds (0, 5, 10, 15, 20, 25, or 30) m/s. After reaching the destination position (650,100), MH will stay there 1sec (pause time) and then again moves back to its initial starting position (50,100) at the same velocity. In this way, the mobile node MH changes its home agent (HA) and foreign agent (FA) several times depending upon its speed. We have used DSDV ad hoc routing protocol. FTP begins transferring packets of size 1000 bytes after 10 second. Each wireless node has a buffer size of 50 packets. Each simulation ends at 150 seconds.

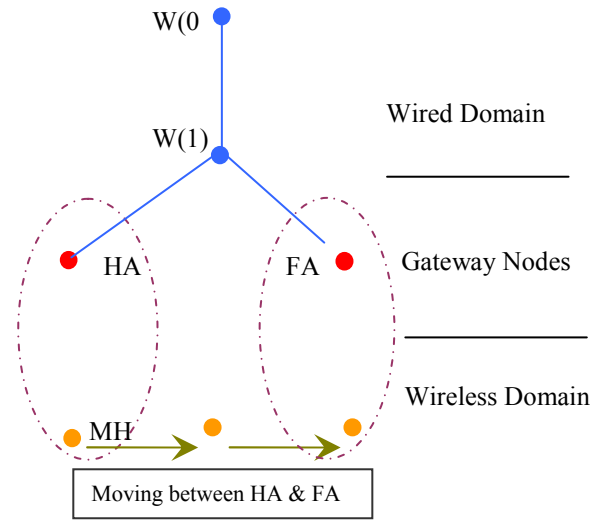


Fig. 1. Topology 'A' for Mobile IP wired-cum-wireless simulation

There are fewer changes in Topology 'B'. We have used three base-station nodes in topology 'B' and call them Home Agent (HA), Foreign Agent-1 (FA1) and Foreign Agent-2 (FA2) respectively. The initial positions, (x,y), of the nodes MH, HA, FA1 and FA2 are (50,100), (200,300), (500,300) and (800,300) respectively in the 1000mX800m topology. In this case, simulation starts at 10 sec and ends at 190 sec. Here the mobile node MH starts moving from position (50,100) at 10 sec towards destination position (950,100) at a fixed speed of (0, 5, 10, 15, 20, 25, or 30) m/s while nodes HA, FA1 and FA2 serve as base station (gateway) nodes. As before, the pause time is 1 sec, that is, the mobile node MH will stay 1 sec at position (950,100) and then moves back to its initial position (50, 100) at the same velocity. Staying there 1 sec, it moves again towards the destination position.

## V. PERFORMANCE ANALYSIS AND DISCUSSION

This section describes a realistic analysis comparing the performance of the various TCP variants over multi-hop Mobile IP wired-cum-wireless networks. We have

combined the TCP senders (New-Reno & Vegas) with the TCP receivers (Base & Delayed-Ack). We have considered the mobile node as a TCP sender.

#### A. Throughput Analysis

We have evaluated the throughput for topology 'A' obtained with the TCP variants in terms of the total number of packets received at the destination node per unit time over intervals of 7 seconds for mobility 0 meter/sec, 15 meter/sec and 30 meter/sec and the results are plotted in figures.

From figure 2, we observe that after the initial connection establishment, all the variants achieve almost fixed throughput to the whole simulation time. It is common that if the window size increases, the throughput will also increase. Among the four variants, Vegas to Delayed-Ack shows the best throughput.

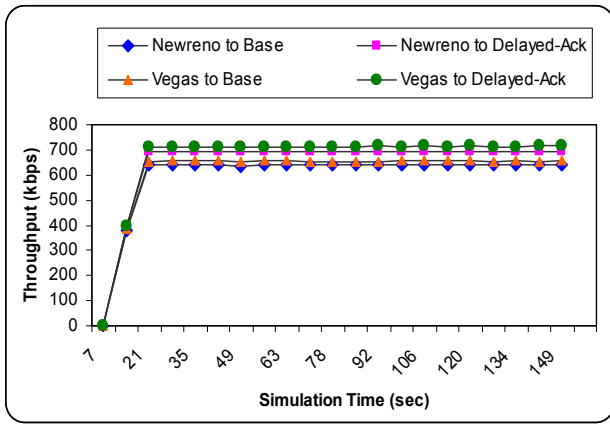


Fig. 2. Throughput (kbps) versus Simulation Time for topology 'A' for mobility 0 (meter/sec)

From figure 3 (a), it is seen that the throughput of each of the TCP variant decreases drastically. By comparing this figure to figure 2, it is observed that the throughput performance is worse in this case. From the start of the simulation to 35 seconds, throughput is the same (almost 700 kbps) as before (figure 2). But after 35 seconds it shows different result. The mobility of the mobile host is the cause of this degradation. In the case of figure 2, the mobile host remains stationary. But in this case the mobile host changes its base station for a number of times during simulation time. When changing base station a mobile host requires a handoff. The mobile host changes base station three times; so require three handoffs. And at the time of each handoff packets are dropped and throughput decreases.

Figure 3(b) shows the similar type of result except the throughput is reduced more times than that of figure 3(a). Since in this case, the mobility is 30 meter/sec, more handoff is required and at the time of each handoff throughput is reduced.

One interesting feature which is observed from figures 3(a) and 3(b) is that during first handoff throughput performance decreases drastically but during the second,

third and successive handoffs throughput decreases slightly. But in such an environment Vegas to Delayed-Ack performs well among the TCP variants. Its throughput reduces less than that of the others.

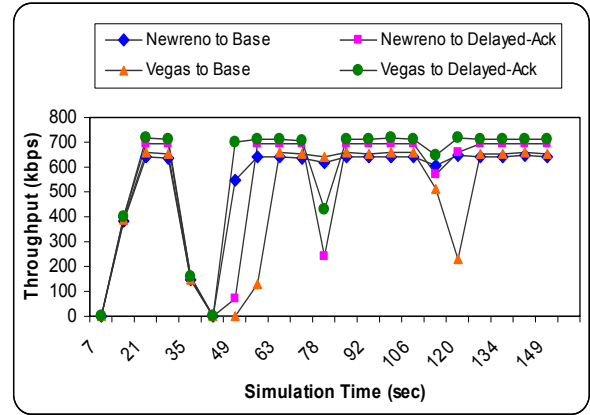


Fig. 3(a). Mobility 15 (m/s)

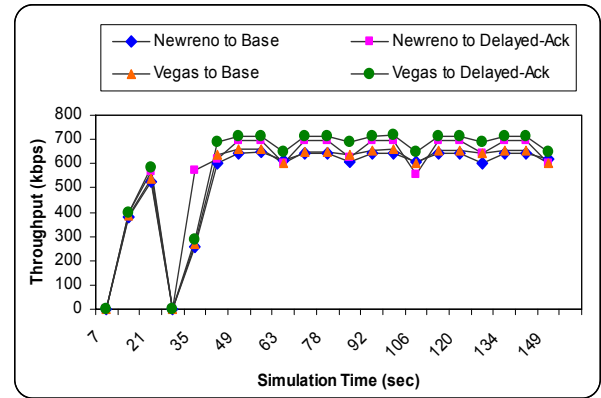


Fig.3(b). Mobility 30 (m/s)

Fig. 3. Throughput (kbps) versus Simulation Time for topology 'A'

Concentrating on the receiver portion in figure 2, 3(a) and 3(b), we see that the performance of the Base TCP receivers is lower than that of the Delayed-Ack TCP receivers; no matter whether the TCP sender is New-Reno or Vegas.

Since all the terminals in Mobile IP network share the same radio channel to send and receive packets, the collision of data packets and acknowledgement (ACK) packets in the radio channel can severely reduce TCP throughput. But Delayed-Ack receiver employs less ACK packets. So, for transferring TCP data packets, TCP sender can utilize the channel more. Hence throughput of any TCP sender synchronized with Delayed-Ack receiver is better than that of TCP sender with Base receiver.

Now with fixed Delayed-Ack receiver, we are going to observe the performance of New-Reno and Vegas sender. By evaluating the throughput from three different mobility scenarios it is seen that TCP Vegas sender performs well. As stated earlier, TCP Vegas estimates the network status to adjust its window value. This property



enriches its performance in such a heterogeneous environment.

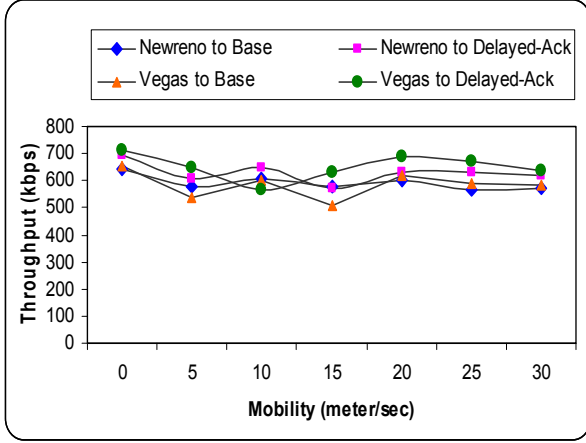


Fig. 4. Throughput vs Mobility for topology 'A'

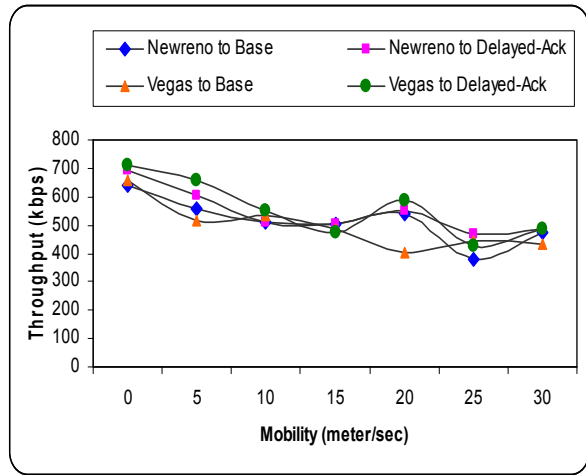


Fig. 5. Throughput vs Mobility for topology 'B'

We also average the throughput over the simulation duration and plot the average throughput obtained with the TCP variants as a function of node mobility for both topology 'A' and 'B' in figure 4 and figure 5 respectively.

According to figure 4 and 5, when the source node is stationary (node mobility 0 meter/sec), all the TCP variants achieve their own highest throughputs in both topology 'A' and topology 'B', because of occurring no handoffs in this situation.

The throughputs achieved with New-Reno to Base TCP and Vegas to Base TCP are lower than those achieved with New-Reno to Delayed-Ack TCP and Vegas to Delayed-Ack TCP in both topology 'A' and topology 'B'. It is due to the fact that ACKs and TCP packets share the same channel; so decreasing the ACKs by using the Delayed-Ack receiver instead of Base receiver it is possible to improve the throughput of TCP.

#### B. Average Delay Analysis

From figure 6 and figure 7, it is seen that the average delay of New-Reno sender (receiver may be Base or Delayed-Ack) is high as compared to Vegas sender in both topologies. Vegas reads and records the system clock

each time a segment is sent. When an ACK arrives, Vegas reads the clock again and does the round trip time (RTT) calculation using this time and a timestamp recorded for the relevant segment. Using an accurate RTT estimates, Vegas leads to a more accurate timeout calculation. This is why Vegas shows very low delay as compared to New-Reno.

The average delay of New-Reno to Delayed-Ack is the worst among all the variants. This is why it has given the lowest packet delivery ratio. The average delay of New-Reno to Base is better than that of New-Reno to Delayed-Ack but worse than Vegas to Base and Vegas to Delayed-Ack. The same thing has happened in case of packet delivery ratio, that is, the packet delivery ratio of New-Reno to Base is better than New-Reno to Delayed-Ack but worse than the Vegas to Base and Vegas to Delayed-Ack.

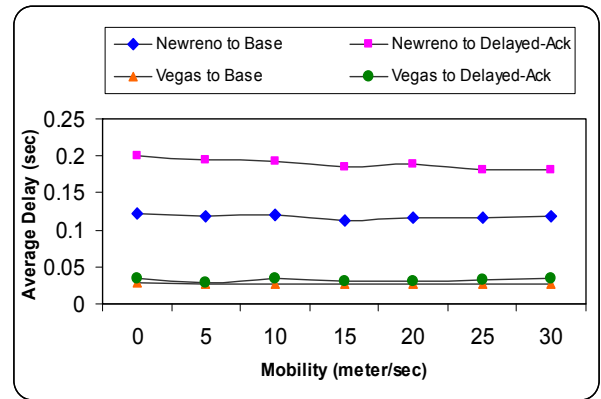


Fig. 6. Average delay vs Mobility for topology A

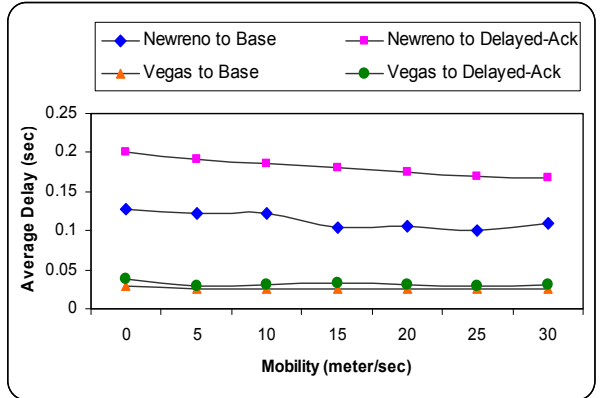


Fig. 7. Average delay vs Mobility for topology B

The delay performance of Vegas to Delayed-Ack TCP and Vegas to Base TCP is relatively similar and they show low delay in all the mobility so they show relatively similar and high packet delivery ratio in both the topologies. Since the average delay of Vegas to Delayed-Ack TCP is low it has given a very good throughput in the simulation time in both topologies.

## VI. CONCLUSION

While TCP is needed for a Mobile IP network, the current variants of TCP are not adequate for the task. To

achieve good TCP performance in Mobile IP wired-cum-wireless networks a variety of solutions have been proposed and in most of the solutions the mobile node acts as a TCP receiver. However, little attention has been paid to study the performance of TCP traffic over combining both TCP senders and receivers. We have investigated the performance of TCP over combining both TCP senders (Newreno & Vegas) and TCP receivers (Base & Delayed-Ack) using simulations in ns-2 for a range of node mobility with a single traffic source. The performance metrics that we have considered include TCP throughput and packet delay.

On throughput performance, Vegas TCP to Delayed-Ack gains the best performance. On the sender side, the effective congestion control technique of Vegas makes it optimal in such a heterogeneous environment. And in the receiver side, the Delayed-Ack receiver makes the proper use of channel by sending fewer ACK packets.

We have also evaluated throughput with respect to node mobility. The result shows that throughput of each TCP variant varies similarly with respect to each other against mobility of node. But Vegas to Delayed-Ack shows the best result as compared with other variants.

Delay performance of each variant shows the confidential result. TCP Vegas has the lowest delay no matter whether the receiver is Base or Delayed-Ack. New-Reno has large delay. But among them New-Reno to Delayed-Ack has the highest delay. After observing packet delivery ratio and delay, it is seen that the variant that gains highest packet delivery ratio has lowest delay.

Though all the variants of TCP suffer in Mobile IP wired-cum-wireless networks, it is clear from the various results that if we use Vegas TCP as a source (sender) and Delayed-Ack as a sink (receiver), then this combination shows the best result among the all TCP variants. Next, we plan to investigate the performance of all the TCP variants with multiple traffic sources in more complex heterogeneous networks and also consider some link layer and network layer feedbacks to improve their performance.

#### REFERENCES

- [1] G. Holland and N. Vaidya, "Analysis of tcp performance over mobile ad hoc networks," *Proceedings of ACM Mobicom*, pp. 219–230, August 1999.
- [2] Bakre, A., Badrinath B.R., "I-TCP: Indirect TCP for Mobile Hosts," *Proceedings of the 15<sup>th</sup> International Conference on Distributed Computing Systems (ICDCS)*, May 1995, p. 136-143.
- [3] Balakrishnan, H., Katz, R., "Explicit Loss Notification and Wireless Web Performance," *Proceedings of IEEE Globecom Internet Mini- Conference*, Sydney, Australia, November 1998.
- [4] Chockalingam, A., Zorzi, A., Tralli, V., "Wireless TCP Performance with Link Layer FEC/ARQ,"

- Proceedings of IEEE ICC'99*, June 1999, p.1212-1216.
- [5] S. Hadjiefthymiades, S. Papayiannis, and L. Merakos, "Using Path Prediction to Improve TCP Performance in Wireless/Mobile Communications," *IEEE Communication Magazine*, Aug. 2002.
- [6] M. C. Chan and R. Ramjee, "TCP/IP Performance over 3G Wireless Links with Rate and Delay variation," *In Proceedings of MOBICOM '2002*, 2002.
- [7] Perkins E. Charles, 'IP Mobility Support', RFC 2002, October 1996.
- [8] Perkins, C., IP mobility support for IPv4, IETF RFC 3344 (2002).
- [9] J. Hoe, "Improving the start-up behavior of a congestion control scheme for TCP," in *ACM SIGCOMM*, Aug. 1996.
- [10] L. S. Brakmo and L. L. Peterson, "TCP Vegas: End to End Congestion Avoidance on a Global Internet," *IEEE Journal on Selected areas in Communications*, vol. 13, no. 8, pp. 1465–1480, Oct. 1995.
- [11] K. Fall and K. Varadhan, "ns Notes and Documentation", Lawrence Berkeley National Laboratory, 28 Jan 2005, <http://www-mash.cs.Berkeley.edu/ns/>.

# A System for Locating Users of WLAN using Dynamic Mapping in Indoor and Outdoor Environment-LOIDS

Karamtot. Krishna Naik and M.N. Giri Prasad

Sree Visvesvaraya institute of Technology & Science, Mahabubnagar, India  
JNTU College of Engineering, Pulivendula, India

**Abstract** — The development of portable, light weight, hand held computing devices and high speed wireless local area networks enable user to remain connected while moving in indoor and outdoor environment. Wireless local area network has a wide range of utilization in moving from one location to another location. When a user is moving from one location to another, how the other user knows about the required station inside WLAN, for that we designed and implemented a system to locate a mobile user inside the wireless local area network based on RSSI. We designed four different architectures and used dynamic radio mapping of indoor and outdoor locations. These architectures are based on dynamic configuration of radio map in indoor and outdoor environment with the help of Sniffer, a specially designed application to locate a station at that time. We found a better location of a mobile user in WLAN with the help of Pamvotis a WLAN simulator. We tested this work in indoor and outdoor environment with different locations practically.

**Index Terms** — WLAN, AP, RSSI, RPM

## I. INTRODUCTION

These days wireless local area network has become very popular, when the initial IEEE802.11 is the standard for providing wireless connectivity to automatic machinery, equipment and stations that require rapid deployment, which may be portable, handheld or which may be mounted on moving vehicles within a local area. IEEE802.11 Wireless local area network is a shared-medium communication network that transmits information over wireless links for all IEEE802.11 stations in its transmission range to receive [3][4][5][10]. The main characteristic of the IEEE802.11 WLAN's is its simplicity, scalability & robustness against failures due to its distributed nature. IEEE802.11 wireless networks can be configured into two different modes: Ad-hoc & infrastructure modes. In Ad-hoc mode, all wireless stations within the communication range can communicate directly with each other, whereas in infrastructure mode, an Access Point is needed to connect all stations to a distribution system, each station can communicate with one another through the Access point.

IEEE802.11 provides the data rates upto 2 Mbps at 2.4 GHz ISM band. Later, IEEE802.11 working group published its enhanced version named IEEE802.11b that extends the data rate upto 11Mbps at this ISM band. Its high-speed version at 5 GHz UNII band (IEEE802.11a) was also defined. IEEE802.11a standard can achieve data rate of upto 54Mbps

by using OFDM (Orthogonal Frequency Division Multiplexing) modulation technique at physical layer. The fundamental building block of the IEEE 802.11 architecture has the Basic Service Set (BSS). A BSS can be defined as a group of stations that are located in the same geographical area. The area covered by the BSS is also known as basic service area (BSA). Any station can establish a direct communication session with any other station in the BSS, without the requirement of channeling all traffic through a centralized AP.

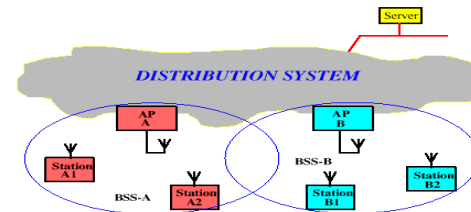


Fig. 1. An Extended Service Set

Infrastructure network is built to provide wireless users with specific services and range extensions. In the context of IEEE 802.11, infrastructure network is established using APs. AP supports geographical range extension by providing the integration points necessary for network connectivity among multiple BSSs to form an Extended Service Set (ESS). The ESS consists of multiple BSSs which are integrated together using a common distributed system as shown in the Figure-1 above. Based on this infrastructure network with location of every station in the network to provide better services must be known to the service provider. That is possible only with the help of creating a map i.e. radio map. This map is based on the geographical area of our institute. We divided this area into few blocks and placed an individual access point which covers the entire area to make the complete network. These access points form a dynamic mapping whenever the user tries to locate the required station, where it/that is.

Here we mainly focus on the signal propagation which can generally attribute to reflection, diffraction and scattering mechanisms. The changes in furniture placement, surrounding structures and occupancy conditions seriously affect signal propagation condition in normal WLAN environment. Large scale fading and small scale fading are the two methods, which are used to explain RSSI (Received Signal Strength Indication) fluctuations. The large scale fading occurs due to the distance



separation between AP (Transmitter) and wireless Sniffer (Receiver), where RSSI decreases as the distance increases. The small scale fading is given by the rapid fluctuations of the RSSI over very short distances or short time durations caused by multipath in the radio environment.

## II. WLAN LOCATION ARCHITECTURE

In the existing WLAN infrastructure without using any additional hardware we can increase efficiency of the network by using suitable network location architecture. With that aspect we are developing a system architecture which estimates the accurate location of the user. Here we are proposing four types of location architectures. One is Source- Destination, another one is Client-Server, third one is Sniffer based and final one is Access Point (AP) based architectures[6][8].

*Source-Destination Architecture:* Source station in one area can search for any other required (destination) station in same or different network. Here a source station sends request to the nearby station that may be the destination or may be connected with the destination. The concerned station which is helping to find the location of the destination records the received signal strength indication from the destination. Based on that RSSI we can identify the location of the destination with help of radio map. The architecture of this is shown in the Figure 2.A.

*Client-Server Architecture:* This architecture has two phases of location estimation, the first one is offline phase and second one is real-time phase. In offline phase infrastructure mode of WLAN access point (AP) provides services to all clients in that area as an ESS. The ESS may have many APs, Client in one AP can receive signal from other or many other APs. But only one AP has to serve the client. That selection of the AP by client is based on the RSSI from other APs. In real-time phase service provided to the clients based on the request made by the respective client with the help of the location based services i.e. as client needed for service made available with that AP. The architecture of this is shown in the Figure 2.B.

*Sniffer Based Architecture:* It has great advantages over all other architectures because it is the application based architecture. In this architecture all stations are preinstalled specially designed application which can identify the location of the system. The application software that monitors all the stations will capture all the traffic in the network. This application software is considered as a Sniffer in the network. For reading the RSSI information from WLAN we will try to connect over the network by using a WNIC (here we are using DWL and NetGare). The WNICs work like passive entities, try to send packets in the wireless medium and capable of capturing all frames transmitted in the same channel. Basically to detect and monitor the wireless medium, Sniffer has to execute two functions continuously. Firstly detecting the wireless clients and recording the available RSSI values of it and second one is measuring RSSI from one or more reference

devices in order to construct the location model in the map. The architecture of this is shown in the Figure 2.C.

*Access point Based Architecture:* As we all are aware of WLAN that, it has two modes of operation one is Ad-Hoc mode and another is Infrastructure mode, these

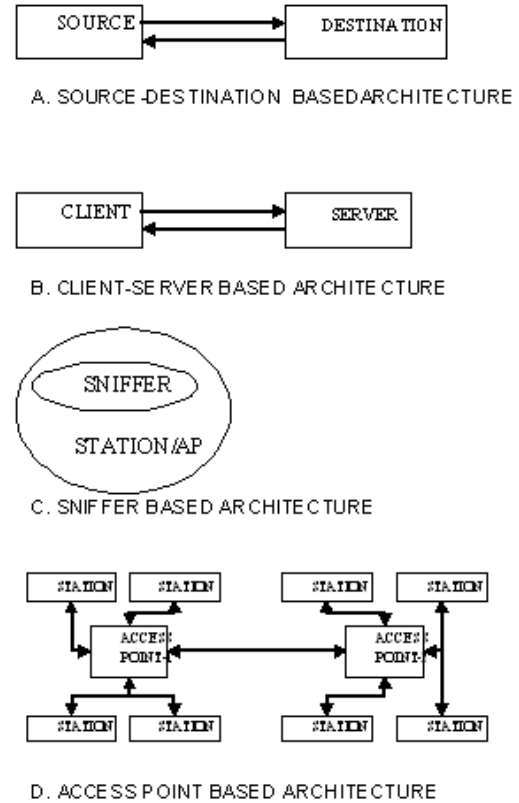


Fig. 2. Architectures for location determination

two modes will be used to form BSS and ESS. Here we are considering ESS to construct a specialized architecture for WLAN location determination system. In ESS basically we use an access point to connect available stations in the coverage area. Suppose any client wants to connect any remote station a client need to contact parent AP to find the location of destination station. The destination station may be available with any other AP or the same AP. In this connection to remember all the stations in the network we are placing a RADIUS server with every AP. RADIUS is a server which will keep information about all stations available with that AP if any AP sends a request asking information about any station, then that AP [1] RADIUS server sends reply to the originated AP with the available information. The architecture of this is shown in the Figure 2.D.

### III. LOIDS- LOCATION OF OUTDOOR AND INDOOR DETERMINATION SYSTEM

Our aim is to discuss and differentiate the Indoor and outdoor locations of WLAN enabled stations and providing the best services to every station in the network. This system (LOIDS) helps to find the required station in the network. A lot of efforts in preparing LOIDS include building of Radio map. There are two methods to consider one is empirical and mathematical methods. In empirical method the radio map created by involving a mobile user in different places and recording physical coordinate at each position for different APs with in the range. In mathematical method the radio map is generated by using the RSSI of RF propagation.

While describing the LOIDS in the indoor and outdoor environment let us assume the physical dimensional space  $L$  where a device is situated, a AP with Sniffer loaded at  $l$ , RSSI measured by  $K$  Sniffer given a tx device located at  $l$  for each position  $l \in L$ . Here we assume  $L$  discrete and also define a signal space  $S$  with  $K$  dimensions where element of this space is a vector of dimension  $K$  in which its positions represent RSSI readings from the  $K$  different Sniffers[6][7][9]. Samples from the signal space  $S$  are denoted by  $s$ , the given RSSI vector  $s=(s_1, s_2, \dots, s_k)$  we are determining the position of the station by  $l \in L$ , this maximizes the probability  $p(l/s)$  of a tx device to be located at  $l$  with a captured RSSI vector measured by  $K$  Sniffers. The required information is then molded into the database where the location server creates the radio propagation map (RPM), where the grid position  $l=(x,y)$  is associated to probability distribution  $p(s/l)$ . The location server builds a RPM for each Sniffer  $I$ , i.e. for  $K$  Sniffers a total of  $K$  RPM  $s$  will be formed at the server. To measure the signal strength  $S_i$  from a tx situated at  $l$  is given by  $p(s/l)$  according to Gaussian distribution:

$$P(s/l) = \frac{1}{\sigma(l)\sqrt{2\pi}} \exp\left(-\frac{s - \mu(l)^2}{2\sigma^2(l)}\right) \quad (1)$$

Here  $\mu(l)$  is RSSI expected values of all stations in the network at a Sniffer loaded station and  $\sigma(l)$  is the distribution standard deviation (SD).  $\mu(l)$  Can be estimated by large scale propagation model,

$$\mu_{(l)}(d) = \mu_0(d_0) - 10_{n0} \log\left(\frac{d}{d_0}\right) - \alpha \quad (2)$$

In this equation  $\alpha$  is attenuation caused by interferences and  $d$  distance between transmitter, receiver and Sniffer placed station. To calculate the distance  $d$  between transmitter, receiver and Sniffer placed station is

$$d = \sqrt{(|x_{\text{SNIFFER}} - x|)^2 + (|y_{\text{SNIFFER}} - y|)^2}$$

in above equation  $x$  and  $y$  are coordinates at location  $l$ . RPM can be constructed when the significant changes in RSSI occur. RSSI measurements are used between the pairs Sniffers and APs to sense whether the WLAN propagation environment has changed and when to rebuilt RPM. If RSSI follows Gaussian distribution where mean and standard deviation between the pair of Sniffers and APs the RPM can be estimated by real-time measurements between the pair of Sniffers and APs. After measuring many times the above said measurements for  $M$  consecutive RSSIs may fall outside the interval( mean and standard deviation), with that a new RPM should be formed. Another method to reconstruct the RPM is for every  $T$  seconds. If found any changes in mean and standard deviation that may not affect the RPM because, it is a pre-defined map. Here any number of users may enter into the scenario and may leave the scenario. The calculated RSSI with Sniffer  $s$  and APs may not show much difference. In radio environment no user would stay at a specific point or location, for that the dynamic formation of the RPM is more comfortable than statistical formation of RPM.

As we discussed above the statistical and dynamic location of RPM the RSSI vector  $s = (s_1, s_2, \dots, s_k)$  we found a position  $l \in L$  which maximizes the probability  $p(l/s)$ . As per Bayes theorem to find position of a user in the location  $l$  is by using this formula

$$P(l/s) = \frac{P(s/l)P(l)}{P(s)} = \frac{P(s/l)P(l)}{\sum_{l' \in L} P(s/l')P(l')} \quad (3)$$

This is for  $l \in L$  ie all possible positions in the RPM grid the sum goes through.  $P(s/l)$  indicates the probability to receive  $s$  from tx position at  $l$   $P(l)$  is the probability a priority of finding a tx in the position  $l$ . Let a wireless transmitter situated at  $l$  and  $k$  Sniffer will calculate a vector  $s = (s_1, s_2, \dots, s_k)$  and for each Sniffer  $P(l/s)$  is determined.

### IV. EXPERIMENTAL EVALUATION

As we discussed in the above sections, we prepared a setup of all four types of WLAN location determination architectures (A, B, C & D architectures) in our laboratory. We used DWL and NetGear network interface card which supports IEEE802.11a/b/g standards. In the institute premises we prepared a radio map in area of 1KM range. Two modes of WLAN Ad-Hoc and Infrastructures are prepared by using DWL and NetGear NICs and APs. In this section we are discussing and showing the differences between Indoor and outdoor locations and determining the location of the user with different positions.

Here we discussed indoor and outdoor location determination of a user in the wireless local area network, with IEEE802.11a/b/g by using DWL and NetGear NICs. The differences between a/b/g are tabulated in the table below in

the table-1. A dynamic mapping of RPM is somewhat difficult than statistical mapping. In dynamic mapping Sniffer has to search and scan the channel all the time in the form of “Refresh Always” method, without which it is not possible to see who is entering the network and who is leaving the network.

TABLE-1

Differences between IEEE802.11a/b and IEEE802.11g

| Properties  | 802.11a/b                                           | 802.11g                                                               |
|-------------|-----------------------------------------------------|-----------------------------------------------------------------------|
| Frequency   | 2.4Ghz / 5Ghz                                       | 2.4 GHz                                                               |
| Modulation  | Direct sequence spread spectrum (CCK, DQPSK, DBPSK) | Orthogonal frequency division multiplexing (64QAM, 16QAM, QPSK, BPSK) |
| Data rate   | 1,2,5.5,11Mbps                                      | 1,2,5.5,6,9, 11,12,18,18, 24,36, 48,54Mbps                            |
| Os platform | Windows Vista                                       | Windows XP                                                            |

Here Figure-3 is a Sniffer application which shows how many users are coming into the network and working in the network by Green indication as shown below, and whoever leaving the network by a Red Cross mark indication. The traffic can also be monitored by a statistics graph.

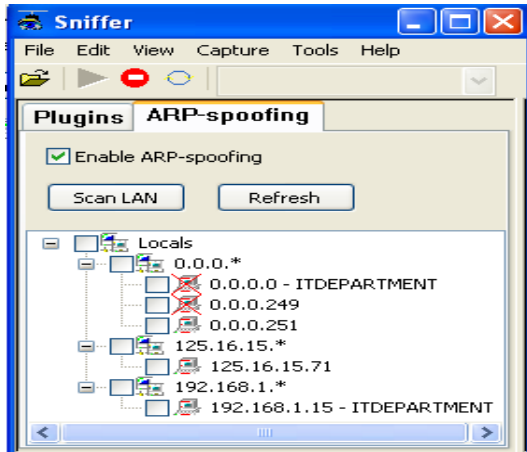


Fig. 3. Sniffer indications

After taking many iterations practically in Indoor and Outdoor locations by using IEEE802.11 a/b/g by connecting DWL and NetGear NICs at client side and server side, readings were recorded in the form of signal strength (RSSI). In indoor we found the location of the user somewhat slow when compared to the outdoor location and it is shown with the graphical notation below. One thing we can say and understand

that, we cannot carry and move AP dynamically from one place to another. For that reason we are proposing a statistical radio map location by equal division area spaces as shown below. We considered all four architectures and verified by taking RSSI indications at client and server sides in both indoor and outdoor environment. Figure-4 and Figure-5 show that the performances of indoor and outdoor environment and performance of Ad-hoc and infrastructure modes differences respectively in different positions i.e. 50mts, 75mts, 100mts, 125mts, 150mts and so on with RSSI indications Excellent, Very good, Good and Poor.

We have clearly seen location of a user in indoor and outdoor environment with Ad-hoc and Infrastructure modes. An infrastructure mode shows better results than Ad-hoc mode in both environments.

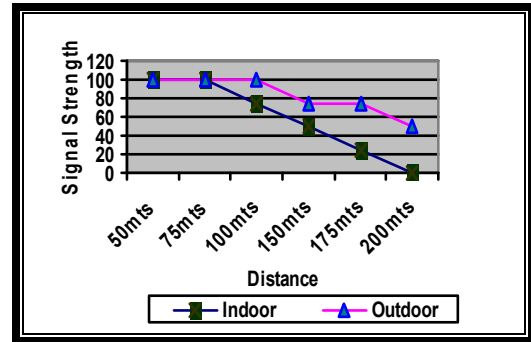


Fig. 4. Performance at Indoor and Outdoor environment

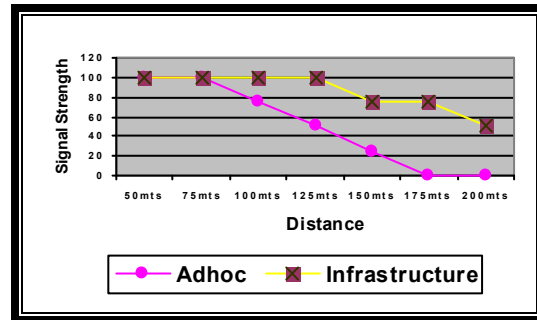


Fig. 5. Performance of Ad-hoc and Infrastructure modes

## V. RELATED AND FUTURE WORK

In this work we basically concentrated to determine location of a user in the environment of indoor and outdoor with dynamic location mapping. We found some difficulty in finding location of a user in statistical method, because of how to locate a position to an access point. Here we are using automatically connect and detect method to connect a system to the network. At the same time we are also focusing to form a radio map with equally divided area in indoor and outdoor locations. Finding an accurate position of a user we used the dynamic method by using Bayes theorem and tried to find the

correct position of the user in both the indoor and outdoor environment. In future we try to design a system, based on Bayes theorem in statistical environment. We are trying to test our work by using some other NIC's and with multiple locations. We also working on continuation of this work with cluster based calibration-free dynamic mapping of signal strength of a user in indoor and outdoor environment.

## VI. CONCLUSION

In this paper we discussed the system LOIDS which helps to locate and finds a user in Indoor and Outdoor environment with the help of dynamic location mapping. Here we developed four architectures and enabled WLAN strategies in those architectures in testbed. We tested all four architectures with different places in indoor and outdoor environment. In source destination architecture, it is somewhat difficult to continue the communication path while searching the location of a user, because of if any node in the sequence or Ad-hoc network is not having next node or station, this system may not work properly. In client server based, Sniffer based and Access point based architectures we found the considerable results shown in the graphical representations in indoor and outdoor environment with Ad-hoc and Infrastructure modes of operation.

## REFERENCES

- [1]Dr. M.N. Giri Prasad and K.Krishna Naik "IAPP for IEEE802.11 WLAN Range Improvements", I-Managers Journal on Software Engineering, Vol. 1 No. 4 April-June 2007.
- [2]K.Krishna Naik and B.M.G.Prasad "QoS in IEEE802.11 Wireless Local Area Networks with Enhanced MAC"- I-Manager's Journal on Engineering & Technology, vol.1 No.3, February 2006.
- [3]K.Krishna Naik "Antenna Design Analysis, Construction and Testing for IEEE802.11 WLAN Range Improvements"- M.Tech Thesis at IIIT-Allahabad-2004.
- [4]Neeli Prasad and Anand Prasad "WLAN System and Wireless IP for next generation communication"
- [5]IEEE 802.11b Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,

<http://standards.ieee.org/getieee802/802.11.html>

- [6]Luis Felipe M.de Moraes and Bruno Astuto A. Nunes "Calibration free WLAN location system based on dynamic mapping of signal strength" Mobiwac'06 october 2006.
- [7]Paramvir Bhal and Venkata N Padmanabhan" A software system for Locating Mobile Users: Design, Evaluation and Lessons".
- [8] P. Bahl, V.N. Padmanabhan, "RADAR: An In-Building RFbased User Location and Tracking System," *Proceedings of IEEE Infocom 2000, Tel Aviv, Israel*, March 2000.
- [9] P. Krishnan, A.S. Krishnakumar, Wen-Hua Ju, C. Mallows, S.Ganu, "A System for LEASE: System for Location Estimation Assisted by Stationary Emitters for Indoor RF Wireless Networks", *Proceedings of IEEE Infocom 2004, Hong Kong, March 2004*.
- [10] Moustafa Youssef, Ashok Agrawala, A. Udaya Shankar, "WLAN Location Determination via Clustering and Probability Distributions," *IEEE Intl. Conf. on Pervasive Comp. andComm. (PerCom) 2003*, March 23-26, 2003.
- [11] W. Stallings. *Wireless Communications and Networks*. Prentice Hall, first edition, 2002

# Evaluation of Traffic Parameters of Multidimensional Traffic of a Combined Link Using a Tabular Method

Md. Imdadul Islam<sup>1</sup>, Md. Shahriar Karim<sup>2</sup>, J. K. Das<sup>3</sup> and M. R. Amin<sup>4</sup>

<sup>1,3</sup>Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh, e-mail: imdad@juniv.edu

<sup>2</sup>Member, IEEE, Support Engineer, Core Network, Ericsson Bangladesh Ltd., e-mail: shahriar.karim@ericsson.com

<sup>4</sup>Department of Electronics and Communications Engineering, East West University, Mohakhali, Dhaka 1212, Bangladesh, e-mail: ramina@ewubd.edu

**Abstract** — In many communication systems, two or more traffic links are merged together to form a common trunk line. Usually, a Markovian technique is used to determine the different traffic parameters of aforementioned combined trunk line. But the technique does not hold good and also the Markov chain takes a complex shape for a network where offered traffics are more than two different types. In this paper, we propose a tabular method instead of a Markovian technique for calculating various traffic parameters of a network with multimedia traffic, applicable in mobile cellular networks. It is found in this work that the proposed tabular method overcomes the complexities of the Markovian approach and yields similar results.

**Keywords** — Call blocking, combined trunk, Markov chain, M/M/n traffic model, multimedia traffic.

## I. INTRODUCTION

Penetration of wireless mobile cellular networks has been tremendous in the last decade, and this will continue in the coming years as well. To remain in harmony with this ever increasing user demand, wireless mobile cellular networks have already been shaped to provide multiple services ranging from normal voice calls to image and video transmission. The intermittent handoff in mobile cellular networks has added another dimension to it. As a result, in recent days, for analyzing the quality of service (QoS) issues for mobile cellular networks, a three-dimensional (3D) traffic scenario is being thought to be the appropriate way to cover all incoming traffics to the network. Generally, probability state transition or Markovian chain is used to design limited and unlimited source traffic of various dimensions. To solve such a chain, cut or node equations are widely used [1]-[3].

In this paper, a model is proposed and analyzed further to realize the same in a normalized tabulated form. For an  $m$ -dimensional traffic ( $m \geq 3$ ), state transition diagram takes a complex shape and also, the large number of

equations for the chain causes additional complexity to gain detail information pertaining to the process. We suggest an alternative way here for the special case of 3D offered traffic system to avoid the complex calculations of state transition or the Markovian chain method. Here in the system, we assume that the offered traffic follows an exponential arrival and service time distribution in a limited server link, i.e. M/M/n traffic. The concept of tabular solution of two-dimensional (2D) Poisson's or Erlang's distribution is adopted here, and normalization of table is done after eliminating unreachable state of combined traffic link, based on the method discussed in [4]-[6]. Though this paper reveals a tabular solution of a three input traffic system, it would become a lengthy work for more than three input trunks or servers. However, it is worth to mention here that the proposed model can be extended for any number of offered traffic cases.

The plan of the paper is as follows. The proposed tabular method is explained in Section II. Section III of the paper deals with the methodology of solution of a 3D offered traffic of a trunk line along with results of the model. In this section, blocking probability of trunk is also represented graphically against offered traffic. Finally, Section IV concludes the paper.

## II. THE PROPOSED MODEL

In this section we describe our model to determine the normalized probability states or QoS of a mobile cellular network in a direct form. One of the convenient ways of presenting a 2D traffic case is in triangular matrix form, where each element of the matrix is obtained by multiplication of specific terms from two series of two-offered traffics.



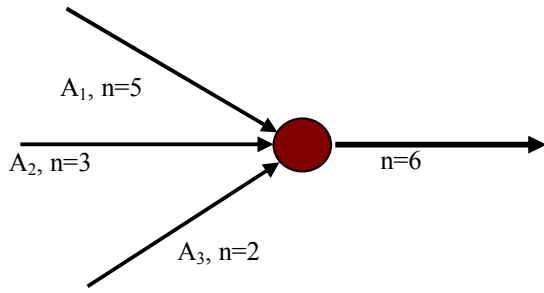


Fig. 1. Synthesis of Three Offered Traffic

Let us first consider that two out of three links are selected to achieve the generalized tabular solution of a 2D traffic situation. Here  $A_1$  and  $A_2$  are selected first and shown in Table 1 after analyzing them based on the method described in [6], [7]. The entries of Table 2 is then multiplied by 1,  $A_3$  and  $A_3^2/2!$ , and at the same time the number of servers is reduced by 0, 1 and 2, respectively. The resulting tables are shown as Tables 3a, 3b, and 3c. The entries of the tables are then normalized in combined forms which are shown as Tables 4a, 4b, and 4c. Call blocking is the sum of dark cells according to the concept given in [6]-[8].

TWO DIMENSIONAL TRAFFIC TABLE OF  $AB_{1B}$  AND  $AB_{2B}$  BEFORE NORMALIZATION

TABLE 2

| $n_1$ | $A_1$    |          |          |          |          |       |
|-------|----------|----------|----------|----------|----------|-------|
| 5     | 0.266667 | 0.266667 | 0.266667 | 0.133333 | 0.044444 |       |
| 4     | 0.666667 | 0.666667 | 0.666667 | 0.333333 | 0.111111 |       |
| 3     | 1.333333 | 1.333333 | 1.333333 | 0.666667 | 0.222222 |       |
| 2     | 2        | 2        | 2        | 1        | 0.333333 |       |
| 1     | 2        | 2        | 2        | 1        | 0.333333 |       |
| 0     | 1        | 1        | 1        | 0.5      | 0.166667 |       |
|       |          | 1        | 1        | 0.5      | 0.166667 | $A_2$ |
|       |          | 0        | 1        | 2        | 3        | $n_2$ |

TABLE 3a:  $m=0$

|       |          |          |          |          |          |
|-------|----------|----------|----------|----------|----------|
| 5     | 8.533333 | 0.266667 | 0.266667 |          |          |
| 4     | 10.66667 | 0.666667 | 0.666667 | 0.333333 |          |
| 3     | 10.66667 | 1.333333 | 1.333333 | 0.666667 | 0.222222 |
| 2     | 8        | 2        | 2        | 1        | 0.333333 |
| 1     | 4        | 2        | 2        | 1        | 0.333333 |
| 0     | 1        | 1        | 1        | 0.5      | 0.166667 |
|       |          | 1        | 5        | 12.5     | 20.83333 |
| $m=0$ |          | 0        | 1        | 2        | 3        |

TABLE 3b:  $m=1$

|       |          |          |          |          |          |
|-------|----------|----------|----------|----------|----------|
| 5     | 8.533333 | 1.333333 |          |          |          |
| 4     | 10.66667 | 3.333333 | 3.333333 |          |          |
| 3     | 10.66667 | 6.666667 | 6.666667 | 3.333333 |          |
| 2     | 8        | 10       | 10       | 5        | 1.666667 |
| 1     | 4        | 10       | 10       | 5        | 1.666667 |
| 0     | 1        | 5        | 5        | 2.5      | 0.833333 |
|       |          | 1        | 5        | 12.5     | 20.83333 |
| $m=1$ |          | 0        | 1        | 2        | 3        |

TABLE 3c:  $m=2$

|       |          |          |          |      |          |
|-------|----------|----------|----------|------|----------|
| 5     | 8.533333 |          |          |      |          |
| 4     | 10.66667 | 8.333333 |          |      |          |
| 3     | 10.66667 | 16.66667 | 16.66667 |      |          |
| 2     | 8        | 25       | 25       | 12.5 |          |
| 1     | 4        | 25       | 25       | 12.5 | 4.166667 |
| 0     | 1        | 12.5     | 12.5     | 6.25 | 2.083333 |
|       |          | 1        | 5        | 12.5 | 20.83333 |
| $m=2$ |          | 0        | 1        | 2    | 3        |

TABLE 4a:  $m=0$

|          |          |          |          |
|----------|----------|----------|----------|
| 0.000848 | 0.000848 |          |          |
| 0.002119 | 0.002119 | 0.00106  |          |
| 0.004238 | 0.004238 | 0.002119 | 0.000706 |
| 0.006358 | 0.006358 | 0.003179 | 0.00106  |
| 0.006358 | 0.006358 | 0.003179 | 0.00106  |
| 0.003179 | 0.003179 | 0.001589 | 0.00053  |

TABLE 4b:  $m=1$

|          |          |          |          |
|----------|----------|----------|----------|
| 0.004238 |          |          |          |
| 0.010596 | 0.010596 |          |          |
| 0.021192 | 0.021192 | 0.010596 |          |
| 0.031788 | 0.031788 | 0.015894 | 0.005298 |
| 0.031788 | 0.031788 | 0.015894 | 0.005298 |
| 0.015894 | 0.015894 | 0.007947 | 0.002649 |

TABLE 4c:  $m=2$

|          |          |          |          |
|----------|----------|----------|----------|
| 0.02649  |          |          |          |
| 0.052979 | 0.052979 |          |          |
| 0.079469 | 0.079469 | 0.039734 |          |
| 0.079469 | 0.079469 | 0.039734 | 0.013245 |
| 0.039734 | 0.039734 | 0.019867 | 0.006622 |

Probability states and blocking probability for individual Tables 4a, 4b and 4c can be evaluated mathematically by applying equations (1)-(4) easily according to [7], [9]-[10]. However, it is not convenient to evaluate ultimate traffic parameters from the combination of the above mentioned three tables based on equations (1) and (2). Hence our tabular solution is preferable for multidimensional traffic, especially, for dimension more than 2.

For example, let us solve the case where  $A_1 = 2$  Erlangs,  $A_2 = 1$  Erlang, and  $A_3 = 5$  Erlangs for  $n_1 = 5$  trunks,  $n_2 = 3$  trunks,  $n_3 = 2$  trunks, and  $n = 6$  trunks. After solving the network, blocking probability comes out to be  $B = 0.18385$ . The numerical value of carried and lost traffic from equations (5) and (6) is evaluated as 6.52916 Erlangs and 1.47084 Erlangs respectively. Blocking probability is also depicted graphically in Fig. 2 against offered traffic  $A_1$  and  $A_3$ .

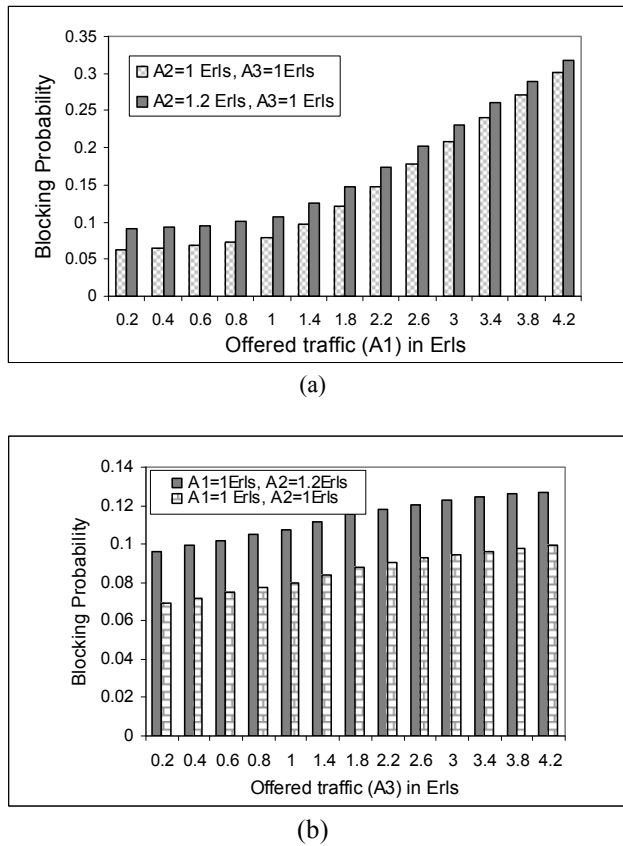


Fig. 2. Variation of call blocking probability against (a) offered traffic  $A_1$ , (b) offered traffic  $A_3$ .

It is noticed from the graphs of Fig. 2 that the blocking probability against  $A_1$  is greater than the blocking probability against  $A_3$ . The traffic  $A_3$  is offered to a link of only two trunks (Fig. 1). Hence, very little carried traffic will reach to the combined trunk line. On

the contrary,  $A_1$  traffic is offered to a link of five trunks. So, less traffic will be lost and almost all of the offered traffic will reach to the combined trunk link  $n$ . Hence, offered traffic  $A_1$  will dominate on call blocking of the combined trunk link compared to  $A_3$ . These are also clearly visualized from the graphs shown in Fig. 2.

#### IV. CONCLUSION

An approach of tabular method is described in this paper for analyzing a multidimensional traffic system of calculating various traffic parameters of a network with multimedia traffic, applicable particularly in the mobile cellular systems. It is found that the tabular method described here overcomes the complexities of the Markovian approach particularly for more than 2D traffic cases. In this paper, the analysis is done for three different offered traffic of similar bandwidth. However, some networks, for example, 3G mobile systems of voice/data integrated services, where voice and data traffic have different bandwidth, the analysis can easily be carried out based on our proposed model described in this paper by using the bandwidth sharing scheme (viz. Eq. (5)).

#### REFERENCES

- [1] G. Dave Morley and Wayne D. Grover, "Strategies to maximize Carried Traffic in Dual-Mode Cellular Systems," *IEEE Trans. on Veh. Tech.*, vol. 49, no. 2, pp. 357-366, March 2000.
- [2] Giselle M. Galvan-Tejada and John G. Gardiner, "Theoretical Model to Determine the Blocking Probability for SDMA Systems," *IEEE Trans. on Veh. Tech.*, vol. 50, no. 5, pp. 1279-1288 September 2001.
- [3] Giselle M. Galvan-Tejada and John G. Gardiner, "Theoretical Blocking Probability for SDMA", *IEE Proc-Commun*, vol.146, no. 5, pp. 303-306, October 1999.
- [4] M. I. Islam and S. Hossain, "Analysis of two dimensional limited source and mixed traffic model for a BTS of a small mobile cellular network", *International Conference on Computer and Information Technology (ICCIT)*, D.U. pp. 190-195, Dhaka, December 2001.
- [5] T. Erke, *Tele-traffic Engineering, Course Lecture Notes*, Telecommunication Program, School of Advanced Technology, AIT, Bangkok, Thailand, 2000.
- [6] D. Bear, *Principle of Telecommunication Traffic Engineering*, London: Peter Peregrinus Ltd., 2<sup>nd</sup> Ed., 1988.



- [7] P. Srisukpibul, *Traffic Measurement on a Cellular Mobile Network*, Masters Thesis, Division of Telecommunications, AIT, Bangkok, Thailand, 1999.
- [8] Shahram Ghaheri Niri, Rahim Tafazoli, "Uniform GoS in a Hierarchical Cell Structure" *The third International Symposium on WPMC*, Bangkok, Thailand, November 2000.
- [9] M. I. Islam and S. Hossain, "A proposed 2-D queuing model of PCT-I traffic" in the *Proceedings of the 6P<sup>th</sup> International Conference on Computer and Information Technology (ICCIT)*, J.U. pp.114-118, Dhaka, December 2003.
- [10] M. I. Islam and S. Hossain, "An analytical model of traffic performance of mobile cellular network in underlay overlay cell system" in the *Proceedings of the 6P<sup>th</sup> International Conference on Computer and Information Technology (ICCIT)*, J.U., Dhaka, pp.230-234, December 2003.

# Analysis and Synthesis of Bangla Dental Consonants

Syed Akhter Hossain, M Lutfar Rahman\* and Farruk Ahmed\*\*

Department of Computer Science and Engineering, University of Liberal Arts Bangladesh

\*Department of Computer Science and Engineering, University of Dhaka

\*\*Department of Computer Science and Engineering, North South University  
Dhaka, Bangladesh

**Abstract** — This paper describes in detail the analysis results of the Bangla Dental consonants /ত/ [t], /থ/ [th], /দ/ [d] and /ধ/ [dh] analyzed through the spectral estimation techniques in four long vowel contexts /আ/ [a], /ই/ [i], /ও/ [o], and /উ/ [u] for an accurate description of their acoustic characteristics and features and the differences between the corresponding cognate sounds. Various parameters like duration of closure/voice bar, duration of burst, voice onset time, duration of aspiration, rate of second formant transition and burst frequencies and amplitudes have been studied in details. For the synthesis of dental sounds, the source and vocal tract parameters of the synthesizer configuration were selected very carefully along with formant frequencies and their relative amplitudes besides the burst amplitude. The quality of synthetic speech was evaluated through subjective listening along with matching the spectra of synthetic speech with original speech.

**Index Terms** — Speech Processing, Synthesis, LPC, Phoneme.

## I. INTRODUCTION

Bangla is an Indo-Aryan language of more than 300 million people in the world concentrating in the eastern region of Indian subcontinent i.e. Bangladesh, Indian states of West Bengal, Tripura, Assam and around. The history of Bangla begins in the late centuries and before that there was only a family of dialects commonly known as Prakrit [1].

Every language has its own set of basic sounds, which constitute the language. Bangla speech contains a set of eleven vowels with eight being the cardinal vowels and 39 consonants, of which about 30 consonants are of frequent usage. The vowels can be pronounced independently, but the consonants need the support of vowels to be pronounced. These consonants can be conveniently classified according to the manner and place of articulation [1].

The Bangla consonants possess certain special features, which are not so common to other European languages and American English. The major discrepancy results from a difference in the number of places at which the stops can be formed by the tip of the tongue making contact with the front portion of the top of the mouth. Consonants differ from vowels in that they had more

energy in the high frequency region compared to the low frequency region. Bangla linguistics classify consonants on the basis of their place of articulation. In fact in Bangla, there are five ways distinction in the place of articulation for the stops sounds, compared to English where there is only four way contrast [2].

Dental sounds involve the upper teeth as the passive articulator. The active articulator may be either the tongue tip or tongue blade and generally diacritic symbols are used to represent the class of dental sounds. In Bangla, the dental sounds are classified in two categories: dental and the labio-dental. In case of dental sounds /ত/ [taa], /থ/ [tha], /দ/ [daa], /ধ/ [dha] – the tip of the tongue is against the back of the upper teeth, or even further forward. In labio-dental sounds /ফ/ [fa], /ব/ [va] – the sound requires the participation of one or both lips and dental ridge. The constriction is formed between lips and dental ridge to generate labio-dental sounds [1-3].

In this work, we investigated spectral properties of Bangla dental sounds from 1800 speech corpora containing vowel-consonant-vowel (VCV) and consonant-vowel-consonant (CVC) syllables produced by male and female speakers of different age groups. The revealed spectral properties of the dental sounds along with the burst amplitude and duration is used in the synthesis of dental sounds. The synthetic sound and spectra was evaluated. The rest of the paper is organized as follows: Section II discusses the analysis procedure and Section III discusses the synthesis procedure with experimental results and discussion in Section IV.

## II. ANALYSIS PROCEDURE

All the consonants were recorded in the form of CVC and VCV syllables spoken by male and female native Bangla speakers in all four vowel contexts /আ/ [a], /ই/ [i], /ও/ [o], and /উ/ [u] at the sampling rate of 16 kHz. These CVC and VCV syllables were then broken into CV and VC syllables for the entire analysis through spectral techniques. The starting and ending times of the vocalic nuclei was measured by hand from high-resolution digital spectrograms. Only the first 10 ms of the release burst or the whole burst was considered.

The spectral analysis of these syllables were carried out using digital spectrograms and other techniques like short time FFT and LPC spectra [4], pitch, formants, waveforms and envelope displayed together. Using the spectrogram most of the important acoustical events were studied in time as well as frequency domain. Duration and sequences of events in each sound/syllable was noted down using time scale of the spectrogram. Unlike other sounds, which can be described largely in terms of steady-state spectra, stops are transient phonemes and thus are acoustically complex. Bangla consonant-vowel (CV) and vowel-consonant (VC) syllables of the type stop plus vowel consist of at least four phonetic segments, viz. the closure or voicebar, burst, voice onset time (VOT), aspiration and a voiced interval for vowel [5, 6].

The burst frequency is the frequency at which the maximum amplitude in the frequency spectrum occurred. The burst frequency was estimated from the frequency spectrum of the release burst. Burst frequency measurements were made using a 20-pole LPC spectrum.

### III. SYNTHESIS PROCEDURE

Bangla dental sound segments were synthesized on the basis of extracted spectral parameters achieved through analysis. Based on the analyzed data, synthetic CV tokens were created from a number of source and tract parameters that were adjusted iteratively in order to achieve a close imitation to natural syllables. These include the source parameters such as amplitude of voicing, frication and aspiration (AV, AF, and AH respectively), Open Quotient (OQ) and vocal tract parameters like formant frequencies, amplitudes, and their respective bandwidths which was fed into the Klatt formant synthesizer [7, 8].

Same formant frequencies of burst (and transitions) were used for consonants belonging to same place of articulation. For example same values of first four formants were used for all four consonants /ত/ [t], /থ/ [th], /দ/ [d] and /ধ/ [dh] in the dental group for a given vowel context.

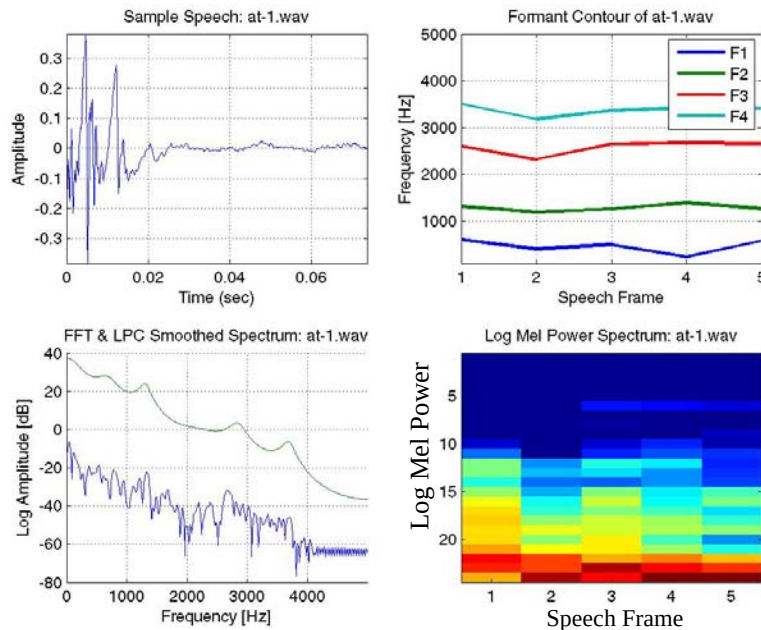


Figure 1: Analysis of Bangla speech segment /আত/ [aat] from the Bangla corpus  
আতা [ata] produced by a male speaker

A half-Hamming window of 20 ms was used that is (only the latter half of the Hamming window was multiplied with the 10 ms of the release burst samples). The LPC spectrum was obtained using a 512-point FFT. The frequency of the maximum amplitude of spectral peak was extracted from LPC spectrum using a global peak picking algorithm. Once this acoustic analysis was done for the clean consonants, the frequencies at which the peaks for the consonants occurred were computed and plotted [6].

### IV. EXPERIMENTAL RESULTS AND DISCUSSION

The spectrogram of Bangla speech corpus আতা [ata] produced by male speaker is shown in the Figure 1 along with the consonant segment of speech corpus আত [aat] in the time domain with other spectral measurements. As can be seen from the spectrogram, at the beginning of the phonation আ [aa], a small voice bar is visible and then followed by a set of formants during the production of the phonation আ [aa]. During the transition phase from the

vowel segment to the consonant segment for ত [ta] and from consonant to vowel segment v [diacritics] a burst frequency range is visible without any noticeable formant frequencies.

As can be seen from the Figure 1, the speech segment /আত/ [aat] is displayed on the top frame and the segment length is about 65msec. The FFT and LPC smoothed spectrum is captured in another frame whereas frame wise formant contour and log Mel power spectrum of the same speech segment is captured in another two frames.

The formant contour clearly demonstrates the transition of formant frequency due to consonant segment /ত/ [ta]. The first formant F1 tend to increase due to the closure of the oral tract. The FFT and LPC smoothed portion of the spectrum shown is the last frame of the speech segment under analysis.

TABLE I  
COMPARISONS OF BURST FREQUENCIES OF DENTAL  
SOUNDS

(a) /আ/ [a] context

| Formant | /ত/ [t] | /থ/ [th] | /দ/ [d] | /ধ/ [dh] |
|---------|---------|----------|---------|----------|
| F1 (Hz) | 600     | 950      | 380     | 750      |
| F2 (Hz) | 1450    | 2000     | 1700    | 1900     |
| F3 (Hz) | 2700    | 2900     | 2700    | 3000     |
| F4 (Hz) | 3400    | 3900     | 3350    | 3850     |

(b) /ই/ [i] context

| Formant | /ত/ [t] | /থ/ [th] | /দ/ [d] | /ধ/ [dh] |
|---------|---------|----------|---------|----------|
| F1 (Hz) | 400     | 450      | 350     | 450      |
| F2 (Hz) | 2150    | 1850     | 2150    | 2150     |
| F3 (Hz) | 2550    | 2950     | 2900    | 2800     |
| F4 (Hz) | 3800    | 4150     | 3250    | 3750     |

(c) /উ/ [u] context

| Formant | /ত/ [t] | /থ/ [th] | /দ/ [d] | /ধ/ [dh] |
|---------|---------|----------|---------|----------|
| F1 (Hz) | 550     | 450      | 375     | 350      |
| F2 (Hz) | 1600    | 1750     | 1650    | 1600     |
| F3 (Hz) | 2750    | 3150     | 2750    | 3000     |
| F4 (Hz) | 3850    | 3550     | 3450    | 3950     |

(d) /ও/ [o] context

| Formant | /ত/ [t] | /থ/ [th] | /দ/ [d] | /ধ/ [dh] |
|---------|---------|----------|---------|----------|
| F1 (Hz) | 550     | 500      | 250     | 450      |
| F2 (Hz) | 1500    | 1900     | 1725    | 1450     |
| F3 (Hz) | 2850    | 2550     | 2550    | 2850     |
| F4 (Hz) | 3550    | 3950     | 3950    | 3900     |

In English, in the production of /t/ and /d/ sounds, the tip of the tongue comes in contact with the alveolar ridge while in Bangla, there is a mixed alveolar sounds and it has either dental consonants or retroflex. These sounds are produced by the tip of the tongue in interaction with the back part of the alveolar ridge.

As can be seen from the Table 1, the place of articulation and the vowel context determine the burst frequencies of dental sounds. The burst and formant transition of the target vowel varies across the stop consonants. The same stop consonant shows variation in the formant frequency transition in different vowel context.

The vowel context /ই/ [i] shows significant difference in formant transition between first formant F1 and second formant F2 for all the retroflex sounds of Bangla. The formant transition between the second and third formant frequency F2 and F3 is consistent throughout different dental sounds. Besides, the context /উ/ [u] and /ও/ [o] shows persistent formant transitions for different retroflex sounds as can be seen from the Table 1.

The relative burst amplitudes of various dental sounds of Bangla in different vowel contexts were measured and all the measurements were recorded in the Table 2 below.

Table 2 postulate the relative burst amplitude of dental sounds in different vowel contexts. As can be seen from the Table 2, the consonant /ত/ [t] in the vowel context /আ/ [aa] shows relatively higher burst amplitude for second formant frequency compared to higher formant frequencies. While the consonant /ত/ [t] in the vowel context /ই/ [i], /উ/ [u] and /ও/ [o] shows relatively higher amplitude for the burst for higher formant frequency compared to lower formant frequencies. The consonants /থ/ [th], /দ/ [d], and /ধ/ [dh] shows higher relative burst amplitude in the vowel context /আ/ [aa] and /ই/ [i].

The formant frequency transition of various dental sounds of Bangla in different vowel contexts were measured through the formant contour observation and all the measurements were recorded in Table 3 as shown below.

As can be seen from the Table 3, for all the dental sounds of Bangla, the first formant frequency F1 transits from high to low during the movement towards the vowel segment and to high during the retraction from the vowel segment. In case of dental sounds, the second formant frequency is steady during the transition but the third formant frequency transits from low to high and the fourth formant frequency transits from high to low while the segment moves to and from the vowel. In case of dental sounds as can be seen from the Table 3, the fourth formant frequency is steady during the transition but the second formant frequency transits from low to high and the third formant frequency from high to low during movement towards the vowel and to high while the segment moves away from the vowel.

TABLE II  
RELATIVE BURST AMPLITUDES OF RETROFLEX SOUNDS IN VOWEL CONTEXTS  
(WITH RELATIVE LEVEL DIFFERENCE; A2F: F2 AMPLITUDE; A3F: F3 AMPLITUDE; A4F: F4 AMPLITUDE)

| Consonant | /আ/ [aa]    | /ই/ [i]     | /উ/ [u]     | /ও/ [o]     |
|-----------|-------------|-------------|-------------|-------------|
| /ত/ [t]   | A2F>A3F>A4F | A4F>A3F>A2F | A4F>A2F>A3F | A4F>A2F>A3F |
| /থ/ [th]  | A4F>A2F>A3F | A4F>A2F>A3F | A2F>A3F>A4F | A3F>A4F>A2F |
| /দ/ [d]   | A4F>A2F>A3F | A3F>A4F>A2F | A2F>A3F>A4F | A4F>A3F=A2F |
| /ধ/ [dh]  | A4F>A2F>A3F | A3F>A4F>A2F | A2F>A3F>A4F | A2F>A3F>A4F |

TABLE III  
FORMANT TRANSITION OF DENTAL SOUNDS TO AND FROM THE VOWEL SEGMENT

| Consonant | F1                | F2     | F3       | F4       |
|-----------|-------------------|--------|----------|----------|
| /ত/ [t]   | High – low – high | Steady | low-high | high-low |
| /থ/ [th]  | High – low – high | Steady | low-high | high-low |
| /দ/ [d]   | High – low – high | Steady | low-high | high-low |
| /ধ/ [dh]  | High – low – high | Steady | low-high | high-low |

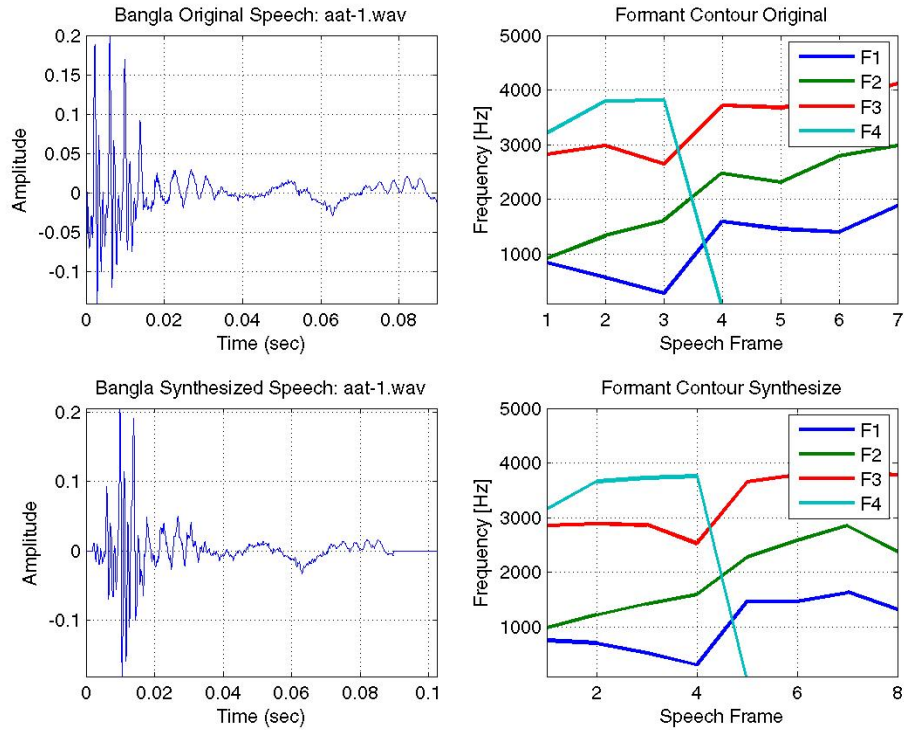


Figure 2: Formant contour of original and synthetic Bangla speech segment আত [aat] of the Bangla word আতা /ata/ produced by a male speaker

The original speech segment of speech corpora /আত/ [aat] is shown with the synthesis counterpart is Figure 2 along with spectral measurements. As can be seen in the Figure 2, the original speech segment /আত/ [aat] containing Bangla consonant /ত/ [t] in the word /আত/ [ata] has duration of about 85 msec. The formant contour is shown on the upper right frame. The formant contour shows the formant transition. The synthesized speech segment is shown on the lower left frame of the display which has almost the same spectral characteristics as well as the duration of the segment. The synthesized segment shows the formant contour same as the original segment with identical formant transition characteristics.

TABLE IV  
FORMANT FREQUENCIES OF SYNTHESIZED BANGLA  
DENTAL SOUNDS IN /আ/ [a] CONTEXT

| Formant | /ত/ [t] | /থ/[th] | /দ/ [d] | /ধ/ [dh] |
|---------|---------|---------|---------|----------|
| F1 (Hz) | 580     | 850     | 400     | 820      |
| F2 (Hz) | 1460    | 1900    | 1580    | 1780     |
| F3 (Hz) | 2690    | 2890    | 2550    | 3150     |
| F4 (Hz) | 3400    | 3490    | 3450    | 3980     |

As can be seen from the Table 4, the formant frequencies of synthesized dental speech segment in vowel context /আ/ [a] follows identical variation compared to the original speech segment.

## V. CONCLUSION

Bangla dental consonants are analyzed for spectral representations and the results are postulated in Table 1 through Table 3 and some of the results are shown in Figure 1 and Figure 2. It has been observed at the burst position that third and the fourth formant contains most of the energy in case of dental sounds while formants four contained most of the energy. Also there is general lowering of the third and fourth formants in case of dentals.

It has also been observed that the bandwidths of the formants are narrower (i.e. the formant peaks are better

defined) in the case of voiced aspirated sounds /দ/ [d], /ধ/ [dh], /ড/ [d], and /ঢ/ [dh] as compared to unvoiced aspirated sounds /ত/ [t], /থ/ [th], /ট/ [t], and /ঠ/ [th]. The formant transitions from burst frequency to the target vowel frequency are part of the aspiration segment. Aspiration duration is about 90-110 ms in case of retroflex sounds, while the duration is 80-100 ms for dental sounds.

Burst frequencies depend on the place of articulation of the stop as well as the vowel context. The spectral difference among the dental consonants across the places of articulation is primarily reflected in the spectra of the burst and formant transitions of the target vowel. The burst frequencies of the same consonants do show slight variation in different vowel contexts. In a very good quality of synthetic sounds, burst frequencies and relative amplitudes along with overall burst intensity plays a major role as the burst frequencies of synthesized retroflex and dental sounds in Bangla vowel context /আ/ [a] is shown in Table 4.

## REFERENCES

- [1]. M.Abdul Hai, *Bengali Language Handbook*, Center for Applied Linguistics, Washington D.C., 1966
- [2]. R. Islam, *An Introduction to Colloquial Bengali*, Chapter 1, Central Board for Development of Bengali, Dhaka, 1970
- [3]. Md. Abdul Hai, *Dhvani Vijnan O Bangla Dhvani-Tattwa (Bangla)*, Mullick Brothers, 2000
- [4]. S. Blumstein and K. Stevens, "Perceptual invariance and onset spectra for stop consonants in different vowel environments," *J. Acoust. Soc. Am.*, vol. 67, pp. 648-662, 1980
- [5]. J.Makhoul, "Spectral Analysis of Speech by Linear Prediction," *IEEE Trans. on Audio and Electroacoustics*, Vol. AU-21, No.3, pp. 140-148, June 1973
- [6]. M. Dorman and P. Loizou, "Relative spectral change and formant transitions as cues to labial an alveolar place of articulation," *J. Acoust. Soc. Am.*, vol. 100, pp. 3825-3830, 1996
- [7]. K. Kleijn, K. Paliwal, *Speech Coding and Synthesis*. Elsevier Science B.V., The Netherlands, 1998
- [8]. Klatt D. H., "Software for a cascade parallel formant synthesizer", *J.Acoust.Soc.Am.*, 67(3), 971-995, 1980

# AN APPROACH TO RECOGNIZE HANDWRITTEN BENGALI NUMERALS FOR POSTAL AUTOMATION

Md. Saidur Rahman, G. M. Atiqur Rahaman, Asif Ahmed and G. M. Salahuddin

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh

E-mail: msrku@yahoo.com, atiq@ku.ac.bd, asif\_cseku@yahoo.com, gms\_far@yahoo.com

**Abstract** — A recognition approach for handwritten Bengali numerals and its application for Bangladesh Postal system are presented in this paper. The approach consists of preprocessing, feature extraction and recognition. Using the canny edge detector the top left corner of the post-code box has been detected and the four handwritten numeral images have been segmented. Each image then goes through the steps of normalization, filtering and thinning. Kirsch mask has been used to get the edges through the horizontal, vertical, right and left diagonal. Principal component analysis (PCA) has been used for dimension reduction as the final feature vector consists of the four directional feature vector along with the normalized image. Then the output of the PCA is passed to a trained Support Vector Machine (SVM) to determine which class the input belongs to. Experiments demonstrate that the average recognition rate, error rate and reliability achieved by the proposed system are 92.5%, 7.5% and 92.5% respectively.

**Index Terms** — Bengali Numeral, Feature Extraction, Kirsch Mask, Principal Component Analysis, Support Vector Machine.

## I. INTRODUCTION

Bengali is the second most popular language in South Asia. But unfortunately we have not yet got the appropriate application towards handwritten Bengali numeral character recognition system as most of the research works have been on retrieving content of printed document [1]-[3]. But for our postal automation a system is really necessary that can recognize Bengali numerals from the envelope. Most of the countries in the world already have been developed the system for their postal automation. Postal automation is a topic of research interest for last two decades and many pieces of published article are available towards postal automation of non-Bengali languages documents [4]-[7]. Several systems are also available for postal automation in USA, UK, France, Canada and Australia. But very few works have been done towards the automation of Bangladeshi postal system [8].

Handwritten character recognition is highly subjective as different people have different art of writing the same numeral character. That is why to get the actual numerals from handwriting is difficult in real time. So the main objective of our work is to recognize handwritten Bengali numerals from the envelope of Bangladesh Postal system to

get the post code with high recognition rate, less error rate, high reliability and comparatively less response time.

## II. LITERATURE SURVEY

Various approaches have been proposed by the researchers for numeral recognition [9] but most of them are used for non-Indian numeral recognition. One of the most widely used approaches is based on neural networks [10]. Here the network architecture is trained by a set of training data and then the input is classified by the trained networks. In structural approach each pattern class is defined by structural description and the recognition is performed according to structural similarities [11]. Statistical approach is also applied to numeral recognition. It is relatively insensitive to pattern noise and distortion but modeling of statistical information is a tedious task [9]. Among others, Hidden Markov Models, Fourier and Wavelet Descriptors [12], Fuzzy rules [13], tolerant rough set [14], are reported in the literature.

Feature extraction plays a vital role for a recognition system. Principal component analysis is a well-known method for feature extraction. By calculating the eigenvectors of the covariance matrix of the original inputs, PCA linearly transforms a high-dimensional input vector into a low-dimensional one whose components are uncorrelated [15]-[17]. Recently, support vector machine (SVM) has become a popular tool for numeral recognition due to its remarkable characteristics such as good generalization performance, the absence of local minima. SVM implements the structural risk minimization principle which seeks to minimize an upper bound of the generalization error rather than minimizing the training error [18].

In this paper, we present a recognition system for handwritten Bengali numerals. The system consists of three main parts: preprocessing, feature extraction and recognition. The input of the system is the image of an envelope and the output is the recognized Bengali numerals of the post-code written upon the envelope. Here we propose a recognition approach based on the direction feature combined with PCA and SVM. The experimental

results show that the recognition performance of the proposed approach is superior to other conventional recognition approaches.

### III. MATERIALS AND METHODS

The overall system is represented in Fig.1. Here the preprocessing phase consists of post-code box detection, segmentation, normalization, filtering, thinning and dilation. In the dimension reduction phase Kirsch mask and the PCA method are applied. Then the SVM is trained up. And in the recognition phase the numeral characters are recognized.

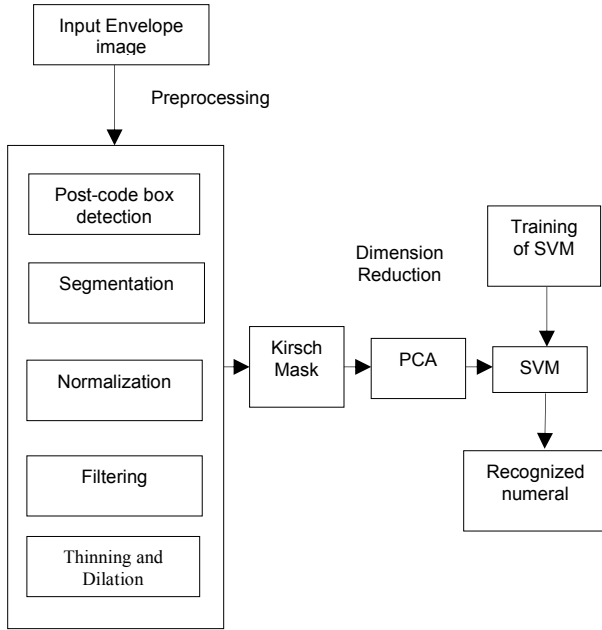


Fig. 1. Block diagram of the proposed system

#### A. Preprocessing

The goal of this part is to extract numeral images of postcode from an envelope image for the subsequent recognition. In Bangladeshi postal documents there are pre-printed boxes to write post-code. We call these boxes as post-code boxes. People generally write the destination post-code of four digits inside these boxes.

**Detection of Post-code Box and Segmentation:** Getting an image of envelope as input, the first task is to locate the post-code box in the envelope. To serve the purpose at first the true color image has been transformed into a binary image and then the edge detection procedure has been applied on it. For edge detection the canny method [19] has been used. The Canny method finds edges by looking for local maxima of the gradient of binary image. The gradient is calculated using the derivative of a Gaussian filter. The

method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This method is therefore less likely than the others to be "fooled" by noise, and more likely to detect true weak edges. The Fig.2 shows an example of the effect of applying Canny Method.

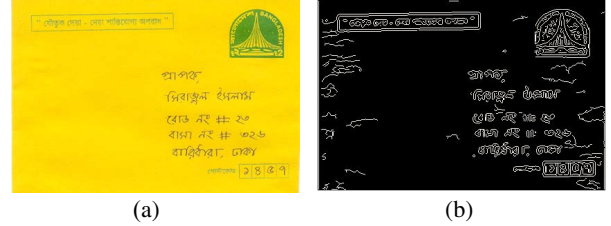


Fig. 2. (a) The input image (b) After applying 'Canny' method for edge detection.

We have started searching for the post-code box from a pre-specified location which is nearby the searching location. During the search for an edge, whenever we have found the top horizontal edge of the post-code box, we also get the left most point. We have noticed that for a 72 dpi image of the envelope, each box is of size 18x20 pixels. So after getting the left-top point of the post-code box, we can easily crop the required 4 images of handwritten Bengali numerals.

**Normalization, Filtering, Thinning and Dilation:** Size normalization is implemented on each character image that has been collected from previous stage. The size is normalized to be 16x16 pixels. Then, we employ the Gaussian filter to remove the noise and then transform the image into binary image. We notice that the thickness of the stroke is varying because different people use different writing tools. Since the thickness of the stroke influences the recognition result, we use the thinning and dilation algorithm to process the normalized numerals and get the same thickness of the stroke.

#### B. Dimension Reduction

**Use of Kirsch Mask on the Normalized Image:** The normalized image can pass through the SVM as input but there will be lot of relevant information missing in that case. That is why, in this experiment, we first obtain four directional feature maps of a numeral image by the Kirsch Mask. The size of each map is 16x16 pixels. The original image is considered as the fifth feature map. Then we arrange the five images to get the feature vectors ( $5 \times 16 \times 16 = 1280$  dimensions). But this is huge dimension to be given as input to the SVM. To solve this problem we have used PCA to decrease the dimension and extract more significant features. In this way, recognition time and training time can be reduced. Fig.3 illustrates this step.



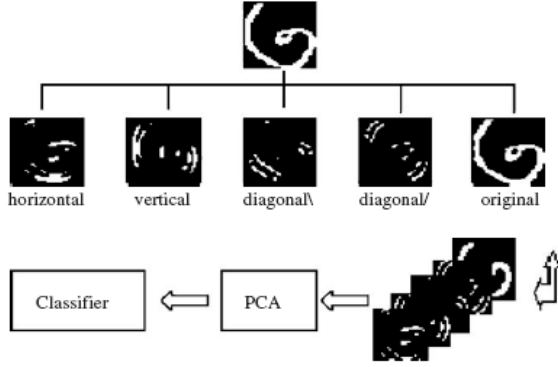


Fig. 3. Applying kirsch mask on the normalized image.

### C. Training of SVM

We have collected hand written Bengali numerals of a number of people. A sample has been shown in Fig.4.

Following the same preprocessing procedure of an input image, we have collected the 10 Bengali numerals of different handwriting. After using Kirsch mask and PCA on these images, we have stored it in a .mat (a MATLAB file extension) file which is used to classify the SVM.

After getting input from PCA for each numeral image, the SVM tries to find the class where the input numeral image should belong.

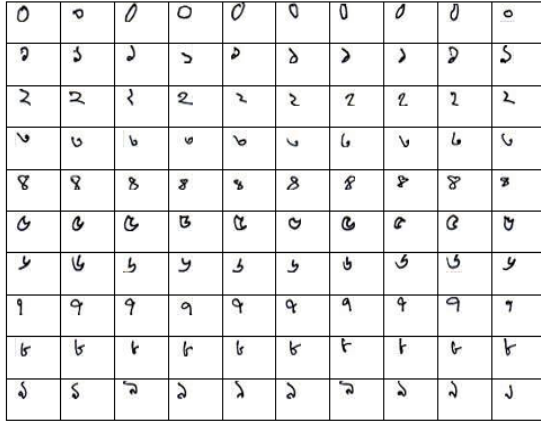


Fig. 4. Samples of handwritten Bengali numeral images (০-৯)

### D. Recognition

Based on PCA and SVM, we propose the recognition approach based on Kirsch mask combined PCA and SVM. Generally, Kirsch edge detector, Prewitt edge detector, Sobel edge detector and so on are the representative edge detectors. Among them the Kirsch edge detector has been known to detect four directional edges more accurately than others because the Kirsch edge detector considers all eight

neighbors. Kirsch defined a nonlinear edge enhancement algorithm as follows [20]:

$$G(i, j) = \max_{k=0}^7 \{1, \max[5S_k - 3T_k]\} \quad (1), \text{ where}$$

$$S_k = A_k + A_{k+1} + A_{k+2} \quad (2)$$

$$T_k = A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7} \quad (3)$$

In Equations (1), (2) and (3)  $G(i, j)$  is the gradient of pixel  $(i, j)$ ,  $A_k$  ( $k = 0, 1, \dots, 7$ ) is eight neighbors of pixel  $(i, j)$  as defined in Fig.5.

|       |          |       |
|-------|----------|-------|
| $A_0$ | $A_1$    | $A_2$ |
| $A_7$ | $(i, j)$ | $A_3$ |
| $A_6$ | $A_5$    | $A_4$ |

Fig. 5. Definition of eight neighbors of pixel  $(i, j)$

In this paper, we calculate directional feature vectors for horizontal (H), vertical (V), right-diagonal (R), and left-diagonal (L) directions as follows:

$$G_H(i, j) = \max(|5S_0 - 3T_0|, |5S_4 - 3T_4|) \quad (4)$$

$$G_V(i, j) = \max(|5S_2 - 3T_2|, |5S_6 - 3T_6|) \quad (5)$$

$$G_R(i, j) = \max(|5S_1 - 3T_1|, |5S_5 - 3T_5|) \quad (6)$$

$$G_L(i, j) = \max(|5S_3 - 3T_3|, |5S_7 - 3T_7|) \quad (7)$$

The Fig.6 shows the Kirsch masks used for calculating directional feature vectors. The normalized image is processed by the pairs of Kirsch masks ( $3 \times 3$  pixels), resulting in four graded feature maps, coding for the presence of horizontal, vertical or diagonal edges. In this way, we get four directional feature maps. We also take the original image as the fifth map in order to compensate the disadvantages of four directional feature maps. The final data representation corresponds to  $5 \times 16 \times 16 (=1280)$  dimensional inputs. The 1280-dimension is too much for input of a classifier so we employ PCA to extract feature and decrease the dimension. Finally, we obtain the new feature presentation as the input of SVM.

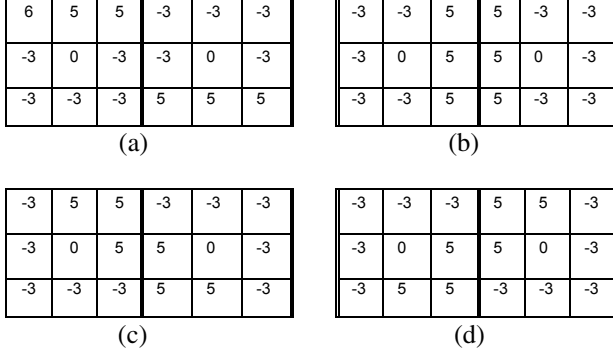


Fig. 6. Kirsch masks, (a) horizontal, (b) vertical, (c) right-diagonal and (d) left-diagonal

#### IV. EXPERIMENTAL RESULTS

##### A. Data Collection

To meet the requirements of a practical usage of the letter sorting machine for Bangladesh Postal System, we should consider not only the recognition rate and the error rate but also the response time. There are various writing styles. For our experiments, 1600 numerals were obtained from the real letters written by Bangladeshis. We randomly select 900 numerals as the training set and the rest as the test set.

For this paper, our experiments were tested on Windows XP, Pentium T(R) D CPU 2.8G and 760 RAM. We employ MATLAB to realize the program.

##### B. Time Performance:

To evaluate the performance of our proposed approach, we have implemented both the proposed method as well as an existing method which takes normalized image as the input of SVM classifier. Table I shows the training and recognition time of our proposed system and the existing system.

TABLE I  
THE RECOGNITION AND TRAINING TIME OF TWO APPROACHES

| Approach               | Recognition time (ms) | Training time (s) |
|------------------------|-----------------------|-------------------|
| Normalized image + SVM | 21.44                 | 67.94             |
| Our Proposed System    | 31.24                 | 146.47            |

##### C. Recognition Performance

The proposed method demonstrates a very high recognition rate. The experimental result shows that close similarity in geometric shape of some numerals led to the chance of false recognition. The Table II represents the recognition rate for different numerals with respect to all numerals.

To evaluate the recognition results, the following measures are employed: recognition rate (Recog.), error rate (Err.), rejection rate (Reject.), and reliability (Reliab.). They are defined as follows:

TABLE II  
THE RATE OF RECOGNITION OF EACH NUMERAL WITH RESPECT TO ALL THE 10 NUMERALS

| Input | Recognition Rate (%)  |      |      |     |     |     |     |     |     |      |
|-------|-----------------------|------|------|-----|-----|-----|-----|-----|-----|------|
|       | Numeral recognized as |      |      |     |     |     |     |     |     |      |
|       | 0                     | 1    | 2    | 3   | 4   | 5   | 6   | 7   | 8   | 9    |
| 0     | 91.7                  | 0    | 0    | 8.3 | 0   | 0   | 0   | 0   | 0   | 0    |
| 1     | 0                     | 91.7 | 8.3  | 0   | 0   | 0   | 0   | 0   | 0   | 0    |
| 2     | 0                     | 16.7 | 83.3 | 0   | 0   | 0   | 0   | 0   | 0   | 0    |
| 3     | 16.7                  | 0    | 0    | 75  | 0   | 0   | 8.3 | 0   | 0   | 0    |
| 4     | 0                     | 0    | 0    | 0   | 100 | 0   | 0   | 0   | 0   | 0    |
| 5     | 0                     | 0    | 0    | 0   | 0   | 100 | 0   | 0   | 0   | 0    |
| 6     | 0                     | 0    | 0    | 0   | 0   | 0   | 100 | 0   | 0   | 0    |
| 7     | 0                     | 0    | 0    | 0   | 0   | 0   | 0   | 100 | 0   | 0    |
| 8     | 0                     | 0    | 0    | 0   | 0   | 0   | 0   | 0   | 100 | 0    |
| 9     | 0                     | 16.7 | 0    | 0   | 0   | 0   | 0   | 0   | 0   | 83.3 |

TABLE III  
RECOGNITION RESULTS BY THE PROPOSED SYSTEM

| Recognition Attributes | ০     | ১     | ২    | ৩  | ৪   | ৫   | ৬   | ৭   | ৮   | ৯    | Avg. |
|------------------------|-------|-------|------|----|-----|-----|-----|-----|-----|------|------|
| Recog.(%)              | 91.67 | 91.67 | 83.3 | 75 | 100 | 100 | 100 | 100 | 100 | 83.3 | 92.5 |
| Err.(%)                | 8.3   | 8.3   | 16.7 | 25 | 0   | 0   | 0   | 0   | 0   | 16.7 | 7.5  |
| Reliab.(%)             | 91.67 | 91.67 | 83.3 | 75 | 100 | 100 | 100 | 100 | 100 | 83.3 | 92.5 |

TABLE IV  
THE RECOGNITION RATES (%) OF TWO APPROACHES

| Numerals                | ০    | ১    | ২    | ৩    | ৪   | ৫    | ৬    | ৭   | ৮   | ৯    |
|-------------------------|------|------|------|------|-----|------|------|-----|-----|------|
| Approach                | ০    | ১    | ২    | ৩    | ৪   | ৫    | ৬    | ৭   | ৮   | ৯    |
| Normalized Image to SVM | 58.3 | 66.7 | 83.3 | 58.3 | 100 | 83.3 | 83.3 | 100 | 100 | 58.3 |
| Proposed System         | 91.7 | 91.7 | 83.3 | 75   | 100 | 100  | 100  | 100 | 100 | 83.3 |

TABLE V  
COMPARISON OF AVERAGE RECOGNITION RATE WITH OTHER EXISTING APPROACHES

|                              | Proposed System | Normalized Image to SVM | Neural Network Based method |                         |
|------------------------------|-----------------|-------------------------|-----------------------------|-------------------------|
|                              |                 |                         | With momentum factor        | Without momentum factor |
| Avg. rate of recognition (%) | 92.5            | 79.1                    | 60.75                       | 89.75                   |

$$\text{Recog.} = \frac{N1 \text{ (correctly recognized)}}{N4 \text{ (tested)}} \times 100\% \quad (8)$$

$$\text{Err.} = \frac{N2 \text{ (error recognized)}}{N4 \text{ (tested)}} \times 100\% \quad (9)$$

$$\text{Reject.} = \frac{N3 \text{ (rejected)}}{N4 \text{ (tested)}} \times 100\% \quad (10)$$

$$\text{Reliab.} = \frac{\text{Recog.}}{100\% - \text{Reject.}} \times 100\% \quad (11)$$

where N1 (correctly recognized) is defined as the number of numerals correctly classified, N2 (error recognized) is defined as the number of numerals falsely classified, N3 (rejected) is defined as the number of numerals rejected by the classifier, and N4 (tested) is the number of input numerals for test.

Table III shows the recognition rate, the error rate and the reliability of different numerals for the proposed system. Finally, we get the average recognition rate; the average error rate and the average reliability are 92.5%, 7.5% and 92.5%, respectively. Experimental results

demonstrate that our presented system performs very well. It meets the response time with the accuracy requirements.

## V. PERFORMANCE COMPARISONS AND DISCUSSIONS

To compare the results of our proposed method we have implemented existing method applying normalized image to SVM and neural network based system using back-propagation algorithm [21]. Table IV presents the recognition results of our proposed approach and other existing method.

The Table V represents the average recognition rate of Bangla numerals. The third method which is based on neural network using back propagation algorithm [21] is tested during the training period with iteration number from 4000 to 7000 with momentum and without momentum.

The experimental results of Table IV and Table V show that the recognition rate of our method is better than other existing system. But from Table I it is found that the time performance of our proposed approach is not better than the other. But it is remarkable that the recognition rate of our method is much higher than the two other existing methods.

From Table II it has been found that the highest reliability is 100% obtained by Bengali numeral “৪”, “৫”, “৬”, “৭”, “৮”. The false recognition or rejection

is due to variability of handwritings as well as bad writing. From the experiments we note that the most confusing numeral pair is “০” and “০” as shown in Fig.7(a). Second confusing pair is “১” and “১” as shown in Fig.7(b). Third confusing pair is “১” and “১” as shown in Fig.7(c). Therefore, error rate or rejection rate for these numerals is higher.

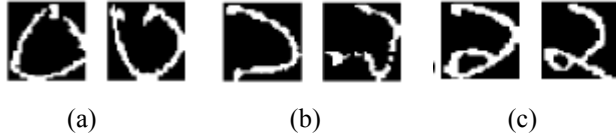


Fig. 7. Examples of some confused handwritten Bengali numeral pairs: (a) “০” and “০” (b) “১” and “১” (c) “১” and “১”.

## VI. CONCLUSIONS

In a recognition system the single most important issue is to ensure a good recognition rate. We have used PCA and SVM to enhance the recognition rate. In order to maximize the performance of numeral recognition, our proposed approach extracts four directional local feature vectors with the Kirsch masks and one global feature vector from a normalized input image. Hence, the important features are retained and the non-important pixels are minimized. The recognition result of our proposed approach outperforms the others. However, our proposed approach has extracted more significant features than the method applies normalized image to SVM. So it needs more time to estimate than that specific approach. But our method also works better with respect to recognition performance than the method for numeral recognition by neural network using back-propagation algorithm with and without momentum factor. As far as the recognition rate is concerned, it is experimentally shown that the recognition approach that we have proposed in this paper is superior to other conventional approaches.

## REFERENCES

- [1] B. B. Chaudhuri, and U. Pal, “A complete Printed Bangla OCR System”, *Pattern Recognition* 31(5): 531–549, 1998.
- [2] U. Pal, “On the development of an optical character recognition (OCR) system for printed Bangla script”, *Ph.D. Thesis*, Indian Statistical Institute, 1997.
- [3] U. Pal, and B. B. Chaudhuri, “Indian Script Character Recognition: a Survey”, *Pattern Recognition*, 37(9): 1887–1899, 2004.
- [4] Y. L. Cun, O. Matan, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, and H. S. Baird, “Handwritten Zip Code Recognition with Multilayer Networks”, *In Proceedings of the 10th International Conference on Pattern Recognition (Atlantic City, NJ)*. IEEE Computer Society Press, vol 2, pp. 35-40, 1990.
- [5] U. Mahadevan, and S. N. Srihari, “Parsing and Recognition of City, State, and ZIP Codes in Handwritten Addresses”, *In Proc. of Fifth ICDAR*, pp. 325-328, 1999.
- [6] X. Wang, and T. Tsutsumida, “A New Method of Character Line Extraction from Mixed unformatted Document Image for Japanese Mail Address Recognition”, *In Proc. of Fifth ICDAR*, pp. 769-772, 1999.
- [7] S. N. Srihari, and E. J. Keubert, “Integration of Hand-Written Address Interpretation Technology into the United States Postal Service Remote Computer Reader System”, *In Proc. Of Forth ICDAR*, pp. 892-896, 1997.
- [8] A. S. M. M. Morshed, M. S. Islam, M. A. Faisal, M. A. Sattar, and M. Ishizuka, “Automatic Sorting of Mails by Recognizing Handwritten Postal Codes using Neural Network Architectures”, *Proc.(CD-ROM) Int'l Conf. on Computer and Information Technology (ICCIT2007)*, session D3, 5 pages, Dhaka, Bangladesh, Dec. 2007.
- [9] R. Plamondon and S. N. Srihari, “On-line and off-line handwritten recognition: A comprehensive survey”, *IEEE Trans. on PAMI*, 22(1): 62-84, 2000.
- [10] K. Kim, J. H. Kim and C. Y. Suen, “Recognition of Unconstrained handwritten Numeral Strings by Composite Segmentation method”, *Proc.15th ICPR*, pp. 594-597, 2000.
- [11] J. Cai and Z. Q. Liu, “Integration of structural and statistical information for unconstrained handwritten numeral recognition”, *IEEE Trans. On PAMI*, 21(3): 263-270, 1999.
- [12] P. Wunsch and A. F. Laine, “Wavelet Descriptors for Multi-resolution Recognition of Hand-printed Digits”, *Pattern Recognition*, 28(8):1237-1249, 1995.
- [13] Z. Chi and H. Yan, “Handwritten numeral recognition using self-organizing maps and fuzzy rules”, *Pattern Recognition*, 28(1):59-66, 1995.
- [14] K. Kim and S. Y. Bang, “A handwritten numeral character classification using tolerant Rough set”, *IEEE Trans. on PAMI*, 22(9): 923-937, 2000.
- [15] M. Turk and A. Pentland, “Eigenfaces for Recognition”, *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.
- [16] K. Ohba and K. Ikeuchi, “Detectability, Uniqueness, and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 1043-1048, 1997.
- [17] H. Murase and S. Nayar, “Visual Learning and Recognition of 3D Objects from Appearance”, *International Journal of Computer Vision*, vol 14, pp. 5-24, 1995.
- [18] V. N. Vapnik, “The Nature of Statistical Learning Theory”, Springer, New York, 1995.
- [19] <http://www.mathworks.com>, November 19, 2007.
- [20] S. W. Lee, “Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network”, *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (6): 648–652, 1996.
- [21] M. W. Islam, M A. M. Hasan, R. C. Debnath, “Handwritten Bangla Numeral Recognition using Back-propagation Algorithm With and Without Momentum factor”, *Proceedings of National Conference on Computer Processing of Bangla (NCCPB 2005)*, Independent University, Bangladesh, pp. 177-182, 2005.

# An Algorithm for Segmenting Modifiers from Bangla Text

Nasreen Akter<sup>1</sup>, Saima Hossain<sup>1</sup>, Md. Tajul Islam<sup>1</sup>, and Hasan Sarwar<sup>2</sup>

<sup>1</sup>Military Institute of Science and Technology (MIST), Dhaka, Bangladesh

<sup>2</sup>United International University (UIU), Dhaka, Bangladesh

Email addresses: [kona\\_mir@yahoo.com](mailto:kona_mir@yahoo.com), [erin\\_saima@yahoo.com](mailto:erin_saima@yahoo.com), [mdhasan70@yahoo.com](mailto:mdhasan70@yahoo.com)

**Abstract** — Script segmentation is an essential and preprocessing task for any OCR system. Bangla is one of the most popular scripts in the world. Since segmentation effects the recognition process, accurate and proper segmentation is necessary to implement Bangla OCR. Many works have been done for both handwritten and printed Bangla text. This paper presents the segmentation process of different bangla modifiers from printed bangla words.

**Index Terms** — Segmentation, Optical Character Recognition, Bangla Modifier.

## I. INTRODUCTION

The structure of bangla language is quite different than any other European language. It consists of 50 basic characters including 11 vowel and 39 consonant characters. Again, there are vowel and consonant modifiers, touching and compound characters which are formed by touching the adjacent characters and by combining 2 or more characters in the word accordingly. So when they are used to form words, we find that thousands of various combinations (both simple and complex) are formed. Segmentation of these complex parts are extremely difficult task. As a result, printed Bangla OCR still remains researchable.

Generally Bangla texts are divided into three zones namely upper zone, middle zone and lower zone [1], [2], [4], [6], [7]. Upper zone contains , , . Middle zone contains the basic character symbols and the lower zone contains the lower modifiers such as , , , etc. Detection of text line, matra line and base line [1], [2], [4], [6], [7] lead towards the character segmentation process. If one of the above three is ignored, it will generate many errors. So it is necessary to detect them accurately.

It is observed that Bangla words have different kinds of complexities like , , , , , , , , , , etc. Some works have been done to solve a few of above the problems such as , , , , , , [1], [2], [4], [6]-[8]. A, font size independent, more efficient technique is proposed in this paper, which can perfectly segment , , , , , , , , , .

## II. LITERATURE REVIEW

Segmentation of a Bangla document into lines, words and words into individual characters is an important and preprocessing task for Bangla OCR. The line segmentation of a document is solved by finding white pixels between two consecutive headlines [1], [2], [4]-[7]. Matra zone detection has been done with the approach that matra line will be that row which contains maximum no of black pixels than any other portion of the text [4], [6], [7] and considering multiple rows that defines matra zone [1], [3]. Baseline is detected by selecting that row which contains highest lowermost points of the components [1,3].

Word segmentation is done by finding some couple of 0 consecutive columns (white pixels) as a word delimiter [5], [6], or n (taken experimentally) consecutive 0 columns [3], or first two 0 consecutive columns [4] or the midpoint of a run of at least k consecutive 0 columns [1], [7].

The most difficult and important part is to segment the characters individually from the words. At first, the adjacent characters are separated by finding the 0 columns between them [1]-[5], [7]. But this process is not sufficient at all due to the problems of overlapping (multiple characters which do not touch each other but enter into the zone of each other like  (   )), touching and compound situations. Even the situation becomes more complex due to many peculiar patterns of the bangla characters and their placement in the word. Sometimes a character (  ) gets divided into sub portions in the character segmentation process and the problem is overcome by joining them to each other [1]. As there are many modifier, , ,  are segmented as matra upper portion character like , ,  [1], [4], [7] and , ,  are segmented after the detection of base line [1], [2], [4], [7] or considering the height, which exceeds the average size of the segmented characters [6],  is picked up as it makes a regular angle with the matra line [2] and  is picked up as it is not touching the matra line [2]. A piecewise linear scanning [1], [3], DFS (Depth First Search) [7], a combination of flood fill and boundary fill algorithm [4] were proposed when the characters overlap with other characters. In paper [2], they detect  when they find that the present

TABLE I  
SUMMARY OF LINE DETECTION TECHNIQUES

| Type                                 | Papers |   |   |   |   |   |   |             |
|--------------------------------------|--------|---|---|---|---|---|---|-------------|
|                                      | 1      | 2 | 3 | 4 | 5 | 6 | 7 | Our process |
| Textline detection                   | Y      | Y | Y | Y | Y | Y | Y | Y           |
| Matraline detection                  | Y      | Y | Y | Y |   |   |   | Y           |
| Base line detection                  | Y      |   | Y |   |   |   | Y | Y           |
| Upper portion of matraline detection |        |   | Y |   |   |   |   | Y           |
| Lower portion of matraline detection | Y      | Y | Y | Y |   |   |   | Y           |

character size is greater than the regular size and apply the segmentation process.

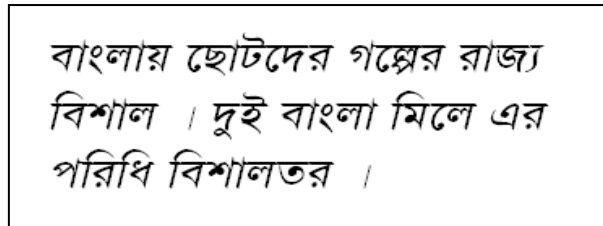
### III. SEGMENTATION PROCESS

This is the most vital and important portion for designing an efficient Bangla OCR because character recognition process depends on this phase to make the recognition process successful. In this paper, the following steps are considered to do the segmentation with the existing techniques along with our approach. The steps are:

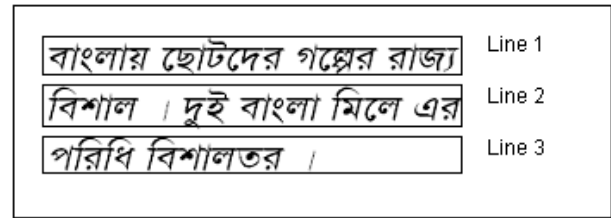
- A. Line Segmentation Process
- B. Baseline Detection Process
- C. Word Segmentation Process
- D. Character Segmentation Process
  - Words with only basic character
  - Words with modifier ি ি ি
    - . Segmentation of মি, টি, ঠি
    - . Segmentation of জী, চী, ঠী
    - . Segmentation of কী, ঠী, ঠী
  - Words With lower Modifier ৲, ৳, ৴

#### A. Line Segmentation Process

The lines of a text block are segmented by finding white pixels between two consecutive matra zone [1], [2], [4]-[7]. Fig: 1 shows the result.



(a)



(b)

Fig. 1. Text line detection process (a) full image, (b) result of text line detection

#### B. Baseline Detection Process

Base line is the row where the middle zone ends and lower modifier zone starts i.e. a separator between middle and lower zone. This is the row where an abrupt change occurs between the previous and next row [7]. Detection of base line is very important.

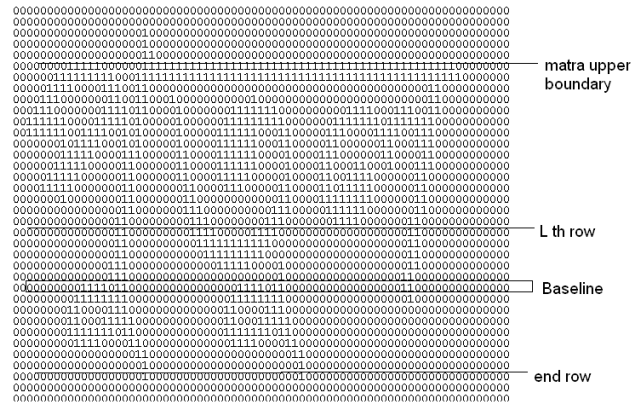


Fig. 2. Process of baseline detection

We proceed with the idea, shown in fig: 2, that there are small no of black pixels at the end portion of the middle zone and from there lower modified zone starts with increasing black pixels. So at first we detect L th row which denotes the lower position of the middle zone. Then we scan from L to end row to find where the black pixels start to increase. That row will be denoted here as the base line.

The difference between the end row and the base line is the lower modified zone. With this technique, sometimes the lower portion of the middle zone may enter into this zone. We can ignore them since they do not touch end row. Lower modifiers always touch end row. Table I compiles the various line detection capability of different works.

### C. Word Segmentation Process

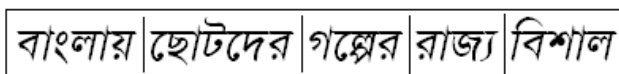


Fig. 3. Process of word segmentation

There are always some white spaces between two words in a text line shown in fig: 3. Vertical scan is done ignoring the matra zone.

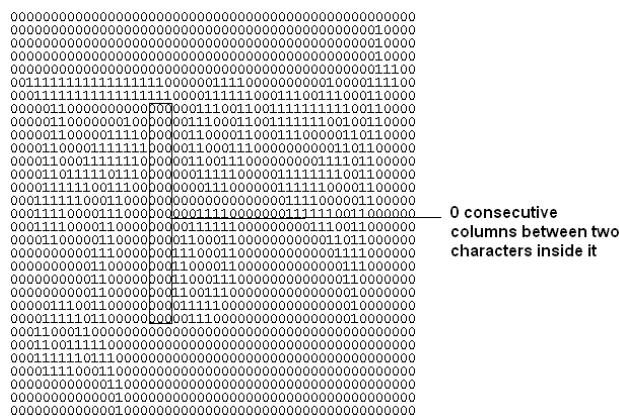


Fig. 4. Word with three 0 consecutive columns between two characters inside it

Sometimes it is also seen that some 0 consecutive columns are formed between two characters in a word shown in fig. 4. Most of the time it occurs when the matra less character is used in a word. If there is a white vertical column then it is treated as a gap between two characters in a word [2]. But this gap is not fixed and can increase with the increase of font size. So these issues require care during word segmentation. We try to solve it with the following technique.

During word segmentation, we start scanning vertically from start row to end row and keep the record of the width

of 0 consecutive columns for each line. Then we take the average of the total widths. If any width is less than half of average, then it is considered as a gap between two characters and easily ignored. Table II shows the capability of different papers considering the above problems.

TABLE II  
SUMMARY OF WORD SEGMENTATION

| Type              | Papers |   |   |   |   |   |               |
|-------------------|--------|---|---|---|---|---|---------------|
|                   | 1      | 2 | 3 | 4 | 5 | 6 | 7 Our process |
| Word segmentation | Y      |   | Y | Y | Y | Y | Y             |

### D. Character Segmentation Process

It is the most difficult and challenging part for building printed Bangla OCR. Many works have been done in order to segment the bangla characters from the words. Different problems are discussed below and some algorithms are made to solve them.

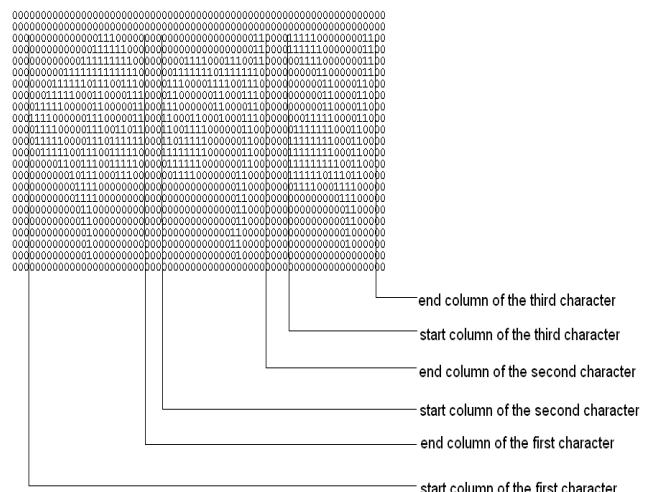


Fig. 5. Character segmentation

### Words with only basic character

The basic idea is that, under the lower boundary of the matra zone, the first black pixel column is fetched and is treated as a starting boundary and then a column of white pixel is searched. When it is found, it is treated as a separator from the next character [1]-[5], [7]. It is shown in fig: 5. The works that use this technique for character separation as well as ours are mentioned in Table III.



TABLE III  
SUMMARY OF WORDS WITH SINGLE CHARACTERS

| Simple | Papers |   |   |   |   |   |             |
|--------|--------|---|---|---|---|---|-------------|
|        | 1      | 2 | 3 | 4 | 5 | 7 | Our process |
| কলম    | Y      | Y | Y | Y | Y | Y | Y           |
| ঐষধ    | Y      | Y | Y | Y | Y | Y | Y           |

Some times this technique fails to segment the characters properly [1]-[7]. Extra care is needed to solve those problems. They are discussed below.

Words with modifier ি, ি, ি

Most of the previous works considered the segmentation of these modifiers as matra upper portion character. In our system মি, জী, টি, ঠি, টী, ঠী, কাঁ, টাঁ, ঠাঁ are segmented properly and efficiently keeping their original shape. The technique is given below.

1) In order to decide whether the current portion contains basic character or character with ি, ি, ি, the system finds first discontinuity point i, e. first white column. Then it searches in the top of the matra zone to get whether the same column contains black pixel or not. If it gets that column of black pixel, then it goes right to end and searches a column of white pixel in the matra upper zone. If gets that column, it searches for the same column contains white pixel under the matra zone. If it gets that column then it will be a basic character. Otherwise, along with not getting the column of white pixels in the matra upper zone, it will be the character with above modifiers.

2) Then the thickness of the characters x and y is determined shown in fig: 6.

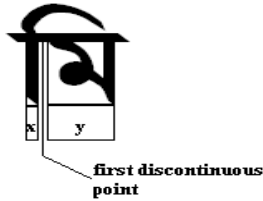


Fig: 6. Determination of the thickness of each character in the syllable

3) If  $x < y$  then there exists ি. For ি, we will get the following situations. They are like মি, টি/ঠি etc. From the immediate upper row of the matra upper boundary we go right to fetch black pixel. After finding that we again go right to find white pixel. After getting that we continue to go to the same direction. During that time if there exist all 0 in that row, the system can understand that the character, which is under the matra zone, has no upper

part. If it is not, then the system can understand that the character has upper part.

a) For the first case, to separate ি it is picked up by removing the adjacent character and that adjacent character is picked up by removing ি. The result is shown below:

$$\text{মি} = \text{মি} + \text{মি}$$

(b) For later case, to separate ি, it is needed to detect whether it is টি or ঠি.

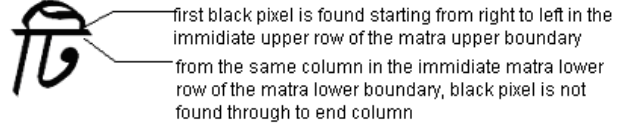


Fig: 7. Segmentation process of টি

At first, first black pixel is searched in the immediate upper row of the matra upper boundary from the end column to left. If it is found, then from the same column of the immediate lower row of the matra lower boundary, next searching is occurred, which lasts at the end column, for finding black pixels. If it is found then there is ঠ otherwise there will be টি[2]. Fig. 7 shows the following process.

i) Segmentation of টি: Form the start row we scan right to get the first black pixel and after getting that we start to pick up those. When we get the first white pixel i, e. 0, we pick up the rest of the image as 0. In this way we pick up the matra upper portion. Matra zone remains identical. Under the matra zone, the first character is picked up by removing the adjacent character. Then the remaining character is picked up from the upper part of matra zone by making the black pixels to white, which was unchanged during the previous one and the rest of the pixels remain unchanged. Under the matra zone, the second character is picked up by removing the previous character. Here, removing means making the 1(black pixel) to 0 (white pixel). But this is not actually done. The result is shown below.

$$\text{টি} = \text{মি} + \text{টি}$$

ii) Segmentation of ঠি: At first the thickness of upper portion of ঠি and a gap are calculated. The system counts the black pixels where the gap becomes zero. Then ঠি is separated by starting to find the start position where summation of the black pixels is greater than the thickness and removing the adjacent portions from the upper zone and as well as from the middle zone. ঠি is separated by picking up the thickness portion and the middle zone portion,



removing rest of the pixels. The result is shown below:

$$\text{ঐ} = \text{ি} + \text{ঔ}$$

4) If  $x > y$  then there exists ঐ or ঔ. For ঐ or ঔ we will get the following situations. They are like ঐ, ঔ, ঐ, ঔ etc. From the startrow if we find that the sum of the black pixels is greater than 3 then there will be ঐ and if not Then ঔ remains there. For the above technique, for ঐ, ঐ is detected as ঐ. We solve this problem later. In the case of জী, টী- From the startrow to half of the upper zone we calculate the maximum gap (no of sequence of 0s between two 1s) between the black pixels.

a) *Segmentation of জী* - If only one gap is found then the system can understand that, the basic character does not have upper part. Then to separate ঐ, it is picked up by removing the adjacent character and that adjacent character is picked up by removing ঐ. The result is shown below.

$$\text{জী} = \text{জ} + \text{ঐ}$$

b) *Segmentation of টী* - From the start row to the matra upper boundary the area is divided into two halves. In the upper half, the system starts from left to right to find the cutting point (the white column after getting the first black pixel). After getting that the rest of the pixels will be removed. This process will continue for this half. In the lower half the system stats from right to left to find the cutting point. The left pixels from the cutting point will remain same by removing others. In this way the upper portion of ট is detected and picked up with the matra zone and middle portion by removing the adjacent one. ঐ is segmented by keeping unchanged the right portion of the cutting point removing the left portion. This will continue to 1<sup>st</sup> half. The second half will remain identical. In this way the upper potion of ঐ is picked up with the matra zone and by removing adjacent characters of the middle portion. The result is shown below

$$\text{টী} = \text{ট} + \text{ঐ}$$

In the case of গী, কী, ণী, ঙী -- গী is isolated from by finding the gaps. There are 2 gaps in the immediate upper row of the matra upper boundary for গী where the others hold one.

c) *Segmentation of গী* - the thickness of the upper portion of ঙ is calculated. From start row, the system scans horizontally to search for black pixels. When the first one is found then the thickness is added with that black pixel's column position called thicknessColumn. Now the system goes left to right and keep the pixels same until it reaches to thicknessColumn making the rest of the pixels 0. In this way the upper potion of ঙ is picked up with the

matra zone and by removing adjacent characters of the middle portion. For ঐ, the system starts from left to right and if the gap is less than or equal to 1, the pixels will remain unchanged. Else the pixels will be zero to thicknessColumn. Then the rest of the pixels will remain unchanged. In this way the upper potion of ঐ is picked up with the matra zone and by removing adjacent characters of the middle portion. The result is shown below.

$$\text{গী} = \text{ঙ} + \text{ঐ}$$

d) *Segmentation of কী, ঙী* - This segmentation process is similar to the segmentation process of মি. The only difference is the process starts from the right side. The result is shown below.

$$\text{কী} = \text{ক} + \text{ঐ}, \text{ঙী} = \text{ঙ} + \text{ঐ}$$

e) *ণী* - the system starts from right to left and continue until the system get white pixel after getting the first black one i.e. cutting point making the rest of the left pixels 0. During this time if the cutting point less than the previous one the current point will be replaced by the previous one.

TABLE IV

SUMMARY OF ঐ, ঐ, ঐ SEGMENTATION

| Complex | Papers |   |   |   |   | Our process |
|---------|--------|---|---|---|---|-------------|
|         | 1      | 2 | 4 | 5 | 7 |             |
| আমি     |        | Y |   |   |   | Y           |
| ঢিয়া   |        | Y |   |   |   | Y           |
| জীব     |        | Y |   |   |   | Y           |
| টাকা    |        | Y |   |   |   | Y           |
| জেঠী    |        | Y |   |   |   | Y           |
| কোটা    |        |   |   |   |   | Y           |
| টা      |        |   |   |   |   | Y           |
| গী      |        |   |   |   |   | Y           |

In this way the upper portion of ট is detected and picked up with the matra zone and by removing adjacent characters of the middle portion. ঐ is picked up by going right to left to the cutting point removing the rest of the left pixels. In this way the upper potion of ঐ is picked up with the matra zone and by removing adjacent characters of the middle portion. The result is shown below.

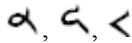
$$\text{টী} = \text{ট} + \text{ঐ}$$

Table IV shows the result for different papers as well as ours.

Words With Lower Modifier ঐ, ঐ, ঐ

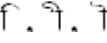
As they occur under the middle zone, they can be segmented properly after the successful detection of base

line [1], [2], [4], [7]. If the base line is not equal to end row of the text then the system can understand that there are lower modifiers. Then the system segments them by finding white pixels after finding first black pixels. Table V shows the use of this technique in different papers as well as ours.

TABLE V  
SUMMARY OF  SEGMENTATION

| Complex | Papers |   |   |   |   |   |             |
|---------|--------|---|---|---|---|---|-------------|
|         | 1      | 2 | 3 | 4 | 5 | 7 | Our process |
| সুখ     | Y      | Y | Y | Y |   | Y | Y           |
| কুল     | Y      | Y | Y | Y |   | Y | Y           |
| কুড়া   |        |   |   |   |   |   | Y           |
| গুন     | Y      | Y | Y | Y |   | Y | Y           |
| রুই/রুই |        |   |   |   |   |   | Y           |
| তুন     | Y      | Y | Y | Y |   | Y | Y           |

#### IV. CONCLUSION

This paper presents an approach to detect Baseline and a new algorithm to segment  in an intact shape i.e. by not separating the upper portions of them like . This algorithm is font size independent and matches best with Sutonny, karnaphuli, Rinkiy fonts etc. We didn't consider any noise in the textual image. We are working to develop segmentation algorithm to separate the overlapping, touching and compound characters.

#### REFERENCES

- [1] B.B.Chaudhury and U.Pal, "A Complete Printed Bangla OCR System", *Pattern Recognition*, vol. 31, pp. 531-549, 1997
- [2] Md. Abdus Sattar, Khaled Mahmud, Humayun Arafat and A F M Noor Uz Zaman, "Segmenting Bangla Text For Optical Recognition", *ICCIT 2007*, pp. 27-29 December, Dhaka, Bangladesh.
- [3] J U Mahmud , M F Rahman and C M Rahman, "A Complete OCR System for Continuous Bengali Characters" *Proc. IEEE TENCON*, 2003, pp. 1372-1376.
- [4] Md. Al Mehedi Hasan, Md. Abdul Alim, Md. Wahedul Islam and M Ganger Ali, "Bangla Text Extraction and Recognition From Textual Image", *Proceedings of 2<sup>nd</sup> national Conference on Computer Processing of Bangla*, pp. 171-176, February 2005, Dhaka, Bangladesh.
- [5] Md. Murad Hossain, Md. Shamsul Alam "BD Fine Reader: A Bangla Character Recognition system for Converting Printed Bangla Text Documents into Editable Electronic Document", pp. 192-199, *ICCPB 2006*, Dhaka, Bangladesh.
- [6] Veena Bansal and R M K Sinha, "Segmentation of Touching and Fused Devnagari Characters", [www.iitk.ac.in/ime/veena/PAPERS/s.pdf](http://www.iitk.ac.in/ime/veena/PAPERS/s.pdf)
- [7] S M Millky Mahmud, Nazib Shahrier A S M Delowar Hossain, Md Tareque Mohmud Chowdhury and Abdus Sattar, "An Efficient Segmentation Scheme For The Recognition Of Printed Bangla Characters", *ICCIT 2004*, pp. 26-28, December 2004, Dhaka, Bangladesh.
- [8] U Garain and B B Chaudhury, "Segmentation Of Touching Characters in Printed Devnagari and Bangla Scripts Using Fuzzy Multifactorial Analysis", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 32, pp. 449-459, Nov.2002.
- [9] Apurba Lal Saha "Performance Evaluation of Neural Network Based Character Recognition System on Different Bengali Fonts", Masters Thesis Paper, IST, 2007, Dhaka, Bangladesh.

# An Empirical Framework for Translating of Phrasal Verbs of English Sentence into Bangla

Mohammed Moshikul Hoque, Md Mokaddem Ahamad, Mohammad Shamsul Arefin, Md. Monjur-Ul-Hasan and Md. Gahangir Hossain

Dept. of Computer Science & Engineering, Chittagong University of Engineering & Technology (CUET), Chittagong-4349, Bangladesh

**Abstract:** Generally Phrasal verbs comprise a verb followed by a preposition that is commonly occurring feature in English. Each of the phrasal verbs acquires absolutely different meanings in different contexts. Having highly context dependent meanings, phrasal verbs may be disambiguated only by devising a technique involving utilization of semantic information pertaining to the context. This paper presented a framework for translating the Phrasal verbs in English sentence into its equivalent Bangla. Semantics is necessary in the field of natural language processing to detect the meaning of the individual words in the sentence. This paper also describes different types of semantic features that perform semantic based disambiguation of phrasal verbs. A translation engine works to find out the exact Bangla meaning of the Phrasal verb from different English sentences. We have tested our model for different type of English sentences containing phrasal verbs that are taken from the English newspaper and English textbook. We got successful results for most of the test cases.

**Index Terms:** Phrasal Verb, Semantic Features, Parse Tree, Example Base, Translation Engine.

## I. INTRODUCTION

Natural language is the most versatile communication medium of mankind. They contain infinite number of different words. Phrasal verbs are then a commonly occurring feature in English and comprise a verb followed by a preposition. Hence, translation of Phrasal verbs requires lexical analysis and syntactic analysis. There must be some semantic features included in the knowledge base for different types of Bangla words. Machine Translation (MT) of English text in terms of occurrence of phrasal verbs poses a great challenge and in case of English of Bangla, it becomes more challenging as not so many remarkable works have been done or seen in this field. The same phrasal verb acquires absolutely different meanings in different contexts. Moreover, no existence of phrasal verb is observed in the Bangla sentences and the Bangla meanings of these phrasal verbs are simply as verbs. Hence, it utmost important to separate the verbs and phrase in machine translation of English sentences to Bangla sentences. In the same way, the special semantics delivered by those phrasal verbs in different sentences should also be analyzed. For this reasons, machine translation of phrasal verbs from

English to Bangla claims importance in the field of natural language processing (NLP). Very few research activities have been done on the Bangla Language Processing and hence translation of Phrasal verbs from English to Bangla has not been done yet. An approach to classify the Phrasal verbs in the Danish language is used in SIMPLE system [1]. Disambiguation of Phrasal Verb form English to Hindi Machine Translation has been described in [2]. The transfer techniques for Bangla to English machine translation system through the transfer component the source language structures are translated into target language structures are focused [3]. An Example based Machine Translation System for translating News Headlines from English to Bengali is also developed [4]. Matrix representation of semantic features of Bangla words with redundancy rules is described in [5]. In Bangla, for semantic analysis we may use Bangla case analysis that is proposed to develop a smart conversation processor [6]. In this paper, we proposed a framework for translating of the occurrences Phrasal verbs in English sentences into its Bangla equivalent. We also find the semantic features of each word that constitutes the given input sentence and specify the exact Bangla meaning delivered by the phrasal verb used in the given sentence.

## II. SEMANTIC FEATURES OF THE ENGLISH WORDS

The features or characteristics that help to recognize a part of speech are known as semantic features [5]. In semantic analysis, many semantic features should be collected to easily identify the tokens.

### A. Noun

Noun is generally used as subject in simple assertive sentence. However, Noun acts both the role of the subject as well as the object in the sentence. Every verb is dependent fully on this subject [7]. To find out semantic features first we have to consider the physical appearance that can help us to easily recognize a noun or pronoun. As for example, “The Cow (গরু)” is noun, which has four legs, two ears, two horns, a tail. It has no wings, no hands and no intelligence [8]. So, “The Cow” comes as the subject of an English sentence, and then it must not read/write or fly in the sky. But as “The Bird (পাখি)” has wings it can fly in the sky. We know that bulb is inanimate and flammable, but inhuman and not

wearable. So, its semantic feature would be: *Bulb Inanimate (1), Flammable (1), Human (0), Wearable (0)*. Again, Noun and Pronoun is also used as the object in the sentence. The correct meaning of a phrasal verb in a sentence depends principally on the semantic feature of the object and the relation between the subject and object of that sentence. For instance, “the Bulb (বাতি)” is a noun of which job is to emit light (জ্বালা). Hence, being used in the sentence “I put on the light” as an object “the light” is going to be “switched on (জ্বালানো)”. Generally, noun in English sentence contains the following semantic features:

[Subject, Object, Animal, Human, Flammable, Wearable, Inanimate]

Although the above features are not completed for all kinds of noun but these are the most common features that can be used in semantic analysis. For some sentences, few features become more dominant than the other; symbolized by one (1) and the inapplicable features are symbolized by zero (0). Semantic features for some noun are shown in the Table I.

TABLE I:  
SEMANTIC FEATURES OF NOUN

| Word  | Semantic Features |        |           |          |         |           |
|-------|-------------------|--------|-----------|----------|---------|-----------|
|       | Human             | Animal | Inanimate | Wearable | Revolve | Flammable |
| Cow   | 0                 | 1      | 0         | 0        | 0       | 0         |
| Shirt | 0                 | 0      | 1         | 1        | 0       | 0         |
| Light | 0                 | 0      | 1         | 0        | 0       | 1         |
| Bulb  | 0                 | 0      | 1         | 0        | 0       | 0         |
| Rahim | 1                 | 0      | 0         | 0        | 0       | 0         |
| Boy   | 1                 | 0      | 0         | 0        | 0       | 0         |
| Fan   | 0                 | 0      | 1         | 0        | 1       | 0         |

#### B. Pronoun

In English sentences, pronoun is also used instead of noun in many situations. Actually, pronoun is the identifier of that noun. Pronoun comes as subject in the sentences generally. So it play vital role in semantic analysis. Generally, semantic features of pronoun are similar to that of noun. But there are also some extra features should be included for identifying pronoun [9]. Semantic features of some pronoun are given in the Table II.

TABLE II  
SEMANTIC FEATURES OF PRONOUN

| Word | Semantic Features |        |           |          |         |
|------|-------------------|--------|-----------|----------|---------|
|      | Human             | Animal | Inanimate | Imitable | Subject |
| I    | 1                 | 0      | 0         | 1        | 1       |
| You  | 1                 | 0      | 0         | 1        | 1       |
| He   | 1                 | 0      | 0         | 1        | 1       |
| They | 1                 | 0      | 0         | 1        | 1       |
| It   | 1                 | 1      | 1         | 0        | 1       |

#### C. Phrasal Verb

When a verb attached with a preposition or adverb, is used as an idiomatic sense, then it is called a Phrasal Verb [10]. The preposition attached with the verb is called as its particle. For example, many verbs take idiomatic senses by connected with their next prepositions. Such as, “I shall call on him” means “আমি তার সাথে দেখা করবো”. Here, “call on” means “দেখা করা”, which is a special meaning where as “call” & “on” individually has separate meanings. In this way, Example of commonly used phrasal verbs are call at, call on, come across, come by, put on, put off etc. The absolute meanings of these phrasal verbs depend on the category of the subject and the object of the sentence in which they are used. Some Phrasal verbs of English are presented in table III with their different possible meanings and Bangla equivalent respectively.

### III. PROPOSED TRANSLATION FRAMEWORK

The overall translating process can be depicted as in Fig. 1. This module can be able to translate the Phrasal verbs from English sentence to Bangla. Lexical analyzer will produce tokens of English words from the input sentence. The Tokens produced by the lexical analyzer are passed as input to the syntax analyzer. Syntax analyzer produces parse tree according to given grammatical rules and this parse tree acts as an input to the semantic analyzer. Semantic analyzer

TABLE III  
DIFFERENT BANGLA MEANINGS OF PHRASAL VERBS

| Phrasal Verb | English Meaning                                    | Equivalent Bangla Meaning                                                                                         |
|--------------|----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| Put on       | 1. Wear<br>2. Enkindle<br>3. Show<br>4. Switch on  | 1.Poridhan_kora (পরিধান করা)<br>2.Jaalaano (জ্বালানো)<br>3.Prodarshan_kora (প্রদর্শন করা)<br>4.Chaalaano (চালানো) |
| Run over     | 1. Occurring accident<br>2. Be overflow<br>3. Read | 1.Chaapaa_deya (চাপা দেয়া)<br>2.Upche_pora (উপচে পড়া)<br>3.Pora (পড়া)                                          |
| Take off     | 1. Put off<br>2. Take flight<br>3. Mock at         | 1.Khule_fela (খুলে ফেলা)<br>2.Uddoyon_kora (উড্ডয়ন করা)<br>3.Upohaash_kora (উপহাস করা)                           |
| Put off      | 1. Take off<br>2. extinguish                       | 1.Khule_fela (খুলে ফেলা)<br>2.Neevaano (নিভানো)                                                                   |
| Look up      | 1. Visit<br>2. Find out<br>3. Improvement          | 1.Dekha_kora (দেখা করা)<br>2.Khonja(উড্ডয়ন করা)<br>3.Unnoti_hoya(উনডুবতি হওয়া)                                  |
| Hold up      | 1. Keep up<br>2. Exposition<br>3. Stopped          | 1.Jaagiye_rakha (জাগিয়ে রাখা)<br>2.Jaahir_kora (জাহির করা)<br>3.Thamiye_rakha (থামিয়ে রাখা)                     |

produces annotated parse tree for analyzing the English sentences semantically. The example base contains the semantic information of English word with their possible Bangla meaning. Finally, the translation engine compares with the output of each node of the semantic analyzer with the example base. If translation engine find a match with semantic features of subject and object then gives the output of the equivalent meaning in Bangla. As for example, a sentence “*I put on cap*” is considered as input.

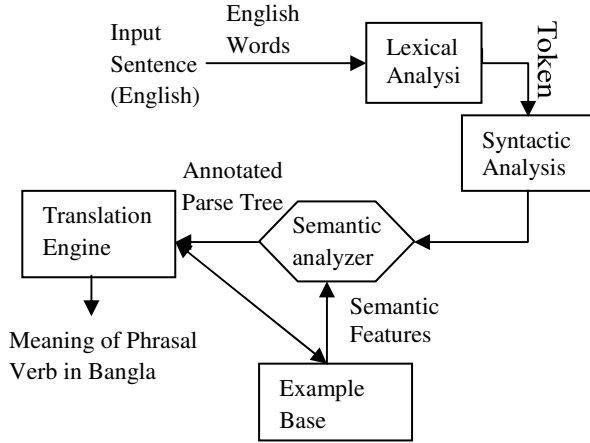


Fig 1: Translation process for translating Phrasal verbs.

#### A. Lexical Analysis

In this analysis, the stream of characters making up the sentence is read from left to right and grouped into token. A set of characters that have a collective meaning is called token. Actually, lexical analyzer separates each meaningful word in the sentence. Unfortunately, many words have several meanings and it may not be possible to choose the correct one just by looking at the word itself. To find out the correct meaning whole sentence should be looked up. And this reflects the semantic feature of machine translation. For example, the output of the lexical analyzer for the input sentence, “*I put on cap*”, is represented as: “*I*”, “*put*”, “*on*”, “*cap*”.

#### B. Syntactic Analysis

Syntactic analysis is the process of taking an input string of expressions and producing representation of kind. In Syntactic analysis, characters or tokens are grouped hierarchically into nested collections with collective meaning. It involves grouping of the tokens into grammatical phrases that are used as the input of semantic analyzer. Usually, a parse tree represents grammatical phrases of the sentence. A parse tree is a graphical representation of a sentence, which shows how the start symbol of a grammar derives the tokens or words in the sentence. Actually syntactic analysis is the process of determining if a string of tokens can be generated by a set of grammatical rules [11]. For the input sentence, “*I put on cap*”, the parse tree is shown as in Fig. 2.

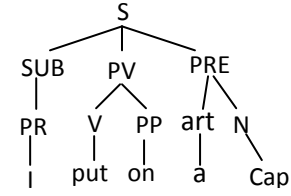


Fig. 2. Parse tree for the sentence, “*I put on cap*”.

#### C. Semantic Analyzer

Semantics is concerned with the meaning of words and a word with semantic features is corresponding to their sense component. The job of the semantic analyzer is to collect the semantic features of the tokens generated by the Lexical analyzer or of the nodes of the parse tree generated by the syntactic analyzer. It also generates the semantic tree i.e. the annotated parse tree. It is a parse tree with semantic features and attributes. The annotated parse tree with semantic features of the input sentence, “*I put on cap*” is shown in Fig. 3. In the Fig. 3, the annotated parse tree of

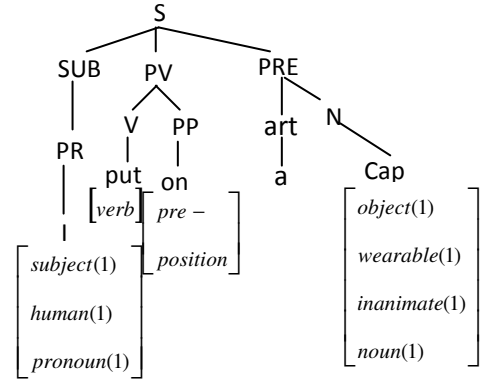


Fig. 3. Annotated parse tree for the sentence, “*I put on cap*”.

the input sentence “*I put on cap*” have been described for detection of phrasal verbs and execution of the semantic features of subject and object. Semantic features of the nouns and pronouns used here are represented as follows:

*I*: [Human(1), Subject(1), Pronoun(1)] *Cap*: [Inanimate(1), Wearable(1), Object(1), Noun(1)]

From the Fig .3. it can easily be noticed that subject “*I*” is a human and object “*cap*” is an inanimate and wearable thing. And there remains a phrasal verb “*put on*”.

#### D. Example-Base

Example-base contains the semantic information of different types of English word. The example base entries incorporate information about context, form and semantics. All possible meaning in different context is stored in the example base. Example base also contains the different possible meanings of Phrasal verb in different context with their equivalent Bangla. Table IV shows the different Bangla meanings of the phrasal verb “*put on*” for four different semantic features of subject and object in the input sentences.

TABLE IV  
DIFFERENT BANGLA MEANING OF PHRASAL VERB, **PUT ON** WITH DIFFERENT SUBJECT AND OBJECT.

| Input Sentence                 | Semantic feature of Subject    | Semantic feature of Object                  | Equivalent Bangla Meaning      |
|--------------------------------|--------------------------------|---------------------------------------------|--------------------------------|
| I <b>put on</b> cap            | I: [Human, Subject, Pronoun]   | Cap: [Inanimate, Object, Noun, Wearable]    | Poridhan_kora (পরিধান করা)     |
| You <b>put on</b> the light    | You: [Human, Subject, Pronoun] | Light: [Inanimate, Object, Noun, Flammable] | Jaalaano (জ্বালানো)            |
| The boys <b>put on</b> a drama | Boy: [Human, Subject, Noun]    | Drama : [Inanimate, Object, Noun, Showable] | Prodorshon_kora (প্রদর্শন করা) |
| He <b>put on</b> the fan       | He: [Human, Subject, Pronoun]  | Fan: [Inanimate, Object, Noun, Revolvable]  | Chaalaaano (চালানো)            |

#### E. Translation Engine

After generating the parse tree and annotated parse tree with semantic features of the input sentence, the translation engine translates the Phrasal verb from English to Bangla. The translation process may be describes as follows:

1. The semantic analyzer gives the annotated parse tree with semantic information of each node including subject and object of the input sentence.
2. The example base contains the several possible meanings of phrasal verbs with semantic features. Example base also contains the semantic features of the subject and object of the input sentence respectively.
3. Identify the Phrasal verb of the input sentence. Generally, a verb followed by a preposition is assigned to a Phrasal verb.
4. The machine compares the semantic features of subject and object of the input sentence with the example base.
5. Search the possible meaning of the Phrasal verb. If the semantic information of each node is matched.
6. With the example base then the equivalent Bangla meaning of that Phrasal verb is found.
7. Else machine shows 'No meaning'.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

We have tested our proposed engine for about 1000 sentences with different word length. The length of the sentences generally varies from three to seven words. All the sentences contained the different Phrasal verbs in different contexts. Most of the examples have been taken from the English newspapers, journal and text books. Translation snapshot of the English phrasal verb “put on” that have different meanings in different context have been represented in Table V as an example. From the Table, we have seen that the proposed module can be able to translate the phrasal verbs of English sentence into its Bangla equivalent successfully. The proposed system may be extended and hence the overall performance of the system will be improved by incorporating with more constituent parts to the existing system.

#### A. Performance Evaluation

The module is able to translate the phrasal verbs with good accuracy. Table VI shows the number of successful and unsuccessful translation of sample sentences as well as the translation rate and error.

TABLE VI  
SUCCESS RATE OF THE PROPOSED TRANSLATION MODULE

| Length of Sentence (No. of words) | No. of Sentences | Successful Translation | Unsuccessful Translation | Translation Rate (%) | Translation Error (%) |
|-----------------------------------|------------------|------------------------|--------------------------|----------------------|-----------------------|
| 3                                 | 150              | 120                    | 30                       | 80%                  | 20%                   |
| 4                                 | 300              | 280                    | 20                       | 93.33%               | 6.67%                 |
| 5                                 | 250              | 240                    | 10                       | 96%                  | 4%                    |
| 6                                 | 200              | 120                    | 80                       | 60%                  | 40%                   |
| 7                                 | 100              | 20                     | 80                       | 20%                  | 80%                   |
| <b>Total</b>                      | <b>1000</b>      | <b>780</b>             | <b>220</b>               | <b>79%</b>           | <b>21%</b>            |

From the Table VI it is observed that translation rate is very high for the sentences having four and five words individually. For three-word sentences, the rate is satisfactory i.e. eighty percent. Fig. 4 represents the translation accuracy and error rate versus word lengths of the system. From three to five words, translation accuracy increase moderately while accuracy starts decreasing with the increase of the length of the sentences from five words and even for seven-word sentences it is only twenty percent. On the other hand, there is a gradual rise in error rates with the number of word lengths.

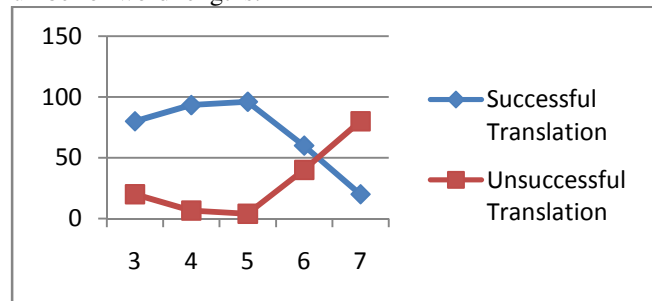
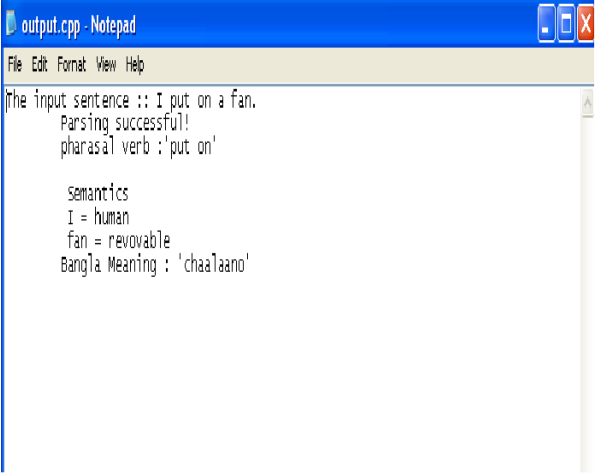
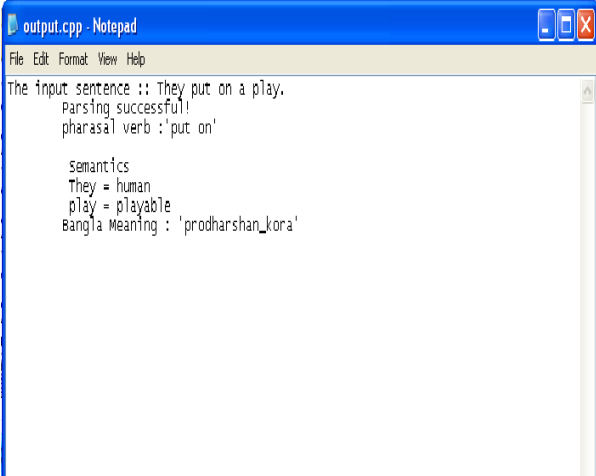
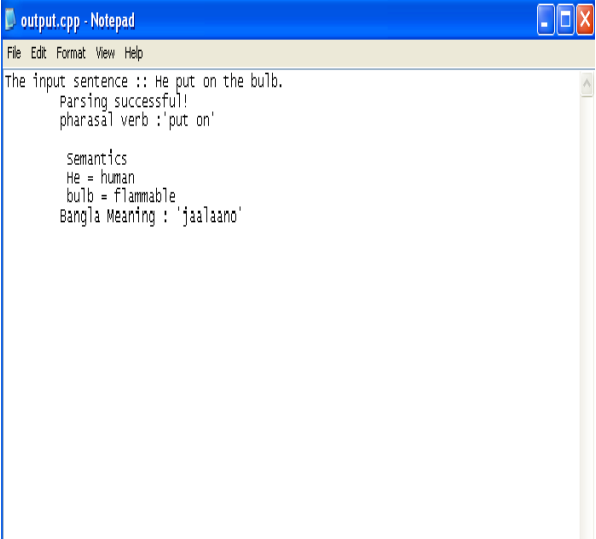


Fig. 4. Word Vs Translation Rate & Translation Error Rate

**TABLE V**  
**EXAMPLE TRANSLATION FOR THE PHRASAL VERB, “PUT ON”**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Example 1</b><br/>Input Sentence: I put on a fan.<br/>Tokens:<br/> I : Pronoun<br/> put : Verb<br/> on : Preposition<br/> a : Article<br/> fan : Noun<br/> Grammar: pronoun + verb + preposition + a + id + full_stop.<br/> Grammar Rule:<br/> S → Suba P_Verb Predicate<br/> Suba → pronoun<br/> P_Verb → verb Pre<br/> Pre → preposition<br/> Predicate → Obj full_stop<br/> Obj → Article Subb<br/> Article → a<br/> Subb → id<br/> Phrasal Verb: put_on<br/> Semantic Features:<br/> I = human<br/> fan = revolvable<br/> Verification Result: Parsing Successful<br/> Meaning: 'Chaalano'</p>                  |  <pre> output.cpp - Notepad File Edit Format View Help The input sentence :: I put on a fan. Parsing successful! pharasal verb : 'put on'  semantics I = human fan = revolvable Bangla Meaning : 'chaalaano' </pre>               |
| <p><b>Example 2</b><br/>Input Sentence: They put on a play.<br/>Tokens:<br/> They : Pronoun<br/> put : Verb<br/> on : Preposition<br/> a : Article<br/> play : Noun<br/> Grammar: pronoun + verb + preposition + a + id + full_stop.<br/> Grammar Rule:<br/> S → Suba P_Verb Predicate<br/> Suba → pronoun<br/> P_Verb → verb Pre<br/> Pre → preposition<br/> Predicate → Obj full_stop<br/> Obj → Article Subb<br/> Article → a<br/> Subb → id<br/> Phrasal Verb: put_on<br/> Semantic Features:<br/> They = human<br/> Play = playable<br/> Verification Result: Parsing Successful<br/> Meaning: 'Pradorshan_kora'</p> |  <pre> output.cpp - Notepad File Edit Format View Help The input sentence :: They put on a play. Parsing successful! pharasal verb : 'put on'  semantics They = human play = playable Bangla Meaning : 'prodharshan_kora' </pre> |
| <p><b>Example 3</b><br/>Input Sentence: He put on the bulb.<br/>Tokens:<br/> He : Pronoun<br/> put : Verb<br/> on : Preposition<br/> the : Article<br/> bulb : Noun<br/> Grammar: pronoun + verb + preposition + the + id + full_stop.<br/> Grammar Rule:<br/> S → Suba P_Verb Predicate<br/> Suba → pronoun<br/> P_Verb → verb Pre<br/> Pre → preposition<br/> Predicate → Obj full_stop<br/> Obj → Article Subb<br/> Article → the<br/> Subb → id<br/> Phrasal Verb: put_on<br/> Semantic Features:<br/> He = human<br/> bulb = flammable<br/> Verification Result: Parsing Successful<br/> Meaning: 'jaalaano'</p>     |  <pre> output.cpp - Notepad File Edit Format View Help The input sentence :: He put on the bulb. Parsing successful! pharasal verb : 'put on'  semantics He = human bulb = flammable Bangla Meaning : 'jaalaano' </pre>         |

## V. FUTURE POSSIBILITIES

There are lots of research scopes of this paper on the field of Bangla NLP. The possible future improvements are mentioned in the following:

- The disambiguation module can be incorporated into the front-end of an existing MT system.
- The example base may be extended incorporating more semantic features for long sentences.
- Translation of Phrasal verbs in English complex and compound sentences may be carried on.
- Another component pertaining to prepositional phrases can be incorporated in the example base.
- Compositional semantics features may be used to identify the English words.
- For disambiguated meaning the concept of neural networks may be employed.

## VI. CONCLUSION

Phrasal verbs have highly context-dependent meanings and may be disambiguated only by devising a technique involving utilization of semantic information pertaining to the context. The machine translation of Phrasal verbs of English into Bangla becomes more difficult and challenging tasks because English and Bangla are widely different in structure and style. The semantic interpretation is the most difficult stage of transformation process. For unambiguous analysis or generation, semantic properties of each word of a sentence must be taken into account for actual translation. For accurate machine translation from Bangla to English or vice versa, semantic features of Bangla words should be structured. Their interpretation may require the use of heuristics, syntactic and semantic constraints, inference, and other forms of object analysis within the discourse content. The proposed framework gives the very high accuracy with the simple assertive sentence for three to five word lengths but the accuracy is degraded for a sentence of long phrase structure with Phrasal verb. If more semantic features are employed with the existing features then the performance of the system will be improved.

## REFERENCES

[1] B. S. Pedersen and S. Nimb, "Semantic encoding Of Danish Verbs in SIMPLE-Adapting a Verb -Framed Model to a Satellite Framed Language", In Proc. *2<sup>nd</sup> International Conference on Language Resources and Evaluation*, LREC 2000, Athens, Greece.

[2] V. Bhandari, R. M. K. Sinha, and A. Jain, "Disambiguation of Phrasal Verb Occurrences for Machine Translation", In Proc. of the *Second Symposium on Translation Support Systems*, STRANS' 02, Kanpur, India, 2002.

[3] M. M. Asaduzzaman and M. M. Ali, "Transfer Machine Translation-An Experience with Bangla English Machine Translation System", In Proc. *International Conference on Computer and Information Technology, ICCIT'04*, Dhaka, Bangladesh, pp. 265-270, vol. 4, 2004.

[4] D. Saha and S. Bandyopadhyay, "A Semantics-Based English-Bengali EBMT System for Translating News Headlines", *MT SUMMIT X*, Phuket, Thailand, 2005.

[5] M. M. Hoque and M. M. Ali, "Semantic Features and Redundancy Rules for Analyzing Bangla Sentences", In Proc. *International Conference on Computer and Information Technology, ICCIT'05*, Dhaka, Bangladesh, pp. 1198-1201, vol. 4, 2005.

[6] G. Hossian, A. Kabir and N. Islam, "Development of Smart Bangla Conversion Processor: A New Approach", In Proc. *National Conference on Computer Processing in Bangla, Dhaka*, Bangladesh, February 2005.

[7] M. M. Hoque and M. M. Ali, "A Parsing Methodology for Bangla Natural Language Sentences", In Proc. *International Conference on Computer and Information Technology, ICCIT'03*, Dhaka, Bangladesh, vol. 2, pp. 277-282, 2003.

[8] "উচ্চতর বাংলা ব্যাকরণ ও রচনা", Globe Library Private Limited, , Dhaka, December 1996.

[9] M. M. Hoque and M. J. Rahman, "Lexical Semantics: A New Approach to Analyze the Bangla Sentence with Semantic Features ", In Proc., *International Conference on Information and Communication Technology, ICICT '07*, Dhaka, Bangladesh, pp. 87-91, 2007.

[10] A. H. M. Mohiuddin and A. F. M. A. Kasem, "A Text Book of Advanced Functional English", Ideal Library, Dhaka, 1999.

[11] M. R. Selim and M. Z. Ikbal, "Syntax Analysis of Phrases and Different types of Sentences in Bangla", In Proc., *International Conference on Computer and Information Technology, ICCIT '99*, Sylhet, Bangladesh, vol. 2, pp. 175-186, 1999.



# Design of an Interactive Bangla Mobile Keypad Based on Ergonomics

Muhammad Raisul Alam<sup>1</sup>, Md. Masum<sup>2</sup>, Shafiul Hasan Md. Tareq<sup>3</sup>  
and Mostafa Wasiuddin Numan<sup>4</sup>

<sup>1</sup>Institute of Microengineering and Nanoelectronics (IMEN), Universiti Kebangsaan Malaysia

<sup>2</sup>Dept. of Computer Science and Engineering, Shah Jalal University of Science and Technology

<sup>3</sup>IT Magnet Dhaka Ltd., Bangladesh

<sup>4</sup>Dept. of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia

**Abstract** — In this paper, a Bangla mobile keypad layout is proposed based on ergonomics. The finger movement across the keypad is analyzed and an optimal keypad layout is proposed for right handed one thumb operation. To minimize the effort of the user's thumb movement, the characters are organized in such a way that the most frequently used characters are positioned in the most comfortable zone of the keypad. A couple of equations are derived to calculate the required time to insert a character in different existing layouts. The performance is analyzed by comparing time which reveals that the proposed layout outperforms the other layouts.

**Index Terms** — Bangla Characters, Bangla, Ergonomics, Mobile keypad, Mobile phone.

## I. INTRODUCTION

Minimizing the time to insert Bangla text in a mobile handset is a challenge to researchers because in Bangla script there are huge number of characters and joint characters. It is even more difficult to accommodate all the characters in a mobile keypad of only 12 buttons. Mobile phone manufacturer used the concept of English sequential keypad layout to introduce Bangla in mobile phones [1], [2], [3] which is not feasible. A frequency based multi-tap keypad layout is also proposed in [4]. Another keypad layout is proposed based on phonetics [5]. All of the designs require more keystrokes and effort to insert Bangla character. To reduce time to write Bangla using mobile keypad one has to consider human factor, Bangla joint character formation and Bangla character frequency. In this paper an innovative approach is considered to propose a Bangla keypad layout where character frequency and user comfort is taken into account and this fulfills all of the above criteria.

## II. ANALYSIS OF MOBILE KEYPAD

The main difficulty to introduce Bangla characters to mobile keypad is that it has to accommodate a large number of characters onto a keypad consisting only 12

keys. To perform this task a number of keypad layouts have been proposed.

### A. The English Keypad

The English keypad is designed using sequential ordering of English characters. So user can easily memorize the position of the characters. To position the characters sequentially the character frequency is not considered.

### B. The common Bangla keypad based on character sequence

The most common Bangla keypad is designed based on the alphabet ordering (Like English). The layout of a typical Bangla keypad used by the mobile phone manufacturer [1]-[3] is given in Table I.

TABLE I  
COMMON BANGLA LAYOUT

|                      |                       |                           |
|----------------------|-----------------------|---------------------------|
| 1<br>Punctuation key | 2<br>ক খ গ ঘ ঙ ১      | 3<br>চ ছ জ ঝ ঞ ২          |
| 4<br>ট ঠ ড ঢ ণ ৩     | 5<br>ত থ দ ধ ন ৪      | 6<br>প ফ ব ভ ম ৫          |
| 7<br>য র ও ল শ ষ স ৬ | 8<br>হ ড় ফ় র় ৭ ৪ ৭ | 9<br>অ ই দ্বি উ ঊ ঋ ঌ ঐ ৮ |
| *<br>† ৿ ৺ ৻ ৼ ৽ ৾ ৿ | 0<br>special symbol   | #<br>joint character      |

It is a very simple layout and there is no complicated using procedure. So the user may feel comfort with this layout. In this layout, many keystrokes are needed for vowels and short vowels because one key is used to represent each group. Each key is used to represent up to eleven characters which is very inefficient. There is no mode change key or any shift key which can easily decrease the key-tapping. Joint characters also consume more time in this design. The keypad is not based on ergonomics.



option of doing the same thing by pressing the special symbol key (# key). It reduces user's burden from memorizing difficult formation of joint characters. As it is an alternative feature, it placed in the least comfortable zone.

6) Other characters are placed following the sequence of Bangla alphabet.

7) The Mobile keypad is too small to draw many characters on it. Using sequential character positioning, several characters are shown using only two characters. For example, the characters ক, খ, গ, ঘ, ঙ can be shown as ক - ঙ .

#### B. Bangla typing procedure

The Bangla typing procedure in the proposed layout is explained with some examples. To write 'ক্ষা', the joint consonant ক্ষ (ক্ষ = ক + খ) can be entered using two ways. Firstly, user has to press key-2 once then press 'Mode' key to change mode from normal to extended, then press key-9 five times. The 'Mode' key automatically switches to normal mode after inserting the joint character. Secondly, user can also insert the joint consonant using 'Special' key (# key) several times. To insert 'খ' user has to press key-8 five times. To insert 'ব' he has to press 'Mode' key once to switch to extended mode and has to press key-4 once. The 'Mode' key automatically switches to normal mode after inserting short vowel. Another example, to insert 'বুক্ষ' one has to press key-9 once, 'Mode' key once, key-4 four times, then key-2 once, 'Mode' key once and finally key-6 once.

#### IV. PERFORMANCE ANALYSIS

In this paper, a couple of equations are derived to calculate the time required to enter a character in mobile

phone. The performance of the proposed layout is analyzed by comparing the time. In the following equations  $T_s$ ,  $T_p$ , and  $T_e$  represent total time to insert character in sequential, phonetic and ergonomic layout respectively.

$$T_s = t_a + c * (t_i * n_r) \quad (1)$$

$$T_p = (t_a + t_i) * s + (t_a + c * (t_i * n_r)) + (t_a + t_i) * l \quad (2)$$

$$T_e = (t_a + t_i) * m + (t_a + c * (t_i * n_r)) \quad (3)$$

In (1), (2) and (3), the following notations are adopted:

|       |                                                                                  |
|-------|----------------------------------------------------------------------------------|
| $t_a$ | average time required to bring the finger to desired key                         |
| $t_i$ | time interval between consecutive keystrokes to search a character in a same key |
| $n_r$ | number of key repressing                                                         |
| $c$   | Comfort factor                                                                   |
| $m$   | mode variable                                                                    |
| $l$   | link variable                                                                    |
| $s$   | shift variable                                                                   |

The value of  $m$ ,  $l$  and  $s$  is equal to 1 when these keys are pressed and otherwise 0. The value of  $c$  depends on the comfort zones. The different values of  $c$  are as follows,

|                        |            |
|------------------------|------------|
| Most comfort zone      | $c = 0.75$ |
| Medium comfort zone    | $c = 0.90$ |
| Least comfortable zone | $c = 1.00$ |

TABLE III  
REQUIRED TIME TO INSERT VOWELS AND SHORT VOWELS

| Character | Phonetics | Sequence | Ergonomics | Character | Phonetics | Sequence | Ergonomics |
|-----------|-----------|----------|------------|-----------|-----------|----------|------------|
| অ         | 1.525     | 0.8      | 0.725      | ই         | 0.725     | 0.725    | 1.525      |
| আ         | 0.725     | 1.525    | 0.95       | ঈ         | 0.95      | 0.95     | 1.75       |
| ই         | 1.175     | 1.1      | 1.175      | ঊ         | 1.175     | 1.175    | 1.975      |
| ঈ         | 1.875     | 1.4      | 1.4        | ঋ         | 1.4       | 1.4      | 2.2        |
| উ         | 1.04      | 1.7      | 1.615      | ঌ         | 1.615     | 1.615    | 2.425      |
| ঊ         | 1.84      | 2.0      | 1.85       | ঍         | 1.85      | 2.525    | 1.525      |
| ঋ         | 1.875     | 2.3      | 0.725      | ঔ         | 2.075     | 1.85     | 1.75       |
| এ         | 1.04      | 2.6      | 0.95       | ও         | 2.3       | 2.075    | 1.975      |
| ঐ         | 1.84      | 2.9      | 1.175      | ঔ         | 2.525     | 2.3      | 2.2        |
| ও         | 1.31      | 3.2      | 1.4        |           |           |          |            |
| ঔ         | 2.11      | 3.5      | 1.615      |           |           |          |            |

If  $t_a = 0.5$  sec and  $t_i = 0.3$  sec, then the time required to insert vowels and short vowels are listed in Table III. These are arbitrary value. As we are only interested to compare between layouts, any arbitrary value according to finger movement is suitable for it.

Fig. 3 illustrates a performance comparison between phonetics, sequential and ergonomics based Bangla keypad layout. Among the 11 vowels, 7 of them need less time in comparison with phonetics and sequential layout. And among 9 short vowels, inserting 4 short vowels need less time than other keypad layouts. As vowels and short vowels are most frequently used characters [6], [7], ergonomics based keypad shows better performance than others. 20 out of 39 consonants require less time than other keypads. Table IV shows the time required to insert consonants in different

layouts. Table V shows some joint characters where ergonomic keypad performs the best.

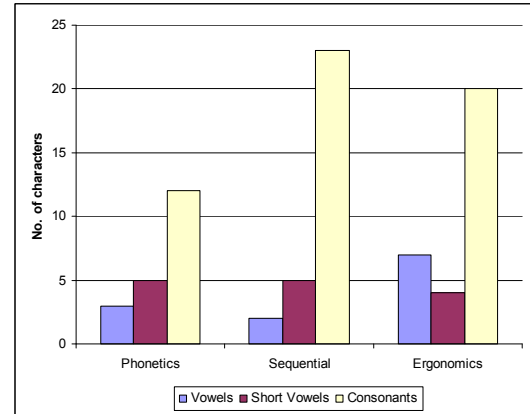


Fig. 3. Required time comparison between different layout

TABLE IV  
REQUIRED TIME TO INSERT CONSTANTS

| Character | Phonetics | Sequence | Ergonomics | Character | Phonetics | Sequence | Ergonomics |
|-----------|-----------|----------|------------|-----------|-----------|----------|------------|
| ক         | 0.95      | 0.725    | 0.725      | প         | 0.725     | 0.77     | 0.77       |
| খ         | 1.75      | 0.95     | 0.95       | ফ         | 1.525     | 1.04     | 1.04       |
| গ         | 0.725     | 1.125    | 1.125      | ব         | 0.95      | 1.31     | 1.31       |
| ঘ         | 1.525     | 1.4      | 1.4        | ভ         | 1.75      | 1.58     | 1.58       |
| ঙ         | 0.8       | 1.625    | 1.625      | য         | 0.77      | 1.85     | 1.85       |
| চ         | 1.175     | 0.77     | 0.77       | র         | 1.525     | 0.725    | 0.8        |
| ছ         | 1.875     | 1.04     | 1.04       | ল         | 1.175     | 0.95     | 1.1        |
| জ         | 0.725     | 1.31     | 1.31       | শ         | 1.175     | 1.125    | 1.4        |
| ঝ         | 1.75      | 1.58     | 1.58       | স         | 2.2       | 1.4      | 1.7        |
| ঞ         | 1.6       | 1.85     | 1.85       | ষ         | 1.57      | 1.625    | 2.0        |
| ট         | 0.77      | 0.725    | 0.725      | ণ         | 1.4       | 1.85     | 2.3        |
| ঠ         | 1.57      | 0.95     | 0.95       | হ         | 0.95      | 0.77     | 0.8        |
| ড         | 0.77      | 1.125    | 1.125      | ড়        | 1.875     | 1.04     | 1.1        |
| ঢ         | 1.57      | 1.4      | 1.4        | ঢ়        | 2.2       | 1.31     | 1.4        |
| ণ         | 1.84      | 1.625    | 1.625      | য়        | 1.4       | 1.58     | 1.7        |
| ত         | 1.1       | 0.725    | 0.725      | ৎ         | 1.175     | 1.85     | 0.8        |
| থ         | 1.9       | 0.95     | 0.95       | ৎ         | 2.5       | 2.12     | 1.1        |
| দ         | 1.31      | 1.125    | 1.125      | ঃ         | 1.875     | 2.39     | 1.4        |
| ধ         | 2.11      | 1.4      | 1.4        |           | 1.75      | 2.66     | 1.7        |
| ন         | 1.04      | 1.625    | 1.625      |           |           |          |            |

TABLE V  
REQUIRED TIME TO INSERT JOINT CHARACTERS

| Character | Phonetic | Sequence                                                         | Ergonomic |
|-----------|----------|------------------------------------------------------------------|-----------|
| ক         | 3.52     | The sequence is undefined. Over head will increase sequentially. | 3.525     |
| খ         | 4.31     |                                                                  | 4.23      |
| গ         | 2.495    |                                                                  | 2.295     |
| ঘ         | 3.62     |                                                                  | 3.05      |
| ঙ         | 3.61     |                                                                  | 2.6       |
| চ         | 3.2      |                                                                  | 2.25      |
| ছ         | 3.6      |                                                                  | 3.025     |
| জ         | 4.17     |                                                                  | 4.65      |

Table VI shows a number of examples which clearly states that the proposed ergonomics based layout outperforms the other two layouts.

TABLE VI  
PERFORMANCE COMPARISON WITH DIFFERENT WORDS

| Words    | Phonetics | Sequence | Ergonomics |
|----------|-----------|----------|------------|
| বাংলাদেশ | 10.935    | 9.655    | 9.91       |
| ঐশ্বর্য  | 5.79      | 6.525    | 5.015      |
| ঋণ       | 23.715    | 3.925    | 2.35       |
| ঈদ       | 3.185     | 2.525    | 2.525      |
| এখন      | 3.83      | 5.175    | 3.525      |

#### IV. CONCLUSION

Efficiency of the keypad can be increased by reordering Bangla vowels and short vowels according to character frequency and comfortable zone. The performance analysis and examples prove that the ergonomics based keypad layout significantly outperforms the phonetics and sequential layouts. In ergonomics layout, 7 out of 11 vowels, 4 out of 9 short vowels and 20 out of 39 consonants require less time than other two layouts because of the placements of the most frequently used characters in the most comfortable zone. The proposed layout performs better than the existing layouts and represents a good solution for Bangla typing in mobile phone.

#### REFERENCES

- [1] Aktel Bangla SMS. <http://www.aktel.com/>.
- [2] CityCell Bangla SMS. <http://www.citycell.com/>
- [3] Grameen Phone Ltd. Bangla SMS. <http://www.grameenphone.com/>
- [4] M. M. Morshed, Y. K. Thu and Y. Urano, "Frequency Based Two-Layer Multitap Bangla Input Method for Mobile Phones," *10th International Conference on Computer and Information Technology, (ICCIT 2007)*, pp. 1-7, Dec. 2007.
- [5] A. K. Pathan, S. R. Choudhury, A. K. M. N. Islam, "Design of an Interactive Bangla Mobile Keypad Based on Phonetics," *7th International Conference on Computer and Information Technology, (ICCIT 2004)*, pp. 658-660, Dec. 2004.
- [6] M. H. Alam, A. K. M. Masum, M. M. Hassan, S. M. Kamruzzaman "Optimal Bangla Keyboard Layout using Association Rule of Data Mining," *7th International Conference on Computer and Information Technology, (ICCIT 2004)*, Dec. 2004.
- [7] M. S. Hossain, K. M. E. Ali, R. C. Debnath, "Frequency Distribution of Bangla Words in Bangla Newspapers", *National Conference on Computer Processing of Bangla (NCCPB 2004)*, Feb. 2005.

# Crowd Behaviour Monitoring on the Escalator Exits

Md. Haidar Sharif, Nacim Ihaddadene, and Chabane Djeraba

Computer Science Laboratory of Lille (LIFL)  
University of Science and Technology of Lille (USTL)  
59655 Villeneuve d'Ascq, France  
{md-haidar.sharif, nacim.ihaddadene, chabane.djeraba}@lifel.fr

**Abstract**—This paper presents an approach to detect an abnormal situation in a crowd scene. The proposed approach estimates sudden changes and abnormal motion variations in a set of interest points. The number of tracked points of interest is reduced by using a mask that corresponds to the hot areas of the built motion heat map. Optical flow technique tracks the points of interest. There are sufficient variations in the optical flow patterns in a crowd scene when there are cases those showing abnormal situations. Statistical treatment of optical flow information has been thresholded. To demonstrate the interest of this approach, we present the results based on the detection of collapsing events in real videos of airport escalator exits.

**Index Terms**—Points of interest, Motion heat map, Optical flow, Entropy, Breakdown situation, Escalator exits.

## I. INTRODUCTION

Escalators, are often installed in pairs, are the mechanized moving stairways, can frequently be seen in airports, shopping malls, cinema halls, museums, hospitals, hotels, multi-story parking garages, train stations, subway stations, or other public and private places. Escalator-related injuries occur infrequently but may result in significant trauma. To evaluate the effect of the safety rules and passenger education, authors in [1] studied escalator-related injury patterns occurring at Taipei Metro Rapid Transit (MRT) stations. In 2000, the accident rate for escalator riding was about 0.815 accidents per million passenger trips through MRT heavy capacity stations. However, their study reported that the majority of escalator injuries were caused by passengers trying to perform other tasks while riding, loss of balance, not holding the handrail, or riders who were struck by other passengers. Automatic video surveillance systems would benefit from a system capable of observing people as well as other moving and interacting objects in a given environment (e.g., escalator in airport, supermarket, subway, etc.) for patterns of normal or abnormal activities, interesting events, and other domain-specific goals.

We will demonstrate an approach which is characterized by optical flow patterns of human behaviours followed by some statistical treatments to detect abnormalities in crowd scenes. We set out by performing a calculation of a motion heat map during a period of time to extract the main regions of motion activity. The use of heat map image improves the quality of the results and reduces processing time which is a great consequence for real-time applications. Points of interest are extracted in the hot regions of the scene. Optical flow is computed on these points of interest, marked off the boundaries by the hot areas of the scene. The optical flow information from video presents

the crowd multi-modal behaviors as optical flow patterns make different in time. There is sufficient agitation in the optical flow pattern in the crowd in case of abnormal and/or emergencies situations (see Fig. 3 (c)). Variations of motions are estimated to make a distinction in favor of potential abnormal events. We consider that we are in a crowd environment with no restrictions of the number of people. To analyze the optical flow patterns of human crowds scenes, we have concentrated on escalator videos to use in our applications. One practical application of our approach is in the detection of real-time collapsing events, which could lead to potentially full of danger or risk situations in exit escalators. The videos used are from camera installed at an airport to monitor the situation of escalator exits. The impression of the application is to have exit escalators continuously perceived to respond rapidly in the event of any collapsing (breakdown). With this aim, cameras are installed in front of the exit locations to view and dispatch the video signal to a control room. In the control room consecrated employees can oversee, take control, and react favorably to a breakdown under the circumstances of an anomalous and/or emergency.

In the remaining parts of the paper, we will start by presenting some related works in Section 2. In Section 3, we will confer our proposed approach. In Section 4, we will discuss the experimental results. Finally, Section 5 presents the conclusion with few clues for further investigation.

## II. RELATED WORKS

We classify the related works into two groups namely *estimation of crowd density* and *aberrant episode detection*.

*Estimation of crowd density*: The methods applied [2], [3], [4], [5] are based on textures and motion area ratio and make an interesting static analysis for crowd surveillance, otherwise than detect abnormal situations.

*Aberrant episode detection*: These works detect irregular events in crowd flows. The general approach consists of modeling normal behaviours, and then estimating the deviant behaviour or attitudes between the normal behaviour model and the observed behaviours. Those deviations are labeled as abnormal. The determining characteristic of the general approach is to exploit the fact that data of standard behaviours are generally available, and data of abnormal behaviours are reasonably less available. For this reason, the divergences from examples of standard behaviours are used to distinguish as the characteristic of abnormality. In this category, authors in [6] encoded optical flow features with Hidden Markov Models (HMM). They used a local modelling

approach to increase the detection sensitivity. The same authors in their another approach [7], where they also used HMM, characterised crowd behaviour by observing the crowd optical flow and used unsupervised feature extraction to encode normal crowd behaviour. The unsupervised feature extraction applied spectral clustering to find the optimal number of models to represent normal motion patterns. The aim of both of their methods is to detect emergency or abnormal events in the crowd. The methods were experimented with simulated data. Authors in [8] proposed a mathematical method based on Lagrangian particle dynamics for crowd flow segmentation and flow instability detection. The method is efficient for segmentation of high density crowd flows (marathons, political events, etc.). Authors in [9] formulated the problem of detecting regularities and irregularities as the problem of composing (explaining) the new observed visual data using spatiotemporal patches extracted from previous visual examples. In the same category, but for low crowded scenes, authors in [10] proposed a visual monitoring system that passively observes moving objects in a site and learns patterns activity from those observations, detect of unusual events in the site that do not fit common patterns using a hierarchical classification.

Our approach contributes to the second category of related works, with succeeding specificities: (i) it considers simultaneously density (motion area ratio) and direction, whereas the previous works methods these factors do not take into account together; (ii) without imposing a restriction of the number of people in the crowd, it detects all events where the variations are important as compared to previous events; (iii) it uses a mask to define a region of interest. This mask is obtained from the hot areas of the motion heat map image. Consequently, useless to analyze the whole scene particularly where there are few motion intensities or no motions. So, it is arousing the attention for real-time applications.

### III. PROPOSED APPROACH

This section shows a flow diagram (see Fig. 1) of the intended approach followed by blow-by-blow implementation procedures.

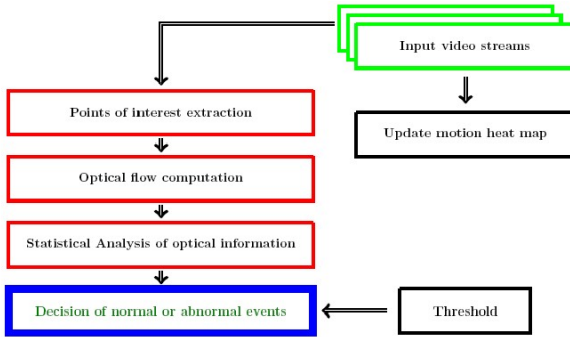


Fig. 1. Block diagram of the proposed framework

#### A. Motion heat map

A heat map is a graphical representation of data where the values taken by a variable in a two-dimensional (2D) map are represented as colours. Motion heat map is a 2D histogram expressing briefly the most important regions of motion activity.

This histogram is built from the accumulation of binary blobs of moving objects, which were extracted following background subtraction method [11]. Assume symbols  $H$  and  $I$  indicate *heat map* and *intensity* respectively. The  $H$  is defined as:

$$H_n(i, j) = H_{n-1}(i, j) + I_n(i, j) \quad (1)$$

$$H_0 = I_0(i, j) \quad (2)$$

where  $n$  is the frame number ( $n \geq 1$ ), and  $i$  &  $j$  are the coordinates (line and column) of the pixel  $(i, j)$  of a frame. The obtained map is used as a mask to define the *region of interest* for the next step of the method. Figure 2 makes noticeable an occurrence of the obtained heat map from an escalator camera view. The use of heat map ameliorates the quality of the results and reduces processing time which is an important factor for real-time applications. The results are more significant when the video duration is long. In practice, even for the same place, the properties of abnormal events may vary depending on the context (day-night, indoor-outdoor, normal-peak time, occasion, vacation, etc.). We build a motion heat map for each set of conditions.

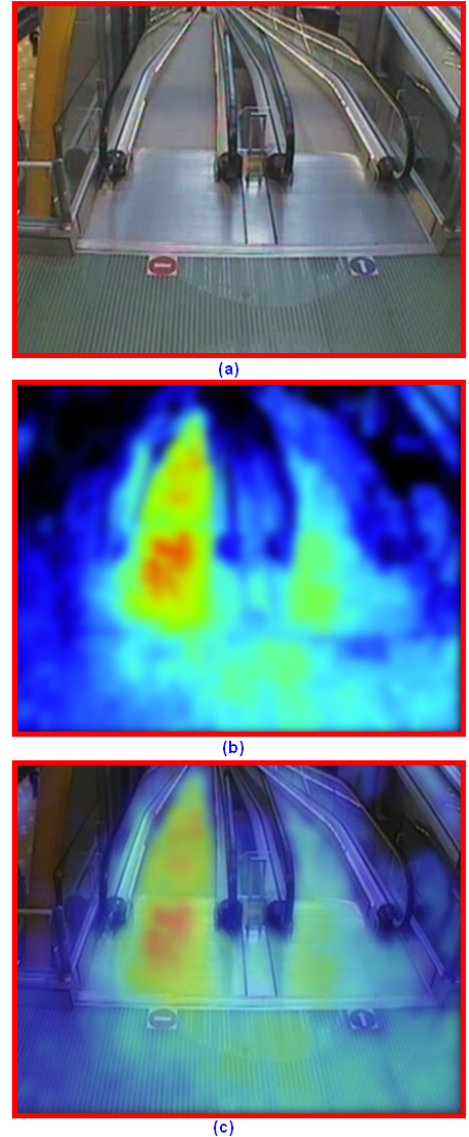


Fig. 2. Motion heat map: (a) camera view, (b) generated motion heat map, and (c) masked view which recommends region of interest.



### B. Points of interest extraction

Moravec's corner detector [12] is a relatively simple algorithm that was used by Moravec and others, but is now commonly considered out-of-date. It is not rotationally invariant (a property prevalent even in more modern corner detectors) as the response is anisotropic (not invariant with respect to direction), is considered to have a noise response, and is susceptible to reporting false corners along edges and at isolated pixels so is sensitive to noise. Nevertheless, it is computationally efficient which was critical for Moravec as he was interested in a real-time application and had minimal computational power at his disposal. The other way around the Harris corner detector [13] is computationally demanding, but directly addresses many of the limitations of the Moravec corner detector. In our approach, we consider Harris corner as a point of interest. The Harris corner detector is a famous point of interest detector due to its strong invariance to rotation, scale, illumination variation, and image noise [14]. It is based on the local auto-correlation function of a signal, where the local auto-correlation function measures the local changes of the signal with patches shifted by a small amount in different directions. Assume a point  $(x, y)$  and a shift  $(\Delta x, \Delta y)$ , then the local auto-correlation function is defined as:

$$c(x, y) = \Sigma_w [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \quad (3)$$

where  $I(\cdot, \cdot)$  denotes the image function and  $(x_i, y_i)$  are the points in the smooth circular window  $w$  centered on  $(x, y)$ . The shifted image is approximated by a Taylor expansion truncated to the first order terms as:

$$I(x_i + \Delta x, y_i + \Delta y) \approx I(x_i, y_i) + [I_x(x_i, y_i)I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (4)$$

where  $I_x(\cdot, \cdot)$  and  $I_y(\cdot, \cdot)$  denote the partial derivatives in  $x$  and  $y$ , respectively. Substituting the right hand site of Eq. 4 into Eq. 3 yields:

$$c(x, y) = \Sigma_w ( [I(x_i, y_i)I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} )^2 \quad (5)$$

$$= [\Delta x \Delta y] M(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (6)$$

where  $M(x, y) =$

$$\begin{pmatrix} \Sigma_w (I_x(x_i, y_i))^2 & \Sigma_w I_x(x_i, y_i)I_y(x_i, y_i) \\ \Sigma_w I_x(x_i, y_i)I_y(x_i, y_i) & \Sigma_w (I_y(x_i, y_i))^2 \end{pmatrix}. \quad (7)$$

The  $2 \times 2$  symmetric matrix  $M(x, y)$  captures the intensity structure of the local neighborhood. Let  $\lambda_1$  and  $\lambda_2$  are the eigenvalues of matrix  $M(x, y)$ . The eigenvalues form a rotationally invariant description. There are three cases to be considered [13]:

- **No point of interest is found:** If both  $\lambda_1$  &  $\lambda_2$  are small, so that the local auto-correlation function is flat, i.e., little change in  $c(x, y)$  in any direction, then the windowed image region is of approximately constant intensity.
- **An edge is found:** If one eigenvalue is high and the other is low, so the local auto-correlation function is rigid shaped, then only shifts along the ridge (i.e., along the edge) cause little change in  $c(x, y)$  and significant change in the orthogonal direction.

- **A point of interest is found:** If both  $\lambda_1$  &  $\lambda_2$  are high, so the local auto-correlation function is sharply peaked, then shifts in any direction results in a significant increase in  $c(x, y)$ .

Figure 3 (a) discloses an example of Harris corner detector. We consider that in video surveillance scenes, camera positions, and lighting conditions allow to get a large number of corner features that can be easily captured and tracked.

### C. Optical flow computation

Optical flow techniques are based on the idea that for most points in the image, neighboring points have approximately the same brightness. To calculate the optical flow between successive video frames the well known combination of feature selection as introduced by Shi and Tomasi [15] and the algorithm of Lucas and Kanade for feature tracking [16] is used. Feature selection finds image blocks which are believed to allow the exact estimation of optical flow translation vector. The Shi-Tomasi algorithm makes use of the smallest eigenvalues of an image block as criterion to ensure the selection of features which can be tracked reliably by the Lucas-Kanade tracking algorithm. This algorithm matches the selected image blocks with blocks in the next frame using an efficient gradient descent technique. A pyramidal implementation of this algorithm is used to deal with larger feature displacements by avoiding local minima in a coarse to fine approach [17]. This combination has proven to allow fast and reliable computation of optical flow information [18], [19]. Once we define the points of interest, we track those points over the next frames using the combination feature tracker. Upon matching points of interest (features) between frames, the result is a set of vectors:

$$V = \{V_1 \dots V_N | V_i = (X_i, Y_i, D_i, \theta_i)\}$$

where  $X_i \mapsto X$  coordinate of the feature  $i$ ,  $Y_i \mapsto Y$  coordinate of the feature  $i$ ,  $D_i \mapsto$  distance between the feature  $i$  in the frame  $f$  and its matched feature in frame  $f + 1$ ,  $\theta_i \mapsto$  direction of motion of the feature  $i$ . Images in figure 3 (b) and (c) give evidence of the set of vectors obtained by optical flow features tracking in two different situations. The image in 3 (b) divulges an orderly vector flow. The image in 3 (c) substantiates a littered vector flow due to the breakdown situation.

*a) Calculation of  $D_i$ :* Figure 4 illustrates the feature  $i$  in the frame  $f$  with its coordinate  $P(X_i, Y_i)$  and its matched in the frame  $f + 1$  with coordinate  $Q(X_i, Y_i)$ . We can easily calculate the Euclidean distance between  $P(X_i, Y_i)$  and  $Q(X_i, Y_i)$  using Euclidean metric as:

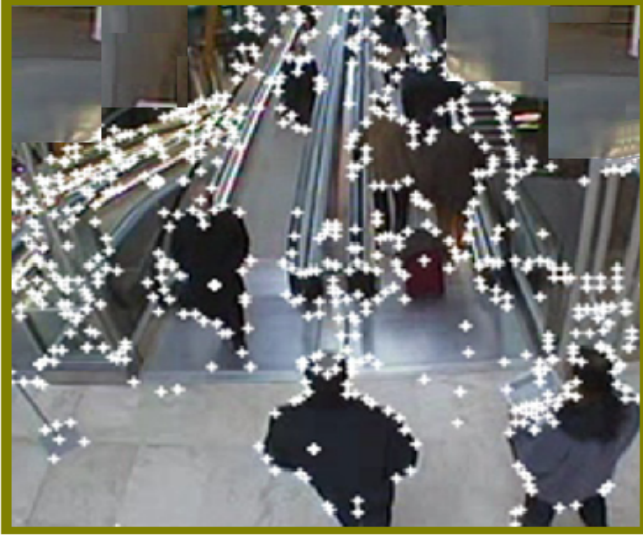
$$D_i = \sqrt{(Q_{X_i} - P_{X_i})^2 + (Q_{Y_i} - P_{Y_i})^2}. \quad (8)$$

*b) Calculation of  $\theta_i$ :* The direction of motion ( $\theta_m$ ) of the feature  $i$  can be calculated using the following trigonometric function:

$$\theta_m = \text{atan}\left(\frac{y_i}{x_i}\right) \quad (9)$$

where  $x_i = Q_{X_i} - P_{X_i}$  and  $y_i = Q_{Y_i} - P_{Y_i}$ . But there are several potential problems if we have a desire to calculate motion direction using Eq. 9, e.g., out of phase, generation





(a)



(b)



(c)

Fig. 3. White points in (a) portray Harris point of interest. Red arrows are optical vector flows: (b) and (c) limn standard and aberrant situations in precisely the given order.

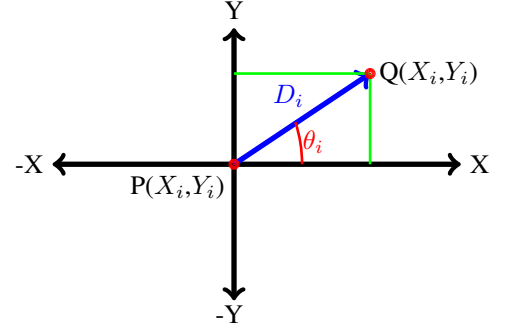


Fig. 4. Cardinal point of the state of change of any feature  $i$ .

of exceptions, etc. To avoid these problems, we apply the  $atan2(y_i, x_i)$  function which takes both  $x_i$  and  $y_i$  as arguments. Henceforth, the accurate direction of motion ( $\theta_i$ ) of the feature  $i$  can be calculated as:

$$\theta_i = atan2(y_i, x_i) = \begin{cases} \phi \cdot \text{sign}(y_i) & \text{if } x_i > 0, y_i \neq 0 \\ 0 & \text{if } x_i > 0, y_i = 0 \\ \frac{\pi}{2} \cdot \text{sign}(y_i) & \text{if } x_i = 0, y_i \neq 0 \\ \text{undefined} & \text{if } x_i = 0, y_i = 0 \\ (\pi - \phi) \cdot \text{sign}(y_i) & \text{if } x_i < 0, y_i \neq 0 \\ \pi & \text{if } x_i < 0, y_i = 0 \end{cases}$$

where  $\phi$  is the angle in  $[0, \pi/2]$  such that  $\phi = atan(|\frac{y_i}{x_i}|)$ . The sign function  $\text{sign}(y_i)$  can be defined as:

$$\text{sign}(y_i) = \begin{cases} -1 & \text{if } y_i < 0 \\ 0 & \text{if } y_i = 0 \\ 1 & \text{if } y_i > 0 \end{cases}$$

As a consequence the function  $atan2(y, x)$  gracefully handles infinite slope, and places the angle in the correct quadrant [for instances:  $atan2(0, 3) = 0$ ,  $atan2(0, -3) = \pi$ ,  $atan2(3, 3) = \pi/4$ ,  $atan2(-3, -3) = -3\pi/4$ , and so forth].

#### D. Statistical treatment of optical flow information

1) *Motion Area Ratio*: In each video frame, the motion area ratio ( $M_R$ ) estimates the ratio between the number of blocks containing motion and the total number of defined blocks. In crowded scenes the area covered by the moving blobs is important as compared to uncrowded scenes. We can use this measure as a density estimator. To estimate  $M_R$ , we divide each video frame into  $N \times M$  blocks, where  $N$  and  $M$  are number of columns and rows respectively. For any block  $(i, j)$ , we define the moving block by means of:

$$\text{movingblock}(i, j) = \begin{cases} 1; & \text{if movement exists} \\ 0; & \text{otherwise} \end{cases}$$

If there are several movements exist in one block, then that block will be enumerated as one moving block. We count out the total number of moving blocks to define  $M_R$  as:

$$M_R = \frac{\sum_{i=1}^N \sum_{j=1}^M \text{movingblock}(i, j)}{N \times M}. \quad (10)$$

2) *Direction Variance-Mean Ratio*: To estimate direction variance ( $\theta_V$ ), it is important to estimate the mean direction  $\bar{\theta}$  of the optical flow vectors in each video frame. The  $\bar{\theta}$  is determined by:

$$\bar{\theta} = \frac{1}{n} \sum_{i=1}^n \theta_i \quad (11)$$

where  $n$  is cardinality of the optical flow vectors in the frame and  $360^\circ \geq \theta_i > 0^\circ$ . Having firsthand knowledge of  $\bar{\theta}$ , we calculate the  $\theta_V$  of those vectors as:

$$\theta_V = \frac{1}{n-1} \sum_{i=1}^n (\theta_i - \bar{\theta})^2 = \frac{1}{n-1} \sum_{i=1}^n \theta_i^2 - \frac{n}{n-1} \bar{\theta}^2. \quad (12)$$

The direction variance-to-mean ratio (coefficient of direction variation) is defined as the ratio of the variance to the mean:

$$\theta_R = \frac{\theta_V}{\bar{\theta}}. \quad (13)$$

3) *Direction Histogram*: The direction histogram ( $\theta_H$ ) gives directions to the direction tendencies and the number of peaks. In the histogram each column puts an address on the number of vectors in a given angle. The  $\theta_H$ , is affiliated to the frame, can be clearly characterized by the way of:

$$\theta_H = \{\theta_H(\theta_i), i = 1 \dots s\} \quad (14)$$

$$\theta_H(\theta_i) = \frac{\sum_{i=1}^n \text{angle}(i)}{s} \quad (15)$$

$$\text{angle}(i) = \begin{cases} 1, & \text{if } \text{angle}(i) = \theta_i \\ 0, & \text{otherwise} \end{cases}$$

where  $\theta_H(\theta_i)$  is the normalized frequency of optical flow vectors those have the same angle  $\theta_i$ . The  $\theta_H$  is a vector of size  $s$  where  $s$  is the total number of angles considering the angle range between  $-\pi$  and  $+\pi$ .

4) *Distance Variance-Mean Ratio*: Observation shows that Distance variance ( $D_V$ ) increases in abnormal situations. With one or many people walking even in different directions, they tend to have the same speed, which means a small value of the motion distance variance. But in case of abnormal observable activities (e.g., collapsing situations, a sudden overwhelming fear, escape circumstances, etc.) that often gives rise to a big value for the  $D_V$ . The mean of distance variance  $\bar{D}$  is clearly delimited by:

$$\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i \quad (16)$$

where  $n$  is the number of optical flow vectors in the frame. Having  $\bar{D}$  it is easy to ascertain  $D_V$  by:

$$D_V = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2 = \frac{1}{n-1} \sum_{i=1}^n D_i^2 - \frac{n}{n-1} \bar{D}^2. \quad (17)$$

The distance variance-to-mean ratio (coefficient of distance variation) is defined as the ratio of the variance to the mean:

$$D_R = \frac{D_V}{\bar{D}}, \quad \text{where } \bar{D} > 0. \quad (18)$$

5) *Entropy Estimation*: We enumerate a function called *Entropy* that depends on motion area ratio, coefficient of direction variation, coefficient of distance variation, and direction histogram characteristics at any frame  $f$ , as:

$$\{Entropy\}_f = P(E_f) \log \frac{1}{P(E_f)} = -P(E_f) \log P(E_f) \quad (19)$$

where  $0 \leq P(E_f) \leq 1$ . In case of  $P(E_f) = 0$ , the  $P(E_f) \log P(E_f)$  will be considered as 0. The  $E_f$  is defined as:

$$E_f = M_R \times \theta_R \times \theta_H \times D_R. \quad (20)$$

To calculate  $P(E_f)$  we use cumulative distribution function (*cdf*) which has strict lower and upper bounds between 0 and 1. Deeming  $\Omega_{\mu,\sigma}(E_f)$  denotes the *cdf* of  $E_f$ . Then  $\Omega_{\mu,\sigma}(E_f)$  can be expressed in terms of a special function called the *error function* (*erf*) or *Gauss error function*, as:

$$\Omega_{\mu,\sigma}(E_f) = \frac{1}{2} [1 + \text{erf} \{ \frac{E_f - \mu}{\sigma \sqrt{2}} \}] \quad (21)$$

where  $\sigma > 0$  is the standard deviation and the real parameter  $\mu$  is the expected value. The *erf* can be defined as a *Maclaurin* series:

$$\text{erf}(E_f) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \{E_f\}^{2n+1}}{n! (2n+1)} \quad (22)$$

$$= \frac{2}{\sqrt{\pi}} \{E_f - \frac{E_f^3}{3} + \frac{E_f^5}{10} - \frac{E_f^7}{42} + \frac{E_f^9}{216} - \dots\}. \quad (23)$$

Since  $E_f$  is skewed to the right (positive-definite) and variances also large, we can use Log-normal distribution. The structure of log-normal distribution of the Eq. 21 yields:

$$P(E_f) = \frac{1}{2} [1 + \text{erf} \{ \frac{\log(E_f) - \mu}{\sigma \sqrt{2}} \}]. \quad (24)$$

By means of Eq. 24 & 23, and having extensive information of the values of  $\mu$  and  $\sigma$  (say  $\mu = 0$ ,  $\sigma = 10$ ) we can expressly estimate the value of  $P(E_f)$  between 0 and 1.

#### E. Decision of Normal or Abnormal events

We examine and note the similarities or differences of each calculated value of *Entropy* with a beforehand defined entropy threshold  $T_E$ , i.e., a deviant frame can be detected if and only if  $Entropy < T_E$ , otherwise standard frame. The hypothetical outlook of reckoning  $T_E$  is that we pay attention to the minimum number of entropies in large videos that keep under control snobbishly standard events:

$$T_E = \min_{k=1 \dots n} \{Entropy\}_k \quad (25)$$

$n$  is the number of frames of the video database. The  $T_E$  (also *Entropy*) depends on the controlled environment (video stream), specifically the aloofness of the camera to the scene, the orientation of the camera, the type and the position of the camera, lighting system, density of the crowd (e.g., motion area ratio, direction & distance variance-mean ratios, etc.), etc. The more is the remoteness between camera and the scene, the less is the considerable amount of optical flows and blobs. In case of escalator,  $T_E$  also places trust on the escalator type and position. Taking into account of these facts, we think carefully that we have at least one threshold by a video stream. If we have  $N$  video streams, then we put forward at least  $N$  thresholds. If the video stream leaves for another, then the threshold should be made over.

#### IV. EXPERIMENTAL RESULTS

The approach is tested on a set of real videos taken from cameras installed in an airport to monitor the situation of escalator exits. The videos were used to provide informative data for the security team who may need to take prompt actions in the event of a critical situation such as collapsing or complete failure or breakdown. The data sets are videos from a video surveillance system on escalator exits. Data sets are 10 different length video streams taken in different duration of different time and seasons. Each video stream consists of standard and abnormal events. The standard situations correspond to crowd flows without breakdown in the escalator exits. Generally, in the videos we have two escalators corresponding to two-way-traffic of opposite directions. Abnormal situations correspond to videos that contain collapsing/breakdown events chiefly in escalator exits. The original video frame size is  $640 \times 480$  pixels. For the features detection and tracking we extract about 1500 features per frame.

Figure 5<sup>1</sup> describes an example of a breakdown situation in an escalator exit point in a video stream. Two persons were standing on the moving escalator, suddenly a trolley rushed out toward them. One person escaped by running (and was not descended under the force of trolley) while other did not. In the scene that followed the non-escapee was rundown by the run-away trolley, and subsequently fell down at the exit point of the moving escalator. The situation was successfully detected by the proposed algorithm. Different video frames in normal and fortuitous circumstances have been differentiated by a threshold label (horizontal red line in FIG. 5). The detection result has been compared with ground truth. Ground truth is the process of manually marking what an algorithm is looked forward to the probable output. The exhaustive evaluation of the proposed algorithm for the provided data-set has been listed in Table I. The algorithm does not put into effective operation of one of the video streams (the 5th video stream listed on the Table I). Syntactically, the anomalous event of that video stream has taken place significantly far distance from the camera, and hence the considerable amount of optical flow vectors is not sufficient to analyze abnormal frames. In Table I *root mean squared error* ( $\Psi$ ) has been also appraised. The symbol  $\infty$  shows deliberately that we could not straightforwardly determine the *squared error* by mathematical calculation in respective cases, because of end of video files. However, it is more important to detect the commencement of the atypical state of affairs than that of termination. The assessed  $\Psi$  performs appointively for many computer vision applications together with escalators.

#### V. CONCLUSION

We came up with an approach that appraises the aberration of a populace flow in escalator exit points. The approach assesses sudden shift and strange locomotion discrepancy of a set of interest points. Optical flow information is computed from those points of interest from the specific parts of the escalator using motion heat map, which primarily helps to make faster the calculation process by providing regions of interest. Statistical

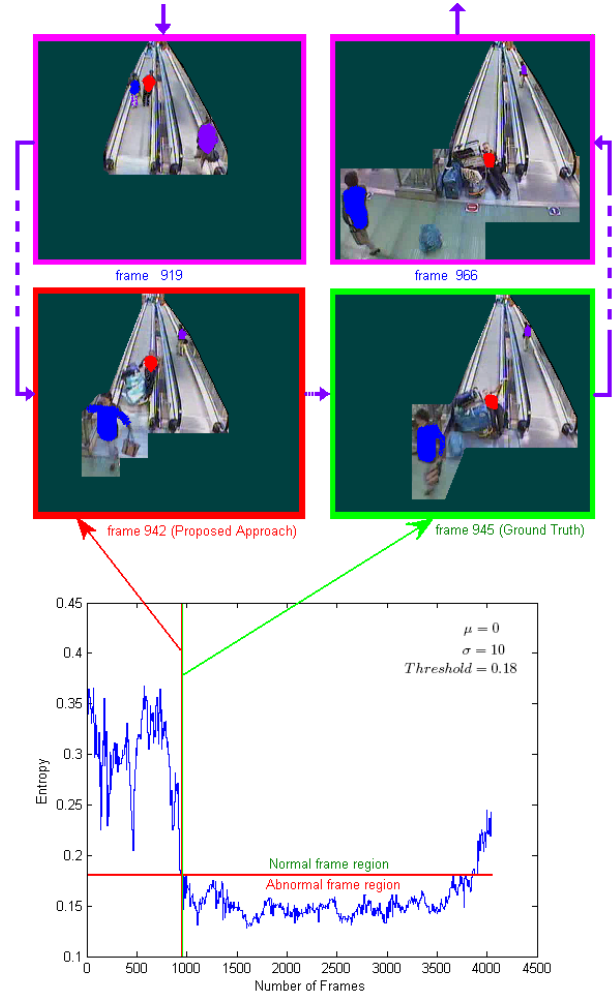


Fig. 5. Person falling situation on the exit point of an escalator. The blue coloured curve is the output of the algorithm.

scrutiny of the optical flow information has been thresholded. The achievement of the algorithm has been assessed by a set of escalator videos obtained by a single camera. The results reveal the efficacy of the proposed procedure for recognizing exceptional events on the escalator exits. It is worth mentioning that, as a short coming of the approach, it is vulnerable to the selection of the appropriate  $T_E$  and region of interest with a view to getting outstanding output.

Future work will take into account the dedication of multiple-camera so that video could be conclusively broken down into its essential features properly in all parts of an elongated escalator (commencement, halfway point, outlet) to proclaim the eccentric event if there will exist any. Future work will also carry on the algorithm (if something indispensable) to bring about the circumstances where crowds are milling like in retail or a surveillance video of a famous public/private space, etc.

#### ACKNOWLEDGMENT

This work was supported by the MIAUCE project, which is the 6<sup>th</sup> Framework Research Programme of European Union (IST-2005-5-033715). [www.miauce.org](http://www.miauce.org).

<sup>1</sup>Videos have been provided by a video surveillance company. Images have been cropped for confidential reason.

TABLE I

The achievement appraisal of the proposed algorithm based on the escalator data sets. The following alphabets and symbols are used for Contextual Information (ConInfo) and curtailment respectively:  $A \mapsto$  Less illumination in escalator,  $B \mapsto$  No shadowing,  $C \mapsto$  Not varying illumination,  $D \mapsto$  Varying illumination,  $E \mapsto$  Light reflection,  $F \mapsto$  Over head light,  $G \mapsto$  Friday morning,  $H \mapsto$  Thursday afternoon,  $I \mapsto$  Saturday night,  $J \mapsto$  Sunday night,  $K \mapsto$  Wednesday afternoon,  $L \mapsto$  Thursday afternoon,  $M \mapsto$  Thursday night,  $N \mapsto$  Sunday afternoon,  $O \mapsto$  Monday night,  $P \mapsto$  Spring,  $Q \mapsto$  Winter;  $\nabla \mapsto$  End of video file,  $\infty \mapsto$  Not straightforwardly determined by mathematical calculation.

| Videos & ConInfo    | Disruption Circumstances                                                       |                         |                             |                         |                      |                                                 |                                               |
|---------------------|--------------------------------------------------------------------------------|-------------------------|-----------------------------|-------------------------|----------------------|-------------------------------------------------|-----------------------------------------------|
|                     | Ground Truth ( $g_t$ )                                                         |                         | Proposed Approach ( $p_a$ ) |                         |                      |                                                 |                                               |
|                     | $S_{gt}$ :<br>Start Frame                                                      | $E_{gt}$ :<br>End Frame | $S_{pa}$ :<br>Start Frame   | $E_{pa}$ :<br>End Frame | $T_E$ :<br>Threshold | $(S_{gt} - S_{pa})^2$                           | $(E_{gt} - E_{pa})^2$                         |
| 1. C, E, G, P       | 945                                                                            | 3886                    | 942                         | 3888                    | 0.18                 | 9                                               | 4                                             |
| 2. D, E, N, P       | 127                                                                            | 1275                    | 130                         | 1278                    | 0.21                 | 9                                               | 9                                             |
| 3. E, F, O, P       | 496                                                                            | $\nabla$                | 492                         | $\nabla$                | 0.15                 | 16                                              | $\infty$                                      |
| 4. E, F, H, P       | 290                                                                            | 2070                    | 294                         | 2075                    | 0.16                 | 16                                              | 25                                            |
| 5. E, F, K, P       | 1161                                                                           | 4158                    | —                           | —                       | —                    | —                                               | —                                             |
| 6. E, F, M, Q       | 132                                                                            | 706                     | 127                         | 712                     | 0.16                 | 25                                              | 36                                            |
| 7. D, E, L, Q       | 103                                                                            | $\nabla$                | 105                         | $\nabla$                | 0.21                 | 4                                               | $\infty$                                      |
| 8. A, C, J, Q       | 100                                                                            | $\nabla$                | 102                         | $\nabla$                | 0.27                 | 4                                               | $\infty$                                      |
| 9. E, F, I, Q       | 199                                                                            | 1446                    | 189                         | 1442                    | 0.15                 | 100                                             | 16                                            |
| 10. B, C, G, Q      | 72                                                                             | 2359                    | 78                          | 2370                    | 0.25                 | 36                                              | 121                                           |
| Effectiveness Study | Root Mean Squared Error $\Psi = \sqrt{\frac{1}{n} \sum_{k=1}^n (g_t - p_a)^2}$ |                         |                             |                         |                      | $\Psi_{start} = \sqrt{\frac{219}{9}} \approx 5$ | $\Psi_{end} = \sqrt{\frac{211}{6}} \approx 6$ |

## REFERENCES

- [1] C. F. Chi, T. C. Chang, and C. L. Tsou, "In-depth investigation of escalator riding accidents in heavy capacity mrt stations," *Accident Analysis and Prevention*, vol. 38, no. 4, pp. 662–670, 2006.
- [2] A. N. Marana, S. A. Velastin, L. F. Costa, and R. A. Lotufo, "Estimation of crowd density using image processing," *Image Processing for Security Applications (Digest No.: 1997/074)*, *IEE Colloquium*, pp. 11/1–11/8, Mar. 1997.
- [3] H. Rahmalan, M. S. Nixon, and J. N. Carter, "On crowd density estimation for surveillance," *International Conference on Crime Detection and Prevention*, 2006.
- [4] S. F. Lin, J. Y. Chen, and H. X. Chao, "Estimation of number of people in crowded scenes using perspective transformation," *Systems, Man and Cybernetics, Part A, IEEE Transactions*, vol. 31, no. 6, pp. 645–654, Nov. 2001.
- [5] R. Ma, L. Li, W. Huang, and Q. Tian, "On pixel count based crowd density estimation for visual surveillance," *IEEE Conference on Cybernetics and Intelligent Systems*, vol. 1, pp. 170–173, Dec. 2004.
- [6] E. L. Andrade, S. Blunsden, and R. B. Fisher, "Hidden markov models for optical flow analysis in crowds," *International Conference on Pattern Recognition*, vol. 1, pp. 460–463, 2006.
- [7] —, "Modelling crowd scenes for event detection," in *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, vol. 1, 2006, pp. 175–178.
- [8] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–6, June 2007.
- [9] O. Boiman and M. Irani, "Detecting irregularities in images and in video," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 17–31, Aug. 2007.
- [10] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, Aug. 2000.
- [11] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian, "Foreground object detection from videos containing complex background," in *Proceedings of the eleventh ACM international conference on Multimedia*, 2003, pp. 2–10.
- [12] H. P. Moravec, "Obstacle avoidance and navigation in the real world by a seeing robot rover," in *Technical Report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University & doct. diss., Stanf. University*, 1980.
- [13] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, 1988, pp. 147–152.
- [14] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.
- [15] J. Shi and C. Tomasi, "Good features to track," in *Proceedings CVPR: Computer Vision and Pattern Recognition*, 1994, pp. 593–600.
- [16] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI)*, 1981, pp. 674–679.
- [17] J. Y. Bouquet, "Pyramidal implementation of the lucas kanade feature traker," in *A part of OpenCV Documentation, Intel Corporation, Microprocessor Research Labs*, 2000.
- [18] J. L. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.
- [19] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, p. 221255, Feb. 2004.

# Solutions to Motion Self-Occlusion Problem in Human Activity Analysis

Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim, and S. Ishikawa

Department of Control Engineering, Kyushu Institute of Technology, Fukuoka, Japan

E-mail: atiqahad@ss10.cntl.kyutech.ac.jp

**Abstract**— Human motion self-occlusion due to motion overlapping in the same region is a daunting task to solve. Various motion-recognition methods either bypass this problem or solve this problem in complex manner. Appearance-based template matching paradigms are simpler and hence faster approaches for activity analysis. In this paper, we concentrate on motion self-occlusion problem due to motion overlapping in various complex activities for recognition. This paper illustrates the directional motion history image concept and compares this motion representation approach with multi-level motion history representation and hierarchical motion history histogram representation to solve the self-occlusion problem of basic motion history image representation. We employ some complex aerobics and find the robustness of our method compared to other methods for this self-occlusion problem. We employ seven higher order Hu moments to compute the feature vector for each activity. Afterwards, k-nearest neighbor method is utilized for classification with leave-one-out paradigm. The comparative results clearly demonstrate the superiority of our method than other recent approaches.

**Keywords**— MHI, DMHI, MMHI, HMHI, motion recognition, complex activity, template matching.

## I. INTRODUCTION

Human motion and behavior analysis is very important area in computer vision. There are many approaches to deal with action or activity recognition and analysis. We concentrated the MHI method [1] as it is an appearance-based template-matching motion recognition approach. It has been widely employed and modified a bit based on the requirements. Various interactive systems have been successfully constructed using motion history template as a primary sensing mechanism. For example, using the MHI method, Davis et al. [19] has developed a virtual aerobics trainer that watches and responds to the user as he/she performs the workout. An interactive art demonstration can be constructed from the motion templates [20]. An interactive and narrative play space for children, called KidsRoom was developed using MHI method successfully. Yau et al. [21] has developed a method for visual speech recognition employing the MHI method. The video data of the speaker's mouth is represented using grayscale images named as motion history image. Automatically localizing and tracking moving person or vehicle for an automatic visual surveillance system was demonstrated in [22-23] by employing MHI method before

employing an extended mean shift approach in [23]. Son et al. [24] calculated MHI and then combined with background model to detect the candidate road image.

Gait history image and gait energy image are created for gait analysis based on the concept of MHI [25-26]. Ref. [27] proposed threat assessment method for automated visual surveillance with the aid of MHI. A PDA-based recognition system based on MHI method is developed by [28]. Petras et al. [29] has devised a flexible test-bed for unusual behavior detection and automatic event analysis using MHI. A recent method by Kellokumpu et al. [12] proposed a recognition method at the top of the MHI method using texture-based description of the movements by employing local binary pattern (LBP) operator. They used HMM later for recognition. Usually, the MHI-based methods consider global image in the scene, rather body parts. However, they split the MHI image into four regional parts to compute histograms of those areas.

Refs. [2-5] cover some dominant methods for activity analysis and recognition. We see that most of the approaches consider simple or one-directional actions (e.g., sitting on a chair, standing from a chair, raise a hand, etc.) for recognition. However, usual activities are complex in nature and not simple in dimension. By complex action or activity, it is meant the action comprises more than one action primitives or basic action elements. Hence, a sit-stand action is a complex action, as it has 'sitting motion' and 'standing motion' in two directions (down and up direction consecutively). Real-life activities are complex in nature. In this paper, we discussed the motion recognition problem of complex action that has motion overwriting problem. Motion overwriting can be defined using the sit-down action as once sitting and then standing will erase or passes the prior motion information due to the sitting down action. In this paper, we started with the Motion History Image (MHI) method [1] as it is found to be robust in recognition and simple in computation. However, this method can not solve the overwriting or motion self-occlusion problem [7,10,14].

Section II covers the basic MHI method and the concept of overwriting problem. Then we present three template-based representations including our method to solve the self-occlusion problem in Section III. The recognition methods based on these motion representations are outlined in Section IV. Section V presents that comparative results and analysis using recognition rates and confusion matrices for various aerobics sequences. Finally, Section VI concludes the paper with some future guidelines.



## II. MOTION HISTORY IMAGE TEMPLATE

The MHI is a scalar-valued image where more recently moving pixels are brighter, and vice versa [1]. The MHI expresses the motion flow or sequence by using the intensity of every pixel in temporal manner. To describe the motion-shape and spatial distribution of motion, [1] proposed Motion Energy Image (MEI) or binary motion region. One of the advantages of MHI representations is that a range of times from frame to frame to several seconds may be encoded in a single frame, and in this way, MHIs span the time scale of human gestures. The MHI  $H_t(x, y, t)$  can be computed from update function  $\Psi(x, y, t)$ :

$$H_t(x, y, t) = \begin{cases} \tau & \text{if } \Psi(x, y, t) = 1 \\ \max(0, H_t(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \quad (1)$$

Here,  $x, y$  and  $t$  show the position and time;  $\Psi(x, y, t)$  signals object presence (or motion) in the current video image;  $\tau$  decides the temporal duration of MHI (in terms of frames); and  $\delta$  is the decay parameter. This update function is called for every new video frame analyzed in the sequence. Some possible image processing techniques for defining this update function are background subtraction, image differencing, optical flow, etc. The MHI is generated from a binarized image, obtained from frame subtraction, using a threshold  $th$ :

$$D(x, y, t) = |I(x, y, t) - I(x, y, t \pm \Delta)| \quad (2)$$

where,  $D(x, y, t)$  with difference distance  $\Delta$  is:

$$B(x, y, t) = \begin{cases} 1 & \text{if } D(x, y, t) \geq th \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$I(x, y, t)$  is the intensity value of pixel location with coordinate  $(x, y)$  at the  $t^{th}$  frame of the image sequence. MEI is cumulative binary motion images. We generate the MEI by thresholding the MHI above zero,

$$E_t(x, y, t) = \begin{cases} 1 & \text{if } H_t(x, y, t) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

For details, we refer to Ref. [1]. However, the MHI can not solve self-occlusion or overwrite problem, since previous motion is deleted or overwritten by the later motion information, if the motion has opposite directions in action (e.g., if a person sits down, and then stands up, the final MHI image should contain brighter pixels in the upper part of the image to represent the stand-up motion). This *self-occlusion* of the moving object or person overwrites the prior information [7]. Hence, for complex actions or activities, the basic MHI method can not recognize properly due to its inherent nature of motion history calculation. We developed a robust method considering four compass directions (namely, up, down, left and right directions) of an action and thereby can simply separate the action into its directional elements. This method is called Directional Motion History Image (DMHI) method. This is based on the computation of optical flow. In this paper, we show that the DMHI method outperforms the basic MHI method. We extend the evaluation with other representation methods that claim to be a solution for such motion

overwriting problem. We evaluate recent methods, called the Hierarchical Motion History Histogram (HMHH) representation [14-16] and the Multi-level Motion History Image (MMHI) method [10-11].

## III. TEMPLATE-BASED REPRESENTATIONS TO SOLVE MOTION SELF-OCCLUSION PROBLEM

### A. Multi-level Motion History Image (MMHI) Representation

The Multiple-level MHI (MMHI) aims at overcoming the problem of motion self-occlusion by recording motion history at multiple time intervals (i.e., multi-level MHIs) [10-11]. It creates all MHIs to have a fixed number of history levels  $n$ . So, each image sequence is sampled to  $(n+1)$  frames. The modified MHI is:

$$H_t(x, y, t) = \begin{cases} s * t & \text{if } \Psi(x, y, t) = 1 \\ H_t(x, y, t-1) & \text{otherwise} \end{cases} \quad (5)$$

where  $s = (255/n)$  is the intensity step between two history levels.  $H_t(x, y, t) = 0$  for  $t \leq 0$ . The final MHI is found by iteratively computing Eq. (5) for  $t = 1, \dots, n+1$ . It focused on automatic detection of facial actions units that compose expressions. It requires sophisticated registration system, because all utilized image sequences must have the faces at the same position and on the same scale. The result does not clearly demonstrate the superiority of MMHI with respect to basic MHI in their experiment on facial action unit detection [11]. However, it is reported that self-occlusion due to motion overwriting problem might be solved using this MMHI [11]. We implemented this method for this proof.

### B. Hierarchical Motion History Histogram (HMHH) Representation

A motion representation and recognition system called Hierarchical Motion History Histogram (HMHH) is proposed to similar overwrite problem [14]. In [15-16], few related other methods are compared [1,6,9] with this HMHH representations for some actions. In basic MHI, at any frame  $k$ ,  $D(x, y, k)$  might have combinations of 1s and 0s (e.g., 00011001000111100010...). Due to the formation of MHI, it keeps the later information of the presence of 1 (i.e., motion) in a specific pixel when there are some earlier movement in the same pixel. For example, for a pixel, the last '1' in the sequence are retained in the MHI at that pixel. It is clear that previous '1' in the sequence, when some action occurred, are not represented, and usually almost every pixel have more than one '1' in their sequence. It can be mentioned as overwriting problem in different form.

To solve this problem, we can define some patterns  $P_i$  in  $D(x, y, :)$  sequences based on the number of connected '1' (e.g.,  $P_1 = 010$ ,  $P_2 = 0110$ ,  $P_3 = 01110$ , ...  $P_M = 01...10$  (M 1s)) [14]. Now define a subsequence  $C_i = b_{n1}, b_{n2}, \dots, b_{ni}$  and denote the set of all sub-sequences of  $D(x, y, :)$  as  $\Omega\{D(x, y, :)\}$ . Then for each pixel  $(x, y)$ , count the number of occurrences of each specific pattern  $P_i$  in the sequence of  $D(x, y, :)$  as shown,

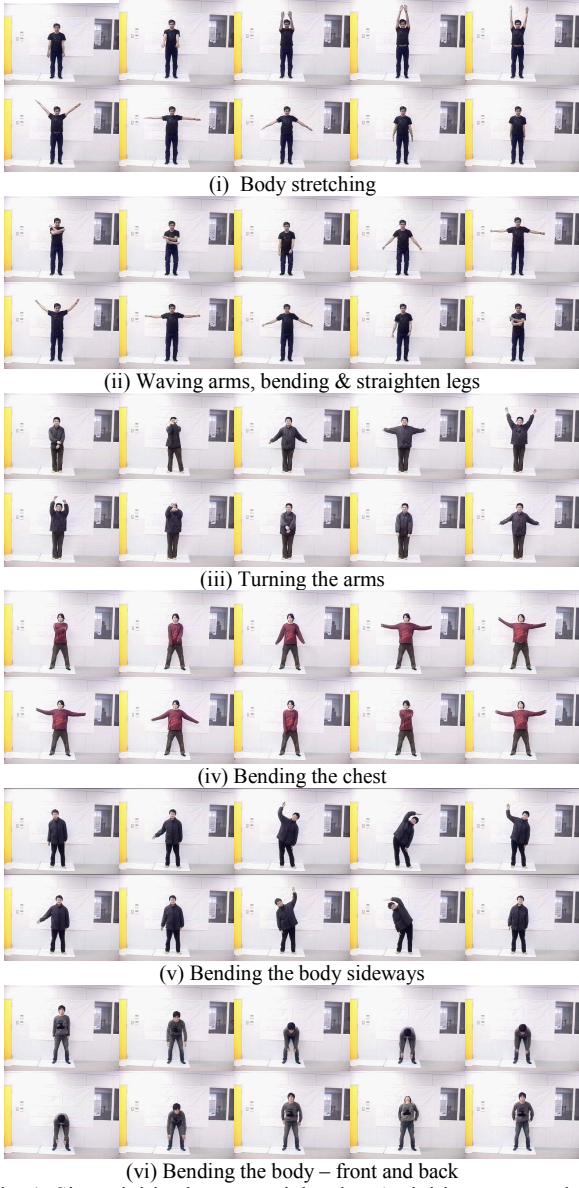


Fig. 1. Six activities in sequential order. Activities are complex.

$$HMHH(x, y, P_i) = \sum_j \mathbf{1}_{\{C_j = P_i | C_j \in \Omega\{D(x, y, :)\}\}} \quad (6)$$

Here,  $\mathbf{1}$  is the indicator function. Hence, from each pattern  $P_i$ , we construct one gray-scale image (called Motion History Histogram (MHH)), and in aggregation, we call all MHH images as Hierarchical MHH (HMHH)). Similar to the concept of the MEI, we can threshold each MHH to compute the corresponding binary energy image. They used these final feature images for classification and then recognition.

### C. Directional Motion History Image (DMHI) Representation

The above-mentioned approaches can not solve the overwriting problem as smartly as our optical flow-based

Directional Motion History Image (DMHI) method [7]. In this approach, instead of background or frame subtraction, gradient-based optical flow is calculated between consecutive two frames and split it into four channels, based on the concept of motion descriptors based on smoothed and aggregated optical flow measurements in four channels [17]. In this way, we get four-directional motion templates for left, right, up and down directions. We calculate the corresponding four history and energy images based on the Eqs. 7 and 8.

$$\begin{aligned} H_{\tau}^{-x}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_{\tau}^{-x}(x, y, t) > th_p \\ \max(0, H_{\tau}^{-x}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \\ H_{\tau}^{+x}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_{\tau}^{+x}(x, y, t) > th_p \\ \max(0, H_{\tau}^{+x}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \\ H_{\tau}^{-y}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_{\tau}^{-y}(x, y, t) > th_p \\ \max(0, H_{\tau}^{-y}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \\ H_{\tau}^{+y}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_{\tau}^{+y}(x, y, t) > th_p \\ \max(0, H_{\tau}^{+y}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \end{aligned} \quad (7)$$

$$\begin{aligned} E_{\tau}^{-x}(x, y, t) &= \begin{cases} 1 & \text{if } H_{\tau}^{-x}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \\ E_{\tau}^{+x}(x, y, t) &= \begin{cases} 1 & \text{if } H_{\tau}^{+x}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \\ E_{\tau}^{-y}(x, y, t) &= \begin{cases} 1 & \text{if } H_{\tau}^{-y}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \\ E_{\tau}^{+y}(x, y, t) &= \begin{cases} 1 & \text{if } H_{\tau}^{+y}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (8)$$

## IV. RECOGNITION METHODOLOGY

From the final template images, the corresponding feature vectors are computed by employing the seven Hu moments [8]. For the MHI method, for each action, seven higher order Hu moments are calculated for history and energy images and therefore, a 14-D feature vector is created for each activity. In the same way, for the MMHI representation, we computed a 14-D feature vector. In this experiment, we implemented the HMHH representation with four patterns because patterns more than the first four have almost no information to employ further for calculating feature vectors.

Therefore, for these four patterns, using the HMHH representation, we computed a 28-D feature vector (for each pattern, we compute seven higher order moments). In another experiment, we also computed the binarized energy images of the four patterns of the history images. But we have not achieved any difference to consider the energy versions for calculating feature vectors. However, for the four directional history and energy components, we need to compute seven Hu moments for each of the eight components. Moreover, to cover the overwriting issue significantly, we also computed the normalized 0<sup>th</sup> order Hu moments. E.g., for  $H_{\tau}^{+x}$  for an image  $f(x, y)$ , the 0<sup>th</sup> order moment is:

$$m_{00}(H_{\tau}^{+x}) = \sum_x \sum_y f(x, y) \quad (9)$$

Then we get normalized 0<sup>th</sup> order moment for  $H_{\tau}^{+x}$  by employing the following equation:

$$\overline{m_{00}(H_{\tau}^{+x})} = \frac{m_{00}(H_{\tau}^{+x})}{m_{00}(H_{\tau}^{+x}) + m_{00}(H_{\tau}^{-x}) + m_{00}(H_{\tau}^{+y}) + m_{00}(H_{\tau}^{-y})} \quad (10)$$

The 0<sup>th</sup> order moments provide the total object area or mass, and hence provide important cues for motion region. Finally, for DMHI, We estimate eight normalized 0<sup>th</sup> order moments for the eight templates (DMHIs and DMEIs) to get a 64-dimensional feature vector (7 Hu moments for each template in addition to normalized 0<sup>th</sup> order moments).

In this experiment, we consider some aerobics which are complex in nature [18]. Fig. 1 shows six activities that are taken from a frontal-view camera without any calibration process or special markers. The actors performed the exercises indoor with almost the same illumination level, without any special arrangements and therefore, the frame numbers are varied for the same activity by various persons. Also the dress, height, etc. of the persons were different.

For classification, we used the nearest neighbor classifier [13]. For partitioning scheme, we employ leave-one-out cross-validation scheme. This means that out of  $N$  samples from each of the  $c$  classes per database,  $N - 1$  of them are used to train (design) the classifier and the remaining one to test it [13]. This process is repeated  $N$  times, each time leaving a different sample out. Therefore, all of the samples are ultimately used for testing. This process is repeated and the resultant recognition rate is averaged.

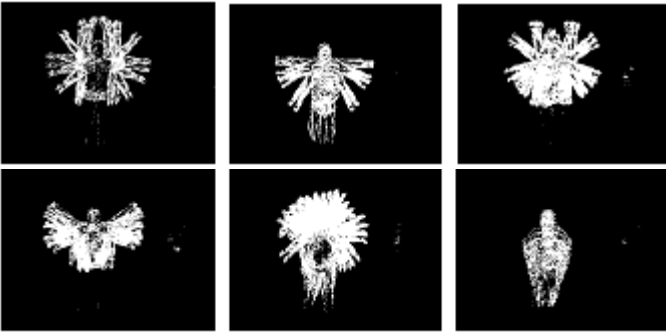


Fig. 2. Energy images of six activities using MMHI representation.

## V. RESULTS AND ANALYSIS

In this paper, we presented the MHI, MMHI, HMHH and DMHI representations for recognizing aerobics dataset. Fig. 2 illustrates the final energy images employing MMHI method for the six activities of Fig. 1. Fig. 3 depicts the four patterns for HMHH representation for six activities. We can note that Pattern-4 of Fig. 3 has almost no information and hence we did not consider pattern-5 or more for this experiment.

Fig. 4 shows the MEIs and MHIs for six activities. Fig. 5 presents the directional history images for four different directions – namely, up, down, left and right for each activity.

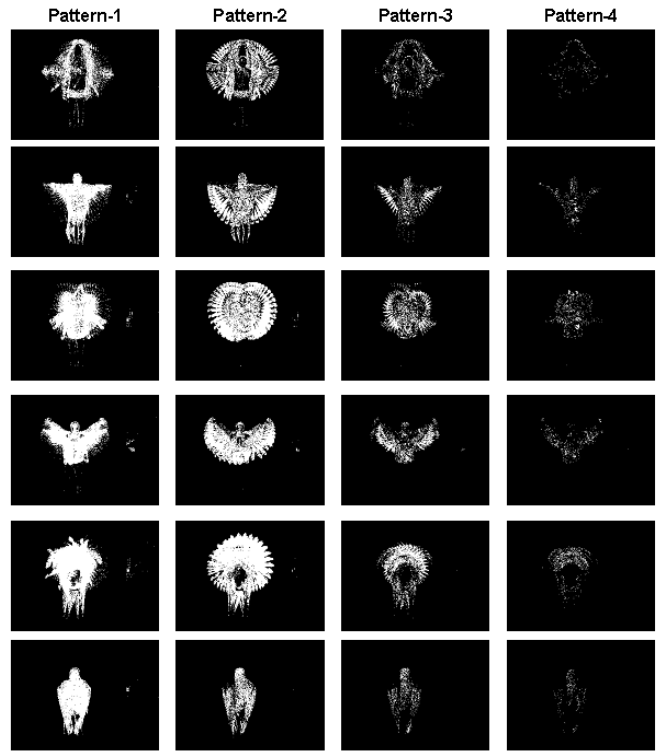


Fig. 3. Four patterns for activities (images in each row represent activity 1 through 6) employing HMHH representation.

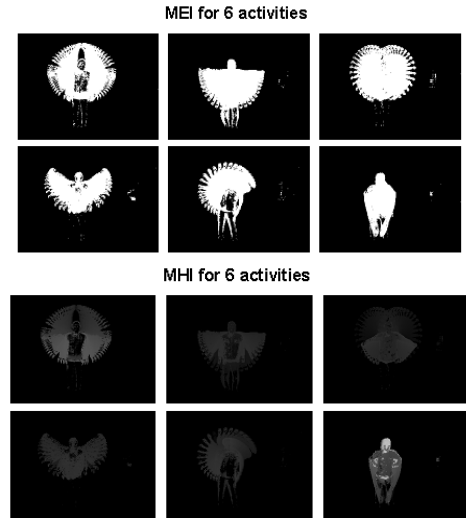


Fig. 4. MEIs and MHIs for six activities sequentially.

Table I. Comparative recognition rates for various aerobic datasets. All results are presented in %.

|    | MHI  |      | MMHI |      | HMHH |      | DMHI |      |
|----|------|------|------|------|------|------|------|------|
|    | K=1  | K=1  | K=1  | K=3  | K=3  | K=3  | K=1  | K=3  |
| A1 | 50.0 | 50.0 | 83.4 | 83.4 | 66.7 | 66.7 | 100  | 100  |
| A2 | 16.7 | 16.7 | 50.0 | 33.4 | 83.4 | 83.4 | 100  | 83.4 |
| A3 | 66.7 | 66.7 | 16.7 | 33.4 | 50.0 | 50.0 | 83.4 | 83.4 |
| A4 | 33.4 | 16.7 | 16.7 | 16.7 | 33.4 | 33.4 | 100  | 100  |
| A5 | 83.4 | 83.4 | 33.4 | 50.0 | 100  | 100  | 83.4 | 83.4 |
| A6 | 83.4 | 83.4 | 50.0 | 50.0 | 66.7 | 66.7 | 100  | 100  |
| RR | 55.6 | 52.8 | 41.7 | 44.5 | 66.7 | 66.7 | 94.4 | 91.7 |



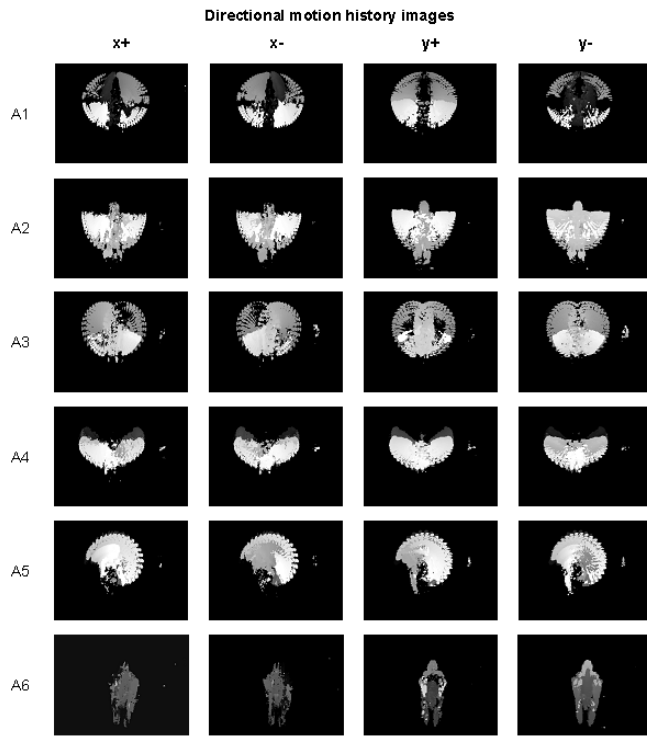


Fig. 5. Directional motion history images for six activities.

Table I shows the comparative recognition results for the four methods for  $k$ 's value as 1 and 3 in the nearest-neighbor classifier. For every activity (in the Tables, these six activities are represented as A1 to A6), the recognition results with our DMHI representation are very satisfactory. With  $k$ 's value as 1, for the DMHI method, we have achieved more than 94% recognition result. This is very good result considering the complexity of the activities. Though the HMHH representation achieved better results than the MHI and MMHI representations, the performance of the former is unacceptable as it achieved about 67% recognition rate.

Table II. CM for the MHI for Aerobics data; traces 333.6.

|    | A1          | A2          | A3          | A4          | A5          | A6          |
|----|-------------|-------------|-------------|-------------|-------------|-------------|
| A1 | <b>50.0</b> | 16.7        | 0           | 16.7        | 16.7        | 0           |
| A2 | 50.0        | <b>16.7</b> | 0           | 16.7        | 16.7        | 0           |
| A3 | 0           | 0           | <b>66.7</b> | 0           | 33.4        | 0           |
| A4 | 33.4        | 16.7        | 0           | <b>33.4</b> | 0           | 16.7        |
| A5 | 16.7        | 0           | 0           | 0           | <b>83.4</b> | 0           |
| A6 | 0           | 0           | 0           | 16.7        | 0           | <b>83.4</b> |

Table III. CM for the MMHI for Aerobics data; traces 250.2.

|    | A1          | A2          | A3          | A4          | A5          | A6          |
|----|-------------|-------------|-------------|-------------|-------------|-------------|
| A1 | <b>83.4</b> | 16.7        | 0           | 0           | 0           | 0           |
| A2 | 0           | <b>50.0</b> | 0           | 33.4        | 0           | 16.7        |
| A3 | 0           | 0           | <b>16.7</b> | 16.7        | 66.7        | 0           |
| A4 | 0           | 33.4        | 33.4        | <b>16.7</b> | 0           | 16.7        |
| A5 | 0           | 0           | 66.7        | 0           | <b>33.4</b> | 0           |
| A6 | 16.7        | 16.7        | 0           | 16.7        | 0           | <b>50.0</b> |

Tables II~V show the confusion matrices (CM) for the MHI, MMHI, HMHH and DMHI methods for  $k$ 's value as 1. The confusion matrix shows the motion label (vertical) versus the classification results (horizontal). Each cell  $(p, q)$  in the

tables shows the percentage of class  $p$  action being recognized as class  $q$ . The trace of the confusion matrix shows the percentage sum of the correctly recognized action while the remaining cells show the percentage of misclassification.

Table IV. CM for the HMHH for Aerobics data; traces 416.9.

|    | A1        | A2          | A3          | A4          | A5         | A6          |
|----|-----------|-------------|-------------|-------------|------------|-------------|
| A1 | <b>50</b> | 0           | 0           | 16.7        | 0          | 33.4        |
| A2 | 16.7      | <b>83.4</b> | 0           | 0           | 0          | 0           |
| A3 | 0         | 0           | <b>66.7</b> | 16.7        | 0          | 16.7        |
| A4 | 16.7      | 16.7        | 16.7        | <b>33.4</b> | 16.7       | 0           |
| A5 | 0         | 0           | 0           | 0           | <b>100</b> | 0           |
| A6 | 16.7      | 0           | 0           | 0           | 0          | <b>83.4</b> |

Table V. CM for the DMHI method for Aerobics data; traces 566.8.

|    | A1         | A2         | A3          | A4         | A5          | A6         |
|----|------------|------------|-------------|------------|-------------|------------|
| A1 | <b>100</b> | 0          | 0           | 0          | 0           | 0          |
| A2 | 0          | <b>100</b> | 0           | 0          | 0           | 0          |
| A3 | 0          | 0          | <b>83.4</b> | 0          | 16.7        | 0          |
| A4 | 0          | 0          | 0           | <b>100</b> | 0           | 0          |
| A5 | 0          | 0          | 16.7        | 0          | <b>83.4</b> | 0          |
| A6 | 0          | 0          | 0           | 0          | 0           | <b>100</b> |

## VI. CONCLUSION

This paper presents various template-based solutions to the motion self-occlusion problem in action analysis and recognition. The basic MHI method is very simple to calculate and hence many researchers employed this method or its variants for various action/gesture recognition, motion analysis, etc., with different datasets.

The motion self-occlusion problem is addressed and solved with satisfactory recognition rate with our DMHI method. We compared the DMHI representation with the MHI, the multi-level motion history image (MMHI) method and the hierarchical motion history histogram (HMHH) approaches. The latter two approaches claim to be better in overwriting problem. However, these representations failed to perform well compared to our method.

Though the performance of the DMHI representations and the employed feature vectors for recognition are satisfactory, we find that this method can not perform well or faces difficulties in some cases. When two or more than two persons are in-view, this method (along with other presented methods in this paper) can not recognize properly, especially when all of them are moving in different directions. Also, like the MHI, it can not recognize properly if the person is walking towards the camera's optical axis or if it moves something like diagonal directions. For these types of actions, we need to consider a bit different feature vector sets. For walking towards or off the camera, if we can exclude the energy images, the history image can separate and recognize the actions properly. These problems will remain for future research so that the method can be made a robust one for different types of motions or activities. Another important pair of activities are running and walking across the optical axis and to recognize these is very crucial for video surveillance. These are difficult to recognize in a simple manner. Currently, we are working by considering time measurement to

differentiate both actions, so that these could be recognized separately. Note that using the DMHI, it is far easier to understand the motion flow, i.e., the direction of the motion – running or walking easily. This information is very useful for behavior understanding and analyzing.

## REFERENCES

- [1] A. Bobick and J. Davis, "The recognition of human movement using temporal templates", *IEEE Trans. on PAMI*, vol. 23, no. 3, pp. 257-267, March 2001.
- [2] Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim, and S. Ishikawa, "Human activity recognition: various paradigms", *Int. Conf. on Control, Automation and Systems (ICCAS)*, Korea, Oct. 2008.
- [3] D. Gavrilă, "The visual analysis of human movement: a survey", *Computer Vision and Image Understanding*, vol. 73, pp. 82-98, 1999.
- [4] M. Pantic, A. Pentland, A. Nijholt, and T.S. Hunag, "Human computing and machine understanding of human behavior: a survey", *8<sup>th</sup> Int. Conf. on Multimodal Interfaces*, pp. 239-248, 2006.
- [5] R. Poppe, "Vision-based human motion analysis: an overview", *Computer Vision and Image Understanding*, vol. 108, no. 1-2, pp. 4-18, Oct. 2007.
- [6] G. Bradski and J. Davis, "Motion segmentation and pose recognition with motion history gradients", *Int. J. of Machine Vision and Applications*, vol. 13, no. 3, pp. 174-184, 2002.
- [7] Md. Atiqur Rahman Ahad, T. Ogata, J.K. Tan, H.S. Kim, and S. Ishikawa, "Motion recognition approach to solve overwriting in complex actions", *8<sup>th</sup> Int. Conference on Automatic Face and Gesture Recognition*, Amsterdam, 6 pages, Sept. 2008.
- [8] M.K. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. on Info. Theory*, pp. 179-187, 1962.
- [9] T. Ogata, J.K. Tan, and S. Ishikawa, "High-speed human motion recognition based on a motion history image and an Eigenspace", *IEICE Trans. Information & Systems*, vol. E89-D, no. 1, pp. 281-289, 2006.
- [10] M. Pantic, I. Patras, and M.F. Valstar, "Learning spatio-temporal models of facial expressions", *Int. Conf. on Measuring Behaviour*, pp. 7-10, Sep. 2005.
- [11] M. Valstar, I. Patras, and M. Pantic, "Facial action recognition using temporal templates", *IEEE Int. Workshop on Robot and Human Interactive Communication*, pp. 253-258, Sep. 2004.
- [12] V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Texture based description of movements for activity analysis" *Third Int. Conf. on Computer Vision Theory and Applications (VISAPP 2008)*, Madeira, Portugal, vol. 1, pp. 206-213.
- [13] A. Khotanzad and Y. H. Hong, "Invariant Image Recognition by Zernike Moments", *IEEE Transaction Pattern Analysis and Machine Intelligence*, Vol. 12, No. 5, pp. 489-497, May 1990.
- [14] H. Meng, N. Pears, and C. Bailey, "A Human Action Recognition System for Embedded Computer Vision Application", *3<sup>rd</sup> Workshop on Embedded Computer Vision (with CVPR)*, pp. 1-6, June 2007.
- [15] H. Meng, N. Pears, and C. Bailey, "Motion Information Combination for Fast Human Action Recognition", *2<sup>nd</sup> International Conference on Computer Vision Theory and Applications*, Spain, Mar. 2007.
- [16] H. Meng, N. Pears, and C. Bailey, "Recognizing Human Actions Based on Motion Information and SVM", *2<sup>nd</sup> IEEE International Conference on Intelligent Environments*, pp. 239-245, Athens, Greece, July 2006.
- [17] A.A. Efros, A.C. Berg, G. Mori, J. Malik, "Recognizing Action at a Distance", *ICCV*, pp. 726-733, 2003.
- [18] <http://gaijin-in-japan.com/2007/08/11/rajo-taiso-radio-exercise/>
- [19] J. Davis and A. Bobick, "Virtual PAT: a virtual personal aerobics trainer", *Proc. Perceptual User Interfaces*, pp. 13-18, Nov. 1998.
- [20] J. Davis and G. Bradski, "Real-time motion template gradients using Intel CVLib", *IEEE ICCV Workshop on Frame-rate Vision*, pp. 1-20, Sept. 1999.
- [21] W.C. Yau, D.K. Kumar, S.P. Arjunan, and S. Kumar, "Visual speech recognition using image moments and multiresolution wavelet", *Int. Conf. on Computer Graphics, Imaging and Visualization*, pp. 194-199, 2006.
- [22] Z. Yin and R. Collins, "Moving object localization in thermal imagery by forward-backward MHI", *Third IEEE Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum (OTCBVS'06)*, pp. 133-140, NY, June 2006.
- [23] L. Li, Q. Zeng, Y. Jiang, and H. Xia, "Spatio-temporal motion segmentation and tracking under realistic condition", *Int'l Symposium on Systems and Control in Aerospace and Astronautics*, pp. 229-232, 2006.
- [24] D. Son, T. Dinh, V. Nam, T. Hanh, and H. Lam, "Detection and localization of road area in traffic video sequences using motion information and fuzzy-shadowed sets", *IEEE Int'l Symp. on Multimedia*, pp. 725-732, Dec. 2005.
- [25] J. Liu and N. Zhang, "Gait history image: a novel temporal template for gait recognition", *IEEE Int. Conf. on Multimedia and Expo*, pp. 663-666, 2007.
- [26] J. Han and B. Bhanu, "Individual recognition using gait energy image", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316-322, 2006.
- [27] T. Jan, "Neural network based threat assessment for automated visual surveillance", *IEEE Int. Joint Conf. on Neural Networks*, vol. 2, pp. 1309-1312, July 2004.
- [28] K. Leman, G. Ankit, and T. Tan, "PDA-based Human Motion Recognition System", *Int. J. Software Engineering and Knowledge*, Vol. 2, Issue 15, pp. 199-205, Apr. 2005.
- [29] I. Petrás, C. Beleznai, Y. Dedeoğlu, M. Pardàs, et al., "Flexible test-bed for unusual behavior detection", *ACM Int. Conf. on Image and video retrieval*, pp. 105-108, 2007.

# Performance Analysis on an Efficient Human Motion Database with Various Motion Representations

S. M. Ashik Eftakhar, Joo Kooi Tan, Hyoungeop Kim, and Seiji Ishikawa

Department of Control Engineering, Kyushu Institute of Technology, Japan

E-mail: {ashik, etheltan, ishikawa}@ss10.cntl.kyutech.ac.jp

**Abstract** — In this paper, our proposed structured human motion database is adopted for different motion representations. The motions are first represented as a sequence of frames of 2D images, which were compressed using three recognized motion representation techniques: Exclusive-OR, MEI (Motion Energy Image), and MHI (Motion History Images). The representation is a 2D feature image. The feature image is compressed by characterizing the eigenvectors. A complete vector space called an Eigenspace is constructed that represents the image feature vectors for the feature image. The motions are indexed using the projections onto the eigenspace. For the purpose of efficient searching within the database, our proposed B-Tree Motion Database is created and maintained. The comparative performance evaluations for the aforesaid representations were investigated and satisfactory performances (about 90% recognition rate and smaller searching time) were realized for all of the cases using our proposed motion database structure.

**Index Terms** — Motion database, motion representation, motion compression, eigenspace, indexing, B-Tree.

## I. INTRODUCTION

A motion is, in a concise sense, any kind of activity performed by any subject. The activity might include any task, or simply any movement of the object. Human motion signifies the actions or movements performed by any human being. The researches involved in the recognition of human motion are considered to be a highly challenging task in computer- and robot-vision areas. With the development of computer vision system, the task of human motion/ activity recognition is getting huge significance in recent times. The motivation behind this kind of researches is the variety of applications in surveillance, discovering persons' physical problems in a medical system, aerobics, video conferencing, motion understanding, man-machine interfaces like robotics, biomechanics, virtual/ mixed reality, and in many other real-life applications[1][2]. There are a wide variety of human motion analysis techniques which are beyond this context. Aggarwal and Cai highlighted various forms of human motion representation and recognition [2]. Analyzing the techniques mentioned in [2], the field of motion representation and recognition can be briefly classified into two classes of methods: model-based and appearance-based methods [3]. Model-based methods

basically deal with the development of human 3D model for the task of representation. 3D human models are generally represented in terms of generalized cones, elliptical cylinders, and spheres [4]. For these geometric modeling a number of parameters is needed to be computed. For the complexity of computation, model-based methods are less preferable than appearance-based methods because of the simple modeling involved in appearance-based method. It deals with the 2D representation of motion rather than 3D. The motions are supposed to be composed of postures (appearances) that are somehow transformed into a suitable form or model for representation and recognition. Some recognized representations include Hidden Markov Model [5], Eigenspace [6], Body parts-based modeling [7], Motion Energy Images and Motion History Images [8], Exclusive-OR Technique [14], Directional MHI [9], and many others. The representation of eigenspace technique is much compact, as well as, suitable and applicable representation. It is identified as an image compression or coding technique. It is based on Principle Component Analysis (PCA), deviation of Karhunen-Loeve Transform [10]-[11]. Until now this technique was adopted successfully in many posture, gait, and motion recognition researches [3][6][12]. Different researchers have adopted different concepts of representation and recognition. For the reliable and real-life application of these representations and recognition strategies, the efficiency of those strategies in terms of recognition rate, as well as, searching complexity is a major issue. Recognition rate is simply the rate of recognized motions, whereas searching complexity is considered as the time that is employed for searching similar motions within the motion database.

However, the motion recognition strategies adopted earlier used to maintain a simple database capable of linear searching within that. But such searching is a brute-force searching approach that takes much time for searching in case of enormous motion data registered in the database. At this point, it is recommended to deduce some technique to reduce the searching by structurizing the searching technique. But in order to structurize the searching, construction of structured database is mandatory. Emphasizing on this point, a structured database is also proposed recently with human poses and motions [13]-

[14]. But the registration of huge amount of data, and high-speed searching with excellent performances is not yet been analyzed fruitfully. In [15]-[16], the balanced tree structure called a B-Tree was proposed as the structured database structure. Moreover, the B-Tree structure was successfully employed in many other researches, e.g., [15]-[16]. The reason behind the employment of B-Tree structure is its data arrangement. However, other database structure was also adopted in the motion analysis researches [17]-[18]; most of those were computationally complex to handle the database management.

In this research, successful analysis of motion database is accomplished by proposing a database capable of handling huge data with the ability of high recognition performance. The performance of the proposed technique is evaluated suitably and lucratively. Various representations are used to validate the performances and flexibility of our proposed method. The database development flowchart of the system is shown in Fig. 1.

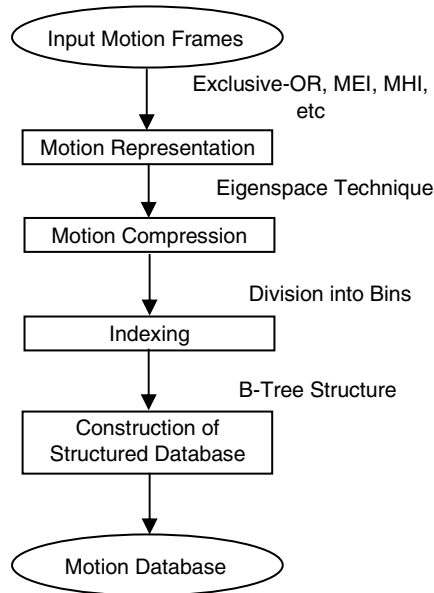


Fig.1. Database Development Flowchart.

## II. MOTION REPRESENTATION

Motion representation is a form of characterizing a motion for the computer to understand and to use the motion for recognition. In this sense it is a very crucial task. However, among many motion representations we have chosen three standard representations: Exclusive-OR, Motion Energy Image (MEI), and Motion History Image (MHI). The overview of these representations will be discussed in this section.

### A. Exclusive-OR (XOR) Representation

Rather taking all the motion frames into account, the Exclusive-OR (XOR) operations are performed between consecutive frames and the cumulative XORed form of the frames are referred to as XOR Image [14]. This is simple, effective, and fast generating motion representation method. Considering  $m$ ,  $h$ ,  $c$  representing motion, persons, camera-directions, and  $f$ ,  $U$  denoting original motion frame and XOR frame, then (1) presents the complete XOR operation on the motion frames.

$$\begin{aligned}
 U_c^{m,h}(2) &= f_c^{m,h}(1) \text{ XOR } f_c^{m,h}(2), \\
 U_c^{m,h}(r) &= U_c^{m,h}(r-1) \text{ XOR } f_c^{m,h}(r), \\
 U_c^{m,h} &\equiv U_c^{m,h}(R).
 \end{aligned} \quad (1)$$

### B. Motion Energy Image (MEI)

Bobbic and Davis [8] represented *Motion Energy Image* (MEI) as the region where there is motion. This information can be used to identify the motion occurrence and viewing condition. It is defined as the cumulative binary image of the motion regions extracted between the consecutive motion frames. Let,  $I(x, y, t)$  be an image sequence and  $D(x, y, t)$  be a binary difference image. Then binary MEI  $E(x, y, t)$  is defined by (2).

$$E_t = \bigcup_{i=0}^{t-1} D(x, y, t-i) \quad (2)$$

Here,  $t$  is the temporal extent which is critical to define. But for the flexibility of the value of  $t$ , it can be taken as the maximum gray level pixel value 255 [9].

### C. Motion History Image (MHI)

Bobbic and Davis [8], like MEI, also represented *Motion History Image* (MHI) as a frame-based temporal template for human motions. As the name implies, this form of motion representation keeps track of the motion history, i.e. representing how the motion is moving along a certain period of time. Let  $H$  be the pixel intensity function of the temporal history of motion at a particular point. The function is represented in a simple way in (3).

$$H_t(x, y, t) = \begin{cases} \tau & \text{if } D(x, y, t) = 1 \\ \max(0, H(x, y, t-1) - 1) & \text{Otherwise} \end{cases} \quad (3)$$

The function returns a scalar value, and according to the function, in the generated image the more recently moving pixels are brighter than past moving pixels. However, both the MEIs and MHIs are represented as vector-images (where and how motion is moving) that can be matched against stored representations of known movements.

In this research, we employed the aforesaid three representations for our proposed technique. However, the motion that is represented as a sequence of 2D frames is transformed into a suitable form to signify the region of interest as well as to eliminate unnecessary and unwanted noise elements. Therefore, the motionless or static region of the motion frames is excluded from each frame. Static region may include motionless part, noise, uneven edge pixels, etc. This portion can be treated as ignorable portion; so it is primarily removed from each motion frames using (4).

$$I_h(n) = \begin{cases} I_h(n) & \text{if } \exists k I_h(n) \neq I_k(n), h \neq k \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

Here,  $I_h(n)$  and  $I_k(n)$  are the  $n$ -th pixel value of  $h$ -th and  $k$ -th frame, respectively;  $n \in N$  and  $k \in F$ , where  $N$  is the total number of pixels and  $F$  is the total number of motion frames.

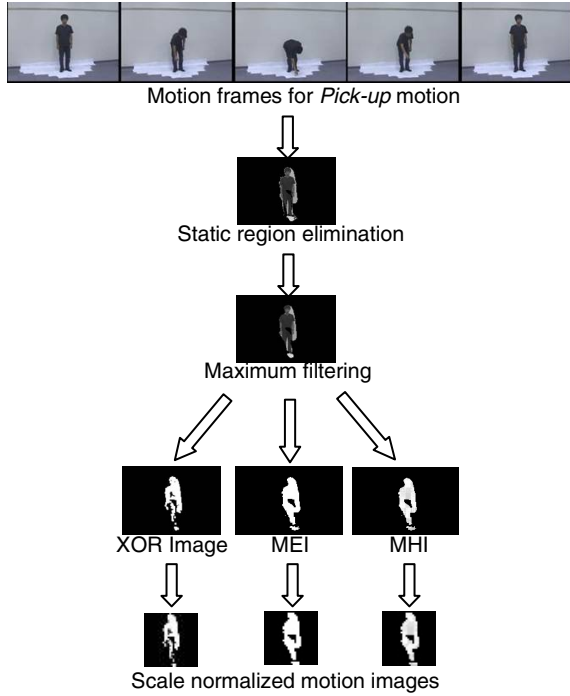


Fig. 2. Generation of different motion Representations.

After the elimination of static portion, the images are filtered using maximum filter to take significant regions into account. Then motion representations are adopted and the resultant motion images are generated step-by-step (Fig. 2). Next, the motion images generated from the input motion frames are *compressed* by projecting the motion images onto the *eigenspaces*.

### III. MOTION COMPRESSION

Motion compression is an important step in the human motion recognition system. It refers to either motion coding, or compressed data that are extracted from the motion. For example, extracting the features and parameter values from the motion are forms of motion compression. The parametric eigenspace representation is considered as an excellent motion compression technique. An *Eigenspace* is a hyperspace or a parametric feature space representing an image as a point on the hyperspace. It is a modified form of Karhunen-Loeve Transform which is basically used to derive relationship among different random variables. It has been used as a standard compression technique for its simplicity of use. In practice, a large set of learning motions is needed to be projected onto the parametric eigenspace by finding featuring eigenvectors. For each *motion image* (*XOR image*, *MEI*, or *MHI*)  $I_m$  ( $m = 1, 2, \dots, M$ ), an image matrix  $\hat{x}_m = (x_1, x_2, \dots, x_N)$  is defined and normalized as  $\|x_m\| = 1$ . A data matrix  $X$  is obtained by subtracting the average image  $c$  from each motion image set as follows,

$$X = (x_1 - c, x_2 - c, \dots, x_M - c) \quad (5)$$

The average image  $c$  is defined by,

$$c = \frac{1}{M} \sum_{m=1}^M x_m \quad (6)$$

The image matrix  $X$  is  $N \times M$ , where  $M$  is the total number of motion images, and  $N$  is the total number of pixels in each image. To compute eigenvectors of the image set (motion set), the *covariance matrix*  $Q$  is defined as (7).

$$Q = XX^T \quad (7)$$

$$\kappa = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \quad (8)$$

The eigenvalues and corresponding eigenvectors are calculated from the  $N \times N$  covariance matrix  $Q$  by solving the eigenvalue problem. However, Principal Component Analysis (PCA) strategy can be used here to reduce the dimensions of the eigenspace. Among  $N$  eigenvectors,  $k$  most prominent eigenvectors ( $e_1, e_2, \dots, e_k$ ) are chosen to create an eigenspace  $ES$  consisting of the learning motions using the metric expressed in (8). In (8), the value of  $\kappa$  is taken to be greater than or equal to 0.80. The eigenvectors for which the variances are more, those are chosen as the prominent eigenvectors. Each learning motion is projected onto the eigenspace  $ES$  as a point  $g_m$  by,

$$g_m = (e_1, e_2, \dots, e_k)^T x_m \quad (9)$$

For each camera direction, separate eigenspaces are created by projecting corresponding directional motions onto the eigenspace using (9). However, a *global eigenspace* is also constructed with all the learning motions. It confirms the decision about the most similar motions.

#### IV. MOTION INDEXING

Indexing is the processing denoting data in a suitable form so that it can be easily allocated into the database for further query. However, in this context, indexing involves the task of generating index from the points inside the eigenspace corresponding to each motion. The indexes are stored within the database afterwards. For the task of indexing the dimensions of the eigenspaces is taken as an important cue. The eigenspaces are uniformly divided into several subdivisions. Each eigen-axis  $e_k$  ( $k = 1, 2, \dots, K$ ) is separated into  $S$  sections each having equal length  $L$ . A hypercube with length  $L$  along each eigen-axis of an eigenspace is referred to as a *bin*. Each bin is assigned a number from 0 to  $S-1$ . Along each eigen-axis, each motion point is assigned a number between 0 and  $S-1$ . Thus each motion point in the eigenspace is represented as  $K$ -digit  $S$ -nary number which is termed as an *index* (Fig. 3). With large number of motions, if the number of bins is small, there is higher possibility that more motion points will have an identical index and need much time to search similar motions. Conversely, if the number of bins is large, less motion points will be sought. The reason behind it is: each bin covers less number of motion points at the time of higher number of bins.

|                  |       |                  |                  |                  |
|------------------|-------|------------------|------------------|------------------|
| Digit-( $N-1$ )  | ..... | Digit-2          | Digit-1          | Digit-0          |
| ( $0 \sim K-1$ ) | ..... | ( $0 \sim K-1$ ) | ( $0 \sim K-1$ ) | ( $0 \sim K-1$ ) |

Fig. 3. Index of  $N$ -digit  $K$ -ary number.

#### V. CONSTRUCTION OF STRUCTURED DATABASE

In this paper, a database is constructed capable of structured organization and fast query. As the large motion-database structure, B-Tree [19] data structure is employed in our proposed method. A B-Tree is a balanced tree that is capable of fast external memory searching. The B-Tree consisting of  $m$  descendents with height  $h$  is defined by  $\tau(m, h)$  and each page can hold up to  $m$  keys. It has the following properties:

- (a) Each path from the root to leaf has the same length  $h$ .
- (b) Root is a leaf or has 2 to  $m$  descendents.
- (c) Each node has  $\lfloor m/2 \rfloor$  to  $m$  descendents, except the root and leaves.

Within each page of the B-Tree, there are two kinds of data: a pointer and a key (Fig. 4). Pointer points to another page within the B-Tree, whereas the key stores the main information, i.e., the index. Considering the above properties, a B-Tree is constructed with the indexes. The generated index is stored in the B-Tree systematically. Those indexes are used for recognizing unknown motions in the recognition phase.

So far discussed database construction steps can be summarized as follows: First, for each camera direction, separate eigenspaces are created, then indexes are generated, and those are stored in a corresponding B-tree. The *global eigenspace* consisting of all the motions is used at the recognition stage. Separate eigenspaces arrange the motion data in an organized way and store much data within each B-Tree, at the same time we can still keep our searching faster. The structure of the page of B-Tree is shown in Fig. 5.

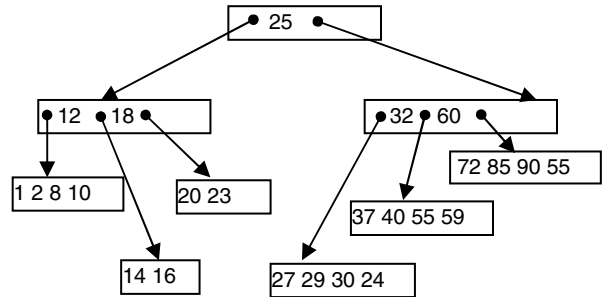


Fig. 4. A B-Tree structure in  $(4, 3)$ .

|       |       |       |       |       |       |              |
|-------|-------|-------|-------|-------|-------|--------------|
| $p_0$ | $x_i$ | $p_i$ | $x_2$ | $p_2$ | ..... | Unused Space |
|-------|-------|-------|-------|-------|-------|--------------|

Fig. 5. Organization of a page of a B-Tree.

#### VI. SEARCHING STRATEGY

The most significant part for developing a recognition system is the searching strategy. Both excellent and faster recognition is the concern here. As the B-Tree is a well-structured database structure, the searching is simpler and less time-consuming. The following conditions [19] maintained by the B-Tree structure guarantees efficient searching within the database:

$$\begin{aligned}
 &(\forall y \in \text{Keys}(p_0))(y < x_1), \\
 &(\forall y \in \text{Keys}(p_i))(x_i < y < x_{i+1}); \quad i = 1, 2, \dots, l-1, \quad (10) \\
 &(\forall y \in \text{Keys}(p_l))(x_l < y).
 \end{aligned}$$

Here, the B-Tree assumes  $l$  keys and  $l+1$  pointers within the page.

With the above structural arrangement, an index is searched within the B-Tree. If the query index resides within the database, it is found by simple *retrieval algorithm*. But for the case of the index not residing within the database, the most similar index corresponding to the most similar motion within the database is found. The pseudo code for finding the most similar index within the B-Tree after using the retrieval algorithm is as follows:

**STEP-1:** If query index  $y > x_i$  and  $Keys(p_i) = \text{NULL}$  and  $Keys(p_{i+1}) = \text{NULL}$ , calculate  $\text{MinDist}(x_{i+1}, y)$  and  $\text{MinDist}(x_i, y)$ . Then return the minimum distance index.

**STEP-2:** If query index  $y > x_i$  and  $Keys(p_i) = \text{NULL}$  and  $Keys(p_{i+1}) \neq \text{NULL}$ , select the last index from  $Keys(p_{i+1})$ . Then calculate  $\text{MinDist}(x_{i+1}, y)$  and  $\text{MinDist}(x_{p_{i+1,l}}, y)$ , and return the minimum distance index.

Besides the above, no other condition holds. The above algorithm finds the most similar index from the B-Tree.

## VII. RECOGNITION

The recognition strategy is also simple, but effective. An unknown motion is tested for recognition. When an unknown motion comes, it is first represented as a sequence of image frames. Then *XOR image*, *MEI*, or *MHI* is obtained from the motion frames, and then projected onto each camera-directional eigenspace. An index is generated from the unknown motion for each eigenspace representing motion identity within the directional database. For each camera direction, the number of similar motions is obtained by searching the corresponding B-Tree by calculating Euclidian distances among the indexes. Thus we get several *candidate* motions. Those candidate motions are then projected onto the global eigenspace as  $g_m (r=1,2,\dots,D)$ , where  $D$  is the number of directions. The unknown is projected as  $g_m$ . The most similar motion is calculated within the large motion database using (11):

$$d_m = \min_r \|g_{m_r} - g_m\| \quad (11)$$

Finally, the global eigenspace is used for finding the most similar motion.

## VIII. PERFORMANCE ANALYSIS

The performance of our proposed structure was analyzed with three different motion representations: Exclusive-OR, MEI, and MHI. The leave-one-out cross-validation method was used with a set of 60 motions, including 4 persons, 5 motions (*Pick-up*, *Carry*, *Walking*, *Headache*, and *Stomachache*) and 3 camera-directions

(*Left*, *Right*, *front*). The concept of bins is introduced here to reduce the amount of searching within the database. It can be assumed that the more the number of bins, the less the amount of searching. Different number of bins along each eigen-axis was taken into account for performance evaluation.

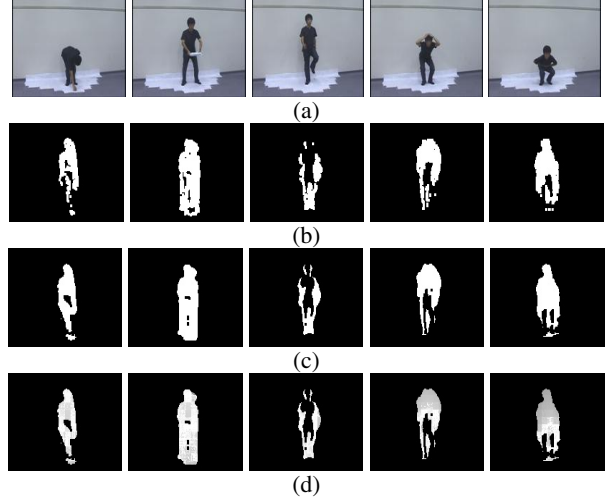


Fig. 6. Motion representations (frontal view) (a) Original motion frames, (b) XOR Image, (c) MEI, (d) MHI.

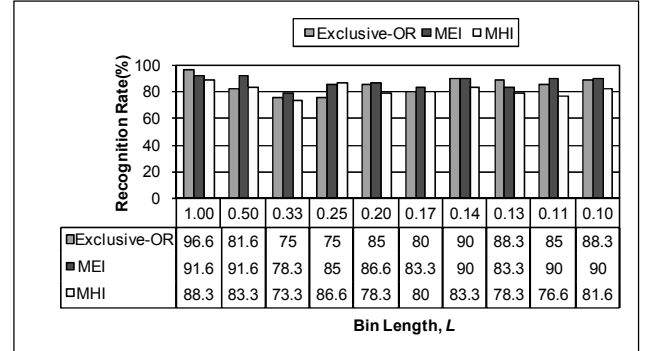


Fig. 7. Performance analysis using recognition rate.

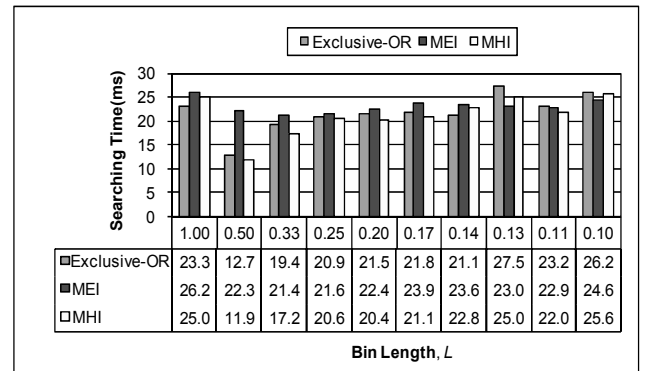


Fig. 8. Performance analysis using searching time.

For each motion representations, 60 motion images were generated with original size 160X107 pixels (Fig. 6) and were scale normalized to 32X32 pixels. Then the eigenspace was created for each camera direction by calculating eigenvalues and eigenvectors using SVD (Singular Value Decomposition) method. The performance was evaluated for each representation. The motions used for learning and testing were quite noisy; even for a single bin the recognition rate is not 100%. Still our proposed method shows excellent result with increased number of bins (Fig. 7). At bin length 0.14, 90% recognition rate was achieved. However, if the database searching time is considered, we see that at bin length 0.14 or less, the searching time is near about 22ms that is surely better than the brute-force search within a single bin (Fig. 8). The average recognition rate and searching time for Exclusive-OR, MEI and MHI are 84.5%, 87%, 81%, and 21.8ms, 23.2ms, 21.2ms, respectively. MEI shows the best recognition rate whereas MHI shows the least searching time with our experimental motions. The experiment was performed by taking the B-Tree parameter values  $m=4$ ,  $h=2$  operated on the PENTIUM IV, 2.8 GHz processor, 384 MB RAM Computer.

## IX. CONCLUSIONS

The main contribution of this research is the achievement of an excellent recognition rate with large number of motions in the motion database keeping searching time shorter. The flexibility of our database structure is also considered by employing different motion representations. For simple motions MEI can show satisfactory results, whereas incorporating MEI and MHI together, the performance will truly be improved [8]. Other representation and compression methods can be fruitfully employed by creating index, and storing into the B-Tree database for recognition by fast searching with accommodating huge number of motions. The proposed method has, indeed, much potential to be applicable in the real-time human motion recognition applications.

## ACKNOWLEDGEMENT

This work was partly supported by KAKENHI under grant 19700175, which is greatly acknowledged.

## REFERENCES

- [1] D. Gavril, "The visual analysis of human movement: a survey", *Computer Vision and Image Understanding*, Vol. 73, pp. 82-98, 1999.
- [2] J. K. Aggarwal and Q. Cai, "Human motion analysis: a review", *Computer Vision and Image Understanding*, Vol. 73, pp. 428-440, 1999.
- [3] T. Ogata, J. K. Tan, and S. Ishikawa, "High-speed human motion recognition based on a motion history image and an eigenspace", *IEICE Transactions on Information and Systems*, Vol. 89, Issue D(1), pp. 281-289, 2006.
- [4] K. Rohr, "Towards model-based recognition of human movements in image sequences", *CVGIP: Image Understanding*, Vol. 59, No. 1, pp. 94 – 115, 1994.
- [5] J. Yamato, J. Ohya, and K. Ishii, "Recognizing human action in time-sequential images using Hidden Markov Model", *In Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 379-385, June, 1992.
- [6] H. Murase and S. K. Nayar, "Learning and recognition of 3D objects from appearance", *IEEE Qualitative Vision Workshop New York*, pp. 39-50, 1993.
- [7] Chih-Chang Yu, Jenq-Neng Hwang, Gang-Feng Ho, and Chaur-Heh Hsieh, "Automatic human body tracking and modeling from monocular video sequence", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 1, pp. 917-920, 2007.
- [8] A. F. Bobick, and J. W. Davis, "The recognition of human movement using Temporal Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 3, pp. 257-267, 2001.
- [9] Md. Atiqur Rahman Ahad, T. Ogata, J.K. Tan, H.S. Kim, and S. Ishikawa, "Performance of multi-directional MHI for human motion recognition in the presence of outliers", *In Proc. Annual Conf. of IEEE Industrial Electronics Society (IECON)*, pp. 2366-2370, Nov. 2007.
- [10] E. Oja, *Subspace Methods of Pattern Recognition*, Research Studies Press, Hertfordshire, 1983.
- [11] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, London, 1990.
- [12] M. M. Masudur Rahman and Seiji Ishikawa, "Representing human postures/ motions using an eigenspace technique", *Proc. of International Conf. on Artificial Intelligence in Engineering & Technology*, pp. 232-236, 2002.
- [13] K. Kouno, J. K. Tan, and S. Ishikawa, "High-speed data retrieval in an eigenspace employing a B-tree structure", *Proc. of SICE-ACCAS International Joint Conference*, pp. 2717-2720, 2006.
- [14] J. K. Tan, K. Kouno, S. Ishikawa, H. S. Kim, and T. Shinomiya, "High speed human motion recognition employing a motion database", *Journal of Image & Electronic Society*, Vol. 36, No. 6, pp. 110-118, 2007.
- [15] T. Arndt, "A survey of recent research in image database management", *Proceedings of the IEEE Workshop on Visual Languages*, Vol. 4, Issue 6, pp. 92-97, 1990.
- [16] E. G. M. Petrakis and C. Faloutsos, "Similarity searching in medical image databases", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 9, pp. 435-447, 1997.
- [17] C. Li and B. Prabhakaran, "Indexing of motion capture data for efficient and fast similarity search", *Journal of Computers*, Vol. 1, No. 3, pp. 35-42, June 2006.
- [18] F. Liu, Y. Zhang, F. Wu, and Y. Pan, "3D motion retrieval with motion index tree", *Computer Vision and Image Understanding*, Vol. 92, Issue 2-3, pp. 265-284, 2003.
- [19] R. Bayer and E. McCreight, "Organization and maintenance of large ordered indexes", *Acta Informatica*, Vol. 1, No. 3, pp. 173-189, 1972.



# Action Recognition with Various Speeds and timed-DMHI Feature Vectors

Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim, and S. Ishikawa

Department of Control Engineering, Kyushu Institute of Technology, Fukuoka, Japan

E-mail: atiqahad@ss10.cntl.kyutech.ac.jp

**Abstract**— Usually, various motion recognition approaches can not perform well on actions that have variable speeds, or part of the dataset is speedy and vice versa. In this paper, we present the timing issue of our Directional Motion History Image method. This method can solve overwriting or motion self-occlusion problem significantly and thereby it performs well for complex and repetitive activities. However, it is important to analyze the method with activities having various speeds to show its robustness. The experimental results demonstrate that it can perform well with variable paces of the actions though the recognition rate is compromised a bit compared to the dataset having usual speed. We also improve the classification method with the incorporation of motion duration in the final feature vector so that for similar type of activities having different pace, it can show better result. Experiments on another dataset show better performance with the timing incorporation. The achieved recognition rates are very encouraging for further research and implementation is some intelligent systems for surveillance and related applications.

**Keywords**— DMHI, appearance-based model, action recognition, feature vector, moment.

## I. INTRODUCTION

Human action recognition with various speeds is very important for real-life video surveillance, aged care home, video annotation and indexing, security, image understanding, sports and video analysis, etc. to name a few [1-6]. Real-life activities vary from person to person, shape, etc. The same action (e.g., walking) may have different nature, speed, shapes based on the person's body structure (height, cloths, etc.) and therefore, it is important to have a global-based motion representation system that can be computed easily and without much computational cost. Most of the existent activity analysis methods concern and consider simple activities [2]. There are various gait analysis methods considering the length of foot-steps, angle, etc. However, gait analysis gives local information of the motion. For gesture or overall motion information of the human activity, we need to consider either human body parts analysis or template-based approaches considering the global human motion [1]. Ref. [23] points out that non-facial gestures must involve the rest of the body to appear natural and convincing. In this paper, we present our directional motion history image (DMHI) method for human motion recognition and activity analysis. However, earlier researches have considered regular pace of the actions. In this paper, to show DMHI method's robustness, we implement actions having various speeds and test with normal speed dataset and faster dataset. We show the robustness of the

DMHI method with actions having various speeds for the same action.

The paper is organized as follows: Section II presents some related works. In Section III, we present the temporal motion descriptors which are employed in this paper. Section IV explains the recognition methodology and timing issues related to feature vector calculations. Based on these concepts, Section V demonstrates some experimental results with two different datasets. Finally, Section VI concludes the paper with some directions on future work based on the present results.

## II. RELATED WORK

Vitaladevuni et al. [7] propose a method for temporally segmenting videos into atomic movements. At the same time, they use the approach on gesture recognition in rapid succession and at high speeds. Other approaches for temporal segmentation methods are available to understand complex actions [19-22]. For example, inspired by the natural language processing, Wang et al. [22] presented an approach for automatically segmenting sequences of natural activities into atomic sections and clustering them. Like words are extracted from long sentences, they extracted primitive gestures from sequences of human motion and clustered using HMM. Vili et al. [8] extract spatially enhanced local binary pattern (LBP) histograms from MHI and MEI temporal templates and model their temporal behavior with HMMs. They choose a fixed frame numbers in order to give a short term motion representation and the MHI is divided into four sub-regions. All the MHI and MEI LBP features are concatenated into one histogram and normalized so that the sum of the histogram equals one.

An adaptive model is constructed automatically by detecting human activity in MHIs using an iterative candidacy-classification-reduction process [12]. Singh et al. [13] used MHI and MEI and developed a Motion Color Image (MCI) by combining of motion and color cues. MCI is constructed by bit-wise OR of the MEI for four previous levels and color localization data. They dynamically controlled the frame differencing. Later they divide the MCI into 9 boxes and compute motion pixels count of motion data in MHI in each of the 9 boxes. Feature vectors are calculated as the sum of motion pixels in each box for classification.

Gorelick et al. [9] analyze actions by considering a video sequence as a space-time volume (XYT). They utilize properties of the solution to the Poisson equation to extract space-time features such as local space-time saliency, action

dynamics, shape structure and orientation. However, this method requires that the background is known ahead. They use the database that contains walk, run and some other motions. Another complex motion patterns' recognition approach considers local spatio-temporal interest points (local features) and integrates such representations with SVM classification schemes [10]. They recognize walking, running, jogging and other three activities. On the contrary to this approach, the MHI and the DMHI approaches consider global motion on the scene or global image measurements. In their experiment, confusion between walking and jogging as well as between jogging and running have high similarities of these classes (running of some people may appear very similar to the jogging of the others). However, this approach shows its robustness with both indoor and outdoor environments.

### III. TEMPORAL MOTION DESCRIPTORS

In this paper, we employ our directional motion history image (DMHI) approach for recognizing some activities with various speeds for complex activities of some aerobics motions. This section briefly presents the MHI and the DMHI representation. To recap what is the MHI method, we can state that the motion *history* image (MHI) can describe *how* the motion is moving in the image sequence [14]. Another representation called motion *energy* image (MEI) can demonstrate the presence of any motion or a spatial pattern in the image sequence. Both MHI and MEI templates comprise the motion history image (MHI) template-matching method. The MHI  $H_\tau(x, y, t)$  can be computed from update function  $\Psi(x, y, t)$ :

$$H_\tau(x, y, t) = \begin{cases} \tau & \text{if } \Psi(x, y, t) = 1 \\ \max(0, H_\tau(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \quad (1)$$

Here,  $x, y$  and  $t$  show the position and time;  $\Psi(x, y, t)$  signals object presence (or motion) in the current video image;  $\tau$  decides the temporal duration of MHI (in terms of frames); and  $\delta$  is the decay parameter. The MHI is a scalar-valued image where more recently moving pixels are brighter, and vice versa, whereas, MEI is a binary image that can be generated by thresholding the MHI above zero:

$$E_\tau(x, y, t) = \begin{cases} 1 & \text{if } H_\tau(x, y, t) \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

We develop the directional-MHI (DMHI) to solve the overwriting or self-occlusion problem of complex motions. A natural real-life motion has more than one atomic movements or basic actions. Using the basic MHI method, we can not solve the overwriting problem (e.g., motions like – sit-down and then stand-up motion; raise hand and then down motion; running both ways motion, etc.) because the former part of the motion (e.g., sit-down) is overwritten by the latter motion elements (e.g., then stand-up) [11, 15-16]. To solve this problem, four directional motion representations are proposed by splitting the optical flow [17] into four different channels and thereby, by creating four motion templates or images [18]. This DMHI approach can solve the motion overwriting problem significantly, and hence performs well for repetitive

or complex activities. Eq. 3 shows four motion history templates for this method based on four-directional optical flow components.

$$\begin{aligned} H_\tau^{-x}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_x^-(x, y, t) > th_p \\ \max(0, H_\tau^{-x}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \\ H_\tau^{+x}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_x^+(x, y, t) > th_p \\ \max(0, H_\tau^{+x}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \\ H_\tau^{-y}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_y^-(x, y, t) > th_p \\ \max(0, H_\tau^{-y}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \\ H_\tau^{+y}(x, y, t) &= \begin{cases} \tau & \text{if } \Psi_y^+(x, y, t) > th_p \\ \max(0, H_\tau^{+y}(x, y, t-1) - \delta) & \text{otherwise} \end{cases} \end{aligned} \quad (3)$$

We also compute four motion energy images (DMEI) after thresholding the motion history templates above zero:

$$\begin{aligned} E_\tau^{-x}(x, y, t) &= \begin{cases} 1 & \text{if } H_\tau^{-x}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \\ E_\tau^{+x}(x, y, t) &= \begin{cases} 1 & \text{if } H_\tau^{+x}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \\ E_\tau^{-y}(x, y, t) &= \begin{cases} 1 & \text{if } H_\tau^{-y}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \\ E_\tau^{+y}(x, y, t) &= \begin{cases} 1 & \text{if } H_\tau^{+y}(x, y, t) > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (4)$$

These eight motion templates are required for further motion representations. Using only history images or energy images for motion representation are not worthy. It has been proved both theoretically and experimentally that both templates are required for better motion representations. Though some researchers employ the history images only for their research, they employ or implement some other features along with the history image to represent the motion sequence in a robust way. However, to use only energy image for representation makes no real sense due to the fact that energy image gives the total motion region in the full video sequence. To make it clear, let's consider the stand up and sit-down motion sequence. In both reverse directions, we will get the same energy image (for sit and then stand sequence; or stand and then sit sequence) and hence energy image only can not give good cue unless the motions or actions are very much different among each other. E.g., actions like walk, jump on a place, raise hands, sleep on the floor, etc., have different dimensions and will give different energy images to separate these. Therefore, to accommodate various activities, as the MHI [14], energy and history templates are utilized [11].

### IV. FEATURE VECTORS COMPUTATION

This paper presents view-based template matching approach. An example of template matching comparisons is the use of distance measurements between the current data record (feature vectors) and each of the data records in the posture set recognizing the posture with the lowest distance measurement [5]. Template matching is the simplest of the hand posture and gesture recognition techniques, and for a

small set of postures, it is appropriate and can be quite accurate [5]. Traditionally, shape descriptors based on moments, which are invariant relatively to rotations, translations and scale, are utilized for feature vector calculation. Shape representation by statistical moments is a classical technique in the literature [25]. We employ Hu moments [24] to develop feature vectors for the MHI and the DMHI representations for each activity. To define the 2D  $(p+q)^{\text{th}}$  order moments of a density distribution function  $\rho(x,y)$  we can write,

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \rho(x,y) dx dy \quad p, q = 0, 1, 2, \dots \quad (5)$$

The central moment  $\mu_{pq}$  can be defined as,

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q \rho(x,y) d(x - \bar{x}) d(y - \bar{y}). \quad (6)$$

Here,  $\bar{x} = m_{10} / m_{00}$ ,  $\bar{y} = m_{01} / m_{00}$ . The first four orders (i.e.,  $(p+q)$  is from 0 to 3) are defined as,

$$\begin{aligned} \mu_{00} &= m_{00} \equiv \mu, \\ \mu_{10} &= \mu_{01} = 0, \\ \mu_{20} &= m_{20} - \mu \bar{x}^2, \\ \mu_{11} &= m_{11} - \mu \bar{x} \bar{y}, \\ \mu_{02} &= m_{02} - \mu \bar{y}^2, \\ \mu_{30} &= m_{30} - 3m_{20}\bar{x} + 2\mu \bar{x}^3, \\ \mu_{21} &= m_{21} - m_{20}\bar{y} - 2m_{11}\bar{x} + 2\mu \bar{x}^2 \bar{y}, \\ \mu_{12} &= m_{12} - m_{02}\bar{x} - 2m_{11}\bar{y} + 2\mu \bar{x} \bar{y}^2, \\ \mu_{03} &= m_{03} - 3m_{02}\bar{y} + 2\mu \bar{y}^3. \end{aligned} \quad (7)$$

For the second and third order moments, following seven orthogonal invariants (Eqs. 8) are achieved [24] to calculate feature vectors and we employ these invariants as feature vector:

$$\begin{aligned} I_1 &= \mu_{20} + \mu_{02}, \\ I_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \\ I_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \\ I_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, \\ I_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + \\ &\quad (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2], \\ I_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}), \\ I_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - \\ &\quad (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]. \end{aligned} \quad (8)$$

Based on these seven invariants, the MHI method calculates the 14-D feature vector (7 for MHI and 7 for MEI). In the DMHI method, we have eight history and energy components. Considering seven invariants for each along with eight normalized 0<sup>th</sup> order moment invariants ( $\overline{m_{00}^\sigma}$  where  $\sigma \in \{up, down, left, right\}$ ), we compute a 64-D feature vector ( $FV_{dmhi}$ ). These normalized 0<sup>th</sup> order invariants confer

the relative mass of the motion area in the scene. We can define  $FV_{dmhi}$  as:

$$FV_{dmhi} = (\forall h)(I_{1 \rightarrow 7}) \cup (\forall e)(I_{1 \rightarrow 7}) \cup (\forall h)(\overline{m_{00}^\sigma}) \cup (\forall e)(\overline{m_{00}^\sigma}). \quad (9)$$

where,  $h$  denotes the motion *history* image components, and  $e$  stands for energy components of the DMHI method.

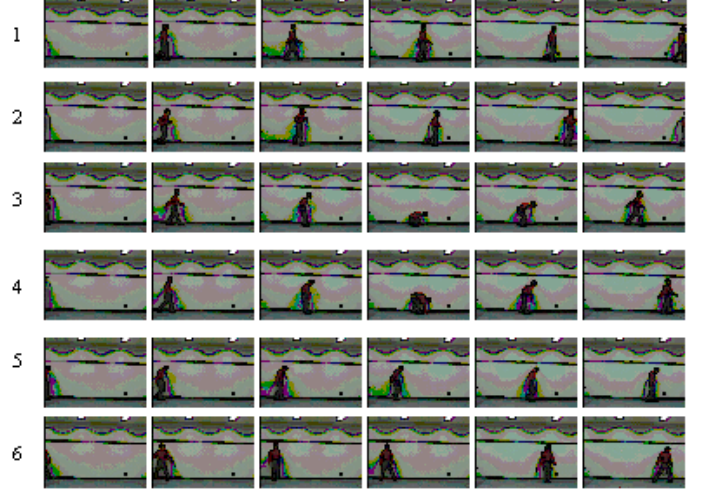


Fig. 1. Walk-type Database: some consecutive frames. Numbers 1 to 6 denotes the actions as – 1 for Walk, 2 for run, 3 for crouch, 4 for pick, 5 for walk as searching and 6 for sideway walk.

In this paper, another feature vector set is proposed considering the timing issue in mind, especially for actions having similarity in the templates but speed variations (Fig. 1). For example, to cover a specific distance by walking and running take different timing and hence, the timing incorporation can significantly improve the recognition rates. We can refer to this feature vector set as *timed-DMHI* ( $\tau$ -DMHI) feature vector,  $FV_{\tau \cup dmhi}$  and achieve a 65-D feature vector set:

$$FV_{\tau \cup dmhi} = (\forall h)(I_{1 \rightarrow 7}) \cup (\forall e)(I_{1 \rightarrow 7}) \cup (\forall h)(\overline{m_{00}^\sigma}) \cup (\forall e)(\overline{m_{00}^\sigma}) \cup (\tau). \quad (10)$$

## V. EXPERIMENTAL RESULTS AND ANALYSIS

For recognition process, we employ nearest-neighbor method for classification. For partitioning scheme, we consider leave-one-out cross-validation partitioning scheme. Two different datasets are employed here to establish the robustness of the DMHI method with various speeds. The achieved recognition results are satisfactory.

### A. Experiment I: Aerobics Datasets

Several aerobics datasets are considered in this experiment. The entire dataset has 10 activities done by 8 subjects. Table I shows the recognition results for various combinations of datasets. With the PA\_66 dataset (six persons' six activities) and 15 fps, we achieve lower recognition result for Activity-6. Without the presence of activity-6, the recognition rate is

better (96.7% for 15 fps). Similarly, with 10 activities for 8 persons (i.e., PA\_810 dataset), activity-6 and activity-7 show poor recognition results comparing to other activities. Due to this factor, the overall recognition rate is not as good as with the dataset of 30 fps.

Table I. Recognition rates (in %) of aerobic datasets with various speeds and combination. Note that the last row is only with the MHI; others are with the DMHI method.

| Dataset  | Testing Condition                                 | Rec. rate |
|----------|---------------------------------------------------|-----------|
| PA_810   | 30 fps                                            | 93.8      |
|          | 15 fps                                            | 78.8      |
| PA_66    | 30 fps                                            | 94.4      |
|          | 15 fps                                            | 86.2      |
| PA_86    | 30 fps                                            | 95.9      |
|          | Training 30 fps; Testing: 20% less data for a1~a4 | 83.4      |
| PA_810   | Training 30 fps; Testing: 20% less data for p1~p4 | 82.5      |
|          | Training and Testing: 20% less data for p1~p4     | 82.5      |
| With MHI | 30 fps                                            | 55.6      |
|          | 15 fps                                            | 38.9      |

We also implement the basic MHI method with 15 fps (thereby, speed is changing by 50%), and the recognition results are dropped than earlier (with 15 fps, the average recognition rate is achieved 38.9% only, whereas with normal speed of 30 fps, the average recognition rate was a bit better and 55.6%). These activities are not slow and the exercise sequences were taken with various subjects without considering any specific pace of the actions. Hence, by considering 15 fps, i.e., taking half of the frames compared to 30 fps for a period, we have lost some important correlations among the pixels. This is a reason to have a bit lower recognition rates with 15 fps. Fig. 2 shows some templates of energy images for 30 fps and 15 fps. For 15 fps, some missing information are visible (as black spots in the bottom row).

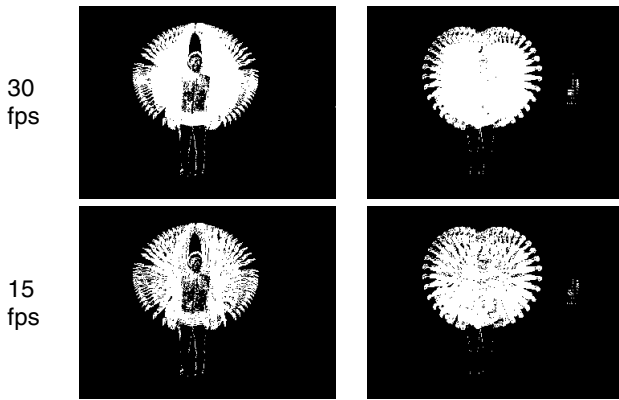


Fig. 2. Motion energy components for two actions with 30 fps and faster 15 fps.

In another experiment, we reduced the frames for activity-1 to activity-4 and rests of the activities are remained with usual speed (means, 30 fps). We reduced it about 20% and we consider training dataset having normal speed with 30 fps, but

for testing, we take reduced dataset. Hence, it is evident that the recognition will be a bit difficult as the speed is changing by one-fifth (<20%) for part of the dataset. In this case, we achieve 83.4% recognition results.

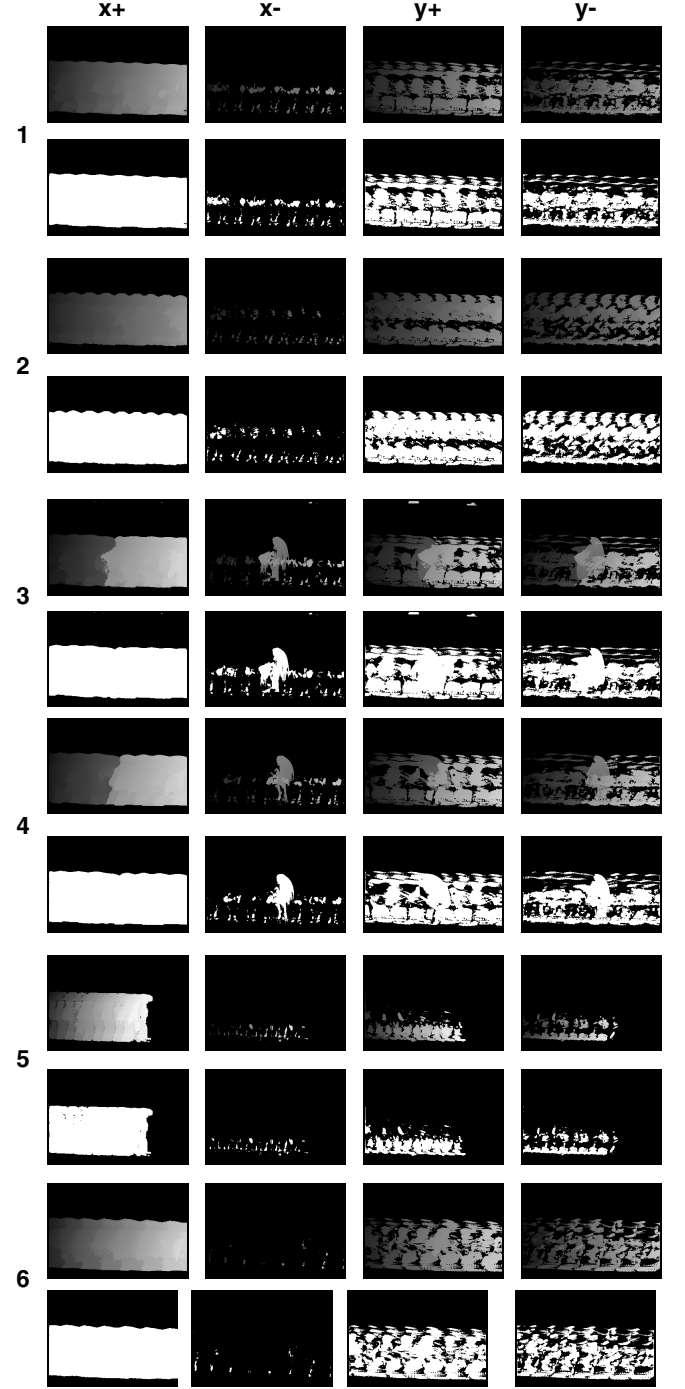


Fig. 3. For six actions (for Walk-type database), two rows represents DMHIs (top row) and the DMEIs (bottom row) respectively.

### B. Experiment II: Walk-type Datasets

Later, we implemented this DMHI method (and few other template-matching methods – not presented here as their performances were poor) for another dataset (we will refer to

it as ‘Walk-type Dataset’), consisting of different six actions (see Fig. 1), namely: normal-walk, run, walk-crouch-walk; walk-pick-walk, walk as if searching, sideways-walk. These actions have variable speeds. It is clear from the action-names that all the actions are very similar (starting from one corner in the scene and finishing the other side with variable pace and style) and to recognize this kind of datasets is quite difficult. There are some complex methods to recognize this kind of actions, but we upgraded the DMHI method with timing incorporation in the feature vectors. This is a simple approach.

Using 64-D feature vectors ( $FV_{dmhi}$ ) of the usual DMHI method for this kind of actions are not suitable as these actions have similarity in the final templates (Fig. 3 shows the final motion history and energy templates for the six actions) but the timing is different.

Table II. Recognition results with two different feature vectors ( $FV_{dmhi}$  and  $FV_{\tau \cup dmhi}$ )

| Action name                    | $FV_{dmhi}$ | $FV_{\tau \cup dmhi}$ |
|--------------------------------|-------------|-----------------------|
| 1. Walk                        | 66.7        | 100                   |
| 2. Run                         | 33.4        | 66.7                  |
| 3. Crouch (with walk)          | 66.7        | 100                   |
| 4. Pick from floor (with walk) | 66.7        | 100                   |
| 5. Search walk                 | 66.7        | 100                   |
| 6. Sideways walk               | 66.7        | 100                   |
| Average Rec. Rate              | 61.2        | 94.5                  |

Table II shows the recognition rates of various Walk-type activities. With the  $\tau$ -DMHI feature vector set ( $FV_{\tau \cup dmhi}$ ), we achieve much better recognition rate. However, to show its robustness, we have to consider various other datasets, which consist of this type of activities. An important pair of activities are running and walking across the optical axis and to recognize these is very crucial for video surveillance. As evident from Table II, the usual DMHI shows a bit similar results for both cases. By considering time measurement, these could be recognized separately.

## VI. CONCLUSIONS AND FUTURE WORKS

Human motion analyses in robust ways are still its infancy. Through various methods we can achieve good results; however most of the cases, these are dataset-specific or environment-specific with various constraints. We demonstrated the robustness of the DMHI method having various speeds and complexity of the actions. We achieved satisfactory recognition results. For Walk-type dataset and with the  $\tau$ -DMHI feature vector set ( $FV_{\tau \cup dmhi}$ ), we achieve much better recognition rate.

However, to show its robustness, we have to consider various other datasets, which consist of this type of activities. The system considers single frontal-view camera and not view-invariant. Moreover, all the experiments are accomplished indoor with almost constant illumination. In future, we need to consider outdoor recognition process as

well as with the presence of more than one person in the view. Currently, we are developing a robust temporal motion segmentation method based on this directional method.

## REFERENCES

- [1] Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim, and S. Ishikawa, “Human activity recognition: various paradigms”, *Int. Conf. on Control, Automation and Systems (ICCAS)*, Korea, Oct. 2008.
- [2] J. K. Aggarwal and Q. Cai, “Human motion analysis: a review”, *Computer Vision and Image Understanding*, vol. 73, pp. 428-440, 1999.
- [3] R. Poppe, “Vision-based human motion analysis: an overview”, *Computer Vision and Image Understanding*, vol. 108, no. 1-2, pp. 4-18, Oct. 2007.
- [4] D. Gavrilala, “The visual analysis of human movement: a survey”, *Computer Vision and Image Understanding*, vol. 73, pp. 82-98, 1999.
- [5] J. LaViola. A survey of hand posture and gesture recognition techniques and technology. *Tech. Report CS-99-11*, Brown University, 80 pages, June 1999.
- [6] T. B. Moeslund, A. Hilton, and V. Kruger, “A survey of advances in vision-based human motion capture and analysis”, *Computer Vision and Image Understanding*, vol. 104, pp. 90-126, Oct. 2006.
- [7] S.N. Vitaladevuni, V. Kellokumpu, and L.S. Davis, “Action recognition using ballistic dynamics”, *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 8 pages, 2008.
- [8] K. Vili, Z. Guoying, and P. Matti, “Texture based description of movements for activity analysis”, *Int. Conf. on Computer Vision Theory and Applications (VISAPP 2008)*, vol. 1, pp. 206-213, 2008.
- [9] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, “Actions as space-time shapes”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2247-2253, Dec. 2007.
- [10] C. Schudt, I. Laptev, and B. Caputo, “Recognizing human actions: a local SVM approach”, *Int. Conf. on Pattern Recognition 2004*, vol. 3, pp. 32-36, Aug. 2004.
- [11] Md. Atiqur Rahman Ahad, T. Ogata, J.K. Tan, H.S. Kim, and S. Ishikawa, “Motion recognition approach to solve overwriting in complex actions”, *8<sup>th</sup> Int. Conference on Automatic Face and Gesture Recognition*, Amsterdam, 6 pages, Sept. 2008.
- [12] J. W. Davis, A. M. Morison and D. D. Woods, “Building adaptive camera models for video surveillance”, *IEEE Workshop on Applications of Computer Vision (WACV’07)*, 6 pages, 2007.
- [13] R. Singh, B. Seth, and U. Desai, “A real-time framework for vision based human robot interaction”, *Proc. of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 5831-5836, Oct. 2006.
- [14] A. Bobick and J. Davis, “The recognition of human movement using temporal templates”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 257-267, March 2001.
- [15] M. Pantic, I. Patras, and M. F. Valstar, “Learning spatio-temporal models of facial expressions”, *Int. Conf. on Measuring Behaviour*, pp. 7-10, Sep. 2005.
- [16] H. Meng, N. Pears, and C. Bailey, “A human action recognition system for embedded computer vision application”, *3<sup>rd</sup> Workshop on Embedded Computer Vision (with CVPR)*, pp. 1-6, June 2007.
- [17] W. J. Christmas, “Spatial filtering requirements for gradient-based optical flow measurement”, *9<sup>th</sup> British Machine Vision Conference (BMVC)*, pp. 185-194, 1998.

- [18] A.A. Efros, A.C. Berg, G. Mori, and J. Malik, "Recognizing action at a distance", *IEEE Int. Conf. on Computer Vision*, pp. 726-733, 2003.
- [19] K. Kahol, P. Tripathi, S. Panchanathan, and T. Rikakis, "Gesture segmentation in complex motion sequences", *Int. Conf. on Image Processing*, vol. 2, pp. 105-108, Sept. 2003.
- [20] Md. Atiqur Rahman Ahad, H. Uemura, J.K. Tan, H.S. Kim, and S. Ishikawa, "A simple real-time approach for action separation into action primitives", *Int. Workshop on Tracking Humans for the Evaluation of Their Motion in Image Sequences (THEMIS2008) with British Machine Vision Conference (BMVC2008)*, pp. 69-78, Sept. 2008.
- [21] J. Barbic, A. Safonova, J.-Y. Pan, C. Faloutsos, J. K. Hodgins, and N. S. Pollard, "Segmenting motion capture data into distinct behaviors", *Graphics Interface*, pp. 185-194, 2004.
- [22] T. Wang, H. Shum, Y. Xu, and N. Zheng, "Unsupervised analysis of human gestures", *IEEE Pacific Rim Conf. on Multimedia*, pp. 174-181, 2001.
- [23] N. Badler, M. Costa, L. Zhao, and D. Chi, "To gesture or not to gesture: what is the question?", *Computer Graphics Int.*, pp. 3-9, 2000.
- [24] M. K. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. on Info. Theory*, pp. 179-187, 1962.
- [25] D. Zhang and G. Lu, "Review of shape representation and description techniques", *Pattern Recognition*, vol. 37, pp. 1-19, 2004.

# Intelligent Video Surveillance for Monitoring Fall Detection of Elderly in Home Environments

Homa Foroughi<sup>1</sup>, Baharak Shakeri Aski<sup>2</sup>, and Hamidreza Pourreza<sup>1</sup>

<sup>1</sup>Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Khorasan, Iran

<sup>2</sup>Islamic Azad University of Ramsar, Ramsar, Mazandaran, Iran

**Abstract** — Video Surveillance is an omnipresent topic when it comes to enhancing security and safety in the intelligent home environments. In this paper we propose a novel method to detect various posture-based events in a typical elderly monitoring application in a home surveillance scenario. These events include normal daily life activities, abnormal behaviors and unusual events. Due to the fact that falling and its physical-psychological consequences in the elderly are a major health hazard, we monitor human activities with a particular interest to the problem of fall detection. Combination of best-fit approximated ellipse around the human body, projection histograms of the segmented silhouette and temporal changes of head position, would provide a useful cue for detection of different behaviors. Extracted feature vectors are fed to a MLP Neural Network for precise classification of motions and determination of fall event. Reliable recognition rate of experimental results underlines satisfactory performance of our system.

**Index Terms** — Elderly Monitoring, Fall Detection, Home Surveillance, Human Shape, MLP Neural Network, Posture Recognition.

## I. INTRODUCTION

Intelligent video surveillance systems are receiving a great deal of interest especially in the fields of personal security and assistance. These systems are built in order to accomplish several tasks from detection of human presence to recognition of their activities. In the past few decades, vision-based surveillance has been extensively applied on industrial inspection, traffic control, security systems, and medical and scientific research.

One of the application areas of video surveillance is monitoring the safety of elderly in home environments. In the case of elderly people living on their own, there is a particular need for monitoring their behavior, such as a fall, or a long period of inactivity. Fall in the elderly is a major public health problem and may lead to injury, restricted activities, fear or death. A fall incident not also causes many disabling fractures but also has dramatic psychological consequences that reduce the independence of the person. It is shown in [12] that 28-34% elderly people in the community experience at least one fall every year, and 40-60% of the falls lead to injury.

Human society is experiencing tremendous demographic changes in aging since the turn of the 20th

century. Thus with the population growing older and increasing number of people living alone, supportive home environments able to automatically monitor human activities are likely to widespread due to their promising ability of helping elderly people.

Nowadays, the usual solution to detect falls is wearable sensors [1], [3]. These autonomous sensors are usually attached under the armpit, around the wrist, behind the ear's lobe or at the waist. These devices integrate accelerometer and/or inclinometer sensors and can monitor velocity and acceleration, vertical posture toward lying posture. However the problem of such detectors is that older people often forget to wear them, indeed their efficiency relies on the person's ability and willingness to wear them. Moreover in the case of noncontact sensors, they often provide fairly crude data that's difficult to interpret. Although, computer vision systems try to extract some considerable features from video sequences of movement patterns to detect falls. The data provided by cameras are semantically richer and more accurate than standard sensors.

In this paper we present a novel video analysis based approach for monitoring human activities with a particular interest to the problem of fall detection. The reminder of the paper is organized as follows: in section II we briefly review some existing vision-based fall detection systems, in section III our proposed system is introduced and then more technical details are described in section IV. Experimental results are represented in section V and finally we conclude in section VI.

## II. RELATED WORK

Recently some research has been done to detect falls using image processing techniques. A simple method was used in [5], [6] based on analyzing aspect ratio of the moving object's bounding box. This method could be inaccurate, depending on the relative position of the person, camera, and perhaps occluding objects. The works in [2], [7] used the normalized vertical and horizontal projection of segmented object as feature vectors. To overcome occluding objects problem, some researchers have mounted the camera on the ceiling: Lee [8] detected a fall by analyzing the shape and 2D velocity of the

person. Nait-Charif [9] tracked the person using an ellipse and inferring falling incident when target person is detected as inactive outside normal zones of inactivity like chairs or sofas. Rougier [4] used wall-mounted cameras to cover large areas and falls were detected using motion history image and human shape variation. Other systems used the audio information or using 3D trajectory and speed of head to infer events [11]. These mechanisms tend to be more complex and need more additional cost. Despite the considerable achievements that has accomplished on this field in the recent years, there are still some clear challenges to overcome:

- Visual fall detection is inherently prone to high levels of false positive as what appears to be a fall might not be a fall, but a deliberate movement towards the ground. In other words, most of current systems [6], [7], [9], [10] are unable to discriminate between real fall incident and an event when person is lying or sitting down abruptly.
- Existent fall detection systems tend to deal with restricted movement patterns and fall incidents are usually detected in contrast with limited normal scenarios like walking; however in real home environments various normal /abnormal motions occur.

### III. PROPOSED SYSTEM OVERVIEW

This paper proposes a novel method for monitoring human posture-based events in a home environment with focus on detecting three types of fall. We tried to simulate real life situations; to this aim, we considered comprehensive movement scenarios consisting normal daily life activities such as walking, running, bending down, sitting down and lying down, some abnormal behaviors like limping or stumbling, and also unusual events like falling. Furthermore several scenarios of falling have been regarded. These ten different postures are described in section V.A in more details.

Since we are interested in analyzing the motion occurring in a given window of time, firstly we need to obtain the segmentation of moving objects. Therefore, a background estimation procedure is performed to separate motion from background. After the silhouettes are acquired, the next step involves extracting features. Since changes in the human shape can discriminate if the detected motion is normal (e.g. the person walks or sits) or abnormal (e.g. the person falls), we analyze the shape changes of the extracted silhouettes in the video sequence. It seems that combination of approximated ellipse around the human body, horizontal and vertical projection histograms and temporal changes of human head position; would provide a useful cue for detecting different behaviors. By using the motion feature, the key motion

information can be encoded without any accurate visual tracking. The extracted features will then be fed to the MLP neural network with back propagation learning scheme for precise classification of motions.

## IV. TECHNICAL DETAILS

### A. Foreground Segmentation

Background subtraction is a particularly popular method for motion segmentation and attempts to detect moving regions in an image by differencing between current image and a reference background image in a pixel-by-pixel fashion. In order to human segmentation and extracting the moving objects; here we use a background subtraction method described in [15], which is fairly robust and gives appropriate results on image sequences with shadows, highlights and high image compression.

### B. Feature Extraction

One of the issues of major importance in a recognition system is feature extraction, i.e., the transition from the initial data space to a feature space that will make the recognition problem more tractable. So we analyze the shape changes of the extracted silhouettes in the video sequence. To this aim, three main features that retain the motion information of the actions and can be easily obtained are selected to form a feature vector.

#### • *Approximated Ellipse*

When a motion is detected, an analysis on the moving object is performed to detect a change in the human shape, more precisely in orientation and proportion. The person is then approximated by an ellipse using moments [4]. An ellipse is defined in a 5D parameter space by its center  $(\bar{x}, \bar{y})$ , its orientation and the length  $a$  and  $b$  of its major and minor semi-axes [13]. The authors believe that this representation of a person is rich enough to support recognition of different events. For a continuous image  $f(x, y)$ , the moments are given by:

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad \text{for } p, q = 0, 1, 2, \dots \quad (1)$$

The center of the ellipse  $(\bar{x}, \bar{y})$  is used to compute the central moment as follows:

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) d(x - \bar{x}) d(y - \bar{y}) \quad (2)$$



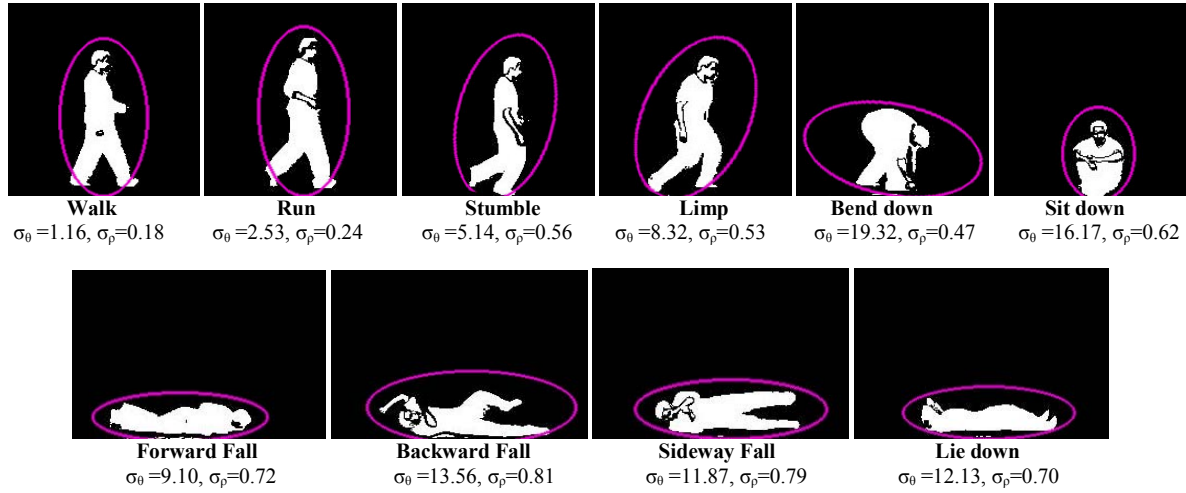


Fig. 1. Approximated Ellipses of Different Human Postures

The angle between the major axis of the person and the horizontal axis  $x$  gives the orientation of the ellipse and can be computed as follows:

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \quad (3)$$

$I_{\min}, I_{\max}$  - the least and greatest moments of inertia-are computed by evaluating the eigenvalues of the matrix:

$$J = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \quad (4)$$

Then, the major semi-axis  $a$  and the minor semi-axis  $b$  of the best fitting ellipse are given by:

$$a = \left(\frac{4}{\pi}\right)^{1/4} \left[\frac{I_{\max}^3}{I_{\min}}\right]^{1/8} \quad b = \left(\frac{4}{\pi}\right)^{1/4} \left[\frac{I_{\min}^3}{I_{\max}}\right]^{1/8} \quad (5)$$

With  $a$  and  $b$ , we can also defined the ratio of the ellipse  $= a/b$  (related to its eccentricity). In order to detecting changes in the human shape discriminate an abnormal behavior from other normal activities, two values are computed:

- **The orientation standard deviation of the ellipse:** If a person falls perpendicularly to the camera optical axis, then the orientation will change significantly and  $\sigma_\theta$  will be high. If the person just walks,  $\sigma_\theta$  will be low.
- **The ratio standard deviation of the ellipse:** If a person falls parallelly to the camera optical axis, then the ratio will change significantly and  $\sigma_p$  will be high. If the person just walks,  $\sigma_p$  will be low.

Note that  $\sigma_\theta$  and  $\sigma_p$  are computed for a 2 second duration. As Fig. 1 demonstrates,  $\sigma_\theta$  and  $\sigma_p$  are completely distinctive for different postures.

#### • Projection Histograms

The shape of a 2D binary silhouette can be represented by its projection histogram [13]. The horizontal and vertical projection histogram of foreground is obtained by calculating the number of foreground pixels row wise and column wise. Since projection histograms vary according to location of object in the scene, it's necessary to perform normalization step first. The widely used approach for normalization is rescaling the detected silhouette to a fixed length  $M$  [2]. However, the normalization parameter is usually obtained by trial and error and also it is posture dependant and strongly sensitive to different postures. To overcome these drawbacks, we propose exploiting *Discrete Fourier Transform* method for normalization.. Let us denote the image size as  $N \times M$ , so we can consider the foreground  $F$  as cloud of  $(x_p, y_p)$  points. Then the horizontal and vertical projection histograms of foreground  $F$  can be defined as follows:

$$Hz(y) = \left| (x_p, y_p) \in F, y_p = y \right| \quad (6)$$

$$Vt(x) = \left| (x_p, y_p) \in F, x_p = x \right|$$

Applying DFT on  $H_z(y)$  and  $V_t(x)$  results in:

$$DFT_{Hz(h)} = \frac{1}{N} \sum_{i=1}^N H_z(i) e^{-2j\pi \frac{hi}{N}} \quad h = 0, \dots, N-1 \quad (7)$$

$$DFT_{Vt(v)} = \frac{1}{M} \sum_{i=1}^M V_t(i) e^{-2j\pi \frac{vi}{M}} \quad v = 0, \dots, M-1$$

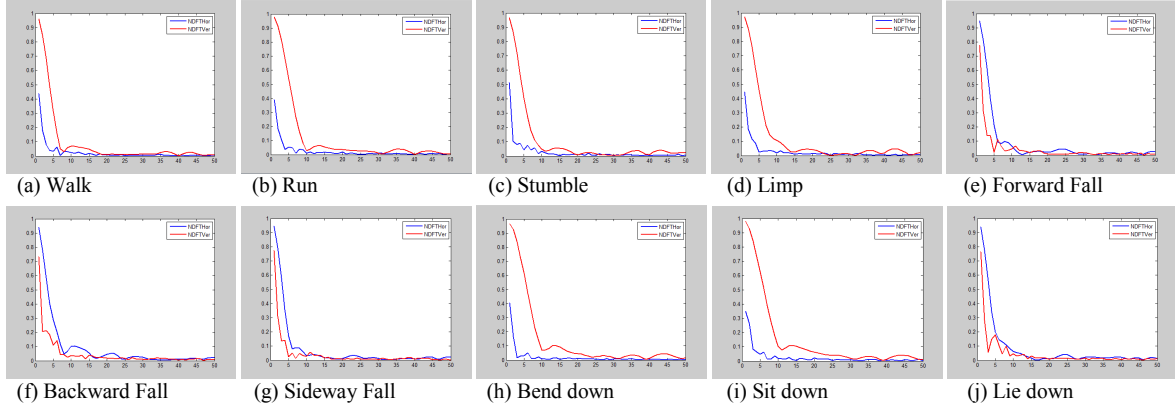


Fig. 2. Projection Histograms

Magnitudes of these DFT coefficients tend to decay for large values of  $h$  and  $v$ . Besides, different postures magnitudes differ mostly in the first few term.

Thus we extract first fifty significant DFT coefficients i.e.  $DFT_{Hz(1)} \dots DFT_{Hz(50)}$  and  $DFT_{Vt(1)} \dots DFT_{Vt(50)}$  are selected and normalized by:

$$NDFT_{Hz(h)} = \frac{|DFT_{Hz(h)}|}{|DFT_{Hz(1)}|} \quad h = 2, \dots, 50 \quad (8)$$

$$NDFT_{Vt(v)} = \frac{|DFT_{Vt(v)}|}{|DFT_{Vt(1)}|} \quad v = 2, \dots, 50$$

Normalized magnitudes of fifty significant DFT coefficients of different postures are shown in Fig. 2.

#### • Temporal Changes of Head Position

Head position is a potentially powerful and intuitive pointing cue if it can be obtained accurately and noninvasively. We choose to track the head because it is usually visible in the scene (no occlusion problems), and has a large movement during a fall. Our proposed method tries to estimate a person's head position in every frame of a video sequence and further infer his attentive behavior. To localize the head of the person, firstly silhouette is enclosed by minimum bounding rectangle and then the topmost detected point of the bounding box is marked in each frame [14]. Then the absolute difference values in successive frames are obtained. The difference in topmost point of the head over successive frames forms this part of our feature vector, and applying appropriate threshold on the vertical displacement of topmost point, results in proper differentiation of various behaviors. The curves obtained from sequence of successive head positions are shown in Fig. 3. Notice that the movement of head is computed for a 3 second duration.

#### C. Posture Classification

Neural Network is an interesting approach for analyzing time-varying data. Here, for each activity which has to be recognized a solitary neural network has been used. So for recognition of  $n$  activities,  $n$  neural networks exist. In this study we use a four-layered MLP network with backpropagation learning schema. MLP is a supervised neural network that can have multiple inputs and outputs and multiple hidden layers with arbitrary number of neurons. The activation function is  $a_j = f(\sum w_{ji} a_i)$ , where the  $w_{ji}$  is the weight of  $j$ th node in current layer to  $i$ th node in previous layer and  $a_i$  is the output of  $i$ th node in previous layer.  $f(x)$  is sigmoid function. All available data are random divided into two subsets, for training and testing respectively.

### V. EXPERIMENTAL RESULTS AND DISCUSSION

#### A. Experimental Setup

In order to evaluate the overall system performance, we apply the proposed approach to a set of videos recorded in our lab. Here we consider indoor environment settings with single fixed camera monitoring static scene. Distance of person to the camera is approximately 4-5 meters. We could expect three different kinds of behavior: Normal, Unusual and Abnormal.

- a) **Normal Daily Activity:** Five different normal daily activities are considered: Walking, Running, Bending down for catching something and rising up, Sitting down on the floor and standing up and Lying down.
- b) **Unusual -Fall-:** As most falls occur during intentional movements initiated by the person, they happen mainly in forward or backward: stumbling on an obstacle during walking, backward slip on wet ground.

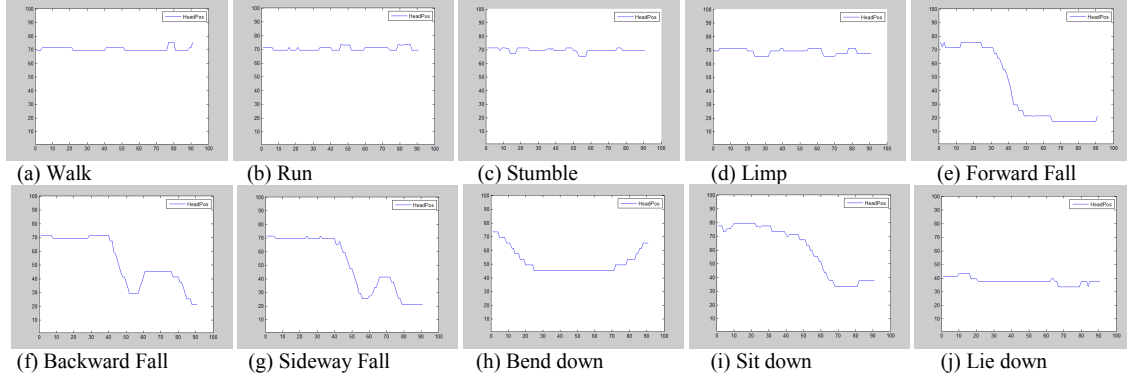


Fig. 3. Vertical Movement of Person's Head

But in some cases the fall occurs sideways, either during a badly controlled sit to stand transfer. In this case, the person frequently tries to grip the wall [1].

#### c) Abnormal

**Stumble:** Subjects were asked to walk in an unusual way, e.g. as if they were suffering a balance deficiency such as dizziness. In these situations person is unable to keep his balance or symmetry and synchrony of his movement. Body movements suggest the person is in a dubious condition.

**Limp:** Limping may be caused by unequal leg lengths, experiencing pain when walking, muscle weakness, disorders of proprioception, or stiffness of joints.

50 subjects with different height, weight and genders whose ages ranged from 20 to 30 were asked to participate in the project. We repeated 10 kinds of activities by 5 times in the experimental space.

#### B. Evaluation of the Proposed Method

TABLE I  
RECOGNITION RATE FOR VARIOUS EVENTS

| Events        | $N_a$ | $N_c$ | $N_f$ | $R$   |
|---------------|-------|-------|-------|-------|
| Walk          | 125   | 114   | 11    | 91.20 |
| Run           | 125   | 112   | 13    | 89.60 |
| Stumble       | 125   | 112   | 13    | 89.60 |
| Limp          | 125   | 110   | 15    | 88.00 |
| Forward Fall  | 125   | 116   | 9     | 92.80 |
| Backward Fall | 125   | 118   | 7     | 94.40 |
| Sideway Fall  | 125   | 114   | 11    | 91.20 |
| Bend down     | 125   | 113   | 12    | 90.40 |
| Sit down      | 125   | 112   | 13    | 89.60 |
| Lie down      | 125   | 118   | 7     | 94.40 |

The experimental results show that the proposed system is able to achieve promising results under most test conditions. Table I represents robust recognition rate of experimental results.  $N_a$  refers to number of actions,  $N_c$

is number of correctly detected events,  $N_f$  is number of falsely detected events and  $R$  is the recognition rate.

To evaluate performance of the proposed system, we use two famous criteria that are widely used in fall detection systems. Sensitivity is the capacity to detect a fall and Specificity is the capacity to detect only a fall:

$$Sensitivity = \frac{TP}{TP + FN}, \quad Specificity = \frac{TN}{TN + FP}$$

The definitions of TP, FP, FN and TN are as follows:

| Fall Incident \ System Recognition | Occurs | Does Not Occur |
|------------------------------------|--------|----------------|
| Positive                           | TP     | FP             |
| Negative                           | FN     | TN             |

According to section V.A dataset of our experiments is composed of video sequences representing 875 normal/abnormal activities and 375 simulated falls. Table II itemizes all the results obtained with our dataset. According to above information, Sensitivity and Specificity of the proposed system would be 92.80 and 97.60 respectively.

TABLE II  
EVALUATION OF RECOGNITION RESULTS

| Fall Incident \ System Recognition | Occurs | Does Not Occur |
|------------------------------------|--------|----------------|
| Positive                           | 348    | 21             |
| Negative                           | 27     | 854            |

#### VI. CONCLUSIONS

In this paper we developed a novel real-time video surveillance system principally dedicated to fall detection. We claim that the proposed system is not just an ordinary

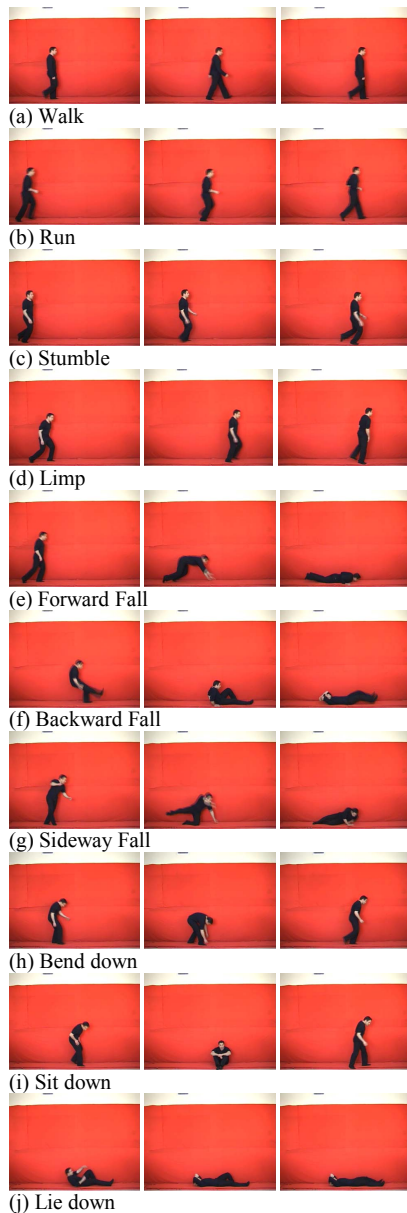


Fig.4. Example of Each Motion

human fall detection system; it has many applicable properties and can be employed in different surveillance systems. Moreover while existing fall detection systems are only able to detect occurrence of fall behavior, the proposed system is able to detect type of fall incident (forward, backward or sideway). Since temporal changes of the human shape give crucial information on human activities, we used combination of approximated ellipse around the human body, horizontal and vertical projection histograms and temporal changes of human head position as feature vectors. Also, our experiments indicate that MLP Neural Network is totally a suitable classifier for human motion recognition. Reliable average recognition

rate of experimental results (91.12%) underlines satisfactory performance and efficiency of our system.

## REFERENCES

- [1] N Noury, A Fleury, P Rumeau, A Bourke, G Laighin, V Rialle, J Lundy "Fall detection - Principles and Methods" Conf Proc IEEE Eng Med Biol Soc. 2007 ;1 :1663-1666
- [2] Arie Hans Nasution and S. Emmanuel; Intelligent Video Surveillance for Monitoring Elderly in Home Environments, International Workshop on Multimedia Signal Processing (MMSP), Greece, October 2007.
- [3] Sixsmith A, Johnson N. "Smart sensor to detect the falls of the elderly", IEEE Pervasive Computing, vol. 3, no. 2, pp. 42-47, April-June 2004.
- [4] Caroline Rougier, Jean Meunier, Alain St-Arnaud, Jacqueline Rousseau, "Fall Detection from Human Shape and Motion History Using Video Surveillance" Advanced Information Networking and Applications Workshops, 2007, 21st International Conference on, Vol. 2 (2007)
- [5] B. Toreyin, Y. Dedeoglu, and A. C. etin. HMM based falling person detection using both audio and video. In IEEE International Workshop on Human-Computer Interaction, Beijing, China, 2005.
- [6] S.-G.Miaou, P.-H. Sung, and C.-Y. Huang, "A Customized Human Fall Detection System Using Omni-Camera Images and Personal Information" Proc of Distributed Diagnosis and Home Healthcare(D2H2) Conference, 2006
- [7] R. Cucchiara, A.Pratti, and R.Vezani, "An Intelligent Surveillance System for Dangerous Situation Detection in home Environments" , in Intelligenza artificiale, vol.1, n.1, pp. 11-15, 2004
- [8] T. Lee and A. Mihailidis. An intelligent emergency response system: preliminary development and testing of automated fall detection. Journal of telemedicine and telecare, 11(4):194-198, 2005.
- [9] Nait CH, McKenna SJ. "Activity summarisation and fall detection in a supportive home environment", Int. Conf. on Pattern Recognition (ICPR), 2004.
- [10] Tao, J., M.Turjo, M.-F. Wong, M.Wang, "Fall Incidents Detection for Intelligent Video Surveillance" , ICICS2005
- [11] C.Rougier, J.Meunier, A. St-Arnaud, J.Rousseau, "Monocular 3D Head Tracking to Detect Falls of Elderly People", International Conference of IEEE Engineering in Medicine and Biology Society, Sept 2006
- [12] Lin,C.-W.,et al., "Compressed-Domain Fall Incident Detection for Intelligent Home Surveillance". Proceedings of IEEE International Symposium on Circuits and Systems,ISCAS 2005,2005:p.2781-3784.
- [13] L. Wang, From blob metrics to posture classification to activity profiling, Proceedings of the 18th International Conference on Pattern Recognition , IV: 736-739, 2006
- [14] A. Madabhushi and J.K. Aggarwal, "A Bayesian approach to human activity recognition," Proc. Second IEEE Workshop on Visual Surveillance, pp.25-32, June 1999.
- [15] M.Shakeri, H.Deldari, H.Foroughi, "A Novel Fuzzy Background Subtraction Method Based on Cellular Automata for Urban Traffic Applications", 9th International IEEE Conference on Signal Processing, ICSP'08

# An Improved Keyword Extraction Method Using Graph Based Random Walk Model

Md. Rafiqul Islam and Md. Rakibul Islam

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh

**Abstract** — Keywords can be considered as condensed versions of documents, which can play important role in some text processing tasks such as text indexing, summarization and categorization. However, there are many digital documents especially on the Internet that do not have a list of assigned keywords. Assigning keywords to these documents manually is a difficult task and requires appropriate knowledge of the topic. Automatic keyword extraction process can solve this problem. In this paper, we introduce a new improved method for keyword extraction using random walk model by considering position of terms within the document and information gain of terms corresponds to the whole set of documents. We also incorporate mutual information (*MI*) of terms with random walk model to extract keywords from documents. The experiments on standard test collections show that our method outperforms the previously proposed methods.

**Index Terms** — Keyword extraction, random walk model, mutual information, term position, information gain.

## I. INTRODUCTION

Keywords can be considered as brief summaries of a text. Therefore, it is possible to think of them as a set of phrases semantically covering most of the text. Although a summary of a text is capable of providing more information about the text than keywords of the text, the summary may not be suitable for some applications due to the complex structure of sentences. Keywords can not be replaced for summarization but these are alternative summary representations that could be consumed by other applications more easily.

Keywords enable readers to decide whether a document is relevant for them or not. They can also be used as low cost measures of similarity between documents. Unfortunately, a great portion of existing documents available today does not have keywords available for them [8]. Considering the fact that it is a hard and time consuming task to assign keywords to documents, it is desirable to automate this task by natural language processing (NLP) techniques.

The random walk ranking algorithm on a graph proposed by Brin and Page in [10], has been used in citation analysis, social networks and analysis of the link structure of the web [1]. A graph-based ranking algorithm is a way of deciding on the importance of a vertex within a graph, by taking into account global information, recursively computed from the entire graph, rather than

relying only on local vertex-specific information. The basic idea implemented by a random walk algorithm is that, when one vertex links to another one, it is basically casting a vote for other vertex [2]. The higher the number of votes that are cast for a vertex, the higher the importance of the vertex. On the other hand, mutual information is used to identify relatedness of two or more terms [7].

Term weighting is one of the most important considerations for keyword extraction. The purpose of a term weighting method is to classify the terms by assigning them weights corresponding to how well they improve both recall and precision. Mihalcea and Tarau propose a modified random walk model to weight terms for keyword extraction which provides better result than state-of-the-art systems developed in these areas [3]. They consider term dependencies with edge weights in their approach. However, they do not take into account terms importance within total corpus and terms positions in a single document to calculate the weight of terms. In our approach to weight a term, we exploit the relationship of local information of a vertex (term position) as well as global information (*MI*) and term dependency. Moreover, taking into account mutual information (*MI*) of co-occurrence terms, we introduce a new formula to weight terms for extracting keywords. Experimental results show the improvement of our graph based random walk-model, providing better result than the work of Mihalcea and Tarau [3] and Jinghua *et al.* [4].

The rest of this paper is organized as follows: In section II we describe the basic random walk model, related works and their problems are discussed in section III. Section IV presents our improved keyword extraction method. The experimental results are discussed in section V. Finally, we present conclusion in section VI.

## II. RANDOM WALK MODEL – TEXTTRANK

The basic idea implemented by a random-walk algorithm is that of *voting* or *recommendation* [2]. When one vertex links to another one, it is basically casting a vote for that other vertex. The higher the number of votes that are cast for a vertex, the higher the importance of the vertex [2]. Moreover, the importance of the vertex casting a vote determines how important the vote itself is, and this

information is also taken into account by the ranking algorithm. There are several random-walk algorithms that have been proposed in the past and between these algorithms, TextRank is previously found successful in a number of applications, including web link analysis, social networks, citation analysis, and more recently in several text processing applications [2].

Given a graph  $G = (V, E)$ , let  $In(V_a)$  be the set of vertices that point to vertex  $V_a$  (predecessors), and  $Out(V_a)$  be the set of vertices that vertex  $V_a$  points to (successors). The TextRank score associated with the vertex  $V_a$  is defined using the equation 1.

$$s(v_a) = (1 - d) + d \times \sum_{b \in In(v_a)} \frac{s(v_b)}{|Out(v_b)|} \quad (1)$$

Where  $s(v_a)$  = score of vertex  $v_a$ ,  $s(v_b)$  = score of vertex  $v_b$ ,  $In(v_a)$  = indegree of vertex  $v_a$ ,  $|Out(v_b)|$  = outdegree of vertex and  $d$  is a parameter that is set between 0 and 1 [2].

The score of each vertex is recalculated upon each iteration based on the new weights that the neighboring vertices have accumulated. Neighbors of a vertex are determined by window size  $N$ . If  $N$  is 3 then consecutive forward and backward three vertices are neighbor. The algorithm terminates when the convergence point is reached for all the vertices, meaning that the error rate for each vertex falls below a pre-defined threshold. This vertex scoring scheme is based on a random-walk model, where a walker takes random steps on the graph, with the walk being modeled as a Markov process [2]. Under certain conditions, the model is guaranteed to converge to a stationary distribution of probabilities associated with the vertices in the graph. Intuitively, the stationary probability associated with a vertex represents the probability of finding the walker at that vertex during the random-walk, and thus it represents the importance of the vertex within the graph [2].

### III. RELATED WORKS

The state-of-the-art in this area is currently represented by supervised learning methods, where a system is trained to recognize keywords in a text, based on lexical and syntactic features. This approach is first suggested in [5], where parameterized heuristic rules are combined with a genetic algorithm into a system for keyphrase extraction that automatically identifies keywords in a document [3].

In Frank *et al.* method where a Naive Bayes learning scheme is applied on the document collection [6], with

improved results observed on the same data set as used in [5].

In [3], it is stated that neither Turney nor Frank report on the recall of their systems, but only on precision: a 29.0% precision is achieved with GenEx for five keyphrases extracted per document, and 18.3% precision achieved with Kea for fifteen keyphrases per document.

In [8], Hulth applies a supervised learning approach that integrates part of speech information into the learning process to keyword extraction from abstracts, using a combination of lexical and syntactic features and proves the improvement significantly over previously proposed methods. Hulth states that, keyword extraction from abstracts is more widely applicable than from full texts, since many documents on the Internet are not available as full-texts, but only as abstracts [3]. Her experiment shows that the accuracy of the system is almost doubled by adding linguistic knowledge to the term representation [3].

In [4], a new algorithm is presented that represents the text as a semantic graph with synset from WordNet, disambiguates the words, and finally extracts keywords from the text based on UW-PageRank. It needs no corpus, moreover it has the ability to disambiguate all the words in the text, and extracts keywords by analyzing the semantic structure of the whole text. The experiment result shows that their algorithm is effective and practical.

The work described in [4], is a corpus independent method. Generally it is accepted that corpus independent method does not yield better method than corpus dependent method [12].

Mihalcea and Tarau introduce the TextRank graph-based ranking model for keyword and sentence extraction from natural language texts [3]. Here, a new formula is introduced for graph-based ranking that takes into account edge weights while computing the score associated with a vertex in the graph [3].

$$WS(v_i) = (1 - d) + d \times \sum_{v_j \in In(v_i)} \frac{w_{ji}}{\sum_{v_k \in Out(v_j)} w_{jk}} WS(v_j) \quad (2)$$

Where  $WS(v_i)$  is the weighting score of vertex  $v_i$ ,  $d$  denotes the damping factor and  $w_{ji}$  is the edge weight of the vertex  $v_i$  and  $v_j$ .

We describe another work of Hassan *et al.* [2] used for text classification, as our approach has similarity with this work. Hassan *et al.* introduce a system that models the weighting problem as a “random-walk” rather than “random choice” [2]. They assume an imaginary reader who steps through the text on a term by term basis. In this setting, the importance of the term is determined by the probability of the random-walker to encounter the target term in the text during the walk. In this paper, the various best random-walk based term weighting models are



TABLE I

RESULT COMPARISON OF OUR METHOD AND MIHALCEA AND TARAU METHOD.

| Method             | Assigned |       | Correct |      | Precision   | Recall | F-measure |
|--------------------|----------|-------|---------|------|-------------|--------|-----------|
| Mihalcea and Tarau | Total    | Mean  | Total   | Mean |             |        |           |
| Undirected, N = 2  | 6784     | 13.7  | 2116    | 4.2  | <b>31.2</b> | 43.1   | 36.2      |
| Undirected, N = 3  | 6715     | 13.4  | 1897    | 3.8  | 28.2        | 38.6   | 32.6      |
| Undirected, N = 5  | 6558     | 13.1  | 1851    | 3.7  | 28.2        | 37.7   | 32.2      |
| Undirected, N = 10 | 6570     | 13.1  | 1846    | 3.7  | 28.1        | 37.6   | 32.2      |
| <b>Our Method</b>  |          |       |         |      |             |        |           |
| Undirected, N = 2  | 6114     | 12.23 | 2386    | 4.77 | <b>39.1</b> | 48.7   | 43.4      |
| Undirected, N = 3  | 6129     | 12.26 | 2145    | 4.29 | 35.2        | 43.7   | 38.9      |
| Undirected, N = 5  | 5985     | 11.97 | 1735    | 3.47 | 29.5        | 35.4   | 23.1      |
| Undirected, N = 10 | 6105     | 12.21 | 1822    | 3.64 | 29.8        | 37.1   | 33.1      |

TABLE II

RESULT COMPARISON OF OUR METHOD AND JINGHUA *et al.* METHOD.

| Coverage                        |                  | Precision                       |                  |
|---------------------------------|------------------|---------------------------------|------------------|
| Jinghu <i>et al.</i> [4] method | Our Method N = 2 | Jinghu <i>et al.</i> [4] method | Our Method N = 2 |
| 0.69                            | 0.72             | 0.34                            | 0.39             |

summarized such as  $rw_e \cdot idf$ ,  $rw_{e.oc}$  etc. Here we describe the following model:

$rw_e \cdot idf$ : This model represents an undirected graph approach that uses the weighted edge version of PageRank with a variable damping factor. The weight of an edge is calculated by the equation 3.

$$E_{v_1, v_2} = tf \cdot idfv_1 \times tf \cdot idfv_2 \quad (3)$$

Where  $E_{v_1, v_2}$  is the edge connecting  $v_1$  to  $v_2$  and  $tf \cdot idfv_1$ ,  $tf \cdot idfv_2$  represent the term frequency multiplied by the inverse document frequency respect to vertices  $v_1$  and  $v_2$ . The damping factor is expressed as a function of the 'incoming edges' weight, calculated as follows:

$$d_{E_{v_1, v_2}} = E_{v_1, v_2} / E_{\max} \quad (4)$$

where  $d_{E_{v_1, v_2}}$  is the damping factor and  $E_{\max}$  represents the highest weight for an edge in the graph. Thus the score of a vertex can be calculated by the formula in [2] is as follows:

$$s'(v_1) = \frac{(1-d)}{|N|} + \sum_{v_2 \in \ln(v_1)} C \times \frac{d_{E_{v_1, v_2}} \times s(v_2)}{|Ou(v_2)|} \quad (5)$$

where  $s'(v_1)$  = score of vertex  $v_1$ ,  $N$  = total number of nodes,  $C$  = scaling constant [2].

Here the new measure of term weighting integrates both the locality of a term and its relation to the surrounding context. Though Hassan *et al.* use the work for text classification, we describe it here as it is related to our work.

#### IV. KEYWORD EXTRACTION METHOD USING GRAPH BASED RANDOM WALK MODEL

This section describes the method we use to extract keywords from documents. To do our work we use two phases - preprocessing phase and post processing phase, in our method. We determine the weight of each vertex in a graph in preprocessing phase. In post processing phase we construct multiword keywords (key phrases) if needed. Description of our method is given bellow:

1) To identify the text units (terms), the document is tokenized using text operations - stopword removals, stemming and each token is annotated with part of speech tags. We use the list of stop words enlisted in the Smart System [11] to remove stop word from graph. Considering only single words as candidates for addition to the graph, we avoid excessive growth of the graph size by adding all

possible combinations of sequences consisting of more than one lexical unit.

2) For each term in the processed documents, we calculate  $tf \times idf$  and random-walk weights ( $rw$ ). To calculate the  $tf \times idf$  of each vertex, we use the method described by Yue-Heng Sun *et al.* in [9], state that terms in the title should have higher importance than those in the abstract. Similarly the terms in the abstract should have higher importance than those in the text, and so on. Therefore in [9], computing method of  $tf_{ij}$  uses some factors to reflect the importance of term position, shown in the equation 6.

$$tf_{ij} = \alpha tf_{ij1} + \beta tf_{ij2} + \gamma tf_{ij3} \quad (6)$$

Where,  $tf_{ijk}$  is the frequency of a term in the  $k^{th}$  area and  $\alpha, \beta, \gamma$  are the factors that can be adjustable according to pre-experiments. Here  $\alpha > \beta > \gamma \geq 1$ .

Moreover, to give more importance to a term that has discriminating power to identify a document, Yue-Heng Sun *et al.* introduce *information gain* ( $IG_j$ ) for a term  $j$  [9]. We calculate  $tf \times idf$  according to the equation 7 given in [9].

$$tf_{ij} \times idf_j = \frac{tf_{ij} \times \log\left(\frac{N}{n_j} + 0.01\right) \times IG_j}{\sqrt{\sum_{i \in d} \left[tf_{ij} \times \log\left(\frac{N}{n_j} + 0.01\right) \times IG_j\right]^2}} \quad (7)$$

Where  $tf_{ij}$  = frequency of term  $j$  within document  $i$ ,  $idf_j$  = inverse document frequency of term  $j$ ,  $n_j$  = number of documents holding term  $j$ ,  $N$  = total number of documents in the corpus. The  $tf_{ij} \times idf_j$  value of each selected term reflects their importance within the document and corresponds to whole documents corpus.

After that, mutual information  $MI_{v_1, v_2}$  of an edge of vertex  $v_1$  and  $v_2$  is determined that express the relatedness of two co-occurrence terms as described in [7] Weight of an edge is calculated using equation 3. Weight of each edge is required to determine the damping factor  $d_{E_{v_1, v_2}}$  in equation 4. The value of damping factor  $d_{E_{v_1, v_2}}$  varies in equation 5 to reflect importance of term. The more the value of damping factor, the term is more important to retrieve the related information. With all of these, our equation will be as follows

$$s'(v_1) = \frac{(1-d)}{|N|} + \sum_{v_2 \in \ln(v_1)} C \times \frac{d_{E_{v_1, v_2}} \times MI_{v_1, v_2} \times s(v_2)}{|Out(v_2)|} \quad (8)$$

All the terms in the document are added as vertices in a graph to represent the document. Identifying relations that connect such terms, we draw edges between vertices in graph. All the terms that fall in the window size of a given term are considered dependent term. This is represented by a set of edges that connect the terms to all other terms in a fixed window size ( $N$ ). After constructing the graph we iterate the random walk algorithm until convergence.

3) Vertices are sorted on their final score. For ranking decision, we use the scores associated with each vertex. Once a final score is obtained for each vertex in the graph, vertices are sorted in reversed order of their score, and the top vertices in the ranking are retained for post-processing. While may be set to any fixed value, usually ranging from 5 to 15 keywords.

Our post-processing phase is similar to work of Mihalcea and Tarau described in [3]. Here all lexical units selected as potential keywords by the graph based ranking algorithm are marked in the text, and sequences of adjacent keywords are collapsed into a multi-word keyword. For instance, in the text *Integrating Logical Operators In Query Expansion*, if *Integrating*, *Logical* and *Operators* are selected as potential keywords by our algorithm, since they are adjacent, they are collapsed into one single keyword *Integrating Logical Operators*.

## V. EXPERIMENTAL RESULT AND DISCUSSION

The data set used in the experiments is a collection of 500 abstracts from the *Inspec* database, and the corresponding manually assigned keywords. This is the same test data set as used in the keyword extraction experiments reported in [3]. The *Inspec* abstracts are from journal papers from Computer Science and Information Technology. Each abstract comes with two sets of keywords assigned by professional indexers: controlled keywords, restricted to a given thesaurus, and uncontrolled keywords, freely assigned by the indexers [3]. We use the uncontrolled set of keywords to compare with the result in [3]. We divided each document in two sections - *title* and *abstract*. The terms in *title* have higher value than the terms in *abstract*.

The result of Mihalcea and Tarau in [3] evaluated using precision, recall, and F-measure is better than the result in [8]. On the other hand, recent corpus independent work of Jinghua *et al.* uses coverage and precision to evaluate their method and achieves better result than  $tf \cdot idf$  method



[4]. So we compare our method with the methods described in [3] and [4].

In [3], experiments are performed with various syntactic filters, including all open class words, nouns and adjectives, and nouns only, and the best performance are achieved with the filter that selects nouns and adjectives only. It is also experimented with a setting in [3], where no part of speech information is added to the text, and all words except a predefined list of stopword are added to the graph. We do our experiments in the similar way. For each method, the Table I list the total number of keywords assigned automatically by each method, the mean number of keywords per abstract, the total number of correct keywords and the mean number of correct keywords. The correctness of keywords is determined against the keywords assigned by professional indexers.

Table II shows comparison between our method and Jinghua *et al.* method. Here coverage is the result of the number of correct keyword divided by the number of keywords the indexers have assigned. Our method achieves better coverage and precision than Jinghua *et al.* method.

We can see from the Table I that we achieve highest precision 39.1 for window size 2 which is greater than the highest precision value 31.2 of Mihalcea and Tarau method. Data of Table II shows that, both coverage and precision values of our method are better than values of Jinghua *et al.* So we can say that our proposed method performs reasonably well than other previous methods.

## VI. CONCLUSION

Automatic keyword extraction is the task of automatically selecting a small set of terms describing the content of a single document. That a keyword is extracted means that it is present verbatim in the document to which it is assigned [8]. In this paper, we describe a new improved way for keyword extraction using random walk model by considering position of terms within the document and information gain of terms corresponds to the whole set of documents. We also incorporate mutual information (*MI*) of terms with random walk model to extract keywords from documents. Using *Inspec* benchmark we test our method to compare with the work of Jinghua *et al.* and Mihalcea and Tarau. The experimental results show that our approach achieves better result than their works. Moreover, our method is language knowledge resources or any manually constructed training data independent.

## REFERENCES

- [1] Roi Blanco, Christina Lioma "Random Walk Term Weighting for Information Retrieval" *SIGIR '07* July 23-27, 2007, Amsterdam, The Netherlands. ACM 978-1-595993-597-7/07/0007.
- [2] Samer Hassan, Rada Mihalcea and Carmen Banea "Random Walk Term Weighting for Improved Text Classification" *In Proceedings of TextGraps: 2nd Workshop on Graph Based Methods for Natural Language Processing. ACL*, 53-60, 2006.
- [3] Rada Mihalcea and Paul Tarau. "TextRank: Bringing Order into Texts" *In Proceedings of Empirical Methods in Natural Language Processing. ACL* 404-411, 2004.
- [4] Jinghua Wang, Jianyi Liu and Cong Wang. "Keyword Extraction Based on PageRank." *In Proceedings of PAKDD 2007*: 738-746.
- [5] P. Turney. "Learning to extract keyphrases from text". *Technical report, National Research Council, Institute for Information Technology*, 1999.
- [6] E. Frank, G. W. Paynter, I. H. Witten, C. Gutwin, and C. G. Nevill-Manning. "Domain-specific keyphrase extraction". *In Proceedings of the 16th International Joint Conference on Artificial Intelligence*, 1999.
- [7] Jian-Yun Nie, Fuman Jin " Integrating Logical Operators In Query Expansion In Vector Space Model." *Workshop on Mathematical/Formal Methods in nformation Retrieval, 25th ACM-SIGIR*, Tampere, Finland, August 2002.
- [8] A. Hulth. "Improved automatic keyword extraction given more linguistic knowledge". *In Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, Japan, August.
- [9] Yue-Heng Sun, Pi-Lian He, Zhi-Gang Chen, "An Improved Term Weighting Scheme For Vector Space Model" *Proceedings of the Third International Conference on Machine Learning & Cybernetics, Shanghai*, 26-29 August 2004.
- [10] S. Brin and L. Page "The anatomy of a large-scale hypertextual Web search engine" *Computer Networks and ISDN system*, 1998
- [11] Salton, G. and C. Buckley. "Term-weighting approaches in automatic text retrieval." *Information Processing and Management: an International Journal* 24(5): 513-523.
- [12] Jianyi LIU, Jinghua WANG, "Keyword Extraction Using Language Network." *In Natural Language Processing and Knowledge Engineering*, 2007.

# Performance Maximization for Question Classification by SubSet Tree Kernel using Support Vector Machines

Muhammad Arifur Rahman<sup>1</sup>, Vitalie Scurtu<sup>2</sup>

<sup>1</sup>Department of Physics, Jahangirnagar University, Dhaka, Bangladesh

<sup>2</sup>Master in Human Language Technology and Interfaces, University of Trento, Trento, Italy  
arifmail@gmail.com, scurtu.vitalie@studenti.unitn.it

**Abstract** — Question answering systems use information retrieval (IR) and information extraction (IE) methods to retrieve documents containing a valid answer. Question classification plays an important role in the question answer frame to reduce the gap between question and answer. This paper presents our research work on automatic question classification through machine learning approaches. We have experimented with machine learning algorithms Support Vector Machines (SVM) using kernel methods. An effective way to integrate syntactic structures for question classification in machine learning algorithms is the use of tree kernel (TK) functions. Here we use SubSet Tree kernel with Bag of words. Trade-off between training error and margin, Cost-factor and the decay factor has significant impact when we use SVM for the mentioned kernel type. The experiments determined the individual impact for Trade-off between training error and margin, Cost-factor and the decay factor and later the combined effect for Trade-off between training error and margin, Cost-factor. Depending on these result we also figure out some hyperplanes which can maximize the performance. Based on some standard data set outcomes of our experiment for question classification is promising.

**Index Terms** — Precision, Recall, SVM, kernel, SST, Question Classification, Question Answering.

## I. INTRODUCTION

The World Wide Web continues to grow at an amazing speed. So, there are also a quickly growing number of text and hypertext documents. Due to the huge size, high dynamics, and large diversity of the web, it has become a very challenging task to find the truly relevant content for some user or purpose. The Open-domain Question Answering system (QA) has been attached great attention for its capacity of providing compact and precise results for users.

The study of Question Classification (QC), as a new field, corresponds with the research of QA. At present the studies on QC are mainly based on the text classification. Though QC is similar to text classification in some aspects, they are clearly distinct in that: Question is usually shorter, and contains less lexicon-based information than text, which brings great trouble to QC. Therefore to obtain higher classifying accuracy, QC has to make further analysis of sentences, namely QC has to

extend interrogative sentence with syntactic and semantic knowledge, replacing or extending the vocabulary of the question with the semantic meaning of every word.

In QC, many systems apply machine-learning approaches [1-3]. The classification is made according to the lexical, syntactic features and parts of speech. Machine learning approach is of great adaptability, and 90.0% of classifying accuracy is obtained with SVM method and tree Kernel as features. However, there is still the problem that the classifying result is affected by the accuracy of syntactic analyzer, which need manually to determine the weights of different classifying features. Some other systems adopting manual-rule [4].

This paper presents our approaches at Question Classification to improve the f1-measure [6] of question categorization. Our experiment tried to determine the optimum value for trade-off between training error and margin, cost-factor by which training errors on positive examples outweighs errors on negative examples. It is used to adjust the rate between Precision and Recall on the development set [5-6] and the decay factor which has the role to penalize larger tree structures by giving them less weight in the overall summation [7]. To represent questions, the paper uses the standard Li and Roth [8] question classification dataset.

After this overview, question classifying method, a brief description about the used algorithm and the kernel methods are introduced, and then impact of different parameters and parameters combination methods has been investigated. The comparisons are testified in experiments based on precision, recall and f1-measure. The last part of the paper is about the conclusion of the present research and about the introduction of the further work could be done on this issue.

## II. QUESTION CLASSIFICATION

Question Classification means putting the questions into several semantic categories. Approaches to question classification can be divided in two broad classes, namely, rule-based and machine learning methods. Most recent studies have been based on machine learning approaches. Li and Roth [8] proposed 6 coarse classes and 50 fine classes for TREC factoid question answering. The UIUC

QC dataset, which they developed, contains 5,500 training questions and 500 test questions, and it is now the standard dataset for question classification [9].

In Table I Each coarse grained category contains a non-overlapping set of fine grained categories. Most question answering systems use a coarse grained category definition. Usually the number of question categories is less than 20. However, it is obvious that a fine grained category definition is more beneficial in locating and verifying the plausible answers.

TABLE 1: THE COARSE AND FINE GRAINED QUESTION CATEGORIES.

| Coarse | Fine                                                                                                                                                                                           |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ABBR   | Abbreviation, expansion                                                                                                                                                                        |
| DESC   | description, definition, manner, reason                                                                                                                                                        |
| ENTY   | body, color, event, food, creation, currency, animal, disease/medical, instrument, language, letter, other, plant, product, religion, sport, substance, symbol, technique, term, vehicle, word |
| HUM    | description, group, individual, title                                                                                                                                                          |
| LOC    | city, country, mountain, other, state                                                                                                                                                          |
| NUM    | code, count, date, distance, money, order, other, percent, period, speed, temperature, size, weight                                                                                            |

### III. SUPPORT VECTOR MACHINE (SVM)

Support vector machine are based on the Structural Risk Minimization principle from statistical learning theory [10]. The idea of structural risk minimization is to find a hypothesis  $h$  for which we can guarantee the lowest true error. The bounds are connected on the true error with the margin of separating hyperplanes. In their basic form support vector machines find the hyperplane that separates the training data with maximum margin. Since we will be dealing with very unbalanced numbers of positive and negative examples in the following, we use cost factors  $C_+$  and  $C_-$  to be able to adjust the cost of false positives vs. false negatives as [11]. Finding this hyperplane can be translated into the following optimization problem:

$$\text{Minimize: } \frac{1}{2} \|\vec{w}\|^2 + C_+ \sum_{i: y_i=1} \xi_i + C_- \sum_{i: y_i=-1} \xi_i$$

$$\text{Subjected to: } \forall k : y_k [\vec{w} \cdot \vec{x}_k + b \geq 1 - \xi_k]$$

Here  $x_i$  is the feature vector of example  $i$ .  $y_i$  equals +1 or -1, if example  $i$  is in class '+' or '-' respectively.

Using support vector machine we tried to determine the optimal value of trade-off between training error and

margin (c), cost-factor (j) by which training errors on positive examples outweighs errors on negative examples. It is used to adjust the rate between Precision and Recall on the development set [11]. Here Lambda ( $\lambda$ ) the decay factor which has the role to penalize larger tree structures by giving them less weight in the overall summation.

### IV. KERNEL METHODS

One of the most difficult tasks on applying machine learning for question classification is the feature design. Feature should represent data in a way that allows learning algorithm to separate positive from negative examples. In SVMs, features are used to build the vector representation of data examples and the scalar product between example pairs quantifiers how much they are similar. Instead of encoding data in the features vectors, kernel functions can be designed that provide such similarity between example pairs without using an explicit feature representation [6].

The kernel we considered in this paper represents trees in terms of their substructure. Such fragments define the feature space which, in turn, is mapped into a vector space. The kernel function measures the similarity between trees by counting the number of common fragments. These functions have to recognize if a common tree subpart belongs to the feature space that we intended to generate. Here we considered SubSet Trees (SSTs). A SubTree is defined as any node of a tree along with all its descendants. A SubSet Tree is more general structure which not necessarily includes all the descendants. The only restriction is that an SST must be generated by applying the same generated the original tree [12].

The main idea of the tree kernel is to compute the number of the common substructures between two trees  $T_1$  and  $T_2$  without explicitly considering the whole fragment space. From [6] given a tree fragment space  $\{f_1, f_2, \dots\} = F$ , here we denote the indicator function  $I_i(n)$  which is equal 1 if the target  $f_i$  is rooted at the node  $n$  and 0 otherwise. It followed that-

$$K(T_1, T_2) = \sum_{n_1 \in N_{T_1}} \sum_{n_2 \in N_{T_2}} \Delta(n_1, n_2) \dots (1)$$

where  $N_{T_1}$  and  $N_{T_2}$  are the sets of the  $T_1$ 's and  $T_2$ 's

nodes, respectively and  $\Delta(n_1, n_2) = \sum_{i=1}^F I_i(n_1) I_i(n_2)$ .

This letter is equal to the number of common fragments rooted at the  $n_1$  and  $n_2$  nodes. We can compute  $\Delta$  as follows-

1. if the production at  $n_1$  and  $n_2$  are different the  $\Delta(n_1, n_2) = 0$ ;

2. if the production at  $n_1$  and  $n_2$  are the same, and  $n_1$  and  $n_2$  have only leaf children (pre-terminal symbols) then  $\Delta(n_1, n_2) = 1$  ;
3. if the production at  $n_1$  and  $n_2$  are the same, and  $n_1$  and  $n_2$  are not pre-terminals then

$$\Delta(n_1, n_2) = \prod_{j=1}^{nc(n_1)} (\sigma + \Delta(c_{n_1}^j, c_{n_2}^j))$$

where  $\sigma \in \{0,1\}$ ,  $nc(n_1)$  is the number of the children of  $n_1$  and  $c_n^j$  is the  $j$ -th child of the node  $n$ . As the production is same  $nc(n_1) = nc(n_2)$ .

When  $\sigma = 0$ ,  $\Delta(n_1, n_2)$  is equal 1 only if  $\forall j \Delta(c_{n_1}^j, c_{n_2}^j) = 1$ , i.e. all the productions associated with the children are identical. By recursively applying the property, it follows the SubTrees in  $n_1$  and  $n_2$  are identical. Thus Eq. 1 evaluate the SubSet Tree kernel .When  $\sigma = 1$ ,  $\Delta(n_1, n_2)$  evaluates the number of SSTs common to  $n_1$  and  $n_2$  as proved in [12].

To include the leaves as fragments it is enough to add, to the recursive rule set for the  $\Delta$  evaluation, the condition is- if  $n_1$  and  $n_2$  are leaves and their associated symbols are equal then  $\Delta(n_1, n_2) = 1$ . By [6] such extended kernel can be referred as SST+bow (*bag-of-words*). Here the decay factor ( $\lambda$ ) is defined as  $\Delta(n_x, n_z) = \lambda$

$$\text{and } \Delta(n_x, n_z) = \lambda \prod_{j=1}^{nc(n_x)} (\sigma + \Delta(c_{n_1}^j, c_{n_2}^j)).$$

## V. RESULTS AND ANALYSIS

We evaluate the performance of Question Classification using SVM with SubSet Tree kernel with Bag of words (SST+BOW). First we investigate the optimum value of trade-off between training error and margin ( $c$ ) keeping cost-factor ( $j$ ) and the decay factor ( $\lambda$ ) constant. Then we investigate the impact of cost factor ( $j$ ) for constant trade-off between training error and margin ( $c$ ) and decay factor ( $\lambda$ ). Then for constant factor ( $\lambda$ ) we determine the hyper plane for trade-off between training error and margin ( $c$ ) and cost-factor ( $j$ ) which have the maximum f1-measure. Finally we determined the impact of decay factor ( $\lambda$ ) for a constant trade-off between training error and margin ( $c$ ) and cost-factor ( $j$ ).

Fig. 1 (a-d). shows the Trade of between Training error and Margin ( $c$ ) versus performance (measured by f1-measure) curve for different types of QC data set using different Cost Factor ( $j$ ) and fixed Decay factor ( $\lambda=0.01$ ). Here it is notable for same set of parameters different data

types have different performance level. For different cost factor highest f1-measure can be achieved by setting the trade-off between training error and margin ( $c$ ) around 2, Afterwards its performance becomes almost constant.

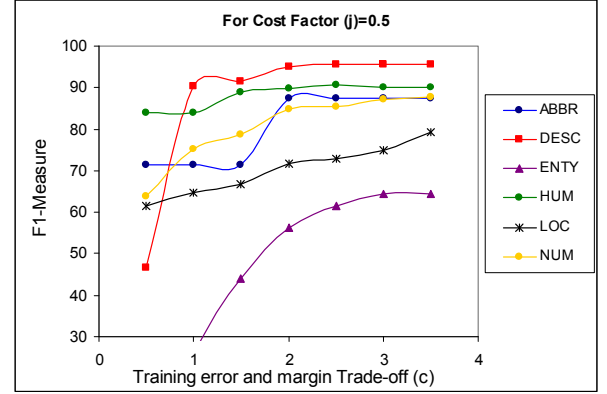


Fig. 1 (a). Training error and Margin Vs F1 Measure for  $j = 0.5$ .

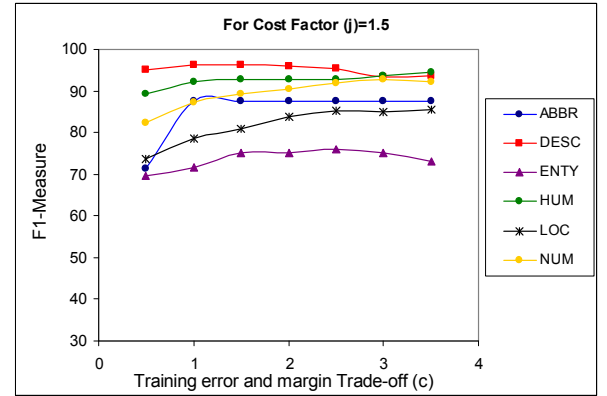


Fig. 1 (b). Training error and Margin Vs F1 Measure for  $j = 1.5$ .

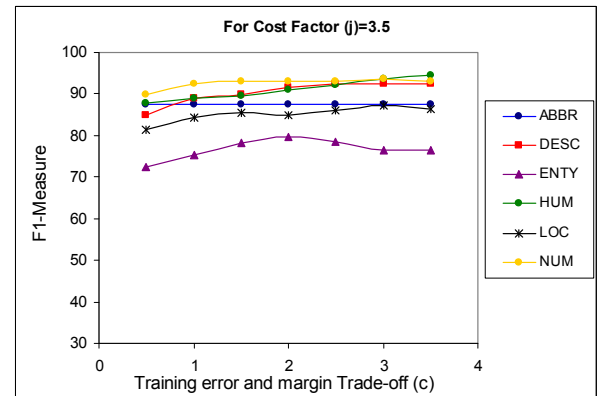


Fig. 1 (c). Training error and Margin Vs F1 Measure for  $j = 3.5$ .

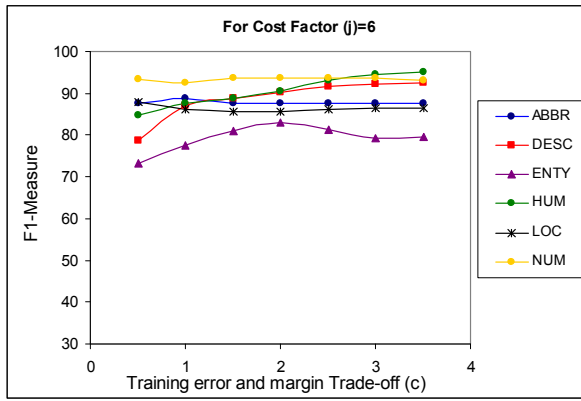


Fig. 1 (d). Training error and Margin Vs F1 Measure for  $j = 6.0$ .

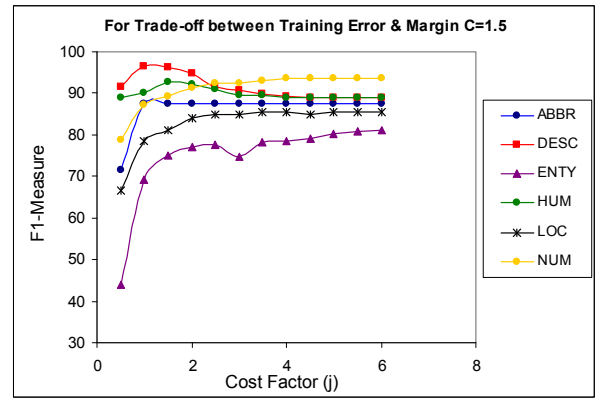


Fig. 2 (c). Cost factor Vs F1 measure for  $C=1.5$ .

Fig. 2. (a-d) shows the cost factor ( $j$ ) versus performance (measured by f1-measure) curve for different types of question classification data set using different of Trade of between Training error and Margin ( $c$ ) and fixed Decay factor ( $\lambda=0.01$ ).

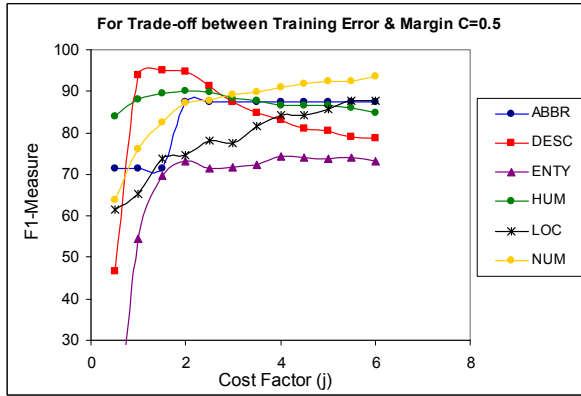


Fig. 2. (a) Cost factor Vs F1 measure for  $C=0.5$ .

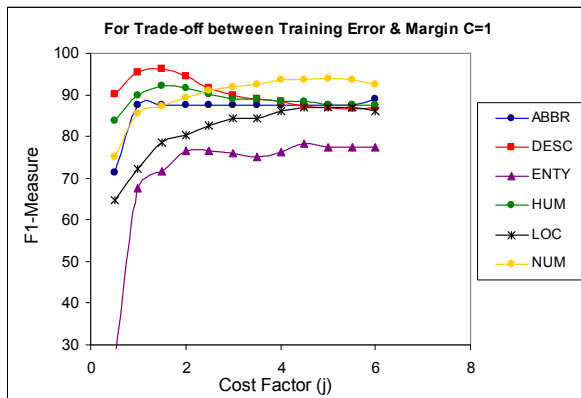


Fig. 2 (b). Cost factor Vs F1 measure for  $C=1$ .

For a constant trade-off between training error and margin different types of questions have highest f1-measure. For different value of  $c$  highest f1-measure can be achieved by setting the cost factor ( $j$ ) less than 2.

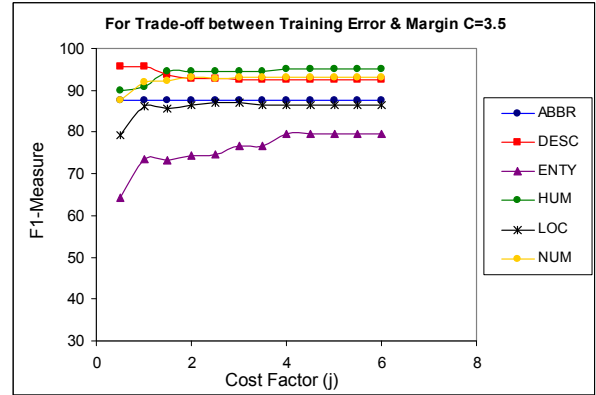


Fig. 2 (d). Cost factor Vs F1 measure for  $C=3.5$ .

If the value of cost factor is increased then the f1-measure is decreased a certain amount and after a stage it becomes almost constant.

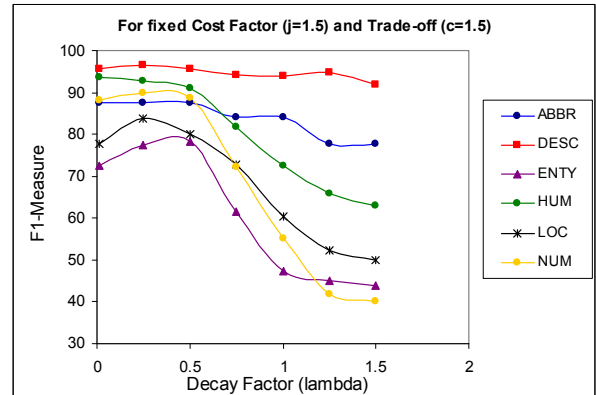


Fig. 3. Impact of Decay Factor.

Fig. 3. shows the Decay factor ( $\lambda$ ) versus f1-measure curve for a constant values of Trade off between Training error and Margin ( $c=1.5$ ) and cost factor ( $j=1.5$ ). To get a very high f1 measure decay factor should have very small. Starting after 0, up to 0.5 its performance is almost steady but if we increase the decay factor more the performance decreased drastically.

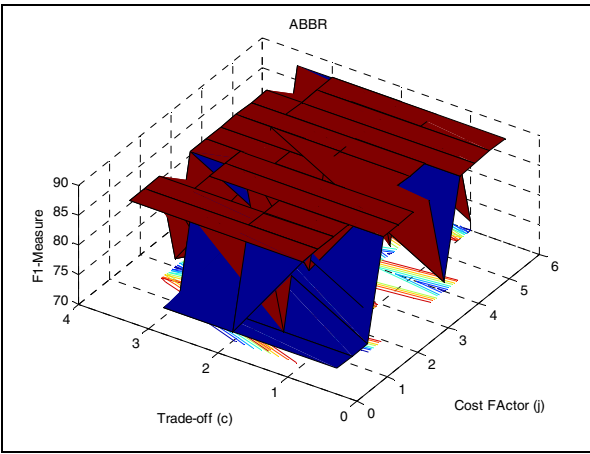


Fig. 4(a). Hyperplane for ABBR data type.

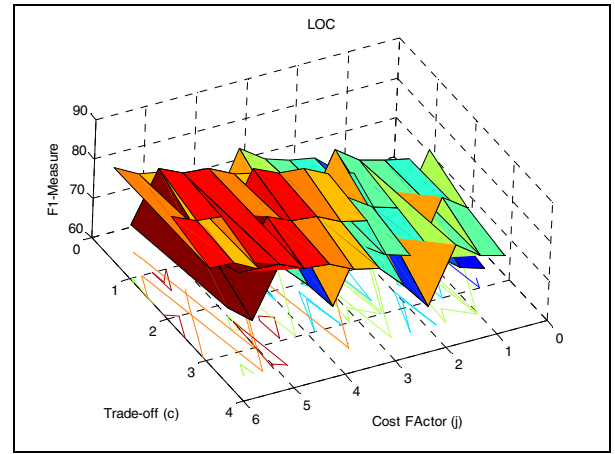


Fig. 4(d). Hyperplane for LOC data type.

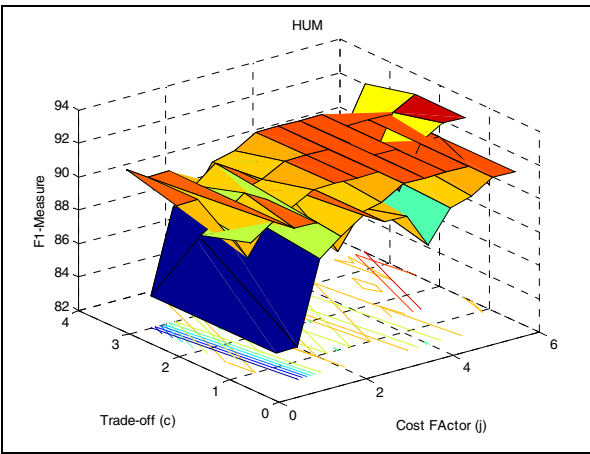


Fig. 4(b). Hyperplane for HUM data type.

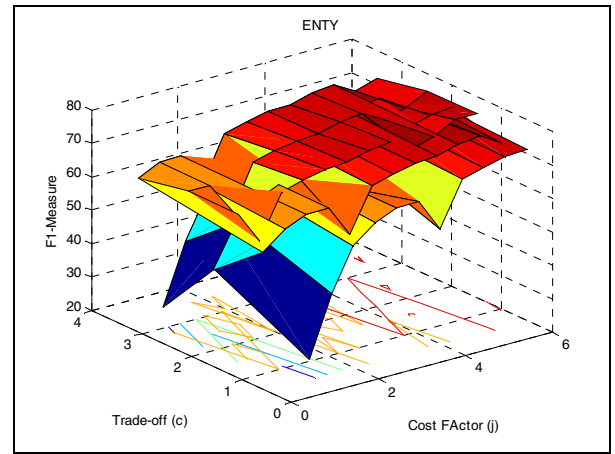


Fig. 4(e). Hyperplane for ENTJ data type.

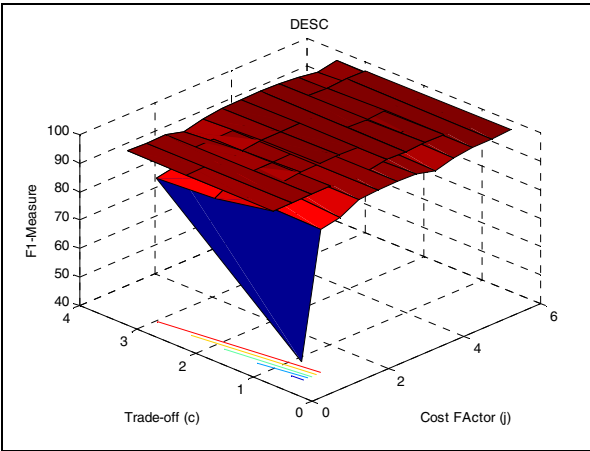


Fig. 4(c). Hyperplane for DESC data type.

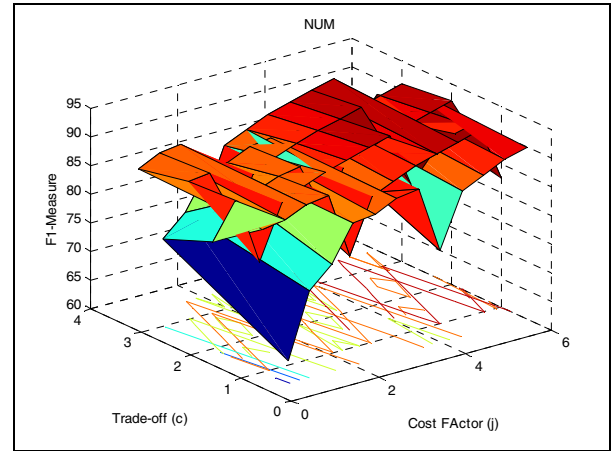


Fig. 4(f). Hyperplane for NUM data type.

Fig. 4(a-f). Shows the hyper plane by which maximum f1-measure can be achieved. For a constant Decay factor ( $\lambda$ ) a set of values of Trade of between Training error and margin (c) and cost factor (j) can provide the highest f1-measure.

Among the six type of data set 'LOC' type of question have to classified by some fixed combination of Trade of between Training error and Margin (c) and cost factor (j) because there is no hyper plane but some peaks.

TABLE 2: MAXIMUM F1- MEASURE OBTAINED USING SUBSET TREE KERNEL-BAG OF WORDS (SST+BOW) FOR  $\lambda=0.01$ .

| Coarse | F1    | P(%) | R(%)  | A(%)  | j   | c   |
|--------|-------|------|-------|-------|-----|-----|
| ABBR   | 88.89 | 99.6 | 88.89 | 88.89 | 6   | 1   |
| DESC   | 96.77 | 98.2 | 95.74 | 97.83 | 1   | 2.5 |
| ENTY   | 82.87 | 93.8 | 86.21 | 79.79 | 6   | 2   |
| HUM    | 95.24 | 98.8 | 98.36 | 92.31 | 4   | 3.5 |
| LOC    | 87.8  | 96   | 86.75 | 88.89 | 5.5 | 0.5 |
| NUM    | 94.01 | 97.4 | 98.08 | 90.27 | 5   | 1   |

Table 2 shows the maximum F1- measure (F1) obtained using SubSet tree Kernel-Bag of words (SST+BOW) for  $\lambda=0.01$  and the corresponding precision (P), recall (R), accuracy (A) for specific cost factor (j) and trade-off between training error and margin (c).

## VI. CONCLUSION

Question classification is one importance role in the QA frame to reduce the gap between question and answer. It can conduct answer choosing and selection. Our question classification method based on the use of linguistic knowledge and machine learning approaches and it exploit different classification features and combination method, also. Though among all the experiments one or two data set didn't provide distinguishable hyperplane in every cases thereafter the outcome of experiments done using the tool SVM\_light on Li and Roth question classification data sets demonstrate some optimal set of values. In the future we aim to investigate the impact of different parameters on constituent trees using different types of kernel.

## REFERENCES

- [1]. Hovy, E., Hermjakob, U., Lin, C.-Y. and Ravichandran, D. "Using Knowledge to Facilitate Factoid Answer Pinpointing", *Proceedings of the 19th International Conference on Computational Linguistics (COLING)*, Taipei, Taiwan, 2002.
- [2]. Ittycheriah, A., Franz, M., Zhu, W.-J. and Ratnaparkhi, A. "IBM's Statistical Question Answering System", *Proceedings of the TREC-9 Conference*, Gaithersburg, MD: NIST, 2000, p. 229.
- [3]. Zhang, D. and Lee, W. S. 2003. "Question Classification using Support Vector Machines", *Proceedings of ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Toronto, Canada, 2003
- [4]. LI Xin, HUANG Xuan-Jing, WU Li-de, "Question Classification by Ensemble Learning", *IJCSNS International Journal of Computer Science and Network Security*, VOL.6 No3, March 2006
- [5]. Alessandro Moschitti, Silvia Quarteroni, Roberto Basili and Suresh Manandhar, "Exploiting Syntactic and Shallow Semantic Kernels for Question/Answer Classification", *Proceedings of the 45th Conference of the Association for Computational Linguistics (ACL)*, Prague, June 2007.
- [6]. Roberto Basili, Alessandro Moschitti, "Automatic Text Categorization From information Retrieval to Support Vector Learning", ARACNE editrice, November 2005.
- [7]. Stephan Bloehdorn and Alessandro Moschitti, "Exploiting Structure and Semantics for Expressive Text Kernels", *Proceeding of the Conference on Information Knowledge and Management*, Lisbon, Portugal, 2007.
- [8]. Li, X., Roth, D., "Learning question classifiers", *Proceedings of the 19th International Conference on Computational Linguistics (COLING)*, Taipei, Taiwan, 2002.
- [9]. M.-Y. Day, C.-H. Lu, C.-S. Ong, S.-H. Wu, and W.-L. Hsu, "Integrating Genetic Algorithms with Conditional Random Fields to Enhance Question Informer Prediction", *Proceedings of the IEEE International Conference on Information Reuse and Integration (IEEE IRI 2006)*, Waikoloa, Hawaii, USA, 2006, pp. 414-419.
- [10]. V. Vapnik, "Statistical Learning Theory", Wiley, New York, USA 1998.
- [11]. K. Morik, P. Brockhausen, and T. Joachims, "Combining statistical learning with a knowledge-based approach – A case study in intensive care monitoring", *International Conference on Machine Learning (ICML)*, Bled, Slovenia, 1999.
- [12]. M. Collins, N. Duffy, "New Ranking algorithm for parsing and Tagging: Kernels over Distance structure, and the Voted Perception", *Association for Computational Linguistics (ACL)*, Philadelphia, USA, 2002.

# On the Design of an Effective Corpus for Evaluation of Bengali Text Compression Schemes

Md. Rafiqul Islam, and S. A. Ahsan Rajon

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh.

**Abstract** — In this paper, we propose an effective platform for evaluation of Bengali Text Compression Schemes. We perform a methodical study on the formulation-approaches of text corpus for data compression and present an effective corpus named *Ekushe-Khul* for evaluating the Bengali text compression schemes, which is the first initiative in the context of Bengali Text Compression. To design the Bengali text compression corpus, we consider Type to Token Ratio as the selection criteria with a number of secondary considerations. This paper also presents a mathematical analysis on data compression performance with structural aspects of corpora. The proposed corpus is effective for evaluating compression efficiency of small and middle sized text files.

**Index Terms** — Corpus, Bengali Text, Bengali Text Compression, Dictionary Coding, Data Management, Evaluation Platform, Compression Efficiency, Type to Token Ratio (TTR).

## I. INTRODUCTION

With a number of languages in the world, Bengali is the only language, which has been established at a cost of lives. However, in the establishment of Bengali as a glittering candidate in the field of research the number of steps or achievements is not so mention-worthy. Enhancing Data Management techniques for Bengali text has not yet got any robust base. Text Compression, which is one of the important aspects of elementary data management and text manipulation schemes, is still now mostly based on generalized universal compression techniques. The remaining sophisticated Bengali text compression schemes also suffer from unavailability of standard evaluation platform, *i.e.* unavailability of text compression corpus. This paper proposes a text compression corpus for evaluating Bengali text compression.

The construction of a data compression corpus is now a demand-of-time. The state-of-the-earth technologies of data management have already equipped the benchmarks and standard collection of experimental data sets for performance evaluation. In Bengali, though a highly impressive step has been initiated in [1] for constructing a linguistic corpus having CIIL corpus [2] as a pioneer, still now any mentionable sophisticated data compression

corpus is not available with text compression benchmarks. This paper focuses on designing a Bengali text compression corpus based on *Type to Token Ratio (TTR)* and provides the mathematical framework for various aspects of choosing the corpus with compression benchmarks.

The paper is organized into eight sections. Section II provides elementary concepts on corpus and text compression schemes. The key-points for constructing a corpus is presented in section III. Section IV is devoted with the literature review presenting various aspects of the existing corpora (Corpuses). Section V describes the proposed approach for formulating the corpus. A brief description on the files comprising the proposed text compression corpus is presented in section VI. Analysis of the proposed corpus depicted in section VII. The paper ends with section VIII providing recommendations and directions for future work.

## II. OVERVIEW OF CORPUS AND DATA COMPRESSION

Text Compression is an elementary aspect of data compression. Compression is the process of reducing the size of a file or data by identifying and removing redundancy in its structure [3],[4]. Data Compression offers an effective approach of reducing communication costs by using available bandwidth effectively. Data Compression technique is generally divided into two categories; namely, Lossless Data Compression and Lossy Data Compression. For Lossless schemes, the recovery of data should be exact. Lossless compression algorithms are to some extent essential for all kinds of text processing, scientific and statistical databases applications, medical and biological image processing, DNA and other biological data management and so on. However, a lossy data compression technique does not ensure the exact recovery of data. For image compression and multimedia data compression, there is a great use of lossy data compression. From the early 1990, certain specific platforms of evaluating English text compression are adapted for standard researches, which are better known as data compression corpus.



In general sense, a corpus is simply a collection of texts. From the point of view of data compression, compression-corpus is a standard collection of texts or data for analyzing and evaluating effectiveness and efficiency *i.e.* performance of compression schemes. The mostly used data compression corpora for evaluating data compression are namely Calgary Corpus, Canterbury Corpus and Project Gutenberg. Though the corpora for English were introduced in the early 1990, still now, for analyzing and evaluating Bengali text compression performance, no such complete and standard corpus is available. In this paper, we present a new Corpus for evaluation of Bengali text compression approaches.

### III. CHARACTERISTICS OF A DATA COMPRESSION CORPUS

The characteristics of a corpus is defined by the purpose of creating the corpus and the evaluation arena for which the corpus is designed for. This criteria give rise to the necessity to field-specific corpora like data compression corpus, information retrieval corpus, corpus for data mining research etc. According to Arnold *et al.* [5], there are six criteria for choosing a corpus.

The mentioned criteria [5] for choosing corpus have successfully established *Canterbury Corpus* as one of the extensively adapted corpus for data compression. However, due to the rapid growth of technology, additional criteria have been evolved in constructing the corpus. Firstly, the criteria were specified for evaluating the compression of middle-sized files. But, with the advancement of embedded devices with lower processing devices and small memory (like mobile-phones) it has become extremely important to provide a new test-bed for evaluation of small text compression schemes. Frequent uses of mobile phones, short text messages and emails have already made a transformation on the use of texts, and hence necessity of field and size specific corpus formation has been evolved. Taking these limitations into concern, we provide specific files regarding the issues with consideration of other traditional characteristics. Moreover, the file structure *i.e.* sentence structure was not considered as any important criteria for choosing corpora. But, compression ratio gradually varies with variations of sentence construction and file structure. Besides these, probability of occurrences of texts-components and frequency of those are important criteria, which fluctuates the text compression performance.

### IV. LITERATURE REVIEW

The history of mostly adapted data compression corpus started a long ago. *Calgary Corpus*, which was collected in 1987 and was published in 1990 is considered as the

pioneer of Data Compression Corpus. The next initiative and the mostly adapted one is *Canterbury Corpus*. It was developed by Arnold *et al.* [5]

The files are solely English text compression evaluation corpus and hence not effective or at all applicable for evaluating Bengali text compression. The most recent and extremely organized analysis on Bengali News Corpus is provided by Majumder *et al.* [1]. This paper proposes a new Corpus on the basis of the well circulated newspaper *prothom-alo* [6] and provides exhaustive analysis regarding *TTR* (Type-to-Token-Ratio), Function Word and other morphological and linguistic analysis. For analysis of data compression efficiency, this corpus requires some specific modification in the sense that, data compression involves the additional criteria of heterogeneity and text-size. Moreover, the news corpus proposed in [1] is extensively tested for linguistic analysis only. In order to use this for data compression, further text-compression related analysis is a must. According to their methodology, they collect html news files from *prothom-alo* websites and converting it to Unicode, they categorize the total collection into twenty seven distinct groups, which are also available in a single file of size three hundred and eighteen mega bytes. For most practical applications and researches on middle-sized or small text compression, it is rarely necessary to have larger files, rather it is essential to have standard small and medium sized corpus.

The first corpus in Bengali developed by Central Institute of Indian Languages (CIIL) in the years 1991-1995 possess a collection of three million Bengali words, which provides valuable linguistic data for research on Bengali language analysis [2]. In evaluating compression efficiency, this impressive corpus also suffers from the limitation of heterogeneity and lack of embedded and small device supportability.

### V. PROPOSED CORPUS FOR EVALUATION OF BENGALI TEXT COMPRESSION SCHEMES

We propose a new Corpus for evaluation of Bengali text Compression Schemes namely *Ekushe-Khul Corpus*. This corpus is intended for evaluating only short and middle-sized text compression schemes rather than evaluation for large text compression. One of the worth-mentionable points is, this corpus is selected by applying a new approach, which takes both *TTR* and *Compression Ratio (CR)* into account. We have adapted the two concepts because; we do not have any standard and sophisticated Bengali text compression corpus available with which we may incorporate and compare the performance and thus perform the selection from candidate files as done for other existing data compression corpora.

To form the corpus we have considered ten groups- Articles, Poems, Advertisements, Speeches, News, SMS, Emails, Particulars, Stories and Reports. The files are of sized 4KB to 1800KB. The files of the proposed Ekushe-Khul Corpus is shown in Table I. The steps of constructing the proposed corpus is depicted in Fig 1.

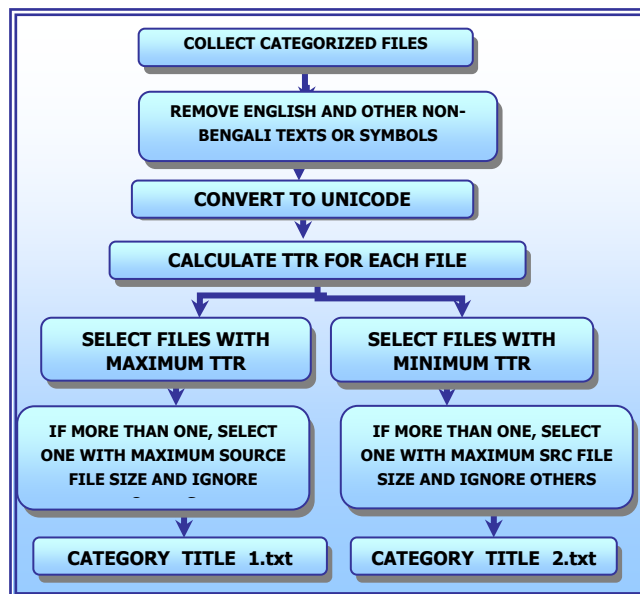


Fig. 1: Steps for constructing the Proposed Bengali text compression corpus.

TABLE I

FILES CONSTITUTING THE PROPOSED *EKUSHEY-KHUL* CORPUS.

| File Name        |                  | Type                         |
|------------------|------------------|------------------------------|
| Article1.txt     | Article2.txt     | <i>Articles in Bengali</i>   |
| Poem1.txt        | Poem2.txt        | <i>Bengali Classic Poems</i> |
| Advertise1.txt   | Advertise2.txt   | <i>Advertisements</i>        |
| Speech1.txt      | Speech2.txt      | <i>Various Speeches</i>      |
| News1.txt        | News2.txt        | <i>News from Newspaper</i>   |
| SMS1.txt         | SMS2.txt         | <i>Short Bengali Message</i> |
| Email1.txt       | Email2.txt       | <i>Emails in Bengali</i>     |
| Particulars1.txt | Particulars2.txt | <i>Particulars</i>           |
| Story1.txt       | Story2.txt       | <i>Bengali Stories</i>       |
| Report1.txt      | Report2.txt      | <i>Academic Reports</i>      |

Collecting the files, the first step was to remove redundancies and other components (e.g. html tags, letters and words from other languages, formatting etc.) and convert them into Unicode as several documents were non-standard encoded and saved as Unicode text files. These files are then forwarded for further analysis.

For each category of files, all the files are passed to a *TTR Calculator* and then *Compression Ratio Evaluator* for selection of final documents. In case of *Canterbury Corpus*, which is the pioneer for data compression corpus was formed by comparing each file with the standard of

the *Calgary Corpus* of same group or category, and the file demonstrating the closest match in terms of compression performance was selected. But, as no such standard exists for Bengali, we are to take account of both the criteria and hence selected two with the maximum and minimum consistence.

*TTR* or *Type to Token Ratio* is a measurement of how many times a previously encountered words repeat themselves before a new word makes its appearance in the text [1]. It is note worthy that though this measure is essential for language engineering, its application in data compression is not mandatory. For calculative approaches, *TTR* is calculated by dividing the total number of word tokens by the total number of distinct words. Texts with a large number of distinct words results into lower *TTR*. For this case, the task of forming a data compression dictionary appears to be more accurate. Conversely, for texts with greater *TTR* i.e. with greater number of frequent words (resulting lower number of distinct words) provides a suitable arena for constructing and evaluating dictionary-based and repetition-analysis oriented text compression schemes. This is the reason, which inspires us to provide peer-files for each category. The files with suffix '2' indicates the file with lower *TTR* (suitable for text compression fluctuation evaluation) and files with suffix '1' indicates greater *TTR* (suitable for dictionary-based text compression scheme evaluation) for the respective group.

The next step involves analyzing the files in terms of compression ratio [7] to establish logical and implementational validity of choosing *TTR* as a benchmark for constructing the corpus. Though a number of approaches exist for compression of English texts, (like Burrows-wheeler Transform [2], Arithmetic Coding, Huffman Coding etc.) it is a sad tale that, still now a little research has been dedicated for Bengali text compression schemes. The uses of non-sophisticated text compression schemes do not provide optimal performance for Bengali text compression purposes. For providing a benchmark of the proposed corpus, we adapt the dictionary based compression scheme in [3] and demonstrate the inter-relation of *TTR* with *Compression Ratio*.

## VI. ASPECTS OF CORPUS SELECTION CRITERIA AND TEXT COMPRESSION PERFORMANCE

Let  $n$  be the total number categories of collections of texts. The categories are representative of texts bearing distinct themes with differences in sentence construction, sentence wording and sentence structure.

Let the categories are  $c_1, c_2, c_3, \dots, c_n$  respectively. For category  $c_1$  the files are

$f_{1,1}, f_{1,2}, f_{1,3}, \dots, f_{1,n}$  with length  $l_{1,1}, l_{1,2}, l_{1,3}, \dots, l_{1,n}$  where, the length indicates total number of characters (including blank spaces, tabs, new lines and other punctuation marks).

If there are total of  $w_{1,1}, w_{1,2}, w_{1,3}, \dots, w_{1,n}$   $p$  words in  $f_{1,1}, f_{1,2}, f_{1,3}, \dots, f_{1,n}$  respectively with number of distinct words be  $d_{1,1}, d_{1,2}, d_{1,3}, \dots, d_{1,n}$  for the same.

Here, for any file  $f_{i,j}$ , where  $i$  denotes *Category Index* and  $j$  denotes *File Index*, and  $w_{i,j} \geq d_{i,j}$ . Assume  $T_{i,j}$  be the Type to Token Ratio (*TTR*) for file  $f_{i,j}$ .

That is,  $T_{i,j} = w_{i,j}/d_{i,j}$ .

For any two files  $p, q$  in category  $i$ , with *TTR*  $T_{i,p}$  and  $T_{i,q}$ , if  $T_{i,p} \geq T_{i,q}$  indicates that,  $T_{i,p}$  contains larger number of distinct words than that of  $T_{i,q}$ .

Applying this rule, we count the *TTR* for the files. Then the file with maximum and minimum *TTR*s are selected. For the file with maximum *TTR*, if the elements comprise the set  $S_d$  and  $S_t$  for distinct words and total words respectively, then for ideal case, set  $S_d - S_t$  would tend towards  $\Phi$  (empty or null) set.

Though *Type to Token Ratio* has been considered as a means of forming linguistic Corpora and Natural Language Processing Corpora including Information Retrieval Corpora, taking it into concern, it is possible to design Data Compression Corpus too. Here, we provide an analysis on the effectiveness of using *TTR* as a criteria of constructing Corpora.

The basic relation between *TTR* and word-distribution is:

- If *TTR* is larger (greater) then the number (frequency) of matching word is also greater; consequently, number of distinct words is smaller.
- If *TTR* is smaller (lower) then, the number (frequency) of matching word is smaller; consequently, number of distinct word is larger.

For designing dictionary based data compression approaches, *TTR* plays a great role. For constructing the dictionary, the prime concern is to analyze the probability of occurrence of each character or syllable or substring or words (or any combination of the mentioned units). The frequencies in which they occur are an important aspect of figuring the probability. As the total number of distinct words are greater in files with lower *TTR* value, the detection of the probability of syllables may be easier as the word distribution is approximately flat in those files. For building word-based dictionaries for data compression, files with higher *TTR* values are more effective as the recurring words may be very easily pointed out. The fluctuation may be easily identified by comparing the performance of the compression scheme using the peer files.

Let us consider a tertiary dictionary based text compression approach *DCA*, which uses a dictionary of  $n$  items. The total  $n$  items consist of  $c$  characters,  $s$  syllables and  $w$  substrings, which save (in average)  $sc$ ,  $ss$ , and  $sw$  bits each. Traditionally,  $sw \geq ss \geq sc$  as words or substrings are generally considered as a collection of syllable and/or symbols and syllables are considered as a collection of characters, but the indices or pointers in most designs require consecutive values spanning against a threshold value.

For a file with greater *TTR* (i.e. Greater number of frequent words and lower number of distinct words), the probability of encoding using word-based coding is high. Consequently, the bit-savings will also be high as the saving weightage is greater for words in compassion with other units. Let the replacement of number of distinct words be  $rw$ . Here,  $rw$  theoretically tends towards the total occurrence in average.

Therefore, the total savings in this phase is  $S_1 = \sum_{y=1}^{rw} (sw_y \times fw_y)$  where  $fw$  is the frequency of corresponding distinct word. That is,  $fw$  indicates the frequency of occurrence of distinct word-indexed at  $y$  and ideally,  $fw \rightarrow tw$ .

The number of distinct syllable  $rs$  in a collection of highly distinct words should be lower. If the weightage is considered as double as distinct words, the total savings would be  $S_2 = \sum_{y=1}^{rs} (ss_y \times fs_y)$  where  $fs$  is the frequency of the syllable.

The rest of the characters should be encoded separately resulting a savings  $S_3 = \sum_{y=1}^{rc} (sc_y \times fc_y)$

If we encode a total of  $p$  characters separately, which saves  $sp$  bits for each characters and if we encode  $q$  words separately with an average of  $d$  characters per word, whenever even  $d = p$ , the savings will be much more greater for word based bit-savings, since the indexing of the constructed code base will be approximately sequential and hence, each single character and each single word based encoding will occupy about same length. Thus, the total saving is  $S = S_1 + S_2 + S_3$

$$S = \sum_{y=1}^{rw} (sw_y \times fw_y) + \sum_{y=1}^{rs} (ss_y \times fs_y) + \sum_{y=1}^{rc} (sc_y \times fc_y) \quad (1)$$

For a file with lower *TTR*, (i.e. greater number of distinct words and lower number of frequent words) there will have non-frequent occurrence of matching words and syllables. This will result extensive application and usage of character based encoding, as the primary unit of encoding is character. That is,  $rw' < rw$  and  $fw' \rightarrow 1$ .

Similarly,  $rs' < rs$  and Therefore, if there are total of  $tw$  words in the file, then,  $fs' \rightarrow l$ . But, because of remaining large number of character units,  $rc' \rightarrow tc'$ .

If there a total of  $tw'$  words in the file with lower  $TTR$  values, then,  $S'_l = \sum_{y=1}^{rw'} (sw_y \times fw'_y)$  and  $fw' \rightarrow tw'$

As low frequent words will be encountered, as much frequent syllables will be faced. However, a minor portion of the syllables may be coded in form of words. That is, we will have a larger number of character units and syllable units ( $ts'$  and  $tc'$ ) left to be coded. For syllable

based coding,  $S'_2 = \sum_{y=1}^{rs'} (ss_y \times fs'_y)$

After the two steps, we will have a greater number of characters left to be coded with distinct characters. Typically this value  $tc'$  may be  $tc' < ts'$  and for usual cases,  $tc' < tw'$ . Let this steps saves,

$$S'_3 = \sum_{y=1}^{rc'} (sc_y \times fc'_y)$$

Again,  $S' = S'_1 + S'_2 + S'_3$

$$S' = \sum_{y=1}^{rw'} (sw_y \times fw'_y) + \sum_{y=1}^{rs'} (ss_y \times fs'_y) + \sum_{y=1}^{rc'} (sc_y \times fc'_y) \quad (2)$$

It has been stated that, the ratio of bit-saving for word based, syllable based and character based encoding is 4:2:1 (i.e. for the ratios  $a : b : c$  we may write,  $a > b > c$ ). The ratio is taken as standard because, for character based dictionary oriented encoding, single character or symbol is considered as the primary unit. Consequently, substrings and syllables ranges from two to any higher value. The number of characters that comprise any word may ideally be considered in the range of greater than four whereas short length words may be thought of as the subgroup of substrings. Hence, we may easily remark that, for the above consequences,

(i)  $S_1 \geq S'_1$ ; since the number of word-matching is greater in  $S_l$  and the difference is a multiplier of four.

(ii)  $S_2 \geq S'_2$ ; since the number of syllable-matching is greater in  $S_2$  and the difference is a multiplier of two.

(iii)  $S_3 \leq S'_3$ ; since the number of matching-character is greater in  $S_2$  and the difference is an unit multiplier.

From the above three consequences, we may easily point that, though  $S'_3 \geq S_3$ , it is nit an average case that,  $S'_3$  will cross the saturation point or threshold value comprising of  $(4 + 2 = 6)$  times higher amplitude. Consequently, we may easily conclude:  $S > S'$ .

We may also prove the same as follows:

From the third clause we may write,  $S_3 + \alpha = S'_3$ , where  $\alpha$  is a constant bit-factor. From clause (i) and (ii), we can write  $S_1 + S_2 \geq S'_1 + S'_2$ .

From (i), (ii), we may deduce that (by adding the same value  $S'_3$  in both sides),

$$S_1 + S_2 + S'_3 \geq S'_1 + S'_2 + S'_3 \quad (3)$$

Being a constant factor  $\alpha = 1, 2, 3, \dots, \text{etc (number of bits)}$ , we may write (3) as,

$$S_1 + S_2 + S_3 + \alpha \geq S'_1 + S'_2 + S'_3 \quad (4)$$

It has been presented that,

$$S = S_1 + S_2 + S_3 \text{ and } S' = S'_1 + S'_2 + S'_3.$$

Hence, from (4),  $S + \alpha \geq S'$ .

Recall from earlier explanation that,  $S$  and  $S'$  are the aggregation of bit savings for matching characters, syllables and words. Again, as the ratio of bit saving for word-based, syllable based and character based encoding is considered 4:2:1, consideration of linear transformation factor  $l$  indicates that,  $S$  will rush towards  $6l$  where  $\alpha$  tends words  $l$  for the average cases.

In better cases, as all the characters are likely to be pre-coded with either syllable based coding or word based coding section. After coding with word based and syllable based coding there will be little portion of source text left for character based coding. That is, the total saving is probable to be derived from word based saving.

Consequently,  $S_1 + S_2 \rightarrow S$ , whereas, character based saving will tend towards zero.  $S_3 \rightarrow 0$ . Again, being  $S_3 + \alpha = S'_3$ ;  $\alpha$  too will tend towards zero. That is,  $\alpha \rightarrow 0$ .

$$\text{Again, from clause (i), we get, } S_1 = S'_1 + \alpha_1 \quad (5)$$

$$\text{Similarly, from clause (ii), } S_2 = S'_2 + \alpha_2 \quad (6)$$

$$\text{We get from (iii), } S_3 = S'_3 - \alpha_3 \quad (7)$$

Where  $\alpha_1, \alpha_2, \alpha_3$  are non negative integer values because they all are equalizing factors in terms of bits.

It has already been stated that,  $S_3 \rightarrow 0$ .

Consequently,  $S'_3 - \alpha_3$  will only tends towards zero if and only if,  $\alpha_3 \rightarrow 0$ . That is, we may clearly deduce that,  $\alpha_1 + \alpha_2 - \alpha_3 = c \geq 0$ .

As a result, we may write,  $\alpha_1 + \alpha_2 - \alpha_3 \geq 0$ .

It has been previously stated that,  $\alpha_3 \rightarrow 0$ .

$$\text{Consequently, } \Rightarrow \alpha_1 + \alpha_2 \geq \alpha_3 \quad (8)$$

Now, (5) + (6) + (7) results-

$$S_1 + S_2 + S_3 = S'_1 + \alpha_1 + S'_2 + \alpha_2 + S'_3 - \alpha_3$$

$$\Rightarrow S_1 + S_2 + S_3 = S'_1 + S'_2 + S'_3 + \alpha_1 + \alpha_2 - \alpha_3$$

$$\begin{aligned} \Rightarrow S_1 + S_2 + S_3 &= S'_1 + S'_2 + S'_3 + c \text{ [Using (8)]} \\ \Rightarrow S_1 + S_2 + S_3 &\geq S'_1 + S'_2 + S'_3 \\ \Rightarrow S &\geq S' \text{ [Since } c \text{ is non-negative]} \end{aligned}$$

That is, the performance of dictionary based data compression would be better for files with larger *TTR* Values and the result will deteriorate for lower *TTR* valued files. This is the aspect, which motivates us to provide peer files of each category to recognize the fluctuation of performance for dictionary based data compression.

## VII. ANALYSIS OF THE PROPOSED CORPUS

In this section, we present statistics regarding *TTR* and *Compression ratio* of the files of each groups. The experimental results demonstrate that, not only in expressing the effectiveness of any corpus, *TTR* may also be employed as a criteria for constructing data compressing corpus with the great scope of presenting the performance fluctuation of dictionary based (*i.e.* repetition analysis oriented) text compression scheme for best case and worse case analysis.

We also analyze the compression ratio for the files using the dictionary-based approach for Bengali text compression provided in [3] and find the following statistics-

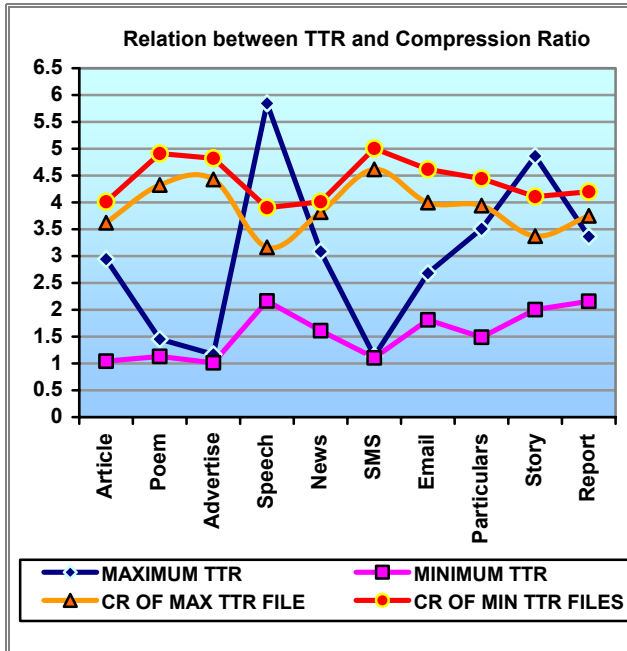


Fig. 2: Relation between Type to Token Ratio and Compression Ratio.

In Fig 2, we have presented a comparative analysis on Compression Ratio with Type to Token Ratio in terms of

applicability of dictionary based data compression techniques. *Compression Ratio (CR)* is a metric of data compression efficiency. Compression Ratio indicates the number of bits required to describe one byte in compressed form. The lower the compression ratio, the better is the compression. In order to evaluate the performance fluctuation of dictionary based Bengali text compression schemes, we propose to use the peer files with assumption that, the best case would be obtained from the files with greater *TTR* values and the worse case will occur for lower *TTR* files. The construction of the dictionary may really be facilitated from the files with larger *TTR*.

## VIII. CONCLUSION

We have proposed a new Corpus for evaluation of Bengali text compression schemes named *Ekushey-Khul*. The methodology is also robust to some extent that, it takes both linguistic and compression-ratio into account. It is also a contribution that, we specify new criteria for choosing text compression corpus, which includes consideration of expected device and the changed attitude of text varying with communication framework for which the compression is intended. Though it is an inaugurating step towards forming a Bengali text compression evaluation corpus, further improvement may be achieved by integrating other types and modes of document like technical documents, Bengali dictionary, and Bengali spreadsheets etc. with consideration of more training files.

## REFERENCES

- [1] Khair Md. Yeasir Arafat Majumder, Md. Zahurul Islam, and Majumder Khan, "Analysis of and Observations from a Bangla News Corpus", *Proceedings of 9th International Conference on Computer and Information technology ICCIT 2006*, pp. 520-525, 2006.
- [2] N. S. Dash, "Corpus Linguistics and Language Technology", 2005.
- [3] S. A. Ahsan Rajon, "A study on Bengali Text Compression Schemes", *Research Report*, Khulna University, Khulna, June 2008.
- [4] Mat Powel, "Evaluating Lossless Compression Algorithms", February, 2001.
- [5] Ross Arnold, and Tim Bell, "A Corpus for the evaluation pf lossless compression algorithms", *Data Compression Conference*, pp. 201-210, IEEE Computer Society Press, 1997.
- [6] Official Web site of Prothom-Alo, [www.prothom-alo.com](http://www.prothom-alo.com)
- [7] M. Burrows and D. J. Wheeler. "A block sorting lossless data compression algorithm". Technical report, Digital Equipment Corporation, Palo Alto, CA, 1994.

# A Corpus-based Evaluation of Lexical Components of a Domain-specific Text to Knowledge Mapping Prototype

Rushdi Shams<sup>1</sup> and Adel Elsayed<sup>2</sup>

Department of Computer Science and Engineering, Khulna University of Engineering & Technology (KUET), Bangladesh<sup>1</sup>

M3C Research Lab, the University of Bolton, United Kingdom<sup>2</sup>

**Abstract** — The aim of this paper is to evaluate the lexical components of a Text to Knowledge Mapping (TKM) prototype. The prototype is domain-specific, the purpose of which is to map instructional text onto a knowledge domain. The context of the knowledge domain of the prototype is physics, specifically DC electrical circuits. During development, the prototype has been tested with a limited data set from the domain. The prototype now reached a stage where it needs to be evaluated with a representative linguistic data set called corpus. A corpus is a collection of text drawn from typical sources which can be used as a test data set to evaluate NLP systems. As there is no available corpus for the domain, we developed a representative corpus and annotated it with linguistic information. The evaluation of the prototype considers one of its two main components- lexical knowledge base. With the corpus, the evaluation enriches the lexical knowledge resources like vocabulary and grammar structure. This leads the prototype to parse a reasonable amount of sentences in the corpus.

**Index Terms** — Corpus, NLP Systems, Knowledge Mapping, Lexicon, Morphology.

## I. INTRODUCTION

Text to Knowledge Mapping (TKM) Prototype [1] is a domain-specific NLP system, the purpose of which is to parse instructional text and to model it with its pre-defined ontology. During development, the prototype has been tested with a limited data set from the domain Physics instructional text on DC electrical circuit. The prototype now reached a stage where its lexical components need to be evaluated with a representative linguistic data set, a corpus- a collection of text drawn from typical sources. Information retrieved during parsing contributes to map and model domain-specific text on its knowledge domain. Therefore, the usability of the TKM prototype as a specialized knowledge representation tool for the domain depends on the evaluation of its lexical components.

An important precondition to evaluate NLP systems is the availability of a suitable set of language data called corpus as test and reference material [2]. With an extensive web-based search, we did not find any corpus for the domain DC electrical circuit. Therefore, we need to develop a representative corpus to evaluate the lexical components of the prototype because a representative

corpus reflects the way language is used in the domain [3]. A usable corpus requires various annotations according to the scope and type of evaluation. As we intend to evaluate the lexical components of the TKM prototype, the corpus should be annotated with linguistic information like Parts of Speech (POS) tagging, phrasal structure annotations, and stem word tagging. These annotations can lead us to adjust the lexical components of the prototype according to the qualitative and quantitative layers [1] [4] of its knowledge model.

In this paper, we proposed a stochastic development procedure of a domain-specific representative corpus that is used to evaluate the lexical components of the TKM prototype. We also represented detail procedure of corpus-based evaluation of an NLP system- that includes enriching the lexicon and morphological database, testing the parsing ability of the prototype, and the adjustment of the lexical components according to the linguistic information in the corpus.

The remainder of this paper is organized as follows. In Section II, corpus-based evaluations of various NLP systems have been discussed. Section III describes the proposed procedure of representative corpus development and annotations. Section IV describes the evaluation of lexical components of the TKM prototype such as the vocabulary and grammar structure. Section V concludes the paper.

## II. RELATED WORK

A text based domain-specific NLP system can be evaluated according to the type, context or discourse of text from the domain although no established agreement has been developed on test sets and training sets [5]. Corpus is not restricted today only for researches on linguistics [6]; it is now becoming the principal resource to evaluate such domain-specific NLP systems. Many NLP systems like Saarbrucker Message Extraction System (SMES) [8] have been tested with a corpus as proper evaluation depends on a representative test set of data like corpus [7]. Corpus contains structured and variable but representative text. A corpus is said representative if the

findings from it can be generalized to language or a particular aspect of language as a whole [3]. Corpus-based evaluations like MORPHIX [13] and MORPHIX++ [7] showed that the evaluation with a representative corpus results in proper adjustments. MORPHIX++ was tested with a corpus and systematic inspection revealed some necessary adjustments like missing lexical entries, discrepant morphology incomplete or erroneous single words.

Now-a-days, many web corpora are present and researchers successfully utilize web data including machine translation [9], prepositional phrase attachment [10] and other anaphora resolution [11] to evaluate NLP systems. One problem in evaluating domain-specific NLP systems with a web corpus is the discrepancy between the domain of the corpus and the domain of the system [14]. Having the ease of access [15], the web corpora development tools like BootCaT [16] and WebBootCaT [17] are used world wide to develop corpus using web data.

NLP systems use either pre-defined or customized grammar rules. For instance, the lexical components of the TKM prototype use Combinatory Categorical Grammar (CCG) [12]. The prototype follows some specific clausal and phrasal structures according to CCG. As it follows a particular grammar, we need to adjust the grammar and phrasal structures according to the structures of text from the domain. For example, TKM prototype, on its early test, was able to parse simple sentences only. This becomes a drawback if majority of text in the domain is written in compound and complex sentences. Therefore, necessary adjustment on CCG can let the prototype parse compound and complex sentences as well. In addition, NLP systems may recognize specific clausal and phrasal structure which maybe absent in domain-specific text. For example, if an NLP system uses grammars that handle one subject and one object, both parsing and knowledge extraction from domain-specific text becomes difficult if majority of the text contains more than one subject and one object. These linguistic properties of domain-specific text bring in the issue of adjustment. The lexicographical resources of such systems can be increased by analyzing linguistic patterns in domain-specific corpus. Statistical data like frequency of words, number of simple, complex or compound sentences, number of subject and object present in the sentences assist to adjust the lexical components of the systems. The grammar structure MORPHIX++ supported was not efficient in its early days. It was adjusted and extended according to the corpus used as its test suite.

### III. CORPUS DEVELOPMENT

In this section, we will discuss regarding the development approach of a domain-specific corpus, proof or its representativeness, and its annotation procedure.

#### A. Development Approach

As we did not find any corpus for the domain DC electrical circuit with extensive web searches, we initiated WebBootCaT to develop a representative corpus. We developed five corpora using the WebBootCaT and analyzed them by comparing the number of distinct domain-specific terms and number of distinct words present. The significant difference between these two numbers and inconsistency on the size of the corpus in Fig. 1 state that web-based tools are not usable to develop domain-specific corpora, but maybe useful for developing domain independent corpora.

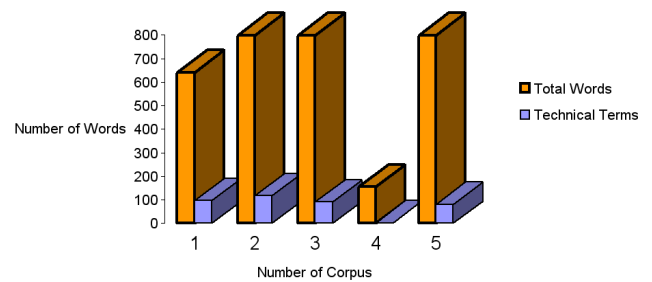


Fig. 1. Inconsistency of WebBootCaT to develop a domain-specific corpus

WebBootCaT works on the basis of searching the web according to user defined keywords called seeds. Researches like [15] showed that search engines like Google provide 67 search results for Speculator and about 82,000 search results for Speculator? which vividly shows the reason behind inconsistent nature of web-based corpus developing tools. Therefore we decided to develop the corpus manually and collected text from 141 web resources containing 1,029 sentences and 18,834 words. During the development, we left the non textual information (e.g., equations and diagrams) as the TKM prototype operates only on text.

#### B. Representativeness of the Corpus

The representativeness of the corpus can be justified with a notion of saturation or closure described by [18]. At the lexical level, saturation can be tested by dividing the corpus into equal sections in terms of number of words or any other parameters. If another section of the identical size is added, the number of new items in the new section



should be approximately the same as in other sections [19].

To find out the representativeness for the corpus, it has been segmented into 15 samples. Each sample is comprised of 1,267 words on average. We plotted the cumulative frequency of the most frequent technical terms in the samples.

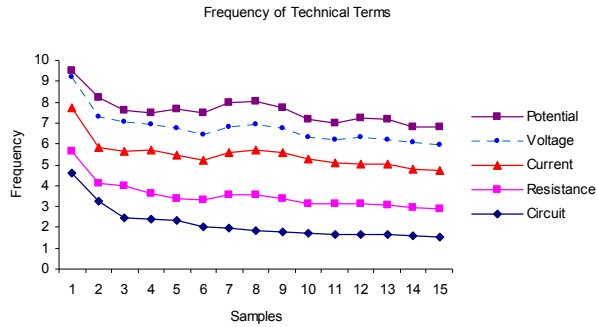
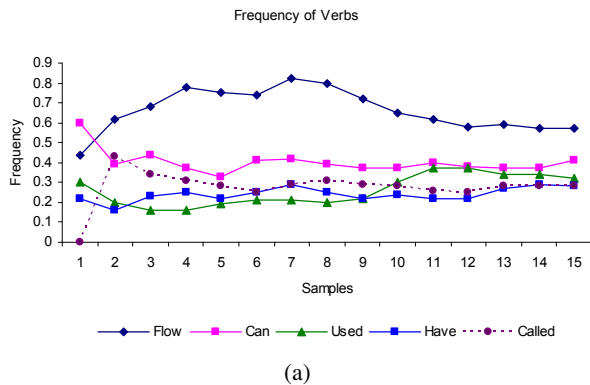


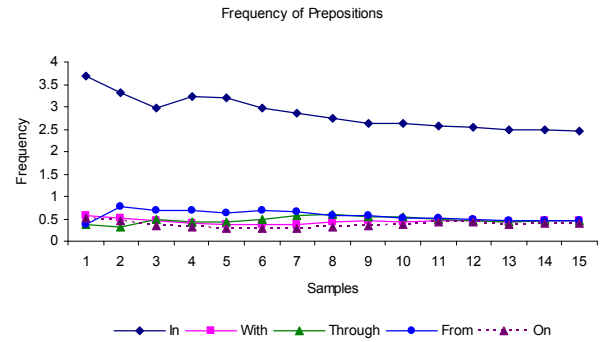
Fig. 2. Representativeness of the corpus with domain-specific technical terms

Fig. 2 depicts that the presence of the domain-specific technical terms becomes stationary after a few samples. This is one of the criteria showing the representativeness of the corpus. After a certain point, no matter how much text we add to the corpus, the frequencies of the terms are becoming stationary.

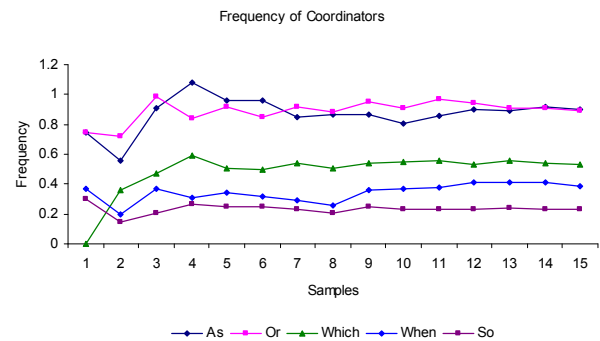
Similarly, we counted the frequency of non-technical words in the corpus and grouped them according to their parts of speech. Statistics on verbs, prepositions and coordinators in Fig. 3 (a), 3 (b) and 3 (c) respectively show that the corpus has been saturated after sample 11.



(a)



(b)



(c)

Fig. 3. Representativeness of the corpus with (a) verbs, (b) prepositions and (c) coordinators

We also counted the frequency of types of sentences in the corpus. As the domain contains instructional text and most of which are simple sentences, it needs to be reflected on the corpus as well. Fig. 4 shows that majority of the text is simple sentence (in percentage).

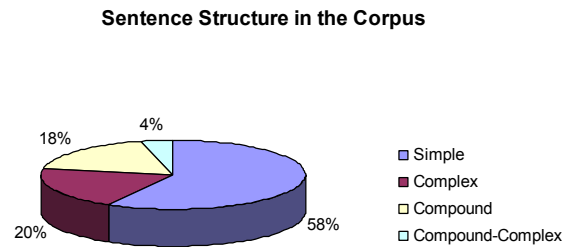


Fig. 4. Sentence structure in the corpus



### C. Corpus Annotation

To annotate the corpus with POS tags, Cognitive Computation Group POS tagger [20] has been used as it works on the basis of learning techniques like Sparse Techniques on Winnows (SNOW). The corpus is annotated with nine parts of speech include noun, pronoun, verb, adverb, adjective, preposition, coordinator, determiner, and modal. The phrasal structure of the corpus has been annotated by the slash-notation grammar rules defined by CCG. We developed seven XML tags including URL, date, sentence, word, stem, pos, and category and created an XML version of the corpus.

## IV. EVALUATION OF LEXICAL COMPONENTS

The evaluation of vocabulary and the grammar structure of the prototype are illustrated in this section. This section also refers to the efficiency in parsing and richness of lexical entries of the prototype.

### A. Evaluation of Vocabulary

The lexicon of the prototype is mapped on the unique words of the corpus. The words present both in the morphology and in corpus are called the vocabulary of the prototype. Initially, only five percent of the vocabulary was covered by the prototype (Table I).

TABLE I  
PRELIMINARY VOCABULARY COVERAGE OF THE TKM  
PROTOTYPE

| Words in Morphology and in Corpus | Unique Words in the Corpus | Vocabulary Covered |
|-----------------------------------|----------------------------|--------------------|
| 101                               | 1,902                      | 5%                 |

MORPHIX++, a second generation NLP system, covered 91 percent of word in the corpus developed to evaluate it. The reason behind this difference is the augmentation of the vocabulary of MORPHIX++ ran parallel with the development of the system. In contrast, the main focus during the development of the TKM prototype was to develop an operational system first rather than increasing its vocabulary.

We used the POS tags of the corpus to populate the lexicon. We retrieved every distinct word for each distinct POS from the corpus and we simply added it if that word was absent in the lexicon. The number of added entries into the lexicon is shown in Table II. On completion of the process, the vocabulary of the prototype covers 90 percent of the corpus (Table III).

TABLE II  
AUGMENTATION OF LEXICAL ENTRIES IN THE TKM  
PROTOTYPE

| POS              | Augmented Entries |
|------------------|-------------------|
| Determiner       | 19                |
| Coordinator      | 5                 |
| Noun and Pronoun | 2,094             |
| Adjective        | 364               |
| Preposition      | 71                |
| Adverb           | 177               |
| Verb             | 264               |

TABLE III  
VOCABULARY OF THE TKM PROTOTYPE

| Words in Morphology and in Corpus | Unique Words in the Corpus | Vocabulary Covered |
|-----------------------------------|----------------------------|--------------------|
| 1,783                             | 1,902                      | 90%                |

### B. Evaluation of Grammar Structure

The TKM prototype struggles to parse *modals* or *auxiliary verb* because CCG does not provide any specification to categorize modals into finite and non-finite [21]. We defined grammar formalisms for modals and adjusted the lexicon that increased the ability of the prototype to parse modals.

CCG does not have any mechanism for phrasal structures like *adjective-adjective-noun* although researches showed that numerous adjectives can be placed before a noun [22]. Moreover, CCG treats adjectives as qualifiers of noun (*np/n* or *n/n*) but in the sentence *when the lamps are connected in series, the more the lamps, the more dimmer they get*, *dimmer* is not particularly preceding any noun or pronoun. Except the regular adjectives, we defined grammar formalisms for noun equivalents (e.g., two *common* types of circuits), participle equivalent (e.g., the *connected* wire), gerund equivalents (e.g., the *conducting* material), and adverb equivalents (e.g., the *above* circuit is series circuit) of adjectives that increased the rate of parsing adjectives.

CCG has specification for pronoun by treating it as a noun without any anaphoric resolves [23]. For example, *it is placed in parallel-* CCG can parse this sentence by treating *it* as a *noun* but cannot obtain *voltmeter* as the subject. Moreover, CCG completely is not able to parse pronouns of other forms (e.g., *them*, *its*, *their*, *that*, *which*) and struggles to represent knowledge when pronouns represent indexicality. The anaphora depends on lexical semantics and its resolve depends on the reasoning through the ontology of the prototype.

CCG is unable to parse sentences that start and end with a prepositional phrase [23]. For example, *in series circuit, the current is a single current*- this sentence is not parsed by CCG. In contrast, *the current is a single current in series circuit*- is sometimes parsed by CCG. The lexicon the prototype is using has nine different types of prepositions. Sometimes, it is difficult to even identify regular prepositions. For instance, *the sum of potential differences in a circuit adds up to zero voltage*- Though in regular grammars, *up* is not treated as adverbs- these are called particles where prepositions have no objects and require specific verbs with them (e.g., throw out, add up). The parsing ability of the prototype increased as we defined grammar rules for such prepositions.

*Complementizer*, although it is a form of preposition, it is not recognized by CCG. *Adverbs*, on the other hand, have a strong coverage by CCG. In many cases, adverbs sit at the end of the sentence- CCG does not provide any category to define these adverbs although it has fully featured adverb categories for other two positions of an adverb in sentence- adverbs that start a sentence or that sit in the middle of a sentence. These issues have been resolved by adding new grammar rules.

The lexicon has two categories for *coordinators*- sitting at the beginning of a sentence (e.g., since, as) and relating two clauses (e.g., and, or). CCG defined that they can be in the middle of two noun phrases only with *np\vp/np* but the sentence *series and parallel circuits are the types of circuits* has the category *n\vp/n* rather than *np\vp/np*. CCG handles adverbs and conjunctions well but it seriously lags in handling sentences having similar verbs as in the sentence *the sum of current flowing into the junction is eventually equal to the sum of current flowing out of the junction*. The identical verbs *flowing* (*gerund*) appear twice with another verb (*be*) is concerning. Moreover, a verb has to be present in a sentence to form predicate argument structure but we discovered that there are sentences which do not have any verbs- *the bigger the resistance, the smaller the current*. Gerund of verb is known as noun. *Gerund* is formed by placing *ing* at the end of the verb. For example, *current flowing into a junction is equal to the current flowing out of the junction*- in this sentence, *flowing* is a gerund. Gerunds are not treated as nouns in CCG. In other words, gerunds, if treated as nouns in CCG, the sentence struggles to be parsed.

After creating grammar rules and phrasal structures and adding them into the lexicon and morphology of the prototype, the parsing ability of the prototype increased to 31 percent (Table IV). Although the prototype was tested with a limited dataset, it was unable to parse any sentence from the corpus before the evaluation.

TABLE IV  
PARSING ABILITY OF THE TKM PROTOTYPE

| State of the Prototype | Total Sentences | Parsed Sentences | Efficiency |
|------------------------|-----------------|------------------|------------|
| Preliminary            | 1,029           | 0                | 0%         |
| Evaluated              | 981             | 300              | 31%        |

We analyzed the 300 sentences parsed by the prototype and figured out the number of subject, object and verb they consist. In Fig. 5 we see that the prototype works well when the number of subjects and objects in a sentence do not exceed two and when the number of verbs does not exceed one.

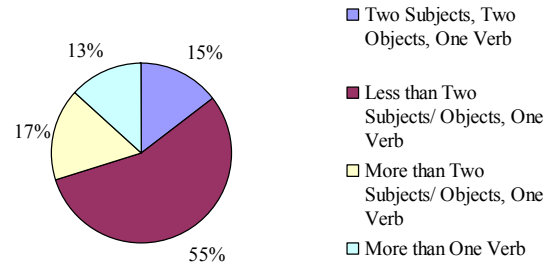


Fig. 5. Number of subjects, objects and verbs in the sentences parsed by the prototype

The inefficiency of the prototype to parse sentence is due to the absence of phrasal structures (hence, the categories). 69 percent of the sentences in the corpus have phrasal structures that are not supported by the CCG structure. It should be noted that the prototype fails to parse sentences even for absence of just one category. For example, *One simple DC circuit consists of a voltage source (battery or voltaic cell) connected to a resistor* – this sentence is not parsed by the prototype for the absence of category of conjunction *or* (*np\vp/np*) and for the category of verb *connected* (*s\vp/pp*). In the corpus, these absent categories are identified so that modification of the lexicon becomes easier.

## V. CONCLUSION

In this paper, we represented a corpus-based evaluation of lexical components of a domain-specific NLP system. We proposed a stochastic approach to develop a domain-specific representative corpus. The linguistic resources of the corpus have been used to evaluate the lexical components of the system. As knowledge representation is the goal of the system and successful parsing and information retrieval is at the heart of knowledge modelling, we adjusted the lexical components of the

system to increase its parsing ability. This evaluation made the TKM prototype an ideal knowledge representation tool for the domain DC electrical circuit as it parses large number of text from the domain and the information retrieval process has been satisfactory. However, the evaluation on the knowledge model and knowledge representation depends on the human cognition of the domain and its representation based on the semantic and linguistic relations in text. We developed a conceptual structure and a framework of semantic relations for the domain [24] and the evaluation of its knowledge model will be accomplished once we are able to tie the significance of linguistic relations in the text with the conceptual structure for the domain.

# REFERENCES

- [1] W. Ou and A. Elsayed, "A Knowledge-Based Approach for Semantic Extraction", *International Conference on Business Knowledge Management*, Macao, 2006.
- [2] T. Declerck, J. Klein, and G. Neumann, "Evaluation of the NLP Components of an Information Extraction System for German", *Proceedings of the first international Conference on Language Resources and Evaluation (LREC)*, Granada, 1998, pp. 293-297.
- [3] D. Evans, "Corpus building and investigation for the humanities", Available: <http://www.humcorp.bham.ac.uk/humcorp/information/corpusintro/Unit1.pdf> [27 September 2008]
- [4] K. D. Forbus, "Qualitative process theory: Twelve years after", *Artificial Intelligence*, volume 59, number 1-2, 1993, pp. 115-123.
- [5] M. Palmer and T. Finin, "Workshop on the evaluation of natural language processing systems", *Computational Linguistics*, volume 16, number 3, 1993, pp. 175-181.
- [6] T. McEnery and A. Wilson, *Corpus Linguistics*, Edinburgh: Edinburgh University Press, United Kingdom, 1996.
- [7] J. Klein, T. Declerck, and G. Neumann, "Evaluation of the Syntactic Analysis Component of an Information Extraction System for German", *Proceedings of the 1st International Conference on Language Resources and Evaluation*, Granada, Spain, 1998, pp. 293-297.
- [8] G. Neumann, R. Backofen, J. Baur, M. Becker, and C. Braun, "An Information Extraction Core System for Real World German Text Processing", *Proceedings of the 5th Conference on Applied Natural Language Processing (ANLP)*, USA, 1997, pp. 209-216.
- [9] G. Grefenstette, "The WWW as a Resource for Example-based MT Tasks", *Proceedings of ASLIB Translating and the Computer Conference*, London, United Kingdom, 1999.
- [10] M. Volk, "Exploiting the WWW as a Corpus to Resolve PP Attachment Ambiguities", *Proceedings of the Corpus Linguistics*, Lancaster, United Kingdom, 2001.
- [11] N. Modjeska, K. Markert, and M. Nissim, "Using the Web in Machine Learning for Other-anaphora Resolution", *Proceedings of EMNLP*, Sapporo, Japan, 2003, pp. 176-183.
- [12] S. Clark, M. Steedman, and J. R. Curran, "Object-Extraction and Question-Parsing using CCG", Available: <http://www.iccs.inf.ed.ac.uk/~stevec/papers/emnlp04.pdf> [27 September 2008]
- [13] W. Finkler and G. Neumann, "Morphix: A fast Realization of a Classification-based Approach to Morphology", *Proceedings der 4. Österreichischen Artificial Intelligence Tagung, Wiener Workshop Wissensbasierte Sprachverarbeitung*, Berlin, 1988, pp. 11-19.
- [14] V. Liu and J. Curran, "Web Text Corpus for Natural Language Processing", *Language Technology Seminar Series*, Sidney, Australia, 2006.
- [15] A. Kilgariff and G. Grefenstette, "Introduction to the special issue on the web as corpus", *Computational Linguistics*, volume 29, 2006, p. 333-347
- [16] M. Baroni and S. Bernardini, "BootCaT: Bootstrapping Corpora and Terms from the Web", *Proceedings of LREC*, 2004.
- [17] M. Baroni, A. Kilgariff, J. Pomikálek, and P. Rychlý, "WebBootCaT: a web tool for instant corpora", *Proceeding of the EuraLex Conference*, Italy, 2006, pp. 123-132.
- [18] T. McEnery, R. Xia, and Y. Tono, *Corpus-Based Language Studies: An Advanced Resource Book*, London: Routledge, 2006.
- [19] D. Biber, "Representativeness in Corpus Design", *Literary and Linguistic Computing*, Volume 8, Number 4, 1993, pp. 243-257.
- [20] University of Illinois at Urbana-Champaign, "The SNoW Learning Architecture", Available: <http://l2r.cs.uiuc.edu/~danr/snow.html> [27 September 2008]
- [21] A. S. Hornby, *Guide to Patterns and Usage in English*, 2<sup>nd</sup> Edition, Oxford University Press, Delhi, 1995, pp. 1-2.
- [22] M. A. Covington, *Natural Language Processing for Prolog Programmers*, 1<sup>st</sup> Edition, Prentice Hall, 1993, pp. 88-90.
- [23] M. Steedman and J. Baldridge, "Combinatory Categorical Grammar", Unpublished Tutorial Paper, 1993.
- [24] R. Shams and A. Elsayed, "Development of a Conceptual Structure for a Domain-specific Corpus", *3<sup>rd</sup> International Conference on Concept Maps (CMC)*, Estonia and Finland, 2008.

# Density Based Clustering Technique For Efficient Data Mining

Md. Asikur Rahman, A.K.M Rasheduzzaman Chowdhury, Daud Md. Jamilur Rahman,  
Abu Raihan Mostofa Kamal

Computer Science & Information Technology (CIT), Islamic university of technology (IUT), Board  
bazaar, Gazipur-1704, Bangladesh

E-MAIL: abir\_cit@yahoo.com, rashed044416@yahoo.com, zamil525@yahoo.com,  
raihan\_rcc@yahoo.com

**Abstract** — Clustering analysis is an important function of data mining. There are various clustering methods in Data Mining. Based on these methods various clustering algorithms are developed. A recent approach for clustering analysis is based on “Swarm Intelligence”. Based on this “Swarm Intelligence” an algorithm was proposed named “Ant-Cluster algorithm”. However, existing “Ant clustering” algorithm has a limitation in finding the value of two constant  $K_1$  and  $K_2$ , which is user defined, for computing the value of the picking up probability  $P_p$  and dropping probability  $P_d$ . In this paper our approach is to gain the value of  $P_p$  and  $P_d$  without giving the user defined value of  $K_1$  and  $K_2$ . We also intend to retain the  $P_p$  and  $P_d$  in between 0 to 1 in order to get optimized result.

**Index Terms** — Ant Clustering methods, Data Mining, Density, Swarm Intelligence.

## I. INTRODUCTION

Data mining is the task of discovering interesting patterns from large amounts of data, where the data can be stored in databases, data warehouses, or other information repositories. It is a young interdisciplinary field, drawing from areas such as database systems, data warehousing, statistics, machine learning, data visualization, information retrieval, and high-performance computing. Other contributing areas include neural networks, pattern recognition, spatial data analysis, image databases, signal processing, and many application fields, such as business, economics, and bioinformatics.

Clustering analysis is a kind of unsupervised learning method that groups a set of data objects into clusters. In these clusters, data objects are similar to one another within the same cluster and are dissimilar to the objects in other clusters [1]. Major clustering methods are classified into five categories, i.e.

- a) partitioning methods
- b) hierarchical methods
- c) density-based methods
- d) grid-based methods and
- e) model-based methods.

Some clustering algorithms integrate the ideas of several clustering methods. But each kind of clustering methods has its own limitation. For example,  $k$ -means algorithm, a widely used partitioning method, is sensitive to the initial objects that can induce local optimum, and needs user to specify the number of clusters.

Swarm intelligence is a kind of evolutionary algorithm inspired by the behaviors of social animals. It has some advantages and characteristics, such as self-adaptation, self-government, and parallel computing, etc.

For example, in these algorithms, the numbers of clusters need not to be specified by user, amount of calculation reduces due to calculating with local objects instead of all objects, and clusters with arbitrary shape can be discovered.

In “Ant-Cluster algorithm”, multi-population of ants with different moving speed is introduced and outlier objects are processed properly. Ants are the agents which can pick up or drop out the data objects or sets based on similarity or density of data objects. Ants are likely to pick up data objects that are either isolated or surrounded by dissimilar and then drop the picked ones in the similar types of data objects. Thus a proper clustering is achieved.

However in our research we found some limitations in “Ant Clustering” algorithm. The next proceedings of this paper will be focusing on basic of existing “Ant Clustering Algorithm” to solve the limitations and hence propose our newly developed method to find out the value of picking up probability  $P_p$  and dropping probability  $P_d$  without giving the user defined value of  $K_1$  and  $K_2$ .

## II. DATA MINING & CLUSTERING

Clustering is also called data segmentation in some applications because clustering partitions large data sets into groups according to their *similarity*. Clustering can also be used for outlier detection, where outliers (values that are “far away” from any cluster) may be more interesting than common cases. Applications of outlier detection include the detection of credit card fraud and the monitoring of criminal activities in electronic commerce. For example, exceptional cases in credit card transactions, such as very expensive and frequent purchases, may be of interest as possible fraudulent activity. As a data mining function, cluster analysis can be used as a stand-alone tool to gain insight into the distribution of data, to observe the characteristics of each cluster, and to focus on a particular set of clusters for further analysis. Alternatively, it may serve as a preprocessing step for other algorithms, such as characterization; attribute subset selection, and classification, which would then operate on the detected clusters and the selected attributes or features.

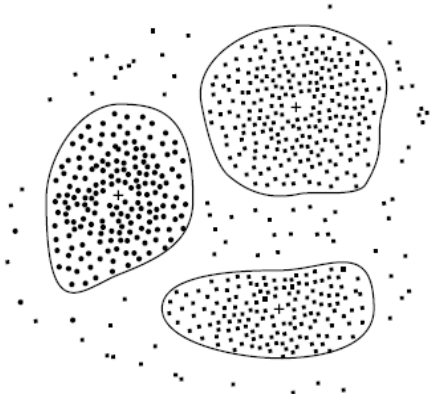


Fig 1 A 2-D form of cluster, where “+” denotes the center of a cluster.

## III. REVIEW OF ANT-CLUSTER ALGORITHM

The parameters and symbols used for ant clustering algorithm are illuminated as follows:

- $\alpha$ : swarm similarity coefficient;
- $r$ : observing radius of each ant;
- $N$ : the maximum of cycle times;
- $size$ : the size of the 2-dimension grid;
- $m$ : the number of ants in each population;
- $p$ : index of populations,  $p = 1, 2, 3$ ;
- $Pp$ : picking-up probability;
- $p_d$ : dropping probability;
- $p_r$ : random probability,  $p_r \in [0, 1)$ ;

$k_1$  and  $k_2$ : threshold constants for computing  $p_p$  and  $p_d$  respectively;

$ant_i$ : the  $i$ th ant;

$o_i$ : the  $i$ th data object;

$loaded$  and  $unloaded$ : state of ant. If there is a data object on an ant, its state is *loaded*; otherwise, its state is *unloaded*;

$v_{high}$ : the speed of ants in high speed population;

$v_{low}$ : the speed of ants in low speed population;

$v_{MAX}$ : the maximal speed in variable speed population;

$l$ : the maximum times of an ant moving with a same data object continuously.

### A. Algorithm :A High-Level Description of Ant-Cluster

Initialization phase: Initialize parameters ( $\alpha$ ,  $r$ ,  $N$ ,  $size$ ,  $m$ ,  $v_{high}$ ,  $v_{low}$ ,  $v_{MAX}$  and  $l$ ). Place data objects on a 2-dimension grid randomly, i.e. assign a pair of coordinates ( $x$ ,  $y$ ) to each data objects. Put three populations of ants with different speed on this 2-dimension grid. Initial state of each ant is *unloaded*;

while (cycle\_time  $\leq$   $N$ )

Adjust  $\alpha$  with specific step;

for ( $p = 1$ ;  $p \leq 3$ ;  $p++$ )

for ( $i = 1$ ;  $i \leq m$ ;  $i++$ )

if (ant encounter a data object)

if (state of ant is *unloaded*)

Compute the swarm similarity of the data object within a local region with radius  $r$ , and compute picking-up probability  $p_p$ . Compare  $p_p$  with a random probability  $p_r$ . if  $p_p > p_r$ , ant pick up this data object, and the state of ant is changed to *loaded*;

else

if (state of ant is *loaded*)

If ant has already moved with the same data object  $l$  steps, the data object is dropped and the state of ant is changed to *unloaded*. Otherwise, compute the swarm similarity of the data object within a local region with radius  $r$ , and compute dropping probability  $p_d$ . Compare  $p_d$  with a random probability  $p_r$ . if  $p_d > p_r$ , ant drops this data object, and the state of ant is changed to *unloaded*.

end

end

end

#### IV. GENERAL DESCRIPTION OF ANT-CLUSTER

In initialization phase, all parameters, including  $\alpha$ ,  $r$ ,  $N$ ,  $size$ ,  $m$ ,  $v_{high}$ ,  $v_{low}$ ,  $v_{MAX}$  and  $l$ , are given values by user. Data objects and three populations of ants with different speed are placed in a 2-dimension grid randomly. There is only one data object and/or one ant in a grid at most. Initial state of each ant is set *unloaded*.

In each of outer loop iteration, i.e. while loop, all ants on the 2-dimension grid move one time. Each of interior loop iteration corresponds to the behavior of one ant. An ant moves one step on the 2-dimension grid randomly with different speed according to different population at a time. When it encounters a data object and its state is *unloaded*, the swarm similarity and picking-up probability are computed for deciding whether or not to pick up the data object. When it does not encounter a data object and its state is *loaded*, the moving times with the same data object is compared with  $l$  at first. If the ant has already moved with the same data object  $l$  times, the data object is dropped. Otherwise, the swarm similarity and dropping probability are computed for deciding whether or not to drop the data object.

The swarm similarity is computed by following formula:

$$f(o_i) = \frac{1}{S} \sum_{o_j \in Neigh(r)} \left[ 1 - \frac{d(o_i, o_j)}{\alpha} \right] \quad (1)$$

where  $f(o_i)$  is a measure of the average similarity of object  $o_i$  with the other objects  $o_j$  present in the neighborhood of  $o_i$ .  $S$  is the number of objects  $o_j$ .  $d(o_i, o_j)$  is the distance between two objects  $o_i$  and  $o_j$  in the space of attributes measured with Euclidean distance. The swarm similarity is transformed to picking-up probability  $p_p$  and dropping probability  $p_d$  by following formulas respectively.

$$P_p(o_i) = \left( \frac{k_1}{k_1 + f(o_i)} \right)^2, \quad p_d(o_i) = \left( \frac{k_2}{k_2 + f(o_i)} \right)^2 \quad (2)$$

Wherein,  $k_1$  and  $k_2$  are two threshold constants assigned by users.

#### V. METHOD TO CALCULATE $P_p$ AND $P_d$ WITHOUT $K_1$ AND $K_2$

We will calculate the swarm similarity for all objects and average the swarm similarity  $f[o_{avg}]$  of all objects.

$$f(o_{avg}) = \{f(o_1) + f(o_2) + \dots \pm f(o_n)\} / n \quad (3)$$

Where,  $n$  is the number of objects in two-dimensional data grid. Here, we use the formula [1] to calculate the swarm similarity of each object.

Then we have to calculate the value of  $IP_i$  and  $ID_i$  for each object  $O_i$  by the following formula :

$$IP_i = f(o_i) / f(o_{avg}) \quad (4)$$

and

$$ID_i = f(o_{avg}) / f(o_i) \quad (5)$$

Here  $i=1, 2, 3, \dots, n$

Then we have to calculate  $IP_{avg}$  and  $ID_{avg}$  by the following formula:

$$IP_{avg} = \sum_{i=1}^n IP_i / n \quad (6)$$

And

$$ID_{avg} = \sum_{i=1}^n ID_i / n \quad (7)$$

Here,  $i=1, 2, 3, \dots, n$ .

Where,  $IP_i$ ,  $ID_i$ ,  $IP_{avg}$ ,  $ID_{avg}$  are the intermediate variables.

Now we have to calculate the picking up probability  $P_p$  and dropping probability  $P_d$  for each object  $O_i$  using the following formula:

$$P_{p_i} = (IP_{avg} - IP_i)^2 \quad (8)$$

and

$$P_{d_i} = (IP_{avg} - ID_i)^2 \quad [9]$$

Here,  $i=1, 2, 3, \dots, n$ .

#### VI. EXPERIMENTAL RESULTS

For experimenting our method we have coded our newly developed formula in Turbo C. Here we have experimented 20 objects in two dimensional Data Grid with their  $x$  and  $y$  co-ordinate.



## ACKNOWLEDGEMENT

Last but not the least we would like to thanks all our faculty members of our university. Also gratitude to ICCIT to give proper guidelines to format this paper.

## REFERENCES

- [4] Cheng-Fa Tsai, Han-Chang Wu, Chun-Wei Tsai, "A new data clustering approach for data mining in large databases", Proceedings of the International Symposium on Parallel Architectures, Algorithms and Networks, Metro Manila, Philippines, pp. 22-24, 2002.
- [5] H.Azzag; N.Monmarché; M.Slimane; C.Guinot; G.Venturini, "AntTree: a new model for clustering with artificial ants", Proceedings of the 7th European Conference on Artificial Life, Dortmund, Germany, pp. 14-17, 2003.
- [6] Wu bin, Shi Zhongzhi, "A clustering algorithm based on swarm intelligence", Proceedings of the International Conferences on Info-tech and Info-net, Beijing, China, pp. 58-66, 2001.
- [7] Wu Bin, Zheng Yi, Liu Shaohui, Shi Zhongzhi, "CSIM: a document clustering algorithm based on swarm intelligence", Proceedings of the Congress on
- [1] Peng jin, Yun-long zhu, Kun-yuan hu, Sixth International Conference on Machine Learning Cybernetics, Hong Kong, 19-22 August 2007
- [2] Eric Bonabeau, Marco Dorigo, Guy Theraulaz, Swarm Intelligence: from Natural to Artificial Intelligence, Oxford University Press, New York, 1999.
- [3] Cheng-Fa Tsai, Chun-Wei Tsai, Han-Chang Wu, and Tzer Yang, "ACODF: a novel data clustering approach for data mining in large databases", Journal of Systems and Software, Vol 73, pp. 133-145,2004.
- Computational Intelligence, Hawaiian, USA, pp. 477-482, 2002.
- [8] Yan Yang, Mohamed Kamel, "Clustering ensemble using swarm intelligence", Proceedings of the 2003 IEEE Swarm Intelligence Symposium, Piscataway, NJ, USA, pp. 65-71, 2003.
- [9] Peng Yuqing, Hou Xiangdan, Liu Shang, "The K-means clustering algorithm based on density and ant colony", Proceedings of the International Conference on Neural Networks and Signal Processing, Nanjing, China, pp. 457-460, 2003.
- [10] E. Lumer and B. Faieta, "Diversity and adaptation in populations of clustering ants", Proceedings of the third international conference on Simulation of adaptive behavior: from animals to animats 3, Brighton, United Kingdom, pp. 501-508, 1994.



# An Intelligent Aviation Obstruction Warning light

Muhammad Mehedi Al Emran Hasan<sup>1</sup>, and Badal Chandra Sarker<sup>2</sup>

Dept. of CSE, East West University, Dhaka, Bangladesh<sup>1</sup>, Dept. of EEE, World  
University of Bangladesh<sup>2</sup>  
mehedi@ewubd.edu, 321badal@gmail.com

**Abstract** -- A scheme for aviation warning light is presented here. Unlike the conventional systems used for this purpose, the system presented here is using intelligent way of sensing daylight sensor fault, intelligent way of sensing light fault, intelligent way of activating standby light in case of the main light fault, ground cable monitoring and responding system with beep for any fault occurs. Microcontroller from 8051 family with tiny operating systems is implemented here for controlling and calculating in a multitasking environment so that the system becomes dependable with advanced feature.

**Index Terms** — Microcontroller, LDR, GCM, RTX51 tiny OS, stand by, line of sight.

## I. INTRODUCTION

The rapid growth of the cellular phone industry and other communication services will lead up to millions of new communication towers all over the Globe. These towers present hazards to flight paths of aircrafts. Aviation Obstruction marking requires especially in Telecom and Cellular Towers, TV and Radio Antennae Towers, High rise Power Transmission cable posts or lines and Pylons, High rise Chimneys, Cooling Towers, High rise Structures, Multistoried Buildings and on surrounding hills near the airports or heliports. The main purpose of providing the Obstruction Lights is for the safety of the flights by identifying any object or structure coming in line of sight and avoids any risk of mishap or accident. Conventional systems are using incandescent bulb which is lack of omni directional equal intensity of light and it fails the aircraft to warn in the line of sight.

Life time of lighting system is another great important parameter for aviation light warning system. Average on time of a light is about 10-12 hours per day. Normally available incandescent bulb is not suitable for this long duration operation. Because incandescent bulb has designed for household application, where it's average on time is about 5-6 hours. It is almost impossible to have a continuous observation on all these conventional warning lights. The warning light may not glow for any reason; it may fail to sense the day or night due to sensor fault, the warning light itself may be faulty or the ground cable may have cut. It would be safe if the system could decide the light fault by itself and glow the stand by light to avoid the risk in some crucial places.

To overcome the above mentioned problems associated with conventional systems we have used power saving LED, advanced features for checking sensors, lights, ground cable, a stand by light and beep against any fault. The aviation light unit consists of LED arrays. The average life time of LED is about 5-7 years with an average on time of a light is about 10-12 hours per day. System can be configured with optional advanced features according to the user for a) Single light normal with a stand by light, b) Single light flashing with a stand by light, c) Multiple light normal d) Multiple light flashing.

The system proposed here for Aviation Warning Lights conform to the specifications of International Civil Aviation Organization (ICAO) [1].

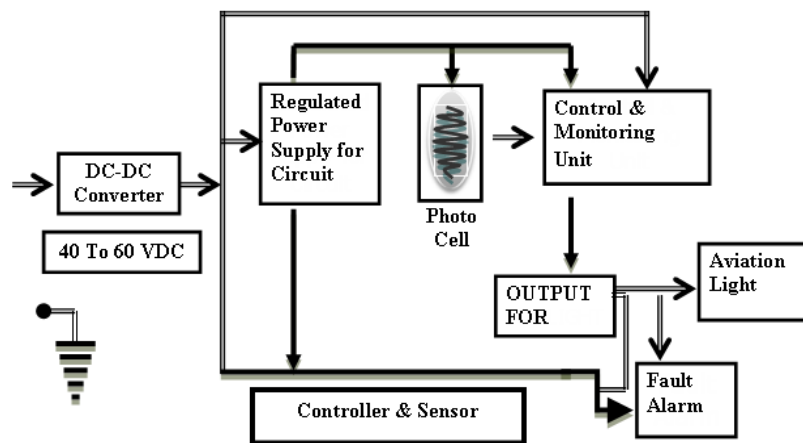


Figure 1: Functional Block Diagram of the System

## II. OVERALL SCHEME

The system comprise with one microcontroller based control circuit along with four sensor circuits, one LED driver circuit, one LCD module, one alarm circuit, and a number of LED arrays. Control circuit consists of one 8051 microcontroller [2], one 12 MHz crystal, some resistors and capacitors. Use of a microcontroller has simplified the design considerably by reducing the control complexity and the number of chips required to implement the system. An LCD segment has been used for displaying the measured data to reduce power consumption. Figure 1 shows the overall system. Day-night sensor circuit consists of photo cell or Light Dependant Resistor (LDR) [3] which will give a signal to the microcontroller if it is night to glow the light. The aviation light is consisting of LED arrays. This light will glow according to the signal generated from the microcontroller. Figure 2 shows a Printed Circuit Board of a actual control circuit.

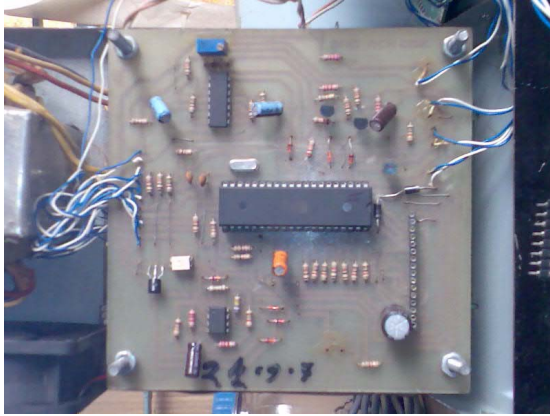


Figure 2: PCB of Control Circuit

LCD module will show the status of the total system including the faults and beep. The fault alarm signal line can be directly connected to the data acquisition system and can be monitored remotely. The system indicates Power on, Tower Light Fault, Sensor Fault, Sensor Active and grounded cable cut signals. Figure 3 shows the controller system.

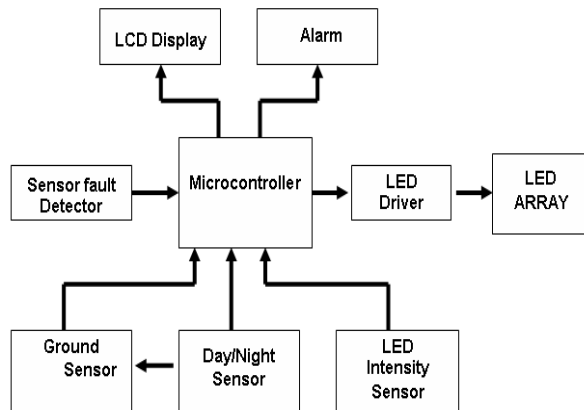


Figure 3: Overall system block diagram of controller

## III. DAY-NIGHT SENSOR

In order to determine day or night a photo conducting device such as Light Dependant Resistor (LDR) has been used in the system. LDR is one kind of semiconductor device in which resistance changes with the changes of light intensity. By using this variation of resistance a signal is produced to control warning light. Day-Night sensor circuit consists of one Light Dependant Resistor (LDR), one operational amplifier, one diode, some resistors and capacitors. Here operational amplifier is used as inverting amplifier. When day light is decreased the resistance of LDR increases and its corresponding voltage drop increases causing the out put of operational amplifier goes low (zero level). This zero level is fed to microcontroller and the microcontroller produces a signal to drive the LED display driver circuit.

To make the system dependable and robust it is very important to check whether the sensor is working. If the Day-Night sensor or Light Dependant Resistor is faulty then the system will be unable to detect day and night automatically and the system will be malfunctioning. In order to solve this problem we use day-Night sensor fault indicator. With the help of this sensor user can monitor the day-night sensor instantaneous status on the LCD module. When the LDR is unserviceable or disconnected, a variation of voltage is occurred in the circuit and by using this voltage we produce a signal which is fed to the microcontroller to generate alarm and also display on the LCD module. This day-night sensor fault detector circuit consists of one operational amplifier two diodes some resistors and capacitors. Here Operational amplifier is used as inverting amplifier.

## IV. WARNING LIGHT

Considering the parameters like power saving, low intensity lights and line of sight we have chosen LED arrays for constructing the warning light. 300 LED has been used in 60 parallel lines, arranging 5 LED in a series to construct the light. To confirm light intensity equal in omni directional, LED arrays are arranged in a hexagonal box so that the light will be in line of sight horizontally form any angle. A hexagonal plate also covers the hexagonal box to glow the light vertically. Figure 4 shows the hexagonal orientation of LED.

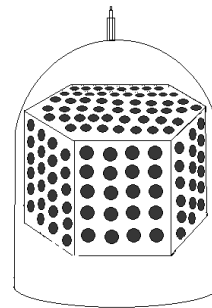


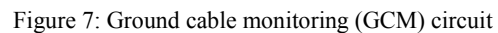
Figure 4: Orientation of LED arrays

If some or total LED of warning light is unserviceable, the intensity of light will be decrease and the visibility will be down. In order to solve this problem a sensor circuit has been used to monitor the instantaneous status of the indicator on the LCD module. If we placed a series resistor with the LED circuit, the total current through the LED is the same as through the resistor. If the intensity of LED is varied, the current through the series resistor also vary and by using this variation of current we produce a signal that signal is fed to the microcontroller and microcontroller produce an alarm signal and display on the LCD module.



Figure 6: A finished warning light

It is difficult to monitor grounding cable daily for any Mobile phone operators. But it is badly needed for telecommunication equipments. Without grounding cable protection, telecommunication equipments stay with serious risk for every moment. A serious damage may occur any time and the whole system remains unsecured. Grounding cable is much expensive that is stolen frequently, to prevent this ignorant and destructive activity a circuit has been used that produce a signal when the grounding cable is cut in any place, this signal is fed to the microcontroller and microcontroller generates an alarm signal and show this message on the LCD module. This signal can be connected in the data acquisition system so that it can be monitored centrally and allow the operator to repair or replace immediately. Figure 7 shows the GCM circuit. This fault detector circuit consists of one opto-isolator [4] along with some resistors and capacitors.



Power supply is a DC to DC converter it consists of four parts Pulse width modulation, out put stage, Error amplifier and Current sense short circuit protector. There are three outputs in this power supply 22 volt, 12 volts and 05 volt. Here we use a DC to DC converter Power supply. This power supply consists of four units Pulse width modulation, out put stage, Error amplifier and Short circuit protection circuit. There are three outputs in this power supply 22 volt, 12 volts and 05 volt. This Power supply itself is very intelligent in both input and output. Considering reverse polarity protection, Input terminal is designed with universal polarity that means you can connect any lead to any terminal of battery. This universal polarity [5] function is done by using a high Ampere rating Bridge Rectifier. Figure 8 shows the universal polarity circuit.

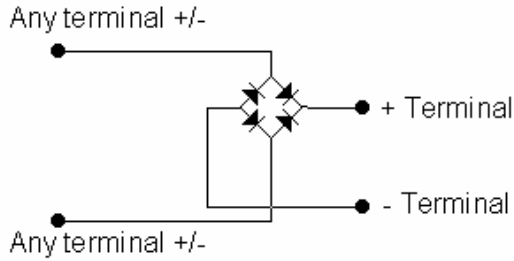


Figure 8: universal polarity circuit

The output is short circuit protected. If output terminal is short circuited by any means then output voltage will be reduced to near about zero so no current flow in the short circuited terminal and there is no damage to power supply. This function is done by using an operational amplifier and a series resistor. When the output current exceeds desired level then the short circuit protection circuit generate a signal that signal will stop the output of this power supply.

## VII. SOFTWARE AND ALGORITHM

Simultaneously checking several units, responding and running tasks in real time lead the software or the firmware for aviation warning light system to a multitasking environment. Many microcontroller applications require simultaneous execution of multiple jobs or tasks. For such applications, a real-time operating system (RTOS) [6, 7, 8] allows flexible scheduling of system resources, like the CPU and memory, and offers communication between tasks. RTX51 Tiny [9] is a powerful RTOS that is easy to use and that works with all 8051 derivatives. RTX51 Tiny programs are written using standard C constructs and compiled with the Keil C51 C Compiler [10]. Figure 9 shows the user system, hard ware and the operating system.

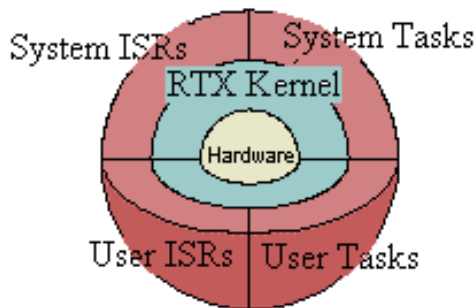


Figure 9: System with RTOS

RTX51 Tiny supports round-robin [9] and the use of signals for task switching which allows quasi-parallel execution of several endless loops or tasks.

Tasks are not executed concurrently but are time-sliced. RTX51 Tiny switches to another task that is ready to run and allows that task to execute for a while. The time slices are very short, usually only a few milliseconds. For this reason, it appears as though the tasks are executing simultaneously.

To starts the system, first system configuration is checked for different mode of light glowing, LCD initialized for displaying the configuration and then current status of the system. To get the status of the system and to respond according to the current status, the overall functionality of this system is divided into four major tasks. They are task1 named check system, task2 named run light, task3 named display and alarm and task4 named check light. RTX51 Tiny does not require you to have a main function in your program. It will automatically begin executing task 0. These entire four tasks have to be created with executing task0. After execution of task0, they will be running in round-robin fashion and with the use of signals between them. Figure 10 shows that the CPU sharing time with the tasks.

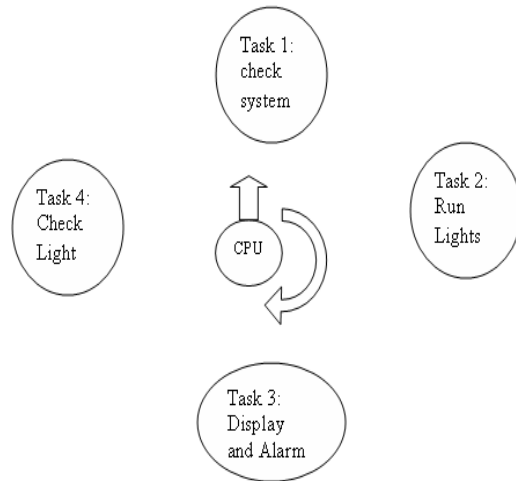


Figure 10: Aviation warning light system in a real time environment

Figure 11 shows the flowchart of task 1 that runs independently. First it checks the day-night sensor. If it is faulty then it starts the light permanently and sends a beep or alarm signal as long as the user is not resetting the beep. If the sensor is not faulty then it goes to check whether it is night. If it is night then it starts the light. But if it is day and there is no flag is on for a schedule testing of light it will stop the light. Subsequently it checks the ground cable. If the cable is cut then it will send a beep or alarm signal. Lastly it will check whether the beep has been reset. User has to press the alarm reset button for a few seconds to stop the beep. If it has not been reset, a flag for alarm or beep will be on and other tasks will be notified through the beep flag.

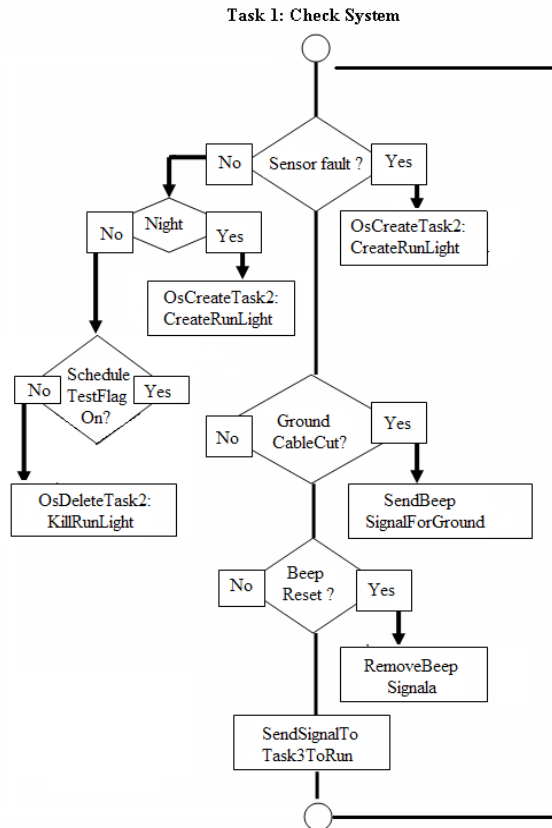


Figure 11: Task 1

If it is night or it is the schedule for checking light in day time, task2 for running light will be started. Light fault checking will be done here. If the light is fault it will immediately send a beep signal and will start the stand by light. Task2 will check the stand by light also. Incase of the failure of the stand by light it will send the alarm or beep signal again. Figure 12 shows the flowchart of task 2 and task3.

Task 3 will be waiting as long as no signal is received from task 1. It is responsible for checking and displaying the status of the system in the LCD module. After displaying it will send signal to task 4 to run.

Task 4 is responsible for a routine wise checking of the light. Before starting first time it will be waiting for a signal from task 3 and will start counting time. If it is night then task 4 has nothing to do but if it becomes day, task 4 will check the light after each thirty minute. This interval can be changed for different users. A flag for schedule test is set on before testing light each time. Figure 13 shows the flowchart of task 4.

For implementing the above mentioned algorithm, total 570 lines of code and 23 methods were written in ANSI C.

For practical implementation, a metal casing has been used for the controller circuit shown in figure 14.

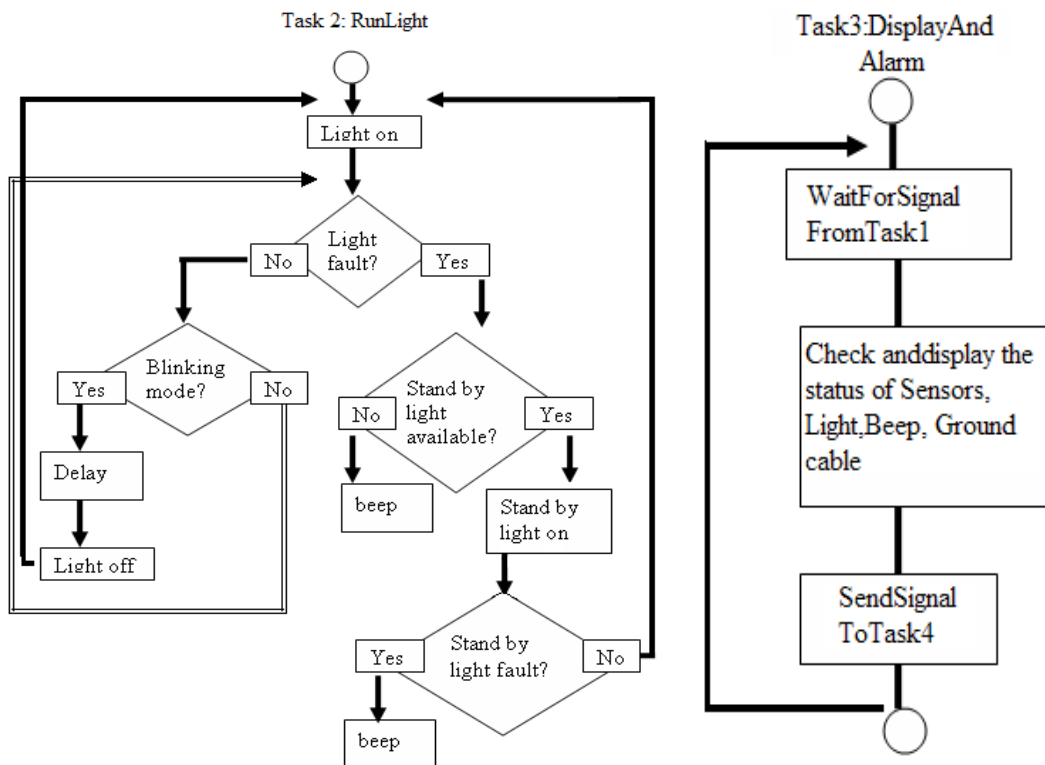


Figure 12: Task 2 and Task 3



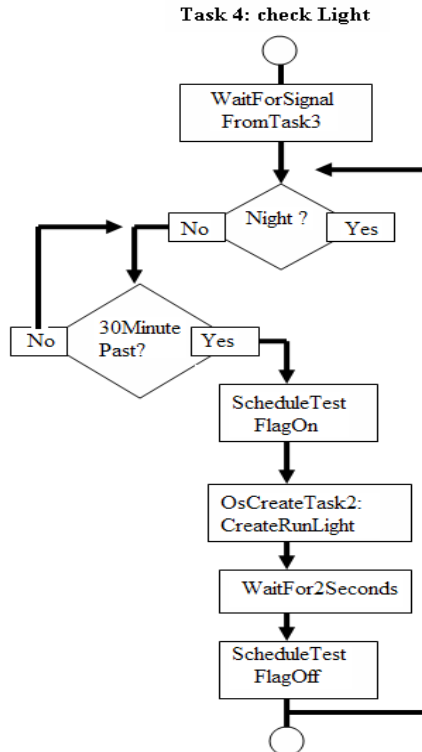


Figure 13: Task 4



Figure 14: Controller casing

## VIII. CONCLUSION

This system is very intelligent if any abnormalities occur such as grounding wire problem, LED intensity down, or problem regarding day-night sensor it will give alarm and notify about that fault in the LCD panel. During day time the system is able to test automatically itself in every 30 minutes and display the self test result in the LCD module. The system has been designed and developed in digital laboratory of East West University, Bangladesh and is being used by well known telecom operators in Bangladesh and is found to be in good

agreement as well. It is worth mentioning here that the total cost of developing the system is approximately US\$ 130 only, which is much lower compared to the devices used for the same purpose.

## REFERENCES

- [1] <http://www.icao.int/atb/sfbranch/index.asp>.
- [2] Mehedi Al Emran, "A Microcontroller Based System for Determining Instantaneous Wind Speed and Direction Using Optical Sensor" *8th International Conference on Computer and Information Technology (ICCIT)*, 2005.
- [3] Soteris A. Kalogirou, "Design and construction of a one-axis sun-tracking system" *Solar Energy*, Volume 57, Number 6, December 1996, pp. 465-469(5).
- [4] Johnston, K.D, "An opto-isolator interface circuit designed for testability", *Sch. Of Electr., Electron. & Syst. Eng., Univ. Coll. of Cardiff*; <http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel3/2101/5415/00209506.pdf?temp=x>.
- [5] <http://www.chris.org/Modifications/reverse-polarity-protection.html>.
- [6] Deitel, H.M., *Operating Systems*, second edition, Addison-Wesley Publishing Company, 1990.
- [7] Ripps, David, *A Guide to Real-Time Programming*, Englewood Cliffs, N.J, Prentice Hall, 1988.
- [8] Allworth, S.T., *Introduction to Real-Time Software Design*, Springer-Verlag Inc., New York.
- [9] <http://www.keil.com/rtx51tiny/>.
- [10] <http://www.keil.com/c51/c51.asp>.

# Low-Cost Realization of Toffoli Gate for the Low-Cost Synthesis of Quantum Ternary Logic Functions

Md. Mehedi Hasan

Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh, Email: mehedibuet@gmail.com

**Abstract** — Reversible quantum computer system is one of the best choices for future computer systems. Multiple-valued logic especially ternary logic is a good candidate for the realization of reversible quantum computer. An efficient logic synthesis mechanism is essential for the low-cost realization. Toffoli gate is an important gate for quantum logic synthesis. It is the basic element for the Galois Field Sum of Product (GFSOP) expression based logic synthesis mechanism. So, for low-cost realization of any ternary logic function, a low-cost implementation of Toffoli gate is very necessary. This paper shows a low-cost, practically realizable, and efficient realization of 3-qutrit ternary Toffoli gate by using ion-trap realizable Muthukrishnan-Stroud gate. This realization is more efficient and less costly than other realizations.

**Index Terms** — Circuit synthesis, logic design, logic functions, minimal realization, multivalued logic.

## I. INTRODUCTION

Quantum computing technology is one of the most promising technologies for future computer systems [1]. As the quantum computer is based on the reversible logic, the power dissipation will be extremely low compared to the counterpart CMOS computer [2]. This computer is able to perform vast amount of computation [1], [3] and also able to solve some of the most unfeasible problems in computer science such as large integer factorization, discrete logarithms etc [4].

For realization of quantum circuits, three-valued logic (ternary logic) is a very good choice as the amount of information per bit for ternary logic is more than that for two-valued (binary) logic. For these reasons, ternary quantum logic synthesis is a very attractive research area in the recent years [5]-[15].

Ternary quantum logic functions can be realized as the network of Toffoli gates. Some synthesis algorithms or methods [5]-[7] are proposed for the realization of functions by using Toffoli gates. Toffoli gates can be realized by using the primitive (basic) gates. The only primitive gate which is physically realizable is Muthukrishnan-Stroud gate (MS gate) [8]. MS gate brings a great momentum in multiple-valued quantum research. Another gate called Generalized Ternary gate (GTG) is proposed in [9] which can be built from a subset of MS gates [15]. Therefore, Toffoli gate which is a very important gate for realization of any ternary logic circuit

must be realized by using physically realizable primitive gates such as MS gates and GTGs. The number of the primitive gates should be as small as possible to reduce the circuit cost.

In literature, several works such as [5], [10], [11] have been done to realize the ternary Toffoli gate using MS gates or GTGs. Khan et al. [5] presents a realization that uses 10 GTGs which is equivalent to 32 MS gates. The work [10] shows a realization which requires 8 GTGs. Another work [11] presents the equivalent MS gate realization of the GTG realization proposed in [10] which requires 42 MS gates.

To reduce the implementation cost, in this paper, we have proposed a very low-cost realization of Toffoli gate. In our design, we have used only 20 MS gates for realization which reduces the cost by 37.5% compared to [5] and 52% compared to both [10] and [11].

The rest of the paper is organized as follows. Section II presents the background on the reversible logic and different quantum gates. Section III presents the proposed realization of ternary Toffoli gate and also presents the explanation of the operation. Finally, section IV provides the conclusion.

## II. BACKGROUND

An  $n$ -input and  $n$ -output multiple-valued logic function is reversible if it maps each input into a unique output. So  $r^n!$  reversible functions is possible for the  $r$ -valued logic. For ternary system,  $3^n!$  reversible functions are possible. But all of them are not realized as a single gate. Muthukrishnan-Stroud gate, Generalized Ternary gate (GTG), Feynman gate, and Toffoli gate are some well-known gates which are used for synthesis.

### A. Muthukrishnan- Stroud gate

Muthukrishnan and Stroud proposed a group of 2-qutrit (quantum digit) primitive ternary gates which are ion-trap realizable [8]. Basically the success in the realization of these gates gives the hopes of building practical quantum computer. This group of gates is known as Muthukrishnan-Stroud (MS) gate.

Fig. 1 shows the MS gate for ternary system. Here,  $A$  and  $B$  are the controlling input and the controlled input respectively while  $P$  and  $Q$  are the outputs. If the logic

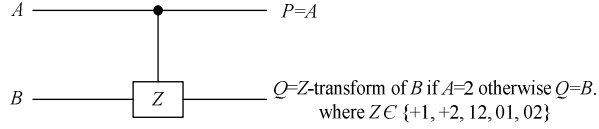


Fig. 1. Quantum ternary MS gate.

TABLE I  
VALUE AFTER QUANTUM Z-TRANSFORM.

| Quantum Z-Transform | Q when B= |   |   |
|---------------------|-----------|---|---|
|                     | 0         | 1 | 2 |
| Buffer              | 0         | 1 | 2 |
| +1                  | 1         | 2 | 0 |
| +2                  | 2         | 0 | 1 |
| 12                  | 0         | 2 | 1 |
| 01                  | 1         | 0 | 2 |
| 02                  | 2         | 1 | 0 |

value of  $A$  equals 2, the value of  $Q$  will be the quantum Z-transform of the input  $B$  and if  $A$  is not equal to 2, then  $Q=B$ . Different quantum Z-transforms are shown in Table I. For example, if Z-transform is +2,  $A=2$  and  $B=0$ ,  $Q$  will be 2.

### B. Generalized Ternary Gates (GTGs)

Fig. 2 shows a Generalized Ternary Gate [9]. Here, input  $A$  is the selector input in a conceptual multiplexer where  $B$  is the input which is processed by Z-transform shown in Table I. When  $A=0$ , the output  $P$  is the  $L$ -transform of the input  $B$ . Similarly, when  $A=1$ ,  $P$  is the  $M$ -transform of  $B$  and when  $A=2$ ,  $P$  is the  $N$ -transform of the  $B$  where  $L, M, N \in \{+1, +2, 12, 01, 02, \text{no-transform or buffer}\}$ . Depending on the six possible transforms for the each position of  $L, M, N$ , there are  $6^3$  or 216 GTGs possible.

However, GTGs can be realized by using MS gates. Fig. 3 shows the realization by using MS gates. Here, if  $A=0$ , the controlling input of the MS gate with  $N$ -transform is 0, so  $N$ -transform is not applied.  $M$ -transform is not also applied as the controlling input is 1. But, as the controlling input of the MS gate with  $L$ -transform is 2,  $L$ -transform is applied and the output  $Q$  becomes the  $L$ -transform of  $B$ . Similarly, if  $A=1$ , the controlling inputs of the MS gates are 1, 2, and 0 with  $N$ -transform,  $M$ -transform, and  $L$ -transform respectively. So  $M$ -transform is applied on input  $B$ . When  $A=2$ , the  $N$ -transform is applied on the input  $B$  and the output  $Q$  becomes the  $N$ -transform of  $B$ .

If  $L$ -transform is *buffer* or *no-transform* (for which the input and the output is the same), there is no need of MS gate for the  $L$ -transform. So, this two-transform GTG gate

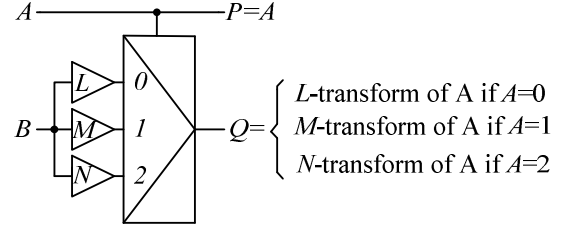


Fig. 2. Generalized Ternary gate

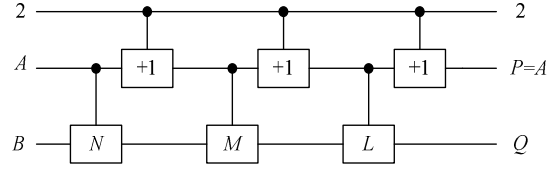


Fig. 3. Realization of Generalized Ternary gate

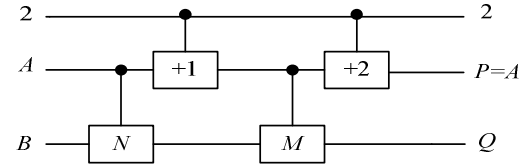


Fig. 4. Realization of GTG with two-transform (L is buffer)

(where two transforms such as  $M$  and  $N$  are non-buffer) can be realized by 4 MS gates (Fig. 4). Similarly, the two-transform GTG where  $M$ -transform is buffer can be realized by using 4 MS gates. And the two-transform GTG with buffer in  $N$ -transform can be realized by using 5 MS gates. Again, one-transform GTG with  $L$ -transform (where  $M$  and  $N$  are buffer) can be realized by using 3 MS gates. One-transform GTG with  $M$ -transform can be realized by using 3 GTGs. And one-transform GTG with  $N$ -transform needs only 1 GTG. This conversion technique is used in the work [11] to convert the GTG realization of ternary Toffoli gate proposed in [10].

### C. Ternary Feynman Gates

Fig. 5 shows a ternary Feynman gate where  $A$  is the controlling input and  $B$  is the controlled input. The output  $P$  is equal to  $A$  and the output  $Q$  is equal to the Galois Field 3 summation (GF3 sum) of  $A$  and  $B$ . Ternary Feynman gate can be realized by using 4 MS gates[11].

### D. Ternary Toffoli Gates

Fig. 6 shows a ternary Toffoli gate. Here, the inputs are  $A, B$ , and  $C$ . The output  $P$  and  $Q$  are equal to the inputs  $A$  and  $B$  respectively. The output  $R$  is equal to  $A.B+C$  where



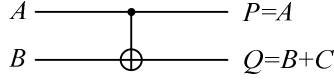


Fig. 5. Ternary Feynman gate

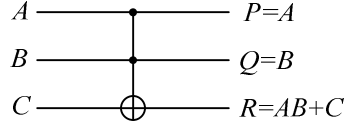


Fig. 6. Ternary Toffoli gate

. and + are the operators for Galois Field 3 (GF3) multiplication and addition respectively. The truth table for the ternary Toffoli gate are shown in Table II.

### III. REALIZATION OF TERNARY TOFFOLI GATE

Fig. 7 shows the proposed realization of ternary Toffoli gate whose symbol is shown in Fig. 6. The operation of the ternary Toffoli gate can be explained based on the understanding of MS gate as discussed in section II. Table III shows the logic conditions of the different nodes for all  $3^3 = 27$  types of inputs. The table also verifies the operation of the realization for all inputs. It also shows that for every input, the realization works correctly. To explain the operation, let us consider an example where  $A=1$ ,  $B=2$ , and  $C=1$ . Here,  $A_1$  becomes 2 due to +1-transform on the input  $A=1$  ( $1+1=2$ ) and the controlling input 2. Similarly,  $A_2$  becomes 0 due to +1-transform on the input  $A_1=2$  ( $2+1=0$ ). By the same way,  $A_3$  and  $A_4$  becomes 2 and 1 respectively.  $A_4$  is equal to  $A$ . For the line of input  $B$ ,  $B_1$  becomes 2 due to 01-transform of  $B=2$ . Then as the controlling input of next 12-transform is  $A_2=1$ , so the output  $B_2=B_1=2$ .  $B_3$  becomes 1 due to 12-transform of  $B_2=2$ . Similarly,  $B_4, B_5, B_6, B_7, B_8, B_9, B_{10}, B_{11}, B_{12}, B_{13}$  and  $B_{14}$  become 2, 2, 2, 0, 2, 2, 1, 2, 2, 0 and 2 respectively. Here  $B_6=2=B$  and  $B_{14}=2=B$ . From the node  $B_9$ , we get  $A.B=2$ . Now, let us consider the line for the input  $C$ . As  $C=1$ ,  $C_1$  becomes 0 due to the +2-transform of  $C$  where the controlling input is  $B_9=AB=2$ . Then,  $C_2$  becomes 0 as the controlling input of the next MS gate is  $B_{10}=1$  and  $C_1=0$ .  $C_2$  indicates the output  $A.B+C$ .

By the same way, the operation of the proposed Toffoli gate is verified for all 27 types of inputs and the logic condition of every node for each input is illustrated in Table III. For every input, the output we get is same as we have shown in the truth table of Table II.

In this design, we have used only 20 MS gates where the work [11] uses 42 MS gates for realization. The work [5] uses 10 GTGs and one Feynman gate. Among them,

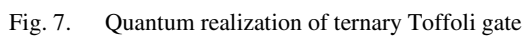
TABLE II  
TRUTH TABLE FOR TERNARY TOFFOLI GATE

| A | B | C | AB | A.B+C |
|---|---|---|----|-------|
| 0 | 0 | 0 | 0  | 0     |
| 0 | 0 | 1 | 0  | 1     |
| 0 | 0 | 2 | 0  | 2     |
| 0 | 1 | 0 | 0  | 0     |
| 0 | 1 | 1 | 0  | 1     |
| 0 | 1 | 2 | 0  | 2     |
| 0 | 2 | 0 | 0  | 0     |
| 0 | 2 | 1 | 0  | 1     |
| 0 | 2 | 2 | 0  | 2     |
| 1 | 0 | 0 | 0  | 0     |
| 1 | 0 | 1 | 0  | 1     |
| 1 | 0 | 2 | 0  | 2     |
| 1 | 1 | 0 | 1  | 1     |
| 1 | 1 | 1 | 1  | 2     |
| 1 | 1 | 2 | 1  | 0     |
| 1 | 2 | 0 | 2  | 2     |
| 1 | 2 | 1 | 2  | 0     |
| 1 | 2 | 2 | 2  | 1     |
| 2 | 0 | 0 | 0  | 0     |
| 2 | 0 | 1 | 0  | 1     |
| 2 | 0 | 2 | 0  | 2     |
| 2 | 1 | 0 | 2  | 2     |
| 2 | 1 | 1 | 2  | 0     |
| 2 | 1 | 2 | 2  | 1     |
| 2 | 2 | 0 | 1  | 1     |
| 2 | 2 | 1 | 1  | 2     |
| 2 | 2 | 2 | 1  | 0     |

there are 2 two-transform GTGs where  $L$  is buffer, 2 two-transform GTGs where  $N$  is buffer, 4 one-transform GTG with  $N$ -transform, 2 one-transform GTGs with  $L$ -transform, and 1 Feynman gate. So, according to the discussion of section II(B) and II(C), the total cost of realization of [5] is  $2 \times 4 + 2 \times 5 + 4 \times 1 + 2 \times 3 + 1 \times 4 = 32$  where cost is represented by the number of the primitive MS gates. Therefore, the proposed design reduces the realization cost by 52% compared to [11] and by 37.5% compared to [5].

### IV. CONCLUSION

In this paper, a new realization of ternary Toffoli gate is proposed. This realization reduces the realization cost significantly. As Toffoli gate is the basic element of realizing any ternary logic function for quantum computer, the reduction of the realization cost of Toffoli gate will reduce the total realization cost of the logic circuits significantly. So, this realization of Toffoli gate is more suitable for practical realization than other realizations.



| A | B | C | A <sub>1</sub> | A <sub>2</sub> | A <sub>3</sub> | A <sub>4</sub> | B <sub>1</sub> | B <sub>2</sub> | B <sub>3</sub> | B <sub>4</sub> | B <sub>5</sub> | B <sub>6</sub> | B <sub>7</sub> | B <sub>8</sub> | B <sub>9</sub> | B <sub>10</sub> | B <sub>11</sub> | B <sub>12</sub> | B <sub>13</sub> | B <sub>14</sub> | C <sub>1</sub> | C <sub>2</sub> |   |
|---|---|---|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|----------------|---|
| 0 | 0 | 0 | 1              | 2              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0              | 0              | 0 |
| 0 | 0 | 1 | 1              | 2              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0              | 1              | 1 |
| 0 | 0 | 2 | 1              | 2              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0              | 2              | 2 |
| 0 | 1 | 0 | 1              | 2              | 1              | 0              | 1              | 2              | 1              | 2              | 1              | 1              | 1              | 0              | 0              | 0               | 0               | 0               | 1               | 1               | 0              | 0              | 0 |
| 0 | 1 | 1 | 1              | 2              | 1              | 0              | 1              | 2              | 1              | 2              | 1              | 1              | 1              | 0              | 0              | 0               | 0               | 0               | 1               | 1               | 1              | 1              | 1 |
| 0 | 1 | 2 | 1              | 2              | 1              | 0              | 1              | 2              | 1              | 2              | 1              | 1              | 1              | 0              | 0              | 0               | 0               | 0               | 1               | 1               | 2              | 2              | 2 |
| 0 | 2 | 0 | 1              | 2              | 1              | 0              | 2              | 1              | 2              | 1              | 2              | 2              | 2              | 2              | 0              | 0               | 0               | 2               | 2               | 2               | 0              | 0              | 0 |
| 0 | 2 | 1 | 1              | 2              | 1              | 0              | 2              | 1              | 2              | 1              | 2              | 2              | 2              | 2              | 0              | 0               | 0               | 2               | 2               | 2               | 1              | 1              | 1 |
| 0 | 2 | 2 | 1              | 2              | 1              | 0              | 2              | 1              | 2              | 1              | 2              | 2              | 2              | 2              | 0              | 0               | 0               | 2               | 2               | 2               | 2              | 2              | 2 |
| 1 | 0 | 0 | 2              | 0              | 2              | 1              | 1              | 1              | 2              | 1              | 1              | 0              | 2              | 2              | 0              | 0               | 0               | 2               | 2               | 0               | 0              | 0              | 1 |
| 1 | 0 | 1 | 2              | 0              | 2              | 1              | 1              | 1              | 2              | 1              | 1              | 0              | 2              | 2              | 0              | 0               | 0               | 2               | 2               | 0               | 1              | 2              | 2 |
| 1 | 0 | 2 | 2              | 0              | 2              | 1              | 1              | 1              | 2              | 1              | 1              | 0              | 2              | 2              | 0              | 0               | 0               | 2               | 2               | 0               | 2              | 0              | 0 |
| 1 | 1 | 0 | 2              | 0              | 2              | 1              | 0              | 0              | 0              | 0              | 0              | 1              | 1              | 1              | 1              | 2               | 1               | 1               | 1               | 1               | 0              | 2              | 2 |
| 1 | 1 | 1 | 2              | 0              | 2              | 1              | 0              | 0              | 0              | 0              | 0              | 1              | 1              | 1              | 1              | 2               | 1               | 1               | 1               | 1               | 1              | 0              | 0 |
| 1 | 1 | 2 | 2              | 0              | 2              | 1              | 0              | 0              | 0              | 0              | 0              | 1              | 1              | 1              | 1              | 2               | 1               | 1               | 1               | 1               | 2              | 1              | 1 |
| 1 | 2 | 0 | 2              | 0              | 2              | 1              | 2              | 2              | 1              | 2              | 2              | 2              | 0              | 2              | 2              | 1               | 2               | 2               | 0               | 2               | 2              | 0              | 0 |

TABLE II  
CONTINUED

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 2 | 0 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 0 | 2 | 2 | 1 | 2 | 2 | 0 | 2 | 0 | 1 |
| 1 | 2 | 2 | 2 | 0 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 0 | 2 | 2 | 1 | 2 | 2 | 0 | 2 | 1 | 2 |
| 2 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 2 | 0 | 2 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 |
| 2 | 1 | 0 | 0 | 1 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 |
| 2 | 1 | 1 | 0 | 1 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 0 | 0 |
| 2 | 1 | 2 | 0 | 1 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| 2 | 2 | 0 | 0 | 1 | 0 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 0 | 1 |
| 2 | 2 | 1 | 0 | 1 | 0 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 |
| 2 | 2 | 2 | 0 | 1 | 0 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 0 |

#### REFERENCES

- [1] M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
- [2] C. H. Bennet, *Logical reversibility of computation*, IBM Res. Develop. 17 (1973) 525-532.
- [3] M. Hirvensalo, *Quantum Computing*, Springer Verlag, 2001.
- [4] P. Shore, "Algorithms for quantum computation: discrete logarithms and factoring," *Proc. 35<sup>th</sup> Ann. Sym. Foundations of Computer Science*, Santa Fe, NM, USA, pp. 124-134, November 1994.
- [5] M. H. A. Khan, M. A. Perkowski, M. R. Khan, and P. Kerntopf, "Ternary GFSOP minimization using Kronecker Decision Diagram and their synthesis with quantum cascades," *J. Multiple-valued Logic and Soft Computing*, vol. 11, pp. 567-602, 2005.
- [6] D. M. Miller, G. Dueck, and D. Maslov, "A synthesis method for MVL reversible logic," *Proc. 34<sup>th</sup> IEEE Int. Symp. Multiple-valued Logic*, Toronto, Canada, pp. 74-80, 2004.
- [7] D. Maslov, G. Dueck, and D. M. Miller, "Toffoli network synthesis with templates," *IEEE Transaction on Comp. Aided Design of integrated circuits and systems*, vol. 24, no. 6, pp. 807- 817, June 2005.
- [8] A. Muthukrishnan and C. R. Stroud, "Multivalued logic gates for quantum computation," *Physical Review A*, vol. 62, pp. 052309/1-8, 2000.
- [9] M. Perkowski, A. Al-Rababi, and P. Kerntopf, "Multiple-valued quantum logic synthesis," *Proc. 2002 International symposium new paradigm VLSI computing*, Sendai, Japan, pp. 41-47, 12-14 December 2002.
- [10] M. H. A. Khan, "Quantum realization of ternary Toffoli gate," *Proc. ICECE 2004*, Dhaka, Bangladesh, pp. 264-266, 28-30 December 2004.
- [11] M. H. A. Khan, "Quantum realization of ternary Toffoli gate using ion-trap realizable Muthukrishnan-Stroud primitive gates," *Proc. ICCIT 2004*, Dhaka, Bangladesh, 26-28 December 2004.
- [12] D. M. Miller, G. Dueck, and D. Maslov, "Synthesis of quantum multiple-valued circuits," *J. Multiple-valued logic and Soft computing*, vol. 12, 2006.
- [13] M. H. A. Khan, "Quantum Realization of Ternary Adder Circuits," *Proc. 3<sup>rd</sup> Int. Conf. Electrical and Computer Engineering*, Dhaka, Bangladesh, pp. 141-144, 2004.
- [14] M. H. A. Khan and M. A. Perkowski, "Quantum ternary parallel adder/subtractor with partially-look-ahead carry," *J. System Architecture*, vol. 53, pp. 453-464, 2007.
- [15] N. Denler, B. Yen, M. Perkowski, and P. Kerntopf, "Synthesis of reversible circuits from a subset of Muthukrishnan-Stroud quantum multi-valued gates," *Proc. IWLS 2004*, Tamecula, California, 2-4 June 2004.

# Design and Implementation of Microprocessor Based Electronic Voting System

Muhammad Raisul Alam<sup>1</sup>, Md. Masum<sup>2</sup>, Md. Mostafizur Rahman<sup>3</sup>, Md. Arifur Rahman<sup>4</sup>

Universiti Kebangsaan Malaysia<sup>1</sup>, Shah Jalal University of Science and Technology<sup>2</sup>,  
Grameenphone Limited<sup>3</sup>, University of Dhaka<sup>4</sup>, Bangladesh.

**Abstract** — In this paper, we described the design and hardware implementation of an electronic voting machine. The main aim of this project is not to design a power efficient perfect device but is to design a mother device that can be adoptable to any recent technologies. The design of the machine and its usage is as simple as an illiterate common people can easily use it. A vote casting procedure is also proposed to use the machine where we prescribed a new procedure called DEMO vote casting which is a testing mechanism to test the device before vote casting and also at the run time. The vote casting system is almost similar to traditional system, only ballot box will be replaced by the device. The machine uses Voter ID to identify a valid voter and restrict multiple vote casting.

**Index Terms** — Election, Electronic Voting Machine, Electronic Voting System, Microprocessor, Vote

## I. INTRODUCTION

The Electronic Voting Machine is an electronic device that receives votes from the voters, stores the votes and counts the votes just after the vote casting process. Many other countries have developed Electronic devices to use in their Election system but those machines only reflect the aspects of their own country. To design such a machine for Bangladesh we have to consider the social facts, voting offences that influence our Election System. Our device considers no further change in the conventional vote casting system, only the ballot box will be replaced by the voting machine. It will partially automate the vote casting system. Its implementation will reduce most of the voting offences like ballot box hijacking and false voting. Higher security and testing mechanism are considered to ensure a free and fair election. The machine can be considered as a product which is customized to be used in third world countries like Bangladesh.

## II. LITERATURE REVIEW

Many countries around the world use electronic voting machine but those machines were developed considering their own country aspect [1]. Those designs are not suitable for developing countries because of the usage of high-tech technologies and cost per machine is too high to use them in national election in Bangladesh. Also while designing a voting machine for our country we have to consider literacy

rate, electricity availability, social facts which are not reflected in those machines.

We have to use the same type of machines that are used in developing countries. India introduced an electronic voting machine that reflects their social facts [2]. They used battery based machine using microprocessor which consists a control unit and a ballot unit. In Bangladesh there developed some computer and software based voting systems which are impossible to use in rural areas where no electricity is available. A microcontroller based voting machine [3] is also proposed but there is no feature of using Voter ID. No initial and run time checking procedure is proposed to check the functionality of the machine. In that system, a half minute delay is considered to prevent device hijacking and false voting that will slow down the process which can easily be solved using Voter ID.

## III. LIMITATIONS OF EXISTING ELECTION SYSTEM

- 1) The main limitation of our vote casting system is that it takes about two hours to count the votes. Sometimes the Authority has to recount the vote which is more time consuming.
- 2) It is also insecure because dishonest people can hijack the ballot box and can give false votes. There is no logical protection like password, voter verification etc.
- 3) The officers of the polling booth have to search the voter information from a list which is time consuming.

## IV. ELECTRONIC VOTING MACHINE

All the data found in the traditional ballot paper based voting system [4]-[6] have been analyzed and an optimized voting machine which will be fully adoptable to the traditional system is designed.

### A. Features

- 1) The Machine is fully adaptable to the existing System. The Symbol Panel acts like the ballot paper and the control panel acts like the ballot box.
- 2) It is more secured than ballot box. The presiding officer must activate the control panel by entering password. The control panel uses a removable storage to store program, database of the voters and candidates. It is easier to ensure

security of this small storage device than the bulky control panel.

3) It can accurately count the votes just after the vote casting process ended.

4) The machine has a DEMO vote casting feature to ensure each key in the symbol panel working properly. The increment of the total number of votes is displayed in the control panel and in an outdoor display which indicate that each key is functioning properly. This procedure can also be run during the vote casting time if the officer noticed that the total number of votes is not incrementing in the display or there is no successful vote casting sound.

5) Privacy in voting information is highly implemented. Neither election authorities nor anyone else can link any ballot to the voter who cast it.

6) NULL voting option is implemented. NULL vote means if a voter is not satisfied with the candidates who are participating in the election, then he can press the NULL key in the Symbol panel which means that he is not casting his vote to anyone. If in a region NULL vote is elected, then it means that the political parties have nominated those types of people who are disliked by the voters. So the Election Commission must rearrange an election again in that region where previously participated candidates cannot take a part again as a punishment and political parties must give nomination to eligible candidates. (It's a proposal)

7) We have found that maximum number of candidates in an election is not more than 22. Average number of candidates does not exceed 10 [4]. So, only 24 keys are placed in the panel for each post.

8) The presiding officer enters the voter ID in the control panel. If the voter has not cast his vote yet, the control panel automatically activates a symbol panel. If the voter has already cast his vote it displays a message on control panel and it does not activate the symbol panel. Thus it prevents false vote casting.

9) There will be an outdoor large display at the vote centre which will display total number of votes cast. Also, the control panel generates a loud sound after a successful vote casting. This will give the voter a mental satisfaction that the system is ok and the entire processing is going well.

### B. Schematic diagram

The Electronic Voting Machine contains a Control panel and one or two Symbol panels. The control panel is used by the presiding officer and assistant presiding officers of a polling booth.

1) The control panel: A display is used to show information to the officer. There are two status LEDs on the panel to show activity of the symbol panel. The panel consists of 10 numeric keys numbered from 0 to 9. These keys are used to enter password and voter ID. There are also some function keys like Start key, Test key, End key, Count Key, Enter key, Del key. (Fig. 1)

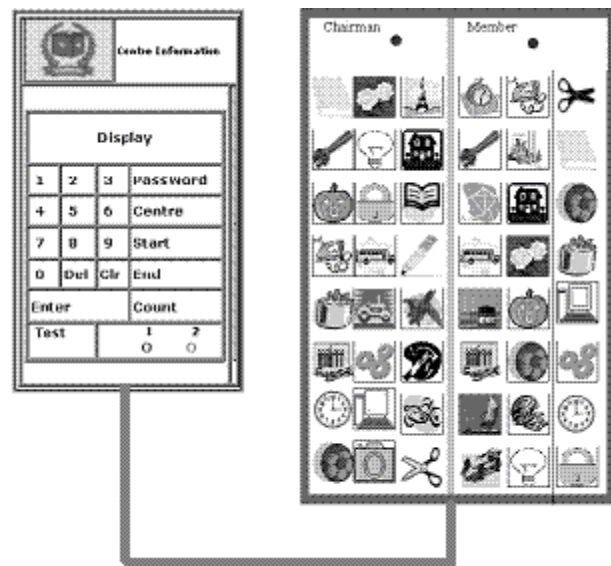


Fig. 1. Schematic Diagram of Electronic Voting Machine

2) Symbol Panel: The symbol panel is used by the voter to cast vote. Each switch is labelled with a symbol of a candidate (Fig. 1). It has two parts; each part indicates different post of the candidate. Each part has 24 keys which are mechanical keys with Encoders. So the total numbers of keys are 48. Each key represents a candidate. The panel is enabled by the control panel and disabled automatically after the voter has cast his vote.

3) Outdoor Display: An Outdoor large display will be used to show the total number of votes cast. The value is automatically incremented after a successful vote casting.

### C. Machine description

We used 8086 microprocessor to execute instructions. Several latches are used to latch address generated from microprocessor. Data buffers are used to control directions of data. Two EEPROMs are used as odd bank and even bank memory to store 16 bit data [7],[8]. These EEPROMs store the main operating system, acts like RAM during instruction execution and also store voting information. The EEPROMs are removable. An address decoder is used to enable display latch, control panel buffer and symbol panel buffer. The block diagram of the device is described in Fig. 2

In control panel and symbol panel key pad, encoders are used to encode key stroke. As we are using mechanical switch the circuit generates incorrect encoded value because of contact bouncing. A Mono stable Multi vibrator IC is used to prevent contact bounce. A Counter is used to indicate new data has arrived. Each keypad design is identical.

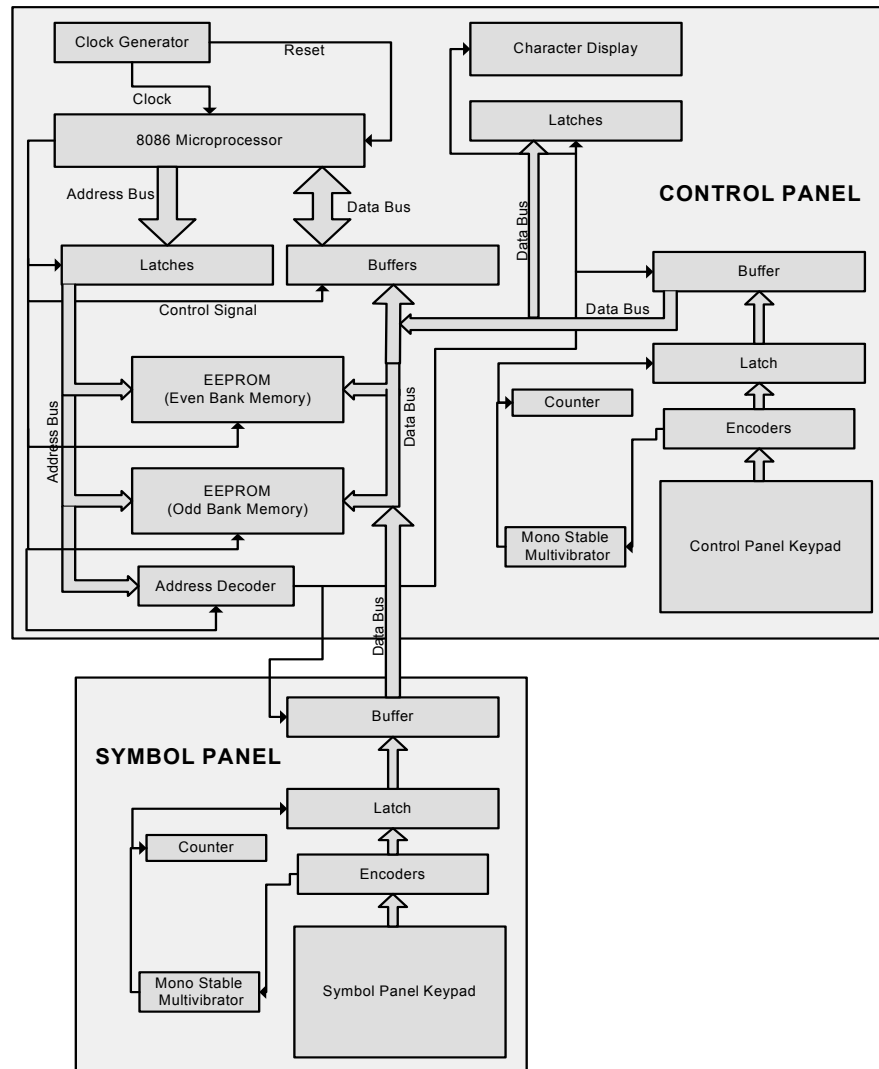


Fig. 2. Block diagram of Electronic Voting Machine

## V. ELECTRONIC VOTING SYSTEM

The electronic voting machine will introduce Electronic Voting System. The procedure in the voting poll slightly changes the traditional vote casting process.

### A. Voter ID

The Assistant presiding officer will insert the ID of the voter in the control panel. The machine will verify the voter according to the voter database. The machine will automatically activate one of the symbol panels if the voter has not given the vote yet.

### B. Vote casting procedure

The presiding officer starts the vote casting procedure. The assistant presiding officer and polling officers help him to run the procedure properly.

1) Start vote casting: At first, the Presiding Officer inserts the removable storage into the control panel. Then he powers on the control panel. The system boots automatically and waits for the Password of the Presiding Officer. After the successful entry of the password, the control panel is activated. Then he checks the centre name. After that, he runs DEMO vote casting process to ensure that the whole system is working properly. After that he starts Vote casting (Fig. 3).

2) DEMO vote casting process: DEMO vote casting process is a testing procedure which ensures the system is functioning properly. It is normally executed in the initial

phase of vote casting procedure. It can also be run to test functionality of Machine at any time.

To start DEMO process the Presiding officer presses the TEST button on the control panel. Then he presses every button on the Symbol panel and checks if it is displayed on the screen of the control panel. The control panel displays each symbol name and the number of DEMO vote of that symbol. Thus the Officer can determine which key is malfunctioning (Fig. 5).

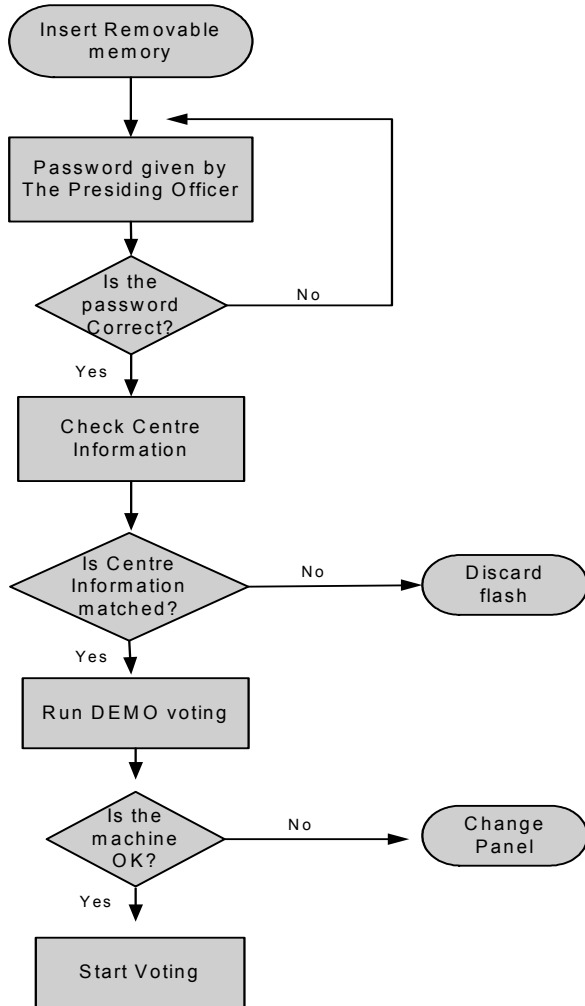


Fig. 3. Initializing and testing the machine before vote casting.

3) Vote casting procedure: The vote casting procedure starts when the presiding officer presses the START button. The polling officers manually verify the voter first. After that the Assistant Presiding Officer enters the ID of the voter in the control panel. If the voter has not cast his vote yet, the control panel activates one of the symbol panels. The voter is then sent to the Symbol panel. After the successful vote casting, the symbol panel generates a sound and automatically disabled (Fig. 4).

4) Stop vote casting: When time is finished the presiding officer end the voting process. The presiding officer has to provide password for this purpose.

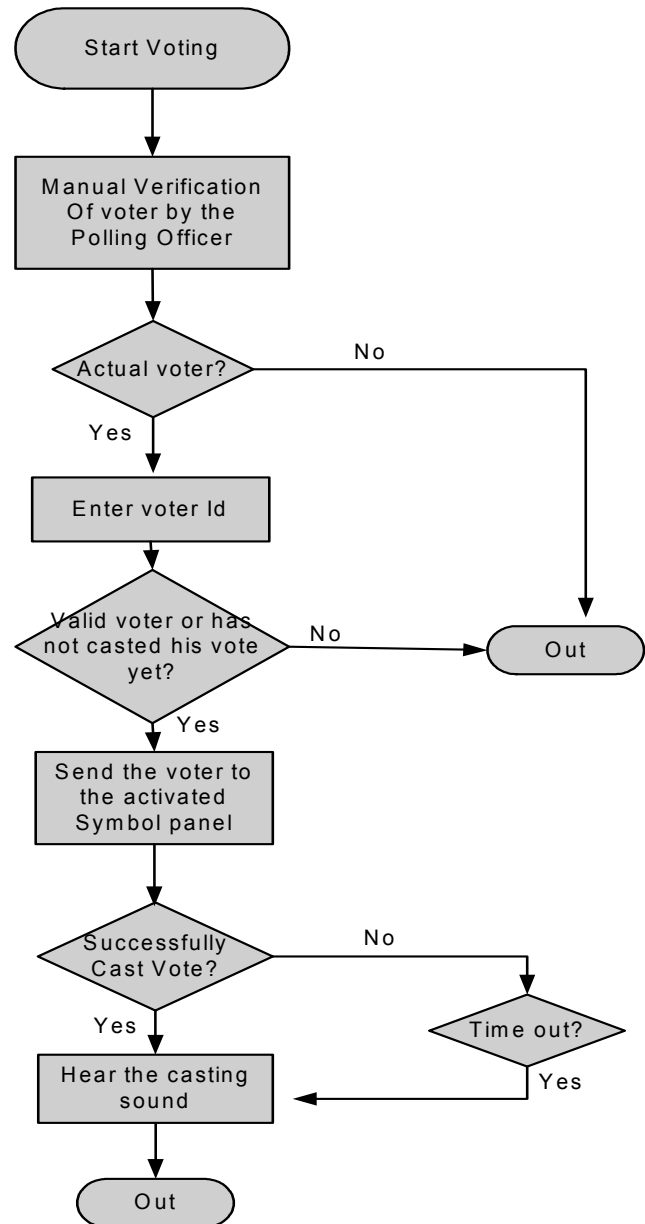


Fig. 4. Vote casting procedure

### C. Vote count

The presiding officer has to provide password to count the vote. After counting the vote no further vote casting process can be executed. All voting information is permanently stored in removable storage.

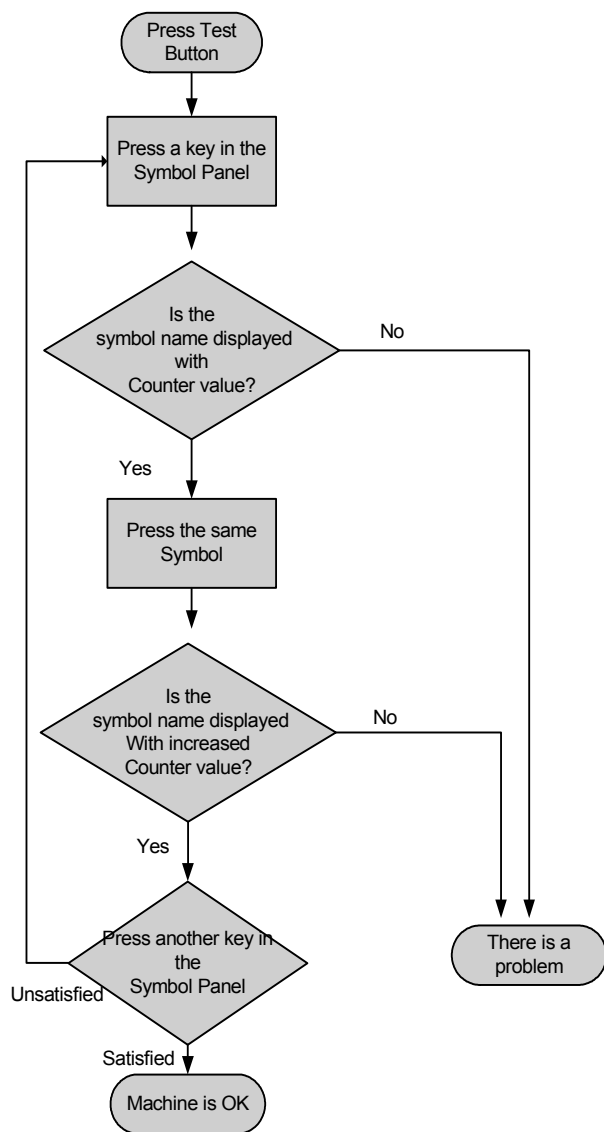


Fig. 5. The DEMO vote casting process.

## VI. COST

Total cost of the system is described in table I. But when large amount of device will be produced in industry the cost must be reduced.

## VII. PERFORMANCE ANALYSIS

Electronic Voting Machine enhances the performance of traditional voting system. It transforms the manual voting system into partially automated one reducing total time.

1) Implementation of the devices turns the Election system into automated system. The major manual process removed. Only a part of voter verification will remain manual. All other parts run automatically.

2) Voter ID number is considered. In our device the officer has to enter the ID manually using the control panel.

TABLE I  
TOTAL COST OF THE MACHINE

| PARTS NAME                                    | NUMBER | COST PER ITEM (TK) | TOTAL (TK)  |
|-----------------------------------------------|--------|--------------------|-------------|
| 8086 Microprocessor                           | 1      | 450                | 450         |
| EEPROM                                        | 2      | 200                | 400         |
| 16x2 Character LCD DISPLAY                    | 1      | 250                | 250         |
| Hex Inverter                                  | 1      | 15                 | 15          |
| Quad 2-Input OR Gate                          | 3      | 15                 | 45          |
| Clock Generator And Driver For 8086 Processor | 1      | 15                 | 15          |
| Monostable Multivibrators                     | 2      | 15                 | 30          |
| 3-To-8-Line Decoders                          | 1      | 15                 | 15          |
| Priority Encoders                             | 4      | 15                 | 60          |
| 3-State Octal Bus Transceiver                 | 4      | 15                 | 60          |
| 4-Bit Binary Counters                         | 2      | 15                 | 30          |
| 3-State Octal D-Type Transparent Latches      | 7      | 15                 | 105         |
| Crystal Oscillator                            | 1      | 20                 | 20          |
| Resistor                                      | 4      | 1                  | 4           |
| Capacitor                                     | 3      | 1                  | 3           |
| Switch                                        | 84     | 5                  | 420         |
| Others                                        |        |                    | 500         |
| <b>Total:</b>                                 |        |                    | <b>2422</b> |

In future, bar code reader may be implemented to automate this.

3) No need of electricity, it is a battery operated device. In case of power failure, the actual information will not be lost because the information is written immediately into the EEPROM. The officer can also be sure about successful vote casting by watching the incrimination of votes in the display of the control panel or hearing the sound of successful vote cast. He can take an immediate decision to recast the vote by the last voter.

4) The symbol panel is expandable up to 64 symbols. So it can accommodate 64 candidates.

5) In the last Parliament Election in 2001 there were 7, 50, 00, 656 voters and total cost was about 70 core taka [4]. Cost per voter was 9 taka .1, 45,000 ballot boxes needed. Each of our devices will not cost more than 2,500 taka. Total cost of 1, 45,000 machine is almost 35 core taka which is almost half of the previous cost



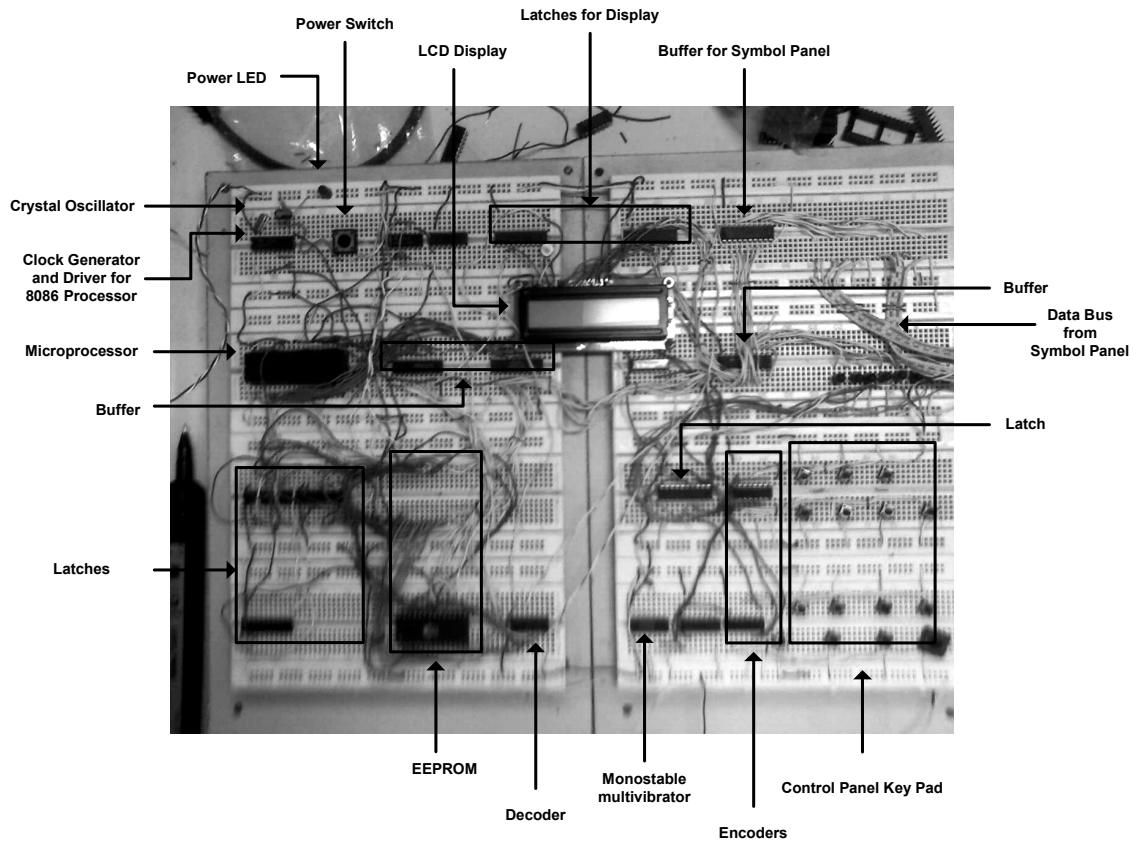


Fig. 6. Prototype of the Electronic Voting machine

- 6) The result can be published immediately. Traditional system takes minimum two hours.
- 7) A new testing system called DEMO voting procedure introduced which ensures accuracy and transparency of the system.
- 8) Another new concept, NULL voting option is added which may increase the freedom of the voters and decrease autocracy of political parties.
- 9) Removable storage improves security. Security is also implemented using password of the presiding officer. In conventional system no this type of security is implemented.

### VIII. CONCLUSION

Creating a user-friendly machine is a very complicated task because one has to design a machine which is more familiar with human behaviour. We proposed a procedure to use the machine which is very easy not only for the voter but also for the officers. In our device voter ID is used to verify the authenticity of the voter so it is usable with our recently introduced voter ID card. Although we were not able to fully implement all the features in software level but our proposed procedures can easily be implemented using Assembly Language by using the EEPROM programmer

which is developed by ourselves. The implementation cost of the machine is very low compared to total cost of an election. Also it improves security by using removable storage where data will be preserved for further usage.

### REFERENCES

- [1] Tadayoshi kohn, Adam stubblefield, Aviel d.. rubin ,dan s. Wallach,"Analysis of electronic voting system" ,*IEEE Symposium on Security and Privacy*, 2004.
- [2] From the official website of Election Commission of India: [http://www.eci.gov.in/Audio\\_VideoClips/presentation.asp](http://www.eci.gov.in/Audio_VideoClips/presentation.asp)
- [3] A.S.S.M.B Ullah, M. Hasan, K.M. Salim,"Enhancement Conventional Voting System by using Microcontroller based Electronic Voting Machine", *8th International Conference on Computer and Information Technology (ICCIT 2005)*,Dec,2005.
- [4] From the official website of Bangladesh Election Commission,<http://www.ecs.gov.bd>
- [5] Bangladesh Election Commission, *Presiding Officer abong Shohokari Presiding Officer der jonno nirdeshika*
- [6] Bangladesh Election Commission, *Returning Officer abong Shohokari Returning Officer der jonno nirdeshika*
- [7] Douglas V. Hall, *Microprocessors and Interfacing: Programming and Hardware*, McGraw-Hill; 2nd edition, 1991.
- [8] Barry B. Brey ,*The Intel Microprocessors*, Prentice Hall, 6<sup>th</sup> Edition.

# Genetic Algorithm Based Synthesis of Ternary Reversible/Quantum Circuit

Rashida Khanom<sup>a</sup>, Tahseen Kamal<sup>b</sup>, and Mozammel H. A. Khan<sup>a</sup>

<sup>a</sup>Department of Computer Science and Engineering, East West University

<sup>b</sup>Department of Electrical and Electronic Engineering, East West University  
43 Mohakhali, Dhaka-1212, Bangladesh

Email: r.khanom@yahoo.com, tahseen@ewubd.edu, mhakhan@ewubd.edu

**Abstract** — Reversible / Quantum circuits are believed to be one of the future computer technologies. In this paper, a Genetic Algorithm (GA) based synthesis of ternary reversible / quantum circuits using Muthukrishnan-Stroud gates is presented. The circuit generated by GA may contain redundant gates. We have used post GA reduction to eliminate these redundant gates. We have experimented with ternary half-adder circuit. The proposed GA converges for many combinations of crossover and mutation..

**Index Terms** — Reversible logic, half-adder, quantum circuit, post GA reduction..

## I. INTRODUCTION

Landaur [1] proved that binary logic circuit built using traditional irreversible gates leads inevitably to energy dissipation, regardless of the technology used to realize the gates. Zhirnov et al. [2] showed that power dissipation in future CMOS leads to impossible heat removal and thus the speeding up of CMOS devices will be impossible. Bennett [3] proved that for power not to be dissipated in a binary logic circuit, it is necessary that the circuit be built from reversible gates. A gate (or circuit) is reversible if it maps each input pattern to a unique output pattern. Bennett's theorem suggests that every future binary technology will have to use reversible gates in order to reduce power dissipation. This is also true for ternary technologies. Quantum technology is inherently reversible and is a promising technology for future computing systems [4].

Reversible/quantum circuit synthesis is more complicated than classical circuit synthesis. Reversible/quantum circuit synthesis is more complicated than classical circuit synthesis. In this paper, we have used GA with real – valued encoding of the chromosome using ordered 3-tuple [5]. GAs are very popular Soft Computing (SC) approaches for solving problems with no identified structure and high level of noise [5]-[7]. The reasons behind this popularity are, (i) a big problem space can be searched, (ii) the size of this search space can be moderated by parameters, (iii) a variety of new solutions can be produced and (iv) with long enough time solution can be obtained that is close to the optimal one. These advantages make GAs useful for synthesizing ternary reversible/ quantum circuit using cascade of

Muthukrishnan-Stroud gates, because the problem structure of such cascade is still unidentified and the search space itself is exponentially large.

## II. TERNARY REVERSIBLE GATE

There are three qudit (quantum digit) states in ternary quantum logic. These are  $|0\rangle$ ,  $|1\rangle$  and  $|2\rangle$ . Ternary 1\*1 gates operate on single qudit state. We represent a 1\*1 gate using the symbol of Fig. 1, where the operation  $y$  is one of the operations listed in Table I. In this paper the ternary operations i.e. the ternary 1\*1 gates are numbered from 0 to 5. For example, the operation +1 is numbered as gate number 1, the operation (12) is numbered as gate number 5 and etc.



where,  $y \in \{0, +1, +2, 12, 01, 02\}$

Fig. 1. Symbol of ternary 1\*1 gate family.

TABLE I  
TERNARY 1\*1 OPERATIONS

| A | A(0)<br>= A | A(+1)<br>= A+1 | A(+2)<br>= A+2 | A(01)<br>= 2A+1 | A(02)<br>= 2A+2 | A(12)<br>= 2A |
|---|-------------|----------------|----------------|-----------------|-----------------|---------------|
| 0 | 0           | 1              | 2              | 1               | 2               | 0             |
| 1 | 1           | 2              | 0              | 0               | 1               | 2             |
| 2 | 2           | 0              | 1              | 2               | 0               | 1             |

A ternary 2\*2 reversible/quantum gate has two inputs and two outputs. The two inputs are a controlling input ( $A$ ) and a controlled input ( $B$ ) as shown in Fig. 2. There are two parameters of the gate, the controlling value ( $x$ ) and the controlled value ( $y$ ). Let us use the notation  $\langle A, B, x, y \rangle$  to represent the gate.  $B$  is passed to the output if  $A \neq x$ , otherwise  $y$ th operation (Table I) is applied to  $B$ . Muthukrishnan et. al. [8] showed that  $\langle A, B, d-1, y \rangle$  gates are practically realizable in quantum ion trap technology for any  $d$ -valued logic. Therefore, according to [8], ternary  $\langle A, B, 2, y \rangle$  gates are practically realizable. Ternary  $\langle A, B, 2, y \rangle$  gates which is known as the Muthukrishnan-Stroud gates have been used and is shown

in Fig. 2, where  $x = 2$  and  $y$  is any operation from Table I.

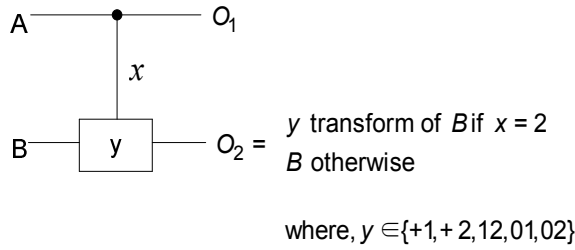


Fig. 2. Symbol of Ternary Muthukrishnan-Stroud gate family.

### III. PROPOSED GENETIC ALGORITHM

Half-adder circuit is an important medium scale building block for digital circuit design. As there is no direct method to construct a ternary circuit using Muthukrishnan-Stroud gates, we have tried with GA based synthesis. Use of GA will allow us to find appropriate combination of the gates that realize (perhaps optimally) ternary logic functions. GA is very popular Soft Computing approach for solving problems with no identified structure and because a variety of new solutions can be produced.

#### A. Problem Encoding

In this work, the generated circuits have four input and four output lines each. The  $3^n$  different combinations (here  $n = 2$  as there are two primary inputs) of 0, 1 and 2 are inputs to the two primary input lines. These are reproduced at the first two output lines. The other two input lines have any constant values 0, 1 or 2 for a particular circuit. The other two output lines are the Sum (S) and Carry out ( $C_o$ ) of the half-adder circuit.

Each of the columns in the circuit is represented by an ordered 3-tuple consisting of controlling line number, controlled line number and controlled gate number (from Table I). Using this notation one chromosome is shown in Fig. 3. In Fig. 3, each of the columns in the circuit is a gene of the chromosome. The circuit corresponding to this chromosome is shown in Fig. 4.

#### B. Fitness Components

In the proposed GA, we have used two fitness components-

##### 1) Output Truth Vector Fitness

-The extent of representation of sum and carry out by the chromosome.

##### 2) Cost Fitness

-Proportional to the number of gates of the resulting circuit.

#### C. Flowchart of Proposed GA

Fig. 5 shows flowchart of proposed GA.

#### D. Genetic Operators

Genetic operators which are used in the proposed GA are discussed below:

##### 1) Selection:

Two different chromosomes are selected randomly as parents.

##### 2) Crossover:

In the proposed GA, based on the crossover probability, one-point classical crossover has been used to generate two offspring from the two selected parents. A crossover point (between one and length-1) is randomly selected.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 333 | 111 | 204 | 225 | 122 | 112 | 132 | 001 | 123 | 223 | 021 | 005 | 033 | 121 | 021 | 021 | 235 | 132 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

Fig 3. Chromosome representing the circuit of Fig. 4.

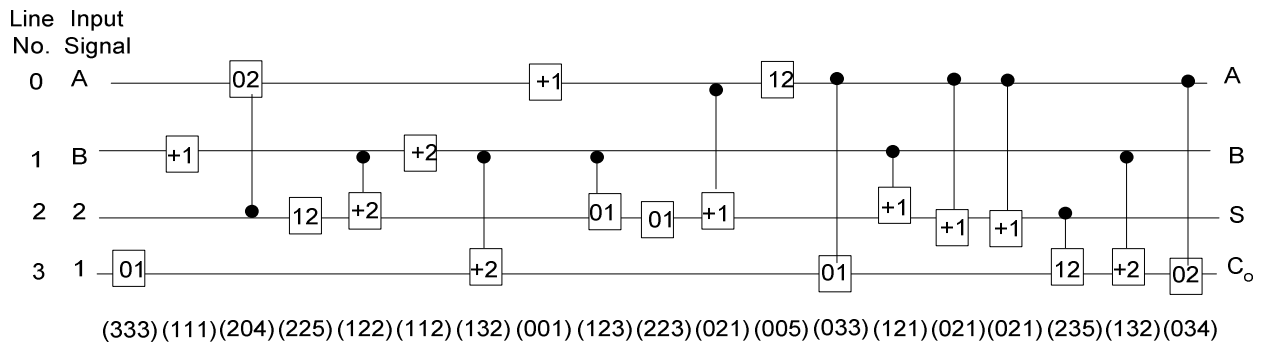


Fig. 4. Ternary half-adder using ternary Muthukrishnan-Stroud gate family (Numeric representation of a column (gene) is controlling line number, controlled line number and controlled gate number, respectively).

### 3) Mutation:

Mutation is done on every gene based on the mutation probability. In the mutation process, a gene is replaced by another randomly generalized gene.

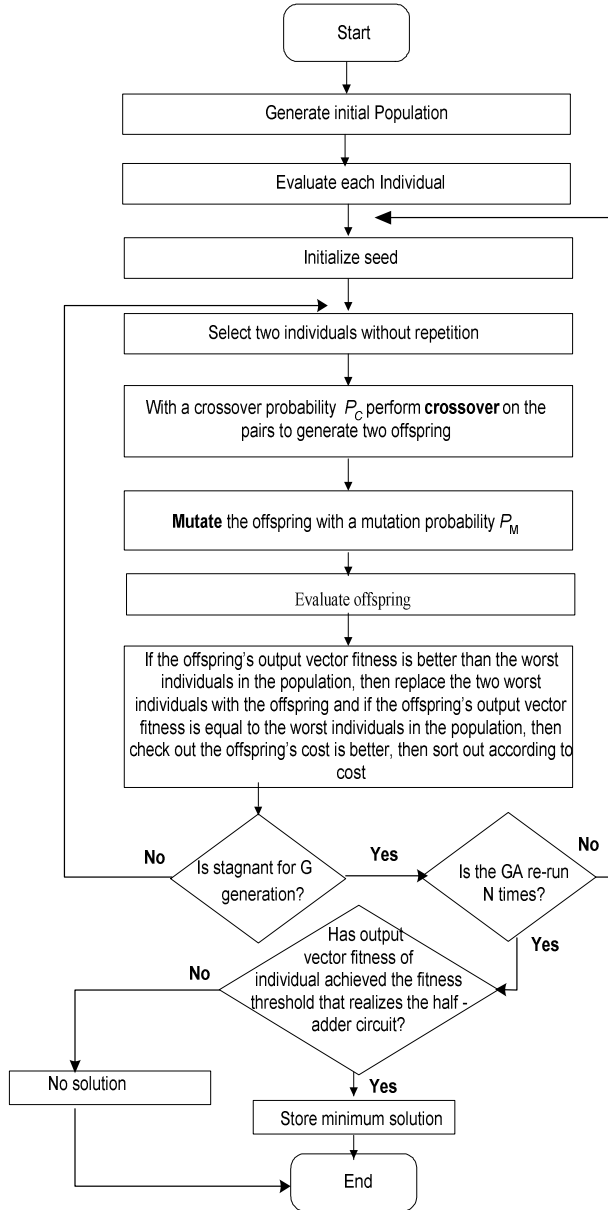


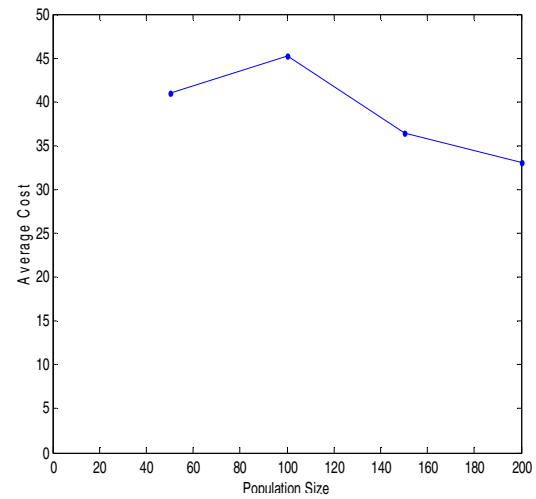
Fig. 5. Flowchart of the proposed GA.

In the proposed GA, if no replacement is done, then we say that the generation is stagnant. If consecutive G generation are stagnant, then we terminate the GA and re-run it with a new seed. The GA is re-run for N times.

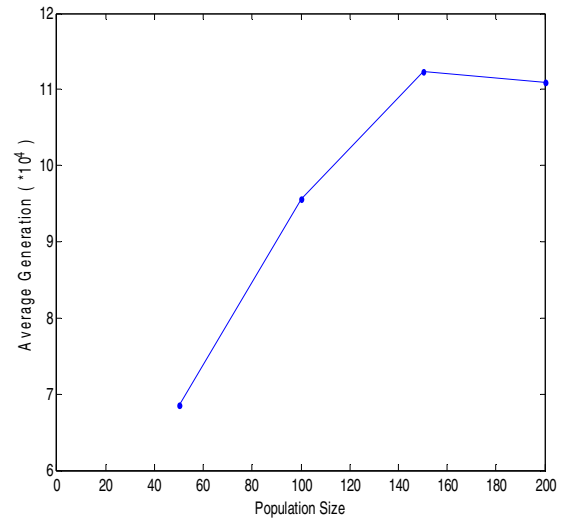
## IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed GA has been implemented in C programming language and experimented with the ternary half-adder circuit with different GA parameter combinations to see the effect of the parameters on the cost of the solution. The experimental results for ternary half-adder circuit are summarized in Table II.

The influences of different GA parameters on the performance of the GA are shown in Fig. 6 and 7. Fig. 6 shows the influence of population size on the average cost and average generation with fixed chromosome length. The influence of crossover probability on the cost with fixed population size and chromosome length is shown in Fig. 7.



(a)



(b)

Fig. 6. The influence of population size on (a) the average cost and (b) the average generation with fixed chromosome length, L=75.

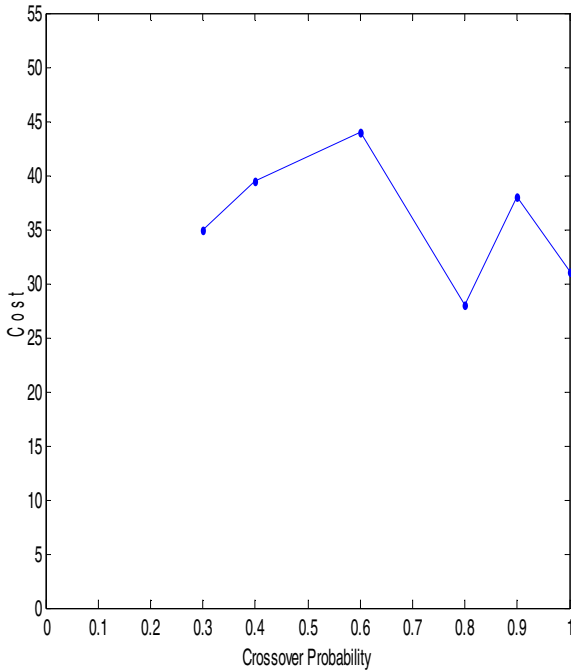


Fig. 7. The influence of crossover probability on the cost with fixed population size,  $P=150$ .

It can be observed from Fig. 6 that population size  $P=200$  gives the best solution of average cost 33. Fig. 7 depicts that the middle crossover probability results in maximum cost, the lower and higher probabilities result in lower costs. Though there is an outlier at the crossover probability 0.8.

## V. POST GA REDUCTION

The circuit generated by the GA may have redundant gates. Post GA reduction is used to eliminate such redundant gates. There are three ways this elimination is done.

1) If two single - qudit gates are in cascade in the circuit, then two single - qudit gates are replaced by one single - qudit gate as shown in Fig. 8. Here,  $O = (A + 1) + 1 = A + 2$ . There are other possible eliminations using this approach [9].

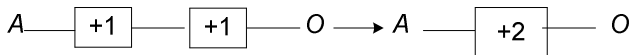


Fig. 8. Two single - qudit gates in cascades is equivalent to one single - qudit gate.

2) If two controlled gates are in cascade in the circuit, then two controlled gates are replaced by one controlled gate as shown in Fig. 9. Here,  $O_2 = (B + 1) + 1 = B + 2$ .

There is other possible elimination using this approach [9].

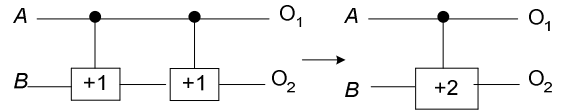


Fig. 9. Two controlled gates in cascade is equivalent to one controlled gate.

3) If there are two single - qudit gates along the control line of a controlled gate - one before the controlled gate and the other after the controlled gate, then the three gates can be reduced to one single - qudit gate before the controlled gate and one controlled gate as shown in Fig. 10.

When  $A = 0$  or  $A = 1$  (that means the controlled gate is inactive), then  $O_2 = (B + 1) + 1 = B + 2$ . When  $A = 2$  (that means the controlled gate is active), then  $O_2 = ((B + 1) + 1) + 1 = (B + 2) + 1$ . This implies that a +2 uncontrolled gate followed by a controlled gate with +1 operation will replace the combination. There are other possible eliminations also [9].

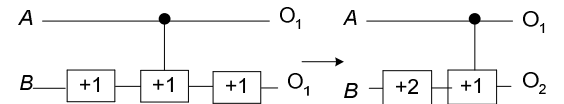


Fig. 10. One single - qudit gate, one controlled gate and one single - qudit gate in cascade is equivalent to one single - qudit gate and one controlled gate.

Ternary half-adder circuit generated by our GA with cost 19 is shown in Fig. 11. There are two controlled gates in cascade in the circuit that is equivalent to one controlled gate. Fig. 12 shows reduced circuit with cost 18.

## VI. CONCLUSION

There are many software methods that synthesize ternary reversible/quantum circuits [10]-[11]. In this work, GA based synthesis method for ternary reversible/quantum circuit using Muthukrishnan-Stroud gates has been proposed. Experimental results show that the proposed GA can produce quantum/reversible circuits. The ranges of the parameter values for which the GA converges and the effect of the GA parameter on the performance of the GA have been shown. Post GA reduction to eliminate redundant gates from the results obtained by the GA has been also shown.

TABLE II  
PARAMETER SETTINGS THAT CONVERGE TO SOLUTION

| Population Size(P) | Chromosome Length(L) | Crossover Probability( $P_c$ ) | Mutation Probability( $P_m$ ) | Cost | No. of Generation(G) |
|--------------------|----------------------|--------------------------------|-------------------------------|------|----------------------|
| 50                 | 75                   | 0.3                            | 0.06                          | 41   | 68532                |
| 100                | 50                   | 0.1                            | 0.06                          | 23   | 67677                |
|                    |                      | 0.1                            | 0.09                          | 23   | 73877                |
|                    |                      | 0.2                            | 0.07                          | 28   | 77726                |
|                    |                      | 0.2                            | 0.09                          | 21   | 80115                |
|                    | 75                   | 0.3                            | 0.04                          | 48   | 79097                |
|                    |                      | 0.5                            | 0.06                          | 46   | 105367               |
|                    |                      | 0.9                            | 0.02                          | 41   | 104897               |
|                    |                      | 0.9                            | 0.05                          | 46   | 92974                |
|                    | 100                  | 0.5                            | 0.01                          | 70   | 132580               |
|                    |                      | 0.7                            | 0.04                          | 51   | 95816                |
|                    |                      | 0.8                            | 0.05                          | 60   | 117895               |
|                    |                      | 1.0                            | 0.02                          | 68   | 89849                |
| 150                | 50                   | 0.1                            | 0.06                          | 22   | 69965                |
|                    |                      | 0.2                            | 0.02                          | 20   | 172623               |
|                    |                      | 0.4                            | 0.04                          | 20   | 95068                |
|                    |                      | 0.4                            | 0.07                          | 24   | 75146                |
|                    |                      | 0.6                            | 0.08                          | 28   | 88189                |
|                    | 75                   | 0.3                            | 0.04                          | 35   | 69707                |
|                    |                      | 0.4                            | 0.05                          | 29   | 95907                |
|                    |                      | 0.4                            | 0.08                          | 50   | 125083               |
|                    |                      | 0.6                            | 0.05                          | 44   | 119054               |
|                    |                      | 0.8                            | 0.04                          | 28   | 90039                |
|                    |                      | 0.9                            | 0.07                          | 38   | 147543               |
|                    |                      | 1.0                            | 0.06                          | 31   | 139506               |
| 200                | 50                   | 0.2                            | 0.03                          | 19   | 121755               |
|                    |                      | 0.4                            | 0.06                          | 28   | 87429                |
|                    | 75                   | 0.5                            | 0.02                          | 33   | 110974               |
|                    | 100                  | 0.7                            | 0.02                          | 67   | 96606                |
|                    |                      | 0.9                            | 0.04                          | 58   | 184552               |

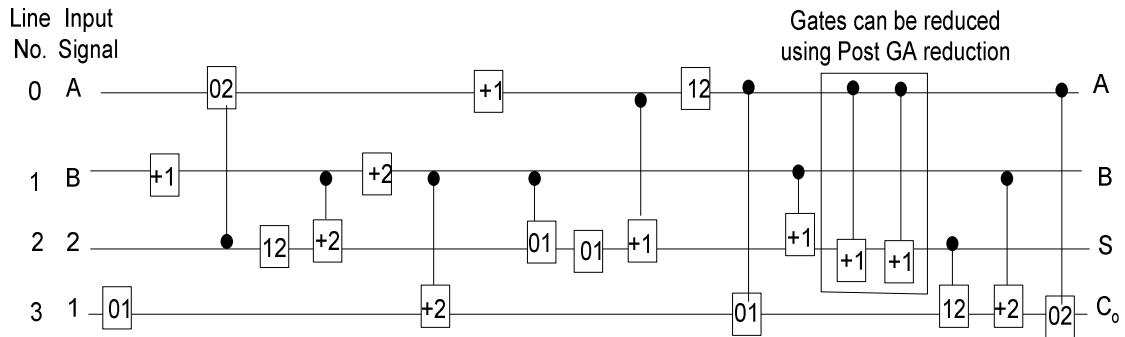


Fig. 11. Ternary half-adder circuit using ternary Muthukrishnan-Stroud gates generated by GA.

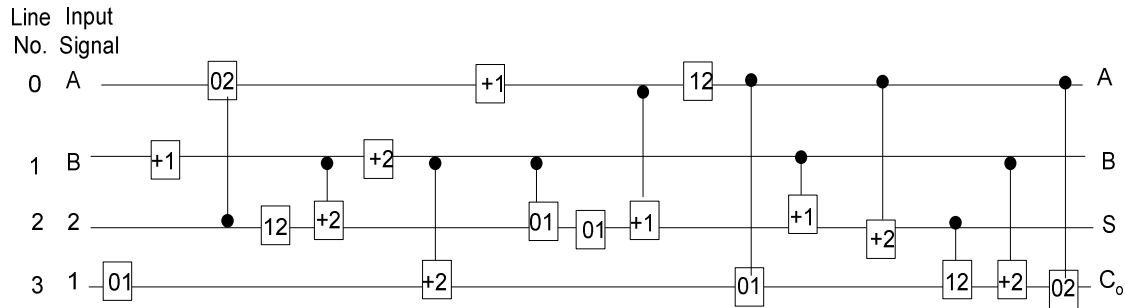


Fig. 12. Reduced ternary half-adder circuit using ternary Muthukrishnan-Stroud gates.

#### REFERENCES

- [1] R. Landauer, "Irreversibility and Heat Generation in the Computational Process", *IBM Journal of Research and Development*, 5, 1961, pp. 183-191.
- [2] V. V. Zhirnov, R. K. Kavin, J. A. Hutchby, and G. I. Bourianoff, "Limits to Binary Logic Switch Scaling - A Gedanken Model", *Proc. of the IEEE*, 91, no. 11, 2003, pp. 1934-1939.
- [3] C. H. Bennett, "Logical Reversibility of Computation", *IBM Journal of Research and Development*, 17, 1973, pp. 525-532.
- [4] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
- [5] Michalewicz Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs* (3<sup>rd</sup> edn). Springer.
- [6] Goldberg D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley.
- [7] Mazumder P., Rudnick E. M. (2002). *Genetic Algorithms for VLSI design, Layout & Test Automation*. Pearson Education Asia.
- [8] A. Muthukrishnan, and C. R. Stroud Jr., "Multivalued Logic Gates for Quantum Computation", *Physical review A*, vol. 62, No. 5, 05309/1-8, 2000.
- [9] R. Khanom, "Genetic Algorithm Based Synthesis of Ternary Reversible/Quantum Circuit", *MS thesis in preparation*, East West University, 2008.
- [10] M. H. A. Khan and M. Perkowski, "Genetic Algorithm based synthesis of multi – output ternary functions using quantum cascade of generalized ternary gates", *Proc. 2004 IEEE Congress on Evolutionary Computation (CEC 2004)*, Portland, Oregon, USA, June 2004.
- [11] Md. M. R. Khan, M. H. A. Khan, and Md. M. Akbar, "Evolutionary Algorithm Based Synthesis of Multi – output Ternary Reversible Circuits Using Quantum Cascades", *Proc. 7<sup>th</sup> Int. Symp. Representations and Methodology of Future Computing Technologies (RM 2005)*, Tokyo, Japan, September 2005.

# New Methodologies for High Level Modeling and Synthesis of Low Density Parity Check Decoders

Syed Mahfuzul Aziz and Sunil Sharma

School of Electrical & Information Engineering, University of South Australia  
Mawson Lakes, SA 5095, Australia

**Abstract**—Low density parity check (LDPC) codes are the error-correcting codes which offer huge advantages in terms of coding gain, throughput and power dissipation. Error correction algorithms are often implemented in hardware in order to ensure fast processing. The hardware implementation of LDPC decoders using traditional hardware description language (HDL) based approach is a complex and time consuming task. This paper investigates new high level approaches to design and synthesis of LDPC decoders using a combination of high level modelling tools. It compares the high level design approaches to traditional HDL-based approach. The results presented in this paper provide some useful insight into the high level design approaches, their efficiencies and possible future directions with a view to develop an efficient design and modelling framework for hardware implementation of complex LDPC decoders.

**Index Terms** – Error correction coding, digital systems, digital communication, logic design, FPGA.

## I. INTRODUCTION

Transaction of information through a practical communication system is not free from noise. Information may rather be corrupted in transit by noise in the channel [1]. Therefore it is of paramount importance for communication systems to have adequate means for the detection and correction of errors in the information received over communication channels. Turbo codes and LDPC (low density parity check) codes are most commonly used for error detection and correction nowadays [2]. Both of these codes provide coding gains [3] close to Shannon's limit [4]. LDPC codes however outperform turbo codes in terms of coding gain for large SNR [5, 6]. LDPC code of length 1 million with a coding rate (ratio of information bits to the sum of information and parity bits) of 0.5 and BER of  $10^{-6}$  provides a capacity which is only 0.13db from Shannon's limit [5]. Further advantages of using LDPC codes are given in [7]. These include less computational complexity as compared to turbo codes, ability to pipeline the decoder to reduce delay at the cost of registers and some latency, and less number of iterations than turbo codes. Increased number of iterations reduces the throughput and increases power dissipation [5, 8]. The studies conducted in [7] and [8] indicate that lesser number of iterations in LDPC codes helps achieving higher throughput and decreases power

dissipation. Another factor which contributes in increasing the throughput of LDPC decoder is the degree of parallelism which is adjustable [9]. Another reason of using LDPC codes is that it is relatively easier to implement than turbo codes [10].

Despite all these advantages of LDPC codes the random parity check matrix makes the wiring between variable and check nodes complex, especially for large matrices. This leads to increased routing congestion in the decoder. Eventually the size of the LDPC decoder increases and speed decreases [11]. Complexity of practical implementation makes high throughput very difficult to achieve [12, 13]. Moreover, designing LDPC decoders in VHDL (Very high speed integrated circuit Hardware Description Language) becomes a cumbersome task when the size of the design increases. Huge amount of time and effort are required to model such large designs in VHDL [14]. It results in decrease in productivity. It becomes a nightmare for the designer to write VHDL code for thousands of connections and to make the changes required. Therefore, hardware implementation of LDPC decoder still remains a challenge.

This paper examines high level modelling and synthesis techniques of LDPC decoders using emerging industry tools. It compares the high level approaches with traditional hardware description language based approaches in terms of modelling complexity, efforts and time. It also compares the results obtained for a representative LDPC decoder design using the high level approaches and using a traditional HDL based approach.

## II. DESIGN APPROACH

This paper investigates a new high-level modelling and synthesis methodology for LDPC decoder using state of the art tools. As opposed to using only hand written VHDL codes for the entire design, this research examines high level design methods using Simulink, involving a combination of blocks designed with predefined library components and with embedded Matlab codes. A VHDL (VHSIC Hardware Description Language) code is generated automatically from the high level Simulink model using Mathwork's Simulink HDL coder (version 1.2). The entire design process is captured in Fig. 1.



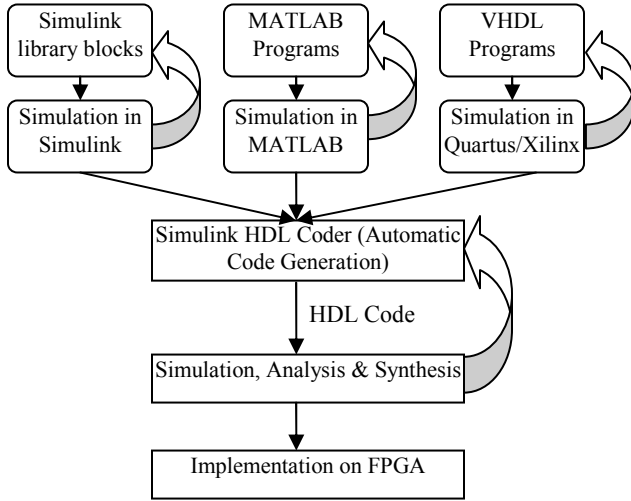


Fig. 1. Proposed design flow.

Simulink HDL coder gives the flexibility of integrating different design approaches, namely Matlab programs, Simulink models and VHDL codes. Simulink and Matlab programs provide a much higher level of abstraction than VHDL. Therefore our design approach utilizes as many Simulink library blocks and Matlab blocks as possible to design various modules of the decoder provided these modules can be successfully processed by the HDL coder for automatic generation of VHDL codes. All the modules are then integrated in a top level Simulink model to produce the overall decoder model. VHDL code is generated automatically using Simulink HDL coder along with an optional test bench. This test bench can be used in ModelSim (a HDL simulator) to check correctness of the design. The auto-generated VHDL code can also be used in Altera's Quartus II or Xilinx ISE (Integrated Software Environment) for synthesis and implementation on desired FPGA (Field Programmable Gate Array). Automatic HDL code generation will lead to reduction in design efforts thereby increasing design productivity.

### III. OVERVIEW OF LDPC CODES AND DECODING ALGORITHM

Low density parity check (LDPC) codes are a class of linear block codes which are used for error detection and correction [15]. LDPC codes were first discovered by Gallager in 1962. In 1981, Tanner worked on LDPC codes and came up with important tanner graphs or bipartite graphs [16]. LDPC codes could not be implemented at the time of invention because the technology was not advanced, though they were studied again after almost three decades because they are more advantageous than any other error-correcting codes. LDPC codes are

extensively used in standards such as 10Gigabit Ethernet (10GBaseT) & digital video broadcasting (DVB-S2) [17].

There are different algorithms which could be used for decoding purposes. We used the min-sum decoding algorithm [12], which is a special case of sum-product algorithm [16]. Sum-product algorithm reduces the computational complexity and makes the decoder numerically stable [16]. Assume that the messages from the host communication system are represented by  $I$  and are passed on to the decoder for error correction. The LDPC decoder consists of a number of variable nodes ( $v$ ) and check nodes ( $c$ ). The operation of the min-sum algorithm can be summarised as follows [12]:

#### A. Variable Node Operation

A variable node performs the operation given in equation (1) and passes the outputs to check nodes.

$$L_{cv} = \sum_{m \in M(v) \setminus c} R_{mv} + I_v \quad (1)$$

where,  $I_v$  is the input to variable node  $v$ , also known as Log Likelihood ratio (LLR),  $L_{cv}$  is the output of variable node  $v$  going to check node  $c$ ,  $M(v) \setminus c$  denotes the set of check nodes connected to variable node  $v$  excluding the check node  $c$ ,  $R_{mv}$  is the output of check nodes going to variable node  $v$ .

#### B. Check Node Operation

A check node receives messages from variable nodes and performs the operation given by equation (2):

$$R_{cv} = \prod_{n \in N(c) \setminus v} \text{sign}(L_{cn}) \times \min_{n \in N(c) \setminus v} |L_{cn}| \quad (2)$$

where,  $R_{cv}$  is the output of check node  $c$  going to variable node  $v$ .

#### C. Parity Check

Every check node also checks whether the parity condition is satisfied by looking at the sign of the messages coming from the variable nodes. Until all the parity checks from all the check nodes are satisfied the messages are sent back to variable nodes and the variable nodes do the operation specified in part (A), otherwise the decoder stops the process.

Min-sum decoding algorithm uses soft decision. However, hard decision is taken on the new LLR ( $I_v$ ). If the new LLR is negative then the output bit would be a 1 otherwise a 0. Interested readers can find the details of the sum-product and min-sum decoder algorithms in [16].

#### IV. REPRESENTATIVE DECODER DESIGNS

LDPC codes can be represented by an  $M \times N$  sparse matrix, usually called  $H$  matrix. The  $H$  matrix contains mostly zeros and a small number of 1s [9]. It can also be represented by a graph called bipartite or tanner graph. Tanner graphs contain variable and check nodes. The decoder prototype designed in this research is based on the matrix shown in Fig. 2. It is a very small matrix compared to the real matrices. This  $10 \times 5$  matrix can be transformed into a tanner graph having 10 variable nodes and 5 check nodes as is shown in Fig. 3.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Fig. 2. A  $10 \times 5$  sparse matrix.

The signals go from the variable nodes to the check nodes and from the check nodes back to variable nodes. Typically different sets of wires are used for the signals going from variable nodes to check nodes and vice versa. The process is iterative and goes until all the parity checks are satisfied. The messages in our prototype designs are 4-bit long.

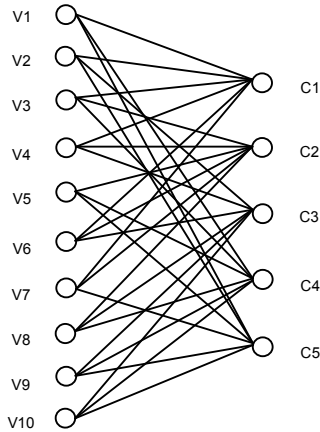


Fig. 3. Tanner graph.

##### A. Design 1

The approach taken in the first decoder design was to use a combination of embedded Matlab blocks and Simulink library blocks. No handwritten VHDL blocks were used. A generic Simulink model was first developed for each of a variable node and a check node.

Variable node design in Simulink: The variable node, shown in Fig. 4, has been designed in Simulink using the basic *add* and *subtract* blocks from Simulink library. The

control block controls the operation of the variable node. This block has been coded in Matlab and incorporated in the Simulink model of the variable node as an embedded Matlab block. The variable node has four inputs, the first three come from various check nodes and the fourth one is the raw LLR, which is an external input supplied by the host communication system. It performs the operation given in equation (1) and passes the outputs to the check nodes to which it is connected. VHDL code can be generated automatically from this Simulink model, which can then be synthesized for a target FPGA.

Check node design in Simulink: The check node has been designed in the same way as the variable node. The Simulink model of the check node is shown in Fig. 5. It uses some of the blocks from Simulink library (absolute and bitwise XOR) and others are the embedded Matlab blocks. The check node finds the minimum of all the inputs and performs the parity checks. It contains a control block, which controls its operation.

10x5 LDPC decoder: An LDPC decoder with 10 variable and 5 check nodes has been designed using instances of the variable and check node components described above. The various nodes are interconnected in a way that corresponds to the tanner graph of Fig. 3.

##### B. Design 2

Design 2 uses a combination of embedded Matlab blocks, Simulink library blocks and VHDL/Link for ModelSim blocks. 'Link for ModelSim' is a utility that enables modules coded in VHDL to be embedded in Simulink models. Design 2 is different from the first design in that it uses serial communication of messages between variable and check nodes. This is achieved using SIPO (serial-in-parallel-out) and PISO (parallel-in-serial-out) registers at each input and output port in all the nodes. This greatly simplifies the interconnections by reducing the number of wires four times at the cost of some extra registers. The variable and check node components are same and are connected in the same way as in Design 1. The PISO and SIPO components are coded in VHDL and are used in the Simulink model with the help of the 'Link for ModelSim' utility (to link the VHDL blocks with ModelSim for simulation and code generation purposes). The way PISO and SIPO are used in a variable node is shown in Fig. 6. Check nodes have been modified in exactly the same manner. The inputs and outputs of all the variable and check nodes are now 1-bit long. The effects of using a VHDL block in a Simulink design are discussed in the next section. The HDL coder does not generate the VHDL code of the Link for ModelSim blocks. The VHDL code for these blocks needs to be added before synthesizing the code.

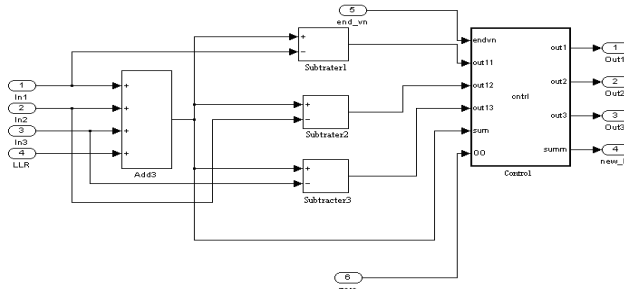


Fig. 4. Variable node design in Simulink.

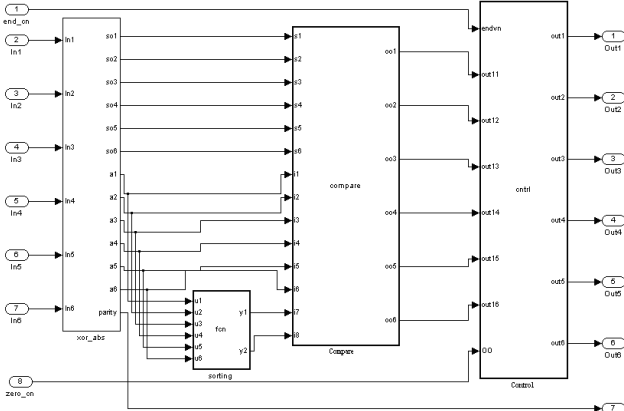


Fig. 5. Check node design in Simulink.

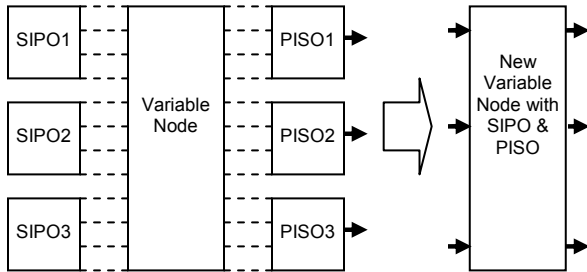


Fig. 6. A variable node with PISO and SIPO.

## V. RESULTS AND ANALYSIS

### A. Convergence

The convergence characteristics of the LDPC decoders presented in this paper have been tested using VHDL testbench and compared with that of a Matlab code for a functionally equivalent decoder. For this purpose the VHDL code automatically generated from the Simulink models were used. Fig. 7 shows the spread of the number of iterations for Design 2 by using plots of (mean number of iterations + standard deviation) and (mean number of iterations • standard deviation) versus SNR. It is clear that convergence is achieved in 6 iterations. This is consistent

with the convergence result of a functionally equivalent decoder shown in Fig. 8, designed using Matlab code.

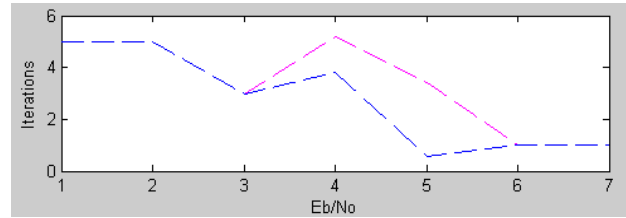


Fig. 7. Spread of the number of iterations for Design 2 obtained from VHDL testbench.

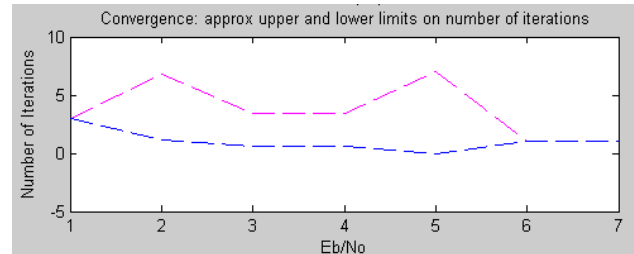


Fig. 8. Spread of the number of iterations for a functionally equivalent Matlab code.

### B. Synthesis

Synthesis results for the high-level designs are given below and are compared with their 'VHDL only' counterparts. The results were obtained using Altera's Quartus II software with Cyclone II EP2C70F672C6 as the target device. Table 1 compares the resources used by Design 1 and its speed with a similar decoder designed using VHDL. Our design uses 2.1% of the FPGA's logical elements, only 120 registers and has a maximum frequency of 42.96MHz. Clearly this Simulink design uses much less resources and is faster than the VHDL design. Similar observations are made from Table 2, which compares our Design 2 with the VHDL only design. Design 2 uses 3.5% of the logic elements, which is slightly higher than the amount used by the VHDL design (2.8%). However, our Design 2 uses nearly half the number of registers used by the VHDL design and still achieves nearly the same maximum frequency ( $F_{max}$ ). Furthermore, our high level designs require much less effort and time as opposed to writing everything in VHDL.

TABLE I  
DESIGN 1 COMPARED WITH VHDL COUNTERPART

|                         | Design 1    | VHDL only   |
|-------------------------|-------------|-------------|
| Logical elements        | 1432 (2.1%) | 1492 (2.2%) |
| Combinational functions | 1432 (2.1%) | 1492 (2.2%) |
| Total registers         | 120 (0.2%)  | 686 (1%)    |
| Maximum frequency       | 42.96MHz    | 37.3 MHz    |

TABLE II  
DESIGN 2 COMPARED WITH VHDL COUNTERPART

|                         | Design 2    | VHDL only   |
|-------------------------|-------------|-------------|
| Logical elements        | 2416(3.5%)  | 1916 (2.8%) |
| Combinational functions | 2416 (3.5%) | 1916 (2.8%) |
| Total registers         | 480 (0.7%)  | 881 (1.3%)  |
| Maximum frequency       | 67.20MHz    | 67.72MHz    |

### C. Behavioural Simulation

Fig. 9 shows the functional simulation result for Design 2. A set of raw LLRs (6, 6, 2, 4, 7, 4, -2, 6, 4, 7) are applied to the variable nodes. The signals *end\_o\_vn* & *end\_o\_cn* are the control signals. The parity becomes 1 in the third iteration when all the parity checks are satisfied and the decoder stops the iterations. The corrected LLRs output by the variable nodes are 7, 7, 7, 7, 7, 7, -8, 4, 5, 7.

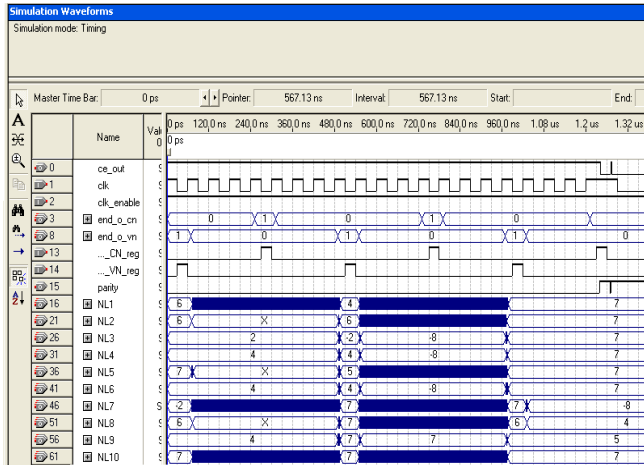


Fig. 9. Simulation results of Decoder Design 2.

### D. Analysis of the High Level Design Methods

The high-level design methodologies presented in this paper reduces design complexity, design effort and time. For example, to design the variable and check nodes of Figures 1 and 2 completely in VHDL a designer has to write behaviours of quite a few modules in VHDL and manually do the port mapping for the top level design. As the design gets larger and larger it becomes very difficult to code everything in VHDL. It not only takes huge effort and time, managing complex designs and reusing the designs are often quite difficult. A great deal of expertise and experience in VHDL is also required. For a large LDPC decoder with hundreds of nodes, the interconnections among the variable and check nodes could easily become a designer's nightmare if the entire design has to be coded in VHDL. An alternative, intuitive and highly productive design method has been a long standing desire of the engineers engaged in LDPC designs.

The high level methodology we have presented in this paper offers a very attractive alternative. The designs can be done intuitively in Simulink using predefined Simulink library blocks and blocks made from high-level Matlab code. The design complexity, effort and time are reduced drastically because the need for manually writing complex VHDL code is almost eliminated. Design reuse requires much less effort because changes are made either at block level in Simulink or in high-level Matlab code. In addition Matlab is a software programming language used much more widely than hardware description languages like VHDL and Verilog. Therefore designers without specific skills in hardware languages are able to design complex digital systems without much problem. Even software engineers and algorithm developers are able to quickly implement and test their high level designs due to the automatic conversion of Simulink models to HDL. This will surely offer great flexibility and efficiency in the design and reuse of complex LPDC decoders. There are some other benefits of the high level design methodology:

- Different parts of the model can be enveloped using the 'create subsystem' property of Simulink. It reduces the design complexity for large designs.
- User created Simulink library blocks provide great flexibility because the revisions made to the user defined library seamlessly propagate through the entire model.

The design methodology we have presented however requires the use of some emerging design tools and library functions, such as Simulink hardware library and embedded Matlab blocks, Link for ModelSim and HDL Coder tools. Because these are very recent developments the library of Simulink blocks supporting various functions a designer needs to implement a design is very limited at the present time. Similarly the capability of HDL Coder is limited to conversion of a few commonly used Simulink blocks and a few embedded Matlab blocks. Some of the specific limitations are discussed below.

### E. Current Limitations

Some common blocks in the Simulink library are not currently supported by the HDL Coder for automatic generation of VHDL code, e.g. flip-flops. Although we could build the PISO and SIPO registers easily in Simulink using the flip flops from its library the registers were not converted to VHDL by the HDL Coder. A careful selection of supported Matlab functions and Simulink library blocks may help addressing this type of problems, but not necessarily always. Some other limitations we experienced are listed below:

- Multiple instances of some modules (components) are used in the LDPC decoder, and in fact in most modular designs utilising a hierarchical design methodology. For example, in our LDPC decoder multiple instances of the check and variable nodes, and PISO and SIPO registers

are used. The HDL Coder dumps the full behaviour of each component as many times as it is instantiated in the design. This produces redundant instances of component behaviour in the generated VHDL code. It is necessary for the designer to edit the auto-generated code.

- There are some functions which are supported by the HDL Coder, but it produces the output in a particular data type. For example, the sign function of Matlab gives output only in 8-bit integer (int8) format. It is not possible to change these default data types in the current version. The only option is to manually edit the auto-generated HDL code.
- Link for ModelSim: This utility enables blocks designed in VHDL to be included in Simulink models. However the ports of the blocks that use Link for ModelSim get interchanged while simulating in Simulink. HDL code cannot be generated for the Simulink model in this situation. The code can be generated only after correcting the design. This makes the design process difficult and time consuming. The other drawback of blocks utilising Link for ModelSim is that these blocks may make changes in the auto-generated code. In case of the LDPC decoder these blocks changed the signals of type signed to unsigned. Once again this requires manually editing the auto-generated VHDL code.

## VI. CONCLUSION

In this paper a new high-level intuitive design methodology based on Simulink for modelling and synthesis of complex LDPC decoders has been presented. It utilises the higher level of abstraction offered by the Simulink modelling environment. The complete modelling, simulation and synthesis process utilises a combination of emerging design tools and associated library functions. These include the Simulink HDL Coder and 'Link for ModelSim' tools, and embedded Matlab and Simulink hardware library blocks. Two versions of 10x5 LDPC decoders have been designed, simulated and successfully synthesised to an Altera FPGA device. The high level design methodology offer great advantages in terms of design complexity, effort and time compared to a HDL-only design method. Designs completed using the new methodology used less FPGA resources and achieved better or similar performance level (Fmax) to those of HDL-only designs. We therefore envisage that with further enrichment of the Simulink Library blocks, and enhancements of HDL Coder and Link for ModelSim tools, design methodologies similar to the one presented in this paper will replace tedious HDL-based design approach. This will pave the way for cost effective design and reuse of a new generation of complex high performance LDC decoders.

## ACKNOWLEDGEMENT

The authors wish to acknowledge Mr Pham Minh Duc and Prof Bill Cowley of UniSA for model optimization and for the very useful suggestions/feedback respectively.

## REFERENCES

- [1] R. G. Gallager, *Low-Density Parity-Check Codes*, Monogram, Cambridge Mass, 1963.
- [2] S. Johnson, *Introducing Low-Density Parity-Check Codes*, Australia: University of Newcastle, 2006.
- [3] S. Lin, D. J. Costello, *Error Control Coding: Fundamentals and Applications*, New Jersey: Prentice Hall, 2004.
- [4] B. Reiffen, "Sequential Decoding for Discrete Input Memoryless Channels," *IRE Trans. Inf. Theory*, vol. 8, no. 3, pp. 208-220, April 1962.
- [5] A. J. Blanksby and C. J. Howland, "A 690- mW 1-Gb/s 1024-b, rate-1/2 low-density parity check code decoder," *IEEE J. Solid State Circuits*, vol. 37, no. 3, pp. 404-412, 2002.
- [6] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, pp. 619-637, Feb. 2001.
- [7] J. Nguyen, B. Nikolic and E. Yeo, "Design of a low density parity check iterative decoder," EECS, College of Engineering, University of California, Berkley, 2002.
- [8] S. Hong and W. Stark, "Design and implementation of a low complexity VLSI turbo-code decoder architecture for low energy mobile wireless communications," *J. VLSI Signal Processing*, vol. 24, pp. 43-57, 2000.
- [9] M. Karkooti, P. Radosavljevic and J. R. Cavallaro, "Configurable, High Throughput, Irregular LDPC Decoder Architecture: Tradeoff Analysis and Implementation," *Proc. Int. Conf. Application Specific Systems, Architectures and Processors*, pp. 360-367, Sept. 2006.
- [10] Y. Lei, L. Hui, R. C. J. Shi, "Code Construction and FPGA Implementation of a low-error-floor multi-rate low-density Parity-check code decoder," *IEEE Trans. Circuits & Systems I*, vol. 53, pp. 892-904, April 2006.
- [11] A. Darabiha, C. A. Carusone, R. F. Kschischang and E. S. Rogers, "Multi\_Gbit/sec Low Density Parity Check Decoders with Reduced Interconnect Complexity," *IEEE Int. Symp. Circuits & Systems*, vol. 5, pp. 5194-5197, 2005.
- [12] J. Sha, M. Gao, Z. Zhang, Li Li and Z. Wang, "An FPGA implementation of array LDPC decoder," *IEEE Asia Pacific Conf. Circ. & Systems*, pp. 1675-1678, December 2006.
- [13] Mauro Cocco, "A scalable architecture of LDPC Decoding," *Proc. Design, Automation & Test in Europe Conf.*, vol. 3, pp. 88-93, Feb. 2004.
- [14] J. A. Wicks, J. R. Armstrong, "Efficiency ratings for VHDL behavioral models," *IEEE Proc. Southeastcon '98*, pp. 401-404, April 1998.
- [15] M. Eroz, F.W. Sun and L.N. Lee, "DVB-S2 low density parity check codes with near Shannon limit performance," *Int. J. Satellite Communications and Networking*, pp. 269-279, 2004.
- [16] W. E. Ryan, *An Introduction to LDPC Codes*, USA: The University of Arizona, 2003.
- [17] T. Mohsenin and B. M. Baas, "Split-Row: A reduced complexity, high throughput LDPC decoder architecture," *Int. Conf. on Computer Design, ICCD 2006*, San Jose, CA, pp. 220-225, 1-4 Oct 2007.

# The Study of Tracking Control of a Robotic Manipulator Actuated by Shape Memory Alloy

S. M. Mahfuzur Rahman and Kyoung Kwan Ahn

Department of Mechanical Engineering, CUET, Chittagong 4349, Email: [smmrahman@cuet.ac.bd](mailto:smmrahman@cuet.ac.bd)  
School of Mechanical and Automotive Engineering, University of Ulsan, Ulsan 680749, Korea.  
Email: [kkahn@ulsan.ac.kr](mailto:kkahn@ulsan.ac.kr)

**Abstract** — Shape Memory Alloy (SMA) composites offer a great potential as actuators because of its excellent power to weight ratio, smooth motion and silent actuation. However, SMA actuated robotic manipulator systems have severe hysteresis, system nonlinearities, model and parametric uncertainties those are often responsible for position inaccuracy in a regulation or tracking system. A simple and upstanding control algorithm for tracking control of a SMA actuated robotic manipulator system is presented in this paper. The controller is developed based on the combination of a variable structure control approach and a heuristic based intelligent control approach like fuzzy logic control (FLC), so that the resulting control algorithms have a superior performance both in stabilization and tracking the desired trajectories. A desktop prototype of the robotic manipulator system is recalled from our previous work [1]. A nonlinear form of proportional – integral – derivative (PID) controller is also employed to the system for comparison of performance. Several experiments have been performed in tracking both stationary and periodically varied input signals. Experimental results from real time control verify the effective and robust performance of the controller.

**Index Terms** — Fuzzy Logic, Manipulator, Position control, SMA actuator, Sliding Mode control.

## I. INTRODUCTION

Shape memory alloy (SMA) is a smart alloy or memory metal that remembers its shape through shape memory effect. Below transformation temperature SMA has very low yield strength and can be deformed easily into any new shape, which it can retain. However, when it is heated above its transformation temperature it undergoes a solid state phase change which causes it to return to its original state. This unique property of SMA is a consequence of the reversible crystallographic phase transformation that occurs between the low temperature Martensite phase and the high temperature Austenite phase. These two phases have the same chemical composition but different crystallographic structures, and thus, different thermal, mechanical and electrical properties. Solid state phase transformation of SMA is accompanied by a large force against the applied stress and recovery of the inelastic strain. The combination of high stiffness, high strength and large recovery strain offers great potential for the use of SMA as actuators in a

variety of applications including robotic arms, space stations, automotive components, surgical devices and so on. Some typical examples can be found as mobile robot actuators [3], micro rotary actuator [5], linear smart actuator [6], and robot manipulator [4].

There is also an increasing demand of SMAs for developing small size and lightweight devices that will be able to produce large forces, large displacements, high speeds, and will be highly energy efficient. In addition, it is easy to drive a SMA actuator simply by passing an electric current through it, which then heats up SMA with resistive electrical (Joule) heating, and thus, SMA offers smooth motion, clean and silent actuation. But, the inherent nonlinear phenomenon associated with SMA is hysteresis [2], which often responsible for controlling the position and motion precisely.

However, tracking control of SMA actuators is essential in many applications such as robotic applications [7], [14] and vibration control [8], but due to the above reasons it is still a challenging task for the researchers. This paper thus motivated to find simple upstanding control algorithm on motion and position control of SMA actuators. Several control methods of SMA actuators have been investigated; these can be categorized as linear control [9], nonlinear control [7], intelligent control [10], [2] and variable structure control [11]. Among these variable structure control approaches have shown to be most effective with the control problems of SMA systems. This paper presents a two steps control algorithm, the first step of the control algorithm (FLC) ensures robustness against high frequency un-modeled dynamics and simultaneously the second step, Sliding Mode Control (SMC) ensures robustness against parametric uncertainty [12]. It is therefore possible to realize smooth and robust control avoiding the system's complexity in real time controls. The fuzzy algorithm is defined for two significant reasons, first, for reducing (eliminating) undesirable chattering of control signal by adjusting boundary layer thickness and second, for making the system modeling very simple. Because FLC works on a linguistic control approach that makes it insensitive to the unmodeled parameters and parameter uncertainties. Finally a nonlinear PID controller is applied to investigate the effectiveness and performance of the controller.



## II. SMA ROBOTIC MANIPULATOR

SMAAs can be conventionally fabricated into different desired shapes with different chemical compositions. SMA wire made of nickel and titanium (also called Nitinol or NiTi) is the most commonly used form of actuator because of its ease of use and convenient electrical activation. As illustrated with schematic diagram in Fig. 1, a SMA wire actuator is used to power the joint of a robot manipulator.

There are two ways of providing the bias force and therefore two types of SMA actuators. The first one, called bias type (or one-way actuator), is composed of a SMA element and a bias spring, and another, called differential type (or two-way actuator), and is composed of two SMA elements. In one-way actuators, the resulting torque is the difference between the bias torque and SMA wire torque. Considering the hardware limitations, in our design the manipulator is actuated by a bias type SMA wire actuator without bias spring, where the bias force (torque) is provided by the weight of the moving arm and payload. A photograph of the experimental setup is given in Fig. 2.

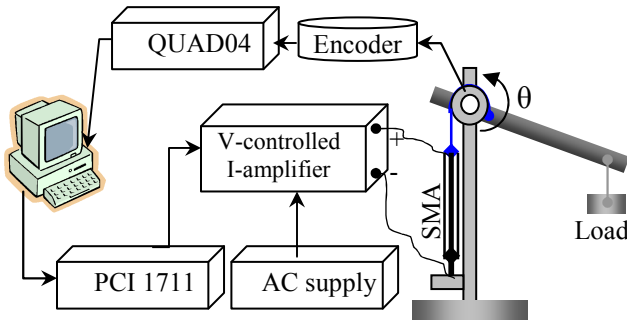


Fig. 1 Experimental setup of SMA manipulator.

TABLE I

LIST OF HARDWARE USED IN THE EXPERIMENTAL SETUP

| Name            | Model        | Manufacturer          |
|-----------------|--------------|-----------------------|
| SMA             | DM01         | MIGA mortor           |
| Rotary encoder  | S30-4-1024ZO | Metronix              |
| D/A board       | PCI 1711S    | Advantech             |
| Wiring Terminal | PCLD-8710    | Advantech             |
| A/D board       | PCI QUAD04   | Computing Measurement |
| Cooling fan     | SP1202512H   | INNOVATIVE            |

The SMA wire is heated with resistive electrical (Joule) heating by applying controlled input voltage, beyond the activation temperature it contracts due to the phase transformation from Martensite to Austenite. Upon cooling the wire's temperature falls and reverse phase transformation (from Austenite to Martensite) occurs, thus SMA wire is elongated under the action of bias force. The manipulator system is actuated using a DM01 Linear Shape Memory Alloy Actuator, such that the moving arm can rotate from  $0^\circ$  to  $150^\circ$ , and a rotary encoder is used to measure the joint angle of the moving arm. Note that, the

controller high and low voltages provided by power supply are  $V_h = 6$  V and  $V_l = 0$  V. A list of hardware used in the experimental setup is given in Table I.

## III. SMA CONTROL SYSTEMS

Shape memory alloy (SMA) miniature actuators have received significant attention in recent years because of their high power to mass ratio, silence and frictionless actuation, and simplicity of their mechanisms. Despite these special features, SMAAs have not been fully exploited by the researchers, particularly in devices employing feedback control systems. These devices are inherently nonlinear, characterized by a hysteresis loop and saturation. Although numerous analytical and phenomenological models have been developed for SMAAs over the last few decades, it is difficult to choose proper models for different applications. Furthermore a given model may contain many unknown parameters, discouraging from further attempts to design a model-based control system [4].

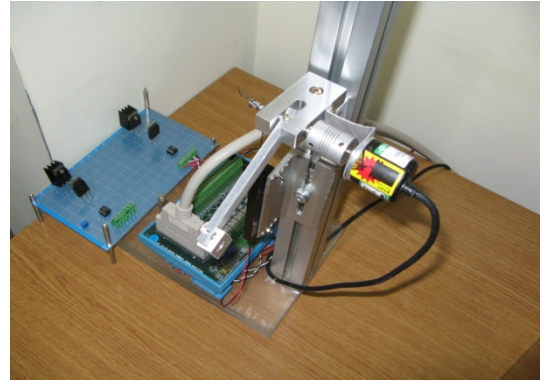


Fig. 2. Photograph of the experimental setup.

In this paper a novel two steps control system is presented for controlling of SMA actuated robotic manipulator. A nonlinear robust sliding mode controller is used to calculate the adjustable input voltage to the SMA and an intelligent fuzzy logic controller is used to estimate the boundary layer thickness for the control input based on position error. Finally a nonlinear PID control is applied to investigate the performance.

### A. Sliding Mode Control

Sliding mode control (SMC), a particular kind of variable structure control, is a robust nonlinear control algorithm in which an  $n$ th order nonlinear and uncertain system is transformed to a 1<sup>st</sup> order system. Considering the  $n$ th order SISO dynamic system,

$$\dot{x}^{(n)} = f(x) + b(x)u \quad (1)$$

Where scalar  $x$  is the output of interest, scalar  $u$  is control input and  $x$  is the state vector.  $f(x)$  and  $b(x)$  are the

unknown nonlinear dynamic function and control gain respectively. The control objective is to follow a desired state vector  $x_d$ , such that  $x_d(0) = x(0)$ . A time-varying surface  $s(t)$  in state space  $R^n$  can be defined as,

$$s(x; t) = \left( \frac{d}{dt} + \lambda \right)^{n-1} \tilde{x} \quad (2)$$

where  $\tilde{x} = x - x_d$  is the tracking error and  $\lambda$  is a positive constant. It can be shown that, tracking the desired state vector is equivalent to that of remaining on the surface  $s(t)$  [12]. The sliding condition to track specific time-varying desired state is,

$$\frac{1}{2} \frac{d}{dt} s^2 \leq -\eta |s|; \quad \eta \geq 0 \quad (3)$$

Further detail on sliding mode control can be found in [12].

### B. Fuzzy Logic Control

Sliding mode controller inherited chattering in applications. The use of boundary thickness in sliding mode control has been a common technique to reduce chattering of the control input. However, the constant width boundary layer design can reduce chattering of the control signals, but it decreases the control accuracy [13]. The compromise between smoothness of control signals and accuracy of control results is dictated by the proper adjustment of the boundary layer thickness. We thus choose a method of using fuzzy logic to control the width of boundary layer. Using error and rate of error as the control input, and the boundary layer thickness ( ) as the control output, the fuzzy control rules are set such that the system can get highest available input (voltage) initially to achieve the fastest possible response. The rule matrix of FLC is shown in Table II. The rules of the FLC relate position error and rate of error with boundary layer thickness so that the method can effectively eliminate chattering and ensure almost perfect control accuracy simultaneously.

TABLE II

FUZZY RULE MATRIX FOR 5 MEMBERSHIP FUNCTIONS

|           |      | $e$ |    |      |    |    |
|-----------|------|-----|----|------|----|----|
| $\dot{e}$ |      | NB  | NM | ZERO | PM | PB |
|           | NB   | VL  | L  | M    | S  | VS |
|           | NM   | L   | M  | L    | M  | S  |
|           | ZERO | M   | L  | M    | S  | M  |
|           | PM   | L   | M  | L    | M  | S  |
|           | PB   | VL  | L  | M    | S  | VS |

Where NB - negative big, NM - negative medium, PM - positive medium, PB - positive big, VL - very large, L - large, M - medium, S - small, VS - very small. We used five trapezoidal membership functions for each of control input and five Gaussian membership functions for control output with 'max-min' reasoning method, and center-of-gravity defuzzification method.

### C. Controller Design

The nonlinear behavior of SMA and its one-way actuation nature complicated the control problem since errors should be eliminated and full state feedback was not available. However, since the controller is model independent, we need to define proper sliding surface as well as the control law. Since only joint angle is measured using a rotary encoder, we can define a proper sliding surface as a function of position error and velocity error as,

$$s = \lambda_p \tilde{\theta} + \lambda_v \dot{\tilde{\theta}} + \lambda_i \int_0^t \tilde{\theta} dt \quad (4)$$

where,  $\tilde{\theta} = \theta_d - \theta$  and  $\dot{\tilde{\theta}} = \dot{\theta}_d - \dot{\theta}$ , subscript d stands for desired values, and an integral term in the sliding surface is added to minimize the steady-state errors. Adjusting the boundary layer thickness properly using FLC, the decision rule or control law for the sliding mode controller can be defined as,

$$u = \begin{cases} V_h & \text{if } \frac{s}{\phi} > 1 \\ s & \text{if } \left| \frac{s}{\phi} \right| \leq 1 \\ V_l & \text{if } \frac{s}{\phi} < -1 \end{cases} \quad (5)$$

where,  $u$  is the controlled input,  $V_h$  highest available voltage and  $V_l$  is the lowest voltage.

As we are using a bias type (one-way) actuator, the application of possible high voltage,  $V_h$  to the SMA actuator will force the trajectory to reach sliding surface expressed in equation (4). Once the sliding surface is reached, sliding mode controller keeps the states on the close neighborhood of the sliding surface (zero tracking error). This surface is asymptotically stable and the stability of the control system can be verified through sliding mode control theory [12]. The block diagram of the controller is shown in Fig. 3.

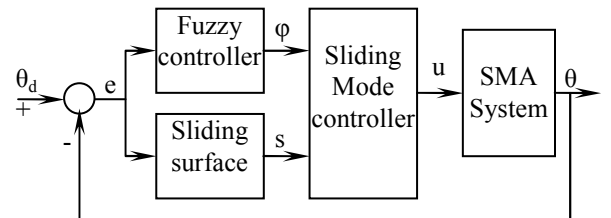


Fig. 3. Block diagram of the controller.

The surface defined in (4) incorporates position, velocity and integral tracking errors. Only position feedback results in a significant steady state error and slow response due to overshoot. The addition of velocity tracking errors to the surface eliminates overshoot and thus obtains fast response in the experiments. It also reduces the steady state errors. Finally integral tracking errors were added to further reduce



the steady state errors. So the controller is now able to follow the trajectories effectively [4].

#### D. Nonlinear PID Controller

Modifying the structure of a linear PID controller shown in equation (6) a novel nonlinear PID (NPID) controller is applied to the system. This controller comprises the sum of nonlinear terms of position error with a gain  $K_T$  that have significant effects in reducing settling time and overshoot of the system. The distinguished feature of this controller is that, for a relatively small amount of error, the cubic term tends to vanish but the quadratic term still produces a nonlinear control effort to minimize the errors. The control law of NPID controller in time domain is expressed in equation (7).

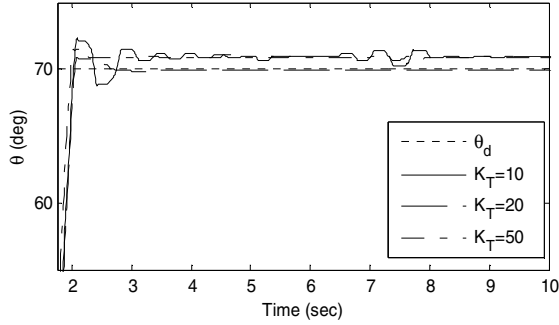


Fig. 4. Effects of control gain  $K_T$  on the NPID controller.

$$u(t) = K_P.e(t) + K_I.\int_0^t e(\tau)d\tau + K_D.\dot{e}(t) \quad (6)$$

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau) d\tau + K_D \dot{e}(t) + K_T \left[ \text{sgn}(e).e^2(t) + e^3(t) \right] \quad (7)$$

The effects of gain  $K_T$  on the performance of NPID controller (7) are shown in Fig. 4. It is clear from Fig. 4 that, with proper tuning of the gain  $K_T$ , the system can achieve better performance.

#### E. Controller Parameters

Experimentally, we have found a suitable range of boundary layer thickness ( $\phi$ ) for the sake of avoiding chattering and overshoot from which the FLC assigned a value to  $\phi$ . The range selection of  $\phi$  is made through numerous iterative studies and some selected results are shown in Fig. 5. This Fig. shows that to avoid overshoot and chattering, the appropriate range of boundary layer thickness should be  $2^\circ$  to  $5^\circ$ .

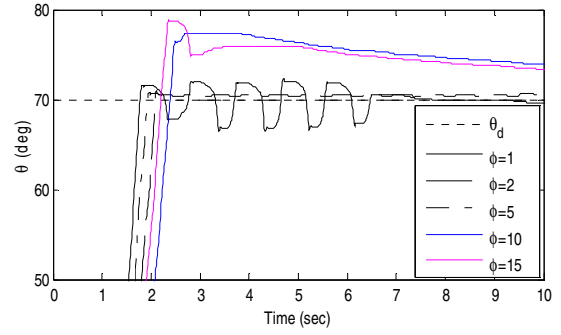


Fig. 5. Effect of boundary layer thickness ( $\phi$ ) on the controller performance.

Numerous iterative studies also made to establish the sliding surface experimentally. However, we only present the final results of our trials. The effects of position, velocity and integral gain on the controller performance obtained experimentally are presented in Fig. 6, 7, 8 respectively. These results suggest that the optimal values of different gains of the surface would be  $\lambda_p \approx 5$ ;  $\lambda_v \approx 0.1$  and  $\lambda_i \approx 0.05$  for desired performance.

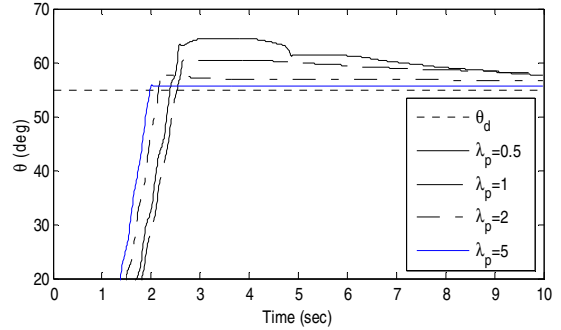


Fig. 6. Effect of position gain ( $\lambda_p$ ) on the controller performance.

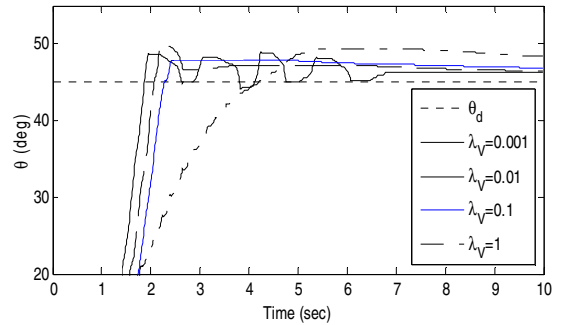


Fig. 7 Effect of velocity gain ( $\lambda_v$ ) on the controller performance.

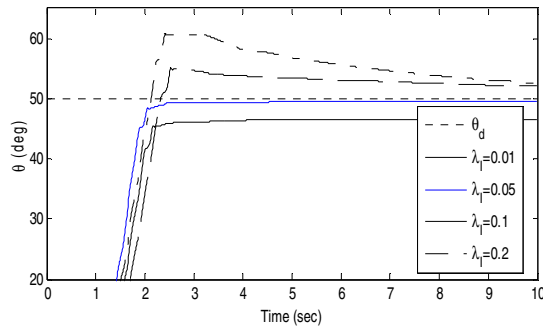


Fig. 8 Effect of integral gain ( $\lambda_i$ ) on the controller performance.

#### IV. RESULTS

A series of experiments have been carried out using both stationary and periodically varied reference trajectories and all experiments have been done through real time control in MATLAB/Simulink environment. Both stationary and periodic variable references are used in experiments. Fig. 9 shows the fuzzy-sliding mode controller can accurately follow the desired step, whereas, the nonlinear PID controller follows the trajectory with small steady state error. Fig. 10-11 verify the effective and superior performance of the both controller in tracking periodically varied reference signals of different frequencies and amplitudes.

Experimental results also confirm the selection process of controller gains as well as estimation of the range of boundary layer thickness. The contribution of FLC toward improved performance of the controller is also justified.

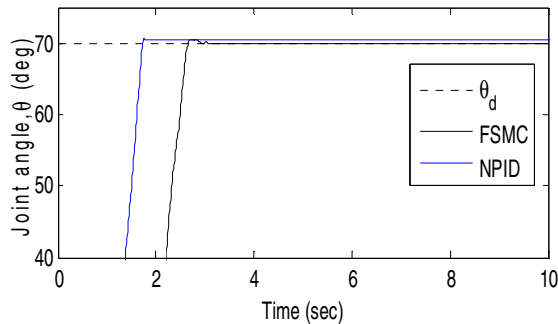


Fig. 9 Controller performance in tracking desired step.

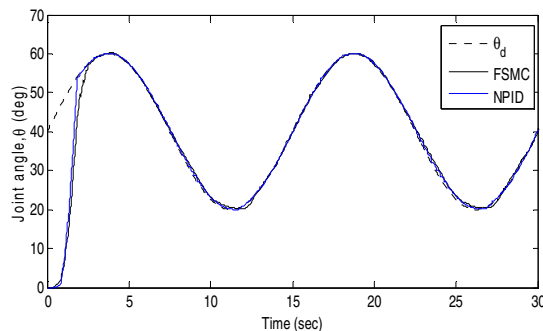


Fig. 10. Controller performance in tracking a sinusoidal trajectory (freq. 1/15 Hz, amp. 20 deg).

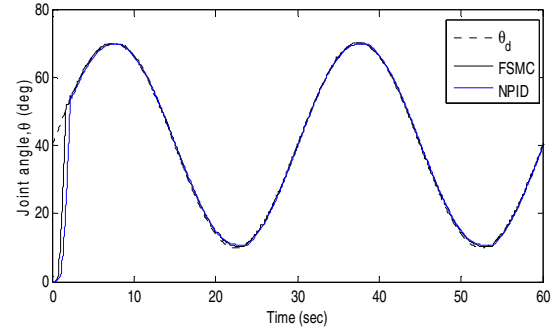


Fig. 11. Controller performance in tracking a sinusoidal trajectory (freq. 1/30 Hz, amp. 30 deg).

#### V. CONCLUSIONS

A model-independent position control algorithm is presented for a SMA actuated robot manipulator in which a heuristic based FLC works in conjunction with a robust SMC. It is thus possible to overcome the complexity in the course of modeling the SMA systems. The two parts of the controller have distinct task: a FLC provides the boundary layer thickness within the selected range and a SMC then calculates the desired input voltage to the SMA. The angular position of the manipulator is measured by a rotary encoder that is given as feedback to the controller. A sliding surface is designed based on the tracking errors, and the effects of different gains of the surface as well as the boundary layer thickness are investigated experimentally. In real time control, step and sinusoidal references are used in experiments, and the results show the effectuality of the controller. The first step of the control algorithm (FLC) ensures robustness against high frequency un-modeled dynamics simultaneously the second step (SMC) ensures robustness against parametric uncertainty. It is therefore possible to realize smooth and robust control avoiding the system's complexity in real time controls. Experimental results also verify the effective performance of nonlinear PID controller.

#### REFERENCES

- [1] Rahman, S.M.M., Ahn, K. K. and Yoon, J. I., 2008, "Experimental Study of the Tracking Control of a SMA Actuated Manipulator" *Proceeding of the International Conference on Smart Manufacturing Application (ICSMA)*, 9-11 April, KINTEX, Gyeonggi-do, KOREA.
- [2] Kha, N.B., Ahn, K.K., Yum, Y.J, Rahman, S. M. M. and Son, J. H., 2007, "Internal Model Control for Shape Memory Alloy Actuators using Fuzzy Based Preisach Model". *IEEE International Conference on Mechatronics and Automation*, Harbin, China.

- [3] Hashimoto, M., Takeda, M., Sagawa, H. and Chiba, I., 1985, "Application of shape memory alloy to robotic actuators". *Journal of Robotic System* 2(1), pp. 3-25.
- [4] Ashrafiuon, H., Eshraghi, M. and Elahinia, M. H., 2006, "Position Control of a Three-link Shape Memory Alloy Actuated Robot". *Journal of Intelligent Material Systems and Structures*, Vol 17, pp. 381-392.
- [5] Gabriel, K, Trimmer, W. and Walker, J., 1988, "A Micro Rotary Actuator using Shape Memory Alloys". *Sensor and Actuators*, vol. 15, no. 1, pp. 95-102.
- [6] Khidir, A. E, Nik, A. M., Nor, M.J. and Mustafa, M. M, 2006, "A new concept of a linear smart actuator", *Sensor and Actuators A* (2006), doi: 10.1016/j.sna.2006.07.010.
- [7] Elahinia M.H. and Ashrafiuon, H, 2002, "Nonlinear control of a shape memory actuated manipulator," *J. Vib. Acoust.*, ASMA Trans; 124:566-75.
- [8] Wang, H., Li Q. and Sun J., 2002, "Active vibration control of rotor with shape memory alloy (SMA) spring," *J. Mech. Strength*; 24:29-31.
- [9] Shameli, E., Alasty, A. and Salaaieh, H., 2005, "Stability analysis and nonlinear control of a miniature shape memory alloy actuator for precise applications," *Mechatronics*, Volume 15, Issue 4, Pages 471-486.
- [10] Song, G., Chaudhury, V. and Batur, C., 2003, "Precision tracking control of shape memory alloy actuators using neural networks and a sliding-mode based robust controller". *Journal of smart materials and structures* 12, pp. 223-231.
- [11] Grant, D. and Hayward, V., 1997, "Variable structure control of shape memory alloy actuators". *IEEE System and Control Magazine* 17(3), pp. 80-88.
- [12] Slotine, J. J. and Li, W., 1990, *Applied Nonlinear Control*, Prentice-Hall Inc, NJ, 1<sup>st</sup> edition.
- [13] Wang, H.R., Chen, W.D., Wang, H.T. and Xiao, J. Z., 2003, "Fuzzy sliding mode control for robotic manipulators based on systematic state," *Proceeding of the 2<sup>nd</sup> International conference on Machine learning and Cybernetics*, Xian 2-5 Nov.
- [14] Waram, T. C., 1993, *Actuator Design using Shape Memory Alloys*, 2<sup>nd</sup> edition, Hamilton, Ontario, Canada.

# A Fuzzy Logic-Based Adaptive Handoff Management Protocol for Next-Generation Wireless Systems

Presila Israt, Namvi Chakma, and M. M. A. Hashem

Department of Computer Science and Engineering

Khulna University of Engineering and Technology (KUET)

Khulna-9203, Bangladesh

presila86@hotmail.com, namvi\_bd@yahoo.com, mma\_hashem@hotmail.com

**Abstract**—In the integrated next-generation wireless systems (NGWS), users are always connected to the best available networks and used to switch between different networks based on their service needs. It is an important and challenging issue to support seamless handoff management in NGWS. The objective of this paper is to develop a seamless handoff management protocol for NGWS. In this paper, a fuzzy logic-based adaptive handoff (FLAH) management protocol is developed which is then integrated with an existing cross layer handoff protocol. Afterward, the handoff performance comparison of the existing protocol and our proposed protocol is carried out. The simulation results exhibit that, proposed fuzzy logic-based handoff management protocol has much better performance than conventional protocols for both intra and intersystem handoffs.

**Index Terms**— Call Admission Control, Fuzzy Logic, Hand-off Management, Intelligent Systems, Mobility Management, Wireless Networks.

## I. INTRODUCTION

In the integrated NGWS, users are always connected to the best available networks and switch between different networks based on their service needs, each of which is optimized for some specific services and coverage area to provide ubiquitous communications to the mobile users. In NGWS, mobile user can roam between a diverse set of wireless architectures like wireless local area network (WLAN), cellular, and satellite network [1]. Call admission control or handoff management of a mobile user while roaming through different wireless network architectures of NGWS is one of the important issues that needed to be solved. Mobility management contains two components: location management and handoff management [2]. Location management enables the system to track the locations of mobile users between consecutive communications. On the other hand, handoff management is the process by which users keep their connections active when they move from one Base Station (BS) to another. In NGWS, two types of handoff scenarios may arise:

- Horizontal Handoff: Handoff between two BSs of the same system. Horizontal handoff can be further classified into
  - Link-Layer Handoff: Horizontal handoff between two BSs that are under the same

Foreign Agent (FA).

- Intrasystem Handoff: Horizontal handoff between two BSs that belong to two different FAs and both the FAs belong to the same system and hence, to same Gateway Foreign Agent (GFA).
- Vertical Handoff (Intersystem Handoff): Handoff between two BSs that belong to two different systems and, hence, to two different GFAs.

However, seamless support of handoff management in NGWS is still an open research issue. The existing handoff management protocols are not sufficient to guarantee handoff support that is transparent to the applications in NGWS. Mobility management protocols operating from different layers of the network protocol stack (e.g., application layer [3], transport layer [4], network layer and link layer [5]) are proposed to provide seamless services. However, none of them alone are capable of supporting seamless mobility in NGWS. Therefore, recently the design of cross-layer mobility management protocols has gained significant attention [2]. Recently, new handoff algorithm are emerging based on advanced techniques like pattern recognition [8], neural networks and fuzzy logic system (FLS) [9]. These complicated algorithms are necessitated by the complexity of the handoff problem and dynamic conditions of wireless networks. In this paper, a fuzzy logic based handoff management protocol is introduced which is integrated with an existing Cross-Layer Handoff Management Protocol (CHMP) [2]. The simulation results reveal that, fuzzy logic based handoff management protocol has much better performance than conventional algorithms as well as CHMP for both intra and intersystem handoffs.

The remainder of the paper is organized as follows: we discuss analysis of handoff process in terms of handoff signaling delay and handoff failure probability in Section II. In Section III, we introduce our proposed fuzzy logic-based adaptive handoff management protocol and carry out its performance evaluation in Section IV. Finally, we summarize performance of FLAH management protocol in Section V.

## II. AN OVERVIEW OF HANDOFF

When a Mobile Terminal (MT) is moving from its serving Old Base Station (OBS) to New Base Station (NBS) with a speed  $v$ , the MT may learn about the possibility of moving into another cell when the Received Signal Strength (RSS) of OBS decreases continuously. When the MT discovers that it may enter into the coverage area of the NBS, the next challenge is to decide the right time to initiate Hierarchical Mobile IP (HMIP) registration procedures with the New Foreign Agent (NFA). The existing link-layer-assisted HMIP protocols [6] propose to initiate the HMIP registration when the RSS from the serving BS, e.g., OBS in Fig.1, drops below a fixed threshold value  $S_{th}$ .

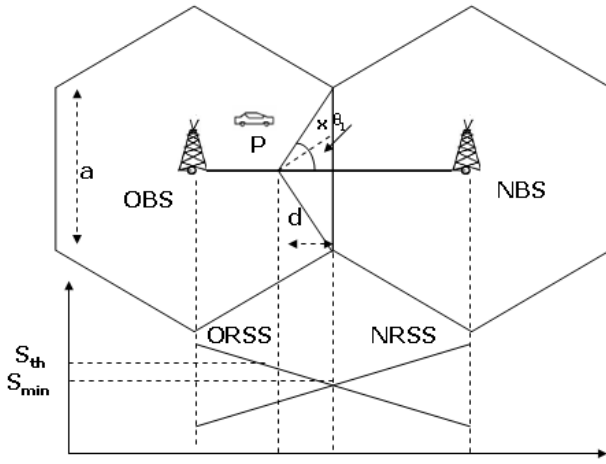


Fig. 1. Analysis of handoff process.

During the course of its movement, when the MT reaches the point  $P$  (the distance of  $P$  from the cell boundary is  $d$ ) as shown in Fig. 1, the RSS from the OBS drops below  $S_{th}$ . Therefore, when the MT reaches  $P$ , it initiates the HMIP registration with the NFA. At this point, the RSS received by the MT from NBS shown as NRSS in Fig. 1 may not be sufficient for the MT to send the HMIP registration messages to NFA through NBS. Hence, the MT may send the HMIP registration messages to NFA through OBS. This is called pre-registration. For a smooth and successful handoff from OBS to NBS, MT's HMIP registration with NFA and link and MAC layer associations with NBS must be completed before the RSS of OBS drops below  $S_{min}$ , i.e., before the MT moves beyond the coverage area of OBS [2].

The need for handoff to NBS arises only if MT's direction of motion from  $P$  is in the range  $[\theta \in (-\theta_1, \theta_1)]$ , where  $\theta_1 = \arctan(\frac{a}{2d})$ . Otherwise, the handoff initiation is a false one. Therefore, the probability of false handoff

initiation is

$$\begin{aligned} p_a &= 1 - \int_{-\theta_1}^{\theta_1} f_{\theta}(\theta) d\theta \\ &= 1 - \frac{2\theta_1}{2\pi} = 1 - \frac{1}{\pi} \tan^{-1}\left(\frac{a}{2d}\right) \end{aligned} \quad (1)$$

The probability of handoff failure [2] is given by

$$P_f = \begin{cases} 1, & \tau > \frac{\sqrt{\frac{a^2}{4} + d^2}}{v} \\ \frac{1}{\theta_1} \cos^{-1}\left(\frac{d}{v\tau}\right), & \frac{d}{v} < \tau < \frac{\sqrt{\frac{a^2}{4} + d^2}}{v} \\ 0, & \frac{d}{v} \geq \tau \end{cases} \quad (2)$$

## III. PROPOSED FUZZY LOGIC BASED ADAPTIVE HANDOFF ALGORITHM

The proposed fuzzy logic based system initiates handoff using fuzzy logic; it uses mobiles speed and distance as input and received signal strength as output. The proposed algorithm is adaptive to velocity interference and distances, resulting in fewer dropped calls, better communication quality, potentially lower MT transmit power requirements, gives good performance at different MT speeds, and decreases handoff failure probability.

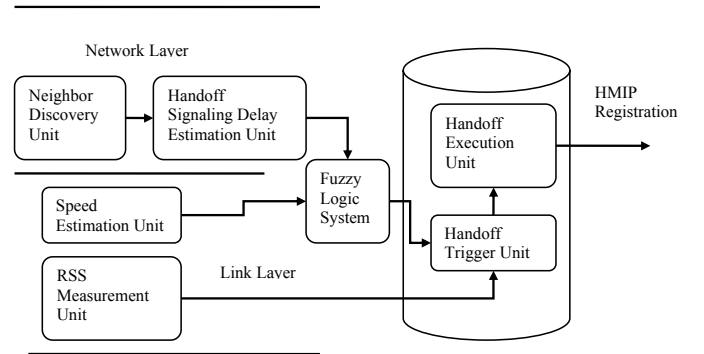


Fig. 2. Proposed fuzzy logic based handoff architecture.

We design an architecture to initiate handoff procedure that is adaptive to the link layer (Layer 2) and network layer (Layer 3) parameters. Then, we develop a handoff management protocol based on the architecture. As the proposed handoff management protocol uses information derived from different layers of TCP/IP protocol stack (e.g., speed information from link layer and handoff signaling delay information from network layer), we call it cross-layer handoff management protocol. The architecture of our proposed FLAH is shown in Fig. 2. Some of these modules collect the link and network layer information useful for handoff management and other modules use this information to decide the appropriate time to initiate and execute the handoff procedures using fuzzy logic system. The functionalities of these units are as follows:

- **The neighbor discovery unit** assists the MT to learn about the neighboring BSs.
- **The handoff signaling delay estimation unit** estimates the delay associated with intra and intersystem handoffs.
- **The speed estimation unit** estimates mobile's speed using VEPSD [7].
- **Fuzzy logic system unit** uses speed and handoff signaling delay information to estimate adaptive RSS threshold ( $S_{ath}$ ) as discussed in Section IV.
- **Handoff trigger unit** collects information from *RSS* measurement unit and when the *RSS* of the serving BS drops below  $S_{ath}$ , the handoff trigger unit sends a trigger to the handoff execution unit to start the HMIP handoff procedures.

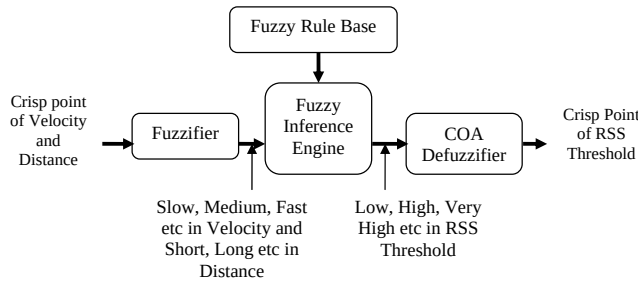


Fig. 3. Fuzzy logic system unit in details.

Finally, the handoff execution unit starts the HMIP registration process at the handoff initiation time calculated by the handoff trigger unit.

#### A. Proposed Fuzzy Logic System

A fuzzy logic rule base is created based on the known sensitivity of handoff algorithm parameters (e.g., RSS threshold, signal to interference, traffic, etc). Fig. 3 shows the fuzzy logic system in details. The task of the fuzzifier is to map crisp value of *distance* and *velocity* to the fuzzy variables. The input fuzzy variable *speed* is assigned to one of the five fuzzy sets, “Slowest”, “Slow”, “Medium”, “Fast” or “Fastest”. This research utilizes the Mamdani FLS. Figure 4 illustrates the membership functions of the input fuzzy variable *velocity* for microcellular system. We used triangular and trapezoidal membership functions and considered 50% overlap of assigned fuzzy sets. Table I shows the rules used in the FLS for microcellular system. Similarly, we defined such rules for macrocellular system. The nonsingleton fuzzifier and the center of area defuzzification method is used.

For example, when the value of *velocity* is “fastest” and the value of *distance* is “short”, this condition indicates that handoff should be encouraged immediately. Hence, output threshold *RSS* is “very high”.



Fig. 4. Membership function of fuzzy variable velocity (microcellular system).

## IV. SIMULATION

### A. Simulation Criteria

Fuzzy logic system determines the value of adaptive *RSS* threshold ( $S_{ath}$ ) to initiate the HMIP handoff procedures using speed and handoff signaling delay information. Then performance is evaluated as follows: first, value of  $d$  is calculated for a desired value of probability of handoff failure  $P_f$  using (2) where  $v$  is MT's speed,  $d$  is MT's distance from the boundary of the serving BS, and  $\tau$  is the handoff signaling delay [2]. Then an approximate value of  $d$  can be calculated using

$$p_f = \frac{\cos^{-1} \frac{d}{v\tau}}{\tan^{-1} \frac{a}{2d}} = \frac{\frac{\pi}{2} - \frac{d}{v\tau}}{\frac{\pi}{2} - \frac{2d}{\sqrt{4d^2 + a^2}}}. \quad (3)$$

And the value of false handoff initiation probability can be calculated using (1).

For CHMP, the value of  $S_{ath}$ , when the MT is at a  $d$  distance from the cell boundary is given by

$$S_{ath} = 10 \log_{10}[P_r(a - d)]. \quad (4)$$

where  $P_r(a - d)$  is the received power at a known distance ( $a - d$ ), given by

$$P_r(x) = P_r(d_0) \left( \frac{d_0}{x} \right)^\alpha + \epsilon. \quad (5)$$

The typical value of  $d_0$  is 1 km for macrocells and 100 m for microcells. The numerical value of  $P_r(d_0)$  depends on different factors, such as frequency, antenna heights, and antenna gains.  $\alpha$  is the path loss exponent and its typical value ranges from 3 to 4 and 2 to 8 for a typical macrocellular and microcellular environment, respectively.  $\epsilon$  is a zero-mean Gaussian random variable that represents the statistical variation in  $P_r(x)$  caused by shadowing. Typical standard deviation of  $\epsilon$  is 8 dB. We simulate and analyze the relationship among *RSS* threshold value, mobile's speed and handoff signaling delay. Then we make comparison between FLAH and CHMP with respect to these relationships.

TABLE I  
FUZZY RULE BASE

| Intrasystem |            |                          | Intersystem |            |                          |
|-------------|------------|--------------------------|-------------|------------|--------------------------|
| Input       |            | Output                   | Input       |            | Output                   |
| Velocity    | Distance   | RSS <sub>threshold</sub> | Velocity    | Distance   | RSS <sub>threshold</sub> |
| Slowest     | Very-short | Near Medium              | Slowest     | Very-short | Low                      |
| Slowest     | Short      | Low                      | Slowest     | Short      | Low                      |
| Slowest     | Medium     | Low                      | Slowest     | Medium     | Low                      |
| Slowest     | Long       | Very Low                 | Slowest     | Long       | Very Low                 |
| Slowest     | Very-long  | Very Very Low            | Slowest     | Very-long  | Very Very Low            |
| Slow        | Very-short | Medium                   | Slow        | Very-short | Medium                   |
| Slow        | Short      | Near Medium              | Slow        | Short      | Near Medium              |
| Slow        | Medium     | Near Medium              | Slow        | Medium     | Near Medium              |
| Slow        | Long       | Low                      | Slow        | Long       | Low                      |
| Slow        | Very-long  | Low                      | Slow        | Very-long  | Low                      |
| Medium      | Very-short | Far Medium               | Medium      | Very-short | Far Medium               |
| Medium      | Short      | Far Medium               | Medium      | Short      | Far Medium               |
| Medium      | Medium     | Medium                   | Medium      | Medium     | Medium                   |
| Medium      | Long       | Medium                   | Medium      | Long       | Medium                   |
| Medium      | Very-long  | Near Medium              | Medium      | Very-long  | Near Medium              |
| Fast        | Very-short | Very High                | Fast        | Very-short | Very High                |
| Fast        | Short      | Very High                | Fast        | Short      | Very High                |
| Fast        | Medium     | High                     | Fast        | Medium     | High                     |
| Fast        | Long       | High                     | Fast        | Long       | High                     |
| Fast        | Very-long  | Far Medium               | Fast        | Very-long  | Far Medium               |
| Fastest     | Very-short | Very Very High           | Fastest     | Very-short | Very Very High           |
| Fastest     | Short      | Very Very High           | Fastest     | Short      | Very Very High           |
| Fastest     | Medium     | Very High                | Fastest     | Medium     | Very High                |
| Fastest     | Long       | Very High                | Fastest     | Long       | Very High                |
| Fastest     | Very-long  | Very High                | Fastest     | Very-long  | Very High                |

### B. Simulation Model

For simulation the following scenarios and parameters are considered: a macro-cellular system with cell size,  $a$ , of 1 km, a micro-cellular system with cell size,  $a$ , of 30m, macro-cell reference distance,  $d_0$ , of 100m, micro-cell reference distance,  $d_0$ , of 1 m were chosen. We assume that the target handoff failure probability,  $P_f$ , is 0.02. We consider that the maximum values of users' speed in micro-cellular and macro-cellular system are 14 km/h and 140 km/h, respectively. Moreover, we assume that the value of  $S_{min}$  is  $-64$  dBm.

### C. Simulation Results

1) *Relationship between  $S_{ath}$  and Speed*: The relationship between  $S_{ath}$  and MT's speed ( $v$ ) for different values of handoff signaling delay ( $\tau$ ) is analyzed. Fig. 5 shows

the relationship between  $S_{ath}$  and  $v$  for different values of  $\tau$  when the serving BS (OBS) belongs to a micro-cellular system. For Macro-cellular system we got similar result.

- Fig. 5 shows that for particular value of  $\tau$ , the value of  $S_{ath}$  increases as MT's speed increases. This implies that for a MT with high speed, the handoff initiation should start earlier compared to a slow moving MT to guarantee the desired handoff failure probability to users independent of their speed.
- Fig. 5 also shows that  $S_{ath}$  increases as  $\tau$  increases. This is because when  $\tau$  is high the handoff must start earlier compared to when  $\tau$  is small. The lower and higher values of  $\tau$  correspond to intra and inter-system handoff, respectively. Therefore, FLS calculates  $S_{ath}$  that is adaptive to both  $v$  and  $\tau$ .

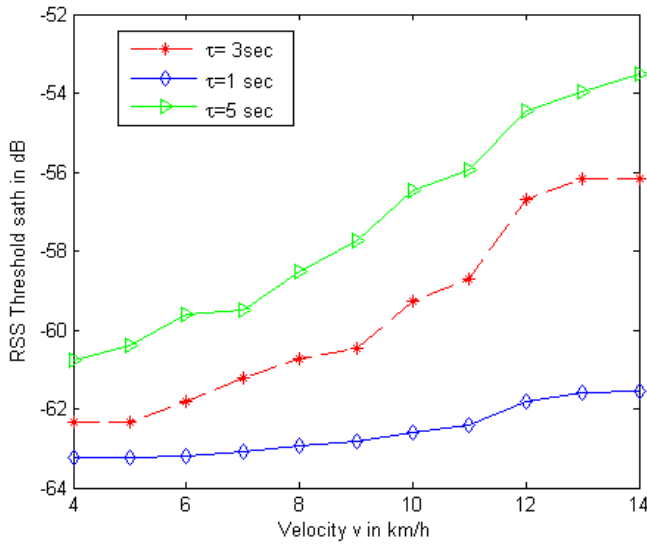


Fig. 5. RSS threshold ( $S_{ath}$ ) for different speed when the OBS belongs to microcellular system.

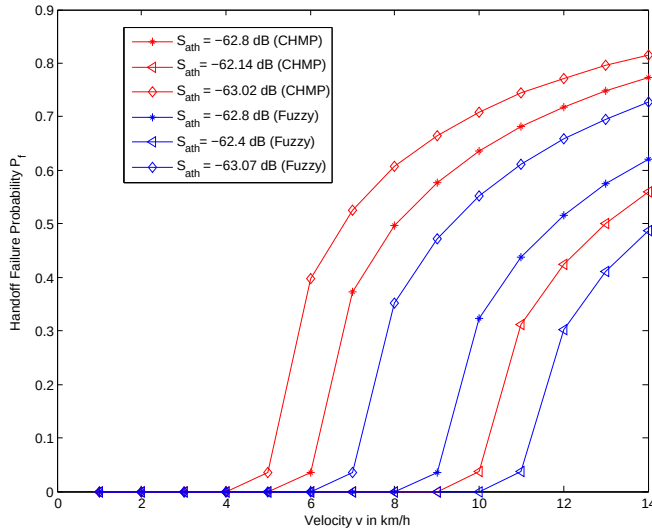


Fig. 6. Relationship between handoff failure probability and speed when OBS belongs to a microcellular intrasystem for fixed  $RSS$ .

2) *Relationship between the Handoff Failure Probability of FLS and Speed:* We calculate  $S_{ath}$  using speed and handoff signaling delay information. Then, we use  $S_{ath}$  to determine the handoff failure probability.

- Fig. 6 shows the handoff failure probability of FLS and CHMP, for different values of speed when the serving BS belongs to a micro-cellular system.
- It shows that reduction in  $P_f$  is achieved in FLS compared to CHMP.
- Fig. 7 shows the results when the serving BS belongs to a macro-cellular intersystem. Similarly we also investigated the results when the serving BS belongs to a macro-cellular intrasystem. Comparison of the

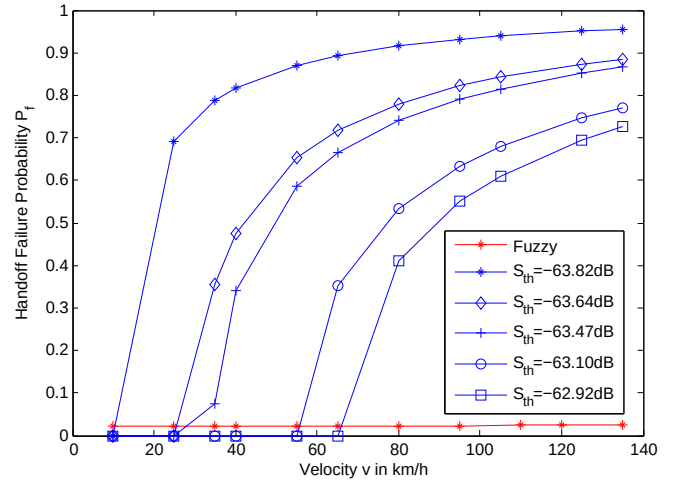


Fig. 7. Relationship between handoff failure probability (Fuzzy) and speed when the OBS belongs to a macrocellular intersystem for fixed  $RSS$ .

results show that when FLS is used and MT's speed is known  $P_f$  is almost constant as the calculated  $RSS$  is adaptive to velocity and distance. On the other hand, for a particular value of fixed  $RSS$  threshold the numerical value of  $P_f$  is different and depends on the numerical value of  $S_{th}$ . This shows that the handoff protocols need to be adaptive irrespective of the type of handoff.

- Fig. 7 also shows that when the speed of the MT is known, 70% to 80% reduction in handoff failure probability is achieved in FLS compared to fixed  $RSS$  based Handoff management system.

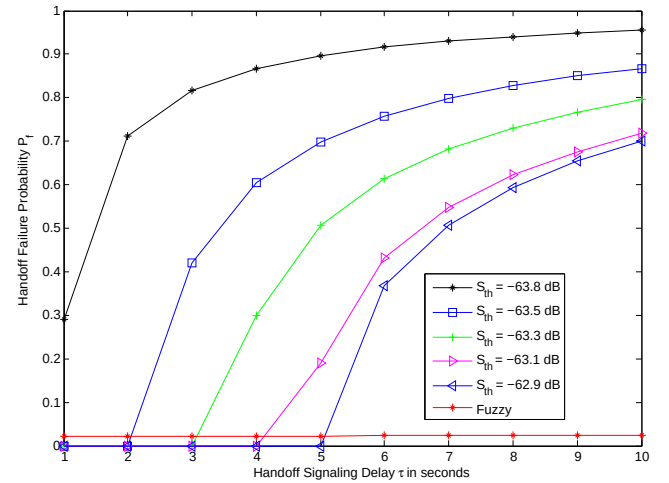


Fig. 8. Relationship between handoff failure probability (Fuzzy) and handoff signaling delay.

3) *Relationship between the Handoff Failure Probability and Handoff Signaling Delay:* Figure 8 shows the handoff failure probability of FLS for different values of handoff



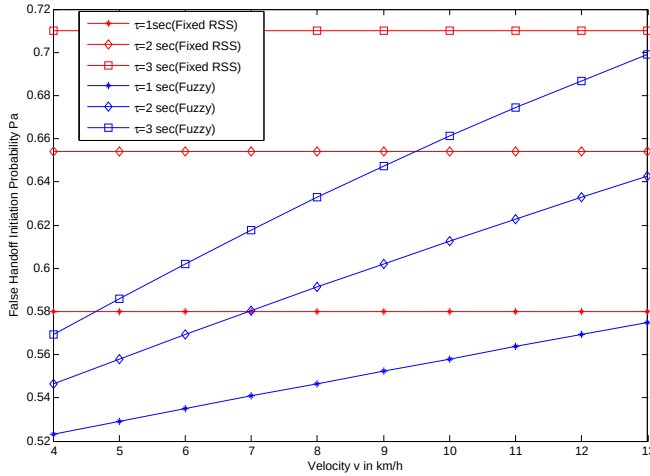


Fig. 9. Relationship between false handoff initiation probability (Fuzzy and Fixed RSS) and speed when the OBS belongs to microcellular system.

signaling delay ( $\tau$ ). The results show that unlike the fixed  $RSS$  based handoff protocols;  $P_f$  remains independent of  $\tau$  in case of FLS. This is because FLS considers  $\tau$  for the tuning membership function of dynamic  $RSS$  threshold. Fig. 8 shows that 70% – 80% reduction in  $P_f$  is achieved in case of FLS compared to the fixed  $RSS$  based handoff protocols. The lower and higher values of  $\tau$  correspond to intra and inter-system handoffs, respectively. Therefore, by incorporating the estimated value of  $\tau$  into dynamic  $RSS$ , the  $P_f$  is limited to the desired value irrespective of users speed and variation of handoff signaling delay.

4) *Relationship between False Handoff Initiation Probability and Speed:* Fig. 9 shows the comparison of the false handoff initiation probability of FLS with the fixed  $RSS$  threshold based algorithms when the serving BS belong to a micro-cellular system. For macro-cellular system similar result is achieved.

- The figure shows that the false handoff initiation probability of FLS is 5 percent to 15 percent less compared to the fixed  $RSS$  threshold based algorithms.
- Thus, FLS achieves up to 15% reduction in the cost associated with false handoff initiation. The use of adaptive  $RSS$  threshold initiates the handoff procedures in such a way that just enough time is there for the successful execution of the handoff.

- Therefore, an adaptive value of  $RSS$  threshold ( $S_{ath}$ ) avoids too early or too late initiation of the handoff process. The former limits the value of handoff failure probability. The later ensures that handoff is carried out smoothly. Thus, FLS optimizes the false handoff initiation probability and handoff failure probability.

## V. CONCLUSION

In this paper, a new fuzzy logic based adaptive handoff management algorithm for next generation heterogeneous wireless system is introduced. It uses mobile's speed and handoff signaling delay information to enhance the performance of hierarchical mobile IP (HMIP) handoff significantly. Performance analysis and simulation results show that FLS significantly enhances the performance of both intra and inter-system handoffs. It can estimate the right time to initiate handoff more precisely. It also significantly reduces the cost associated with the false handoff initiation because it achieves lower false handoff initiation probability than existing handoff protocols.

## REFERENCES

- [1] I. F. Akyildiz, J. McNair, J.S.M. Ho, H. Uzunalioglu, and W. Wang, "Mobility management for next generation wireless systems", *Proc. of IEEE*, Vol. 87, No. 8, pp. 1347-1384, Aug. 1999.
- [2] I. F. Akyildiz and S. Mohanty, "A Cross-Layer (Layer 2+3) handoff management protocol for next-generation wireless systems", *IEEE Trans. on Mobile Computing*, Vol. 5, No.10, pp. 1347-1360, Oct. 2006.
- [3] R. Hsieh, Z. G. Zhou, and A. Seneviratne, "S-MIP: A seamless handoff architecture for mobile IP", *Proc. IEEE INFOCOM '03*, Apr. 2003.
- [4] David A. Maltz, P. Bhagwat, "MSOCKS: An architecture for transport layer mobility", *Proc. IEEE INFOCOM '98*, pp. 1037-1045, Mar.-Apr. 1998.
- [5] Yokota, H., Idoue, A. Hasegawa, T., and Kato T., "Link layer assisted mobile IP fast handoff method over wireless LAN networks", *Proc. ACM MOBICOM '02*, pp. 131-139, Sept. 2002.
- [6] Yaw Nkansah-Gyekye, and Johnson I Agbinya, "Vertical Handoff Decision Algorithm for UMTS-WLAN", *The 2nd International Conference on Wireless Broadband and Ultra Wideband Communications, 2007*, pp. 37, Publication Date: 27-30 Aug. 2007.
- [7] Mohanty, S., "VEPSD: A novel velocity estimation algorithm for next generation wireless systems", *IEEE Trans. on Wireless Communications*, Vol. 4, No.6, pp. 2655-2660, Nov. 2005.
- [8] H. Maturino-Lozoya, D. Munoz-Rodriguez, and H. Tawfik, "Pattern recognition techniques in handoff and service area determination", *Proc. 44th IEEE VTC*, vol. 1, pp. 96-100, Jun. 1994.
- [9] Tripathi N.D., Reed J.H., and VanLandingham H.F., "Fuzzy logic based adaptive handoff algorithms for microcellular systems", *IEEE 49th Vehicular Technology Conference*, Vol. 2, pp. 1419 - 1424, Jul. 1999.

# Automatic Life-Logging: A novel approach to sense real-world activities by environmental sound cues and common sense

Mostafa Al Masum Shaikh<sup>1</sup>, Md. Khademul Islam Molla<sup>2</sup>, and Keikichi Hirose<sup>1</sup>

<sup>1</sup>Dept. of Information and Communication Engineering, The University of Tokyo, Tokyo, Japan

<sup>2</sup>Dept. of Computer Science and Engineering, The University of Rajshahi, Bangladesh

Email: {almasum, molla, hirose}@gavo.t.u-tokyo.ac.jp

**Abstract** — There are many studies that collect and store life log for personal memory. The paper explains how a system can create someone's life log in an inexpensive way to share daily life events with family or friends through social-network or messaging. In the modern world where people are usually busier than ever, family members are geographically distributed due to globalization of companies and humans are inundated with more information than they can process, ambient communications through mobile media or internet based communication can provide rich social connections to friends and family. People can stay connected to their loving ones ubiquitously that they care about by sharing awareness information in a passive way. For users who wish to have a persistent existence in a virtual world – to let their friends know about their current activity or to inform their caretakers – new technology is needed. Research that aims to bridge real life and the virtual worlds (e.g., Second Life, Face book etc.) to simulate virtual living or logging daily events, while challenging and promising, is currently rare. Only very recently the mapping of real-world activities to virtual worlds has been attempted by processing multiple sensors data along with inference logic for real-world activities. Detecting or inferring human activity using such simple sensor data is often inaccurate, insufficient and expensive. Hence, this paper proposes to infer human activity from environmental sound cues and common sense knowledge, which is an inexpensive alternative to other sensors (e.g., accelerometers) based approaches. Because of their ubiquity, we believe that mobile phones or hand-held devices (HHD) are ideal channels to achieve a seamless integration between the physical and virtual worlds. Therefore, the paper presents a prototype to log daily events by a mobile phone based application by inferring activities from environmental sound cues. To the best of our knowledge, this system pioneers the use of environmental sound based activity recognition in mobile computing to reflect one's real-world activity in virtual worlds.

**Index Terms** — Ambient Communication, Life Logging, Virtual World, Activity Recognition, Sound Cues, Auditory Scene Analysis.

## I. INTRODUCTION

Life logs include people's activities performed in specific locations collected from various sources. We envision that with the proliferation of computing power of hand held

devices (HHD), availability of the Internet connectivity and improvements in communication technologies ambient communication will find a universal place at our daily life and allow us to create a vivid and intelligent online social network. Let's consider the following scenario.

*Scenario 1:* Rahman family (Mr. and Mrs. Rahman) lives in Khulna, one of the metropolitan cities of Bangladesh. They have three sons living overseas, one in Texas, another in Ottawa and the youngest one in Bonn of Germany. The Rahman family is now at their age of over 50 and Mr. Rahman had a massive heart operation last year. Mrs. Rahman is also ailing from several sicknesses like diabetics, high blood pressure, etc. The three sons are always worried regarding the well beings of their parents and consequently they often talk to their parents over the phones to know their whereabouts. Though calling to Bangladesh from USA, Canada, and Germany is relatively cheaper now-a-days than before, but having a phone conversation with their parents is not always possible due to various reasons, for example, due to inconvenience in time differences (e.g., when it is 10 am in Khulna it is 11:00 pm in Texas, 12:00 am in Ottawa and 6:00 am in Bonn) that is, when the sons have convenient time to call, their parents are usually sleeping or resting and vice versa. But they are often worried to know at least how their parents are doing everyday. Therefore, let's imagine that Rahman family has internet connectivity at their house and installed an inexpensive system capable of doing the following that may provide mental peace to their sons by providing ambient communications through the internet. The system makes life logging of daily activities by detecting and recognizing sound cues from the surrounding environments and sends email message(s) to their sons reporting their daily life-sketch. In this case, an example email message containing life log for a particular day as following might be very relieving to the sons. "Your parents woke up at 7:30 am today and they had breakfast around 8:15 in the morning. They watched TV several times in the day. Went out of home for two times and walked in the roads and parks. They took lunch and dinner at around 2 PM and 8 PM. Your mother went to

*toilet for 5 times and father went to toilet 6 times in a day. They had communicated with each other or other people by talking. It seems they are doing fine."*

### A. Background

The idea of a "life-log" or a personal digital archive is a notion that can be traced back at least 60 years [1]. Since then a variety of modern projects have spawned such as the *Remembrance Agent*[2], the *Familiar*[3,4], *myLifeBits*[5], *Memories for Life*[6] and *What Was I Thinking* [7]. In [8] the authors evaluate the user's context in real time and then use variables like current location, activity, and social interaction to predict moments of interest. Audio and video recordings using a wearable device can then be triggered specifically at those times, resulting in more interest per recording. Some previous examples of this approach are the *Familiar* and *iSensed* systems [3,4,8] which structure multimedia on the fly; the *eyeBlog* system [9] which records video each time eye contact is established; and the *SenseCam*[10] which records images and sound whenever there's a significant change in the user's environment or the user's movement. Life log includes people's experiences which are collected from various sensors and stored in mass storage device. It is used to support user's memory and satisfy user's needs for personal information. If he wants to inform other people of his experience, he can easily share his experience with them by means of providing his life log.

To collect life log (e.g., GPS based location, SMS, call, charging, MP3, photos taken, images viewed, and weather information, etc) smart phones (e.g., iPhone 3G) are usually used. Smart phone is a mobile device that includes color LCD screen, mass storage, large memory, and communicative function by using Wi-Fi, Bluetooth, and infrared. It also has a variety of software such as scheduler, address book, media player, and e-book. Mika Raento developed a framework for collecting contexts from smart phone [11], which collects GSM Cell ID, Bluetooth, GPS data, phone data, SMS data, and media information that are transmitted to the server. The contexts could be provided for other contents as additional information. Panu collects log data from mobile devices, and extracts features by pre-processing the log data [12]. The mobile device uses GPS log, microphone, temperature, moisture, and light sensor. MyLifeBits Project is one of the implementations of personal record database system [5]. Personal information is collected by PC, SenseCam and so on, and stored in database server with relationships among personal information. However, user faces difficulties to explore and search contents because of large amount of personal data. KeyGraph-based mobile contents management system was suggested to manage user's information in mobile device, which extracted important information using KeyGraph algorithm and provided searching or exploring contents [13]. The problem of the system is using only log data. If

analysis and inference of the data was added to the system, it would give better performance.

### B. Our Approach

Our approach is far different from others in terms of logging daily events by detecting real world activities (e.g., laughing, talking, traveling, cooking, sleeping, etc.) and situational aspects of the person's (e.g., inside a train, at a park, at home, at school, etc.). We focus how to perform daily life logging and to provide a virtual representation of humans in terms of activity and their surrounding by inferring human-activity from environmental and object-interaction related sound cues as well as common sense knowledge. We have collected 114 types of acoustic sounds that are usually produced during object interaction (e.g., cooking pan jingling sound while cooking) or by the environment itself (e.g., bus/car passing sound while on a road) or by a deliberate action of a person (e.g., laughing, speaking). These sample sounds serve as the underlying clues to infer a person's activity and one's surroundings with the help of common sense knowledge (e.g., while the system identifies cooking pan's jingling and chopping board sound as consecutive cues and system's local time indicates evening then from common sense database the system infers this activity as 'cooking'). As an implementation platform we plan to use Second Life (SL) [14] to simulate living and log daily life activities. SL is one of the most popular Internet-based virtual worlds in terms of subscribers with over 500,000 active users. Virtual world simulators represent one of the upcoming successful niches for online services for entertainment, advertisement, business, but also health and elderly care. However, in the usual interaction model, there is still a gap (between the virtual world and the real one) for users who wish to have a persistent existence in a virtual world to let their friends know about their current activity or to inform their caretakers. Specifically speaking, a user cannot automatically mirror/reflect his current movements, activities or surrounding environment (e.g., park, shopping mall, etc.) in his real life to the virtual life of his avatar. Only very recently [15] the mapping of real-world activities to virtual worlds has been attempted by processing multiple sensors data along with inference logic for real-world activities. Detecting or inferring human activity using such simple sensor data is often inaccurate and insufficient. Moreover deploying a ubiquitous sensor network at outdoor environment is expensive and not feasible.

## II. SYSTEM ARCHITECTURE

Fig. 1. serves as the top-level pipelined architecture of the system. Because of their ubiquity we plan to use mobile phones (e.g., iPhone) to deploy this application that will

capture environmental sound at some intervals to be processed. According to Fig. 1., a mixed signal is passed through a robust signal processing and sound classes are detected by trained HMM classifiers. Based on the detected sounds and common sense knowledge regarding human activity, object interaction, ontology of human life (e.g., daily life of a student, or a salary man etc.) and

temporal information (e.g., morning, noon etc.) are applied to infer both the activity and the surrounding of the person. This information is sent to the client application of SL as a text message and this message is parsed by SL client to map the activity of the person after rendering an appropriate virtual world for that activity.

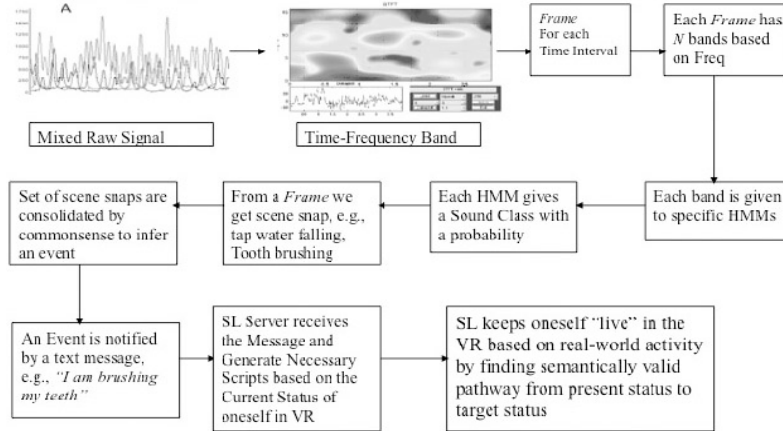


Fig. 1. System Architecture (in brief)

#### A. Development Challenges

In order to develop this system we have the following challenges and concerns:

##### *Environmental sound corpus and features*

For simplicity our collected sound corpus is limited to sound cues related to certain genres like, cooking, bathroom activity, human physiological action, outdoor, home, etc. Since the recognition rate is the only performance measure of such system, it is hard to judge how suitable the selected feature is. To solve this problem, it is essential to analyze the feature itself, and measure how good the feature itself is. It is concluded that MFCC may be the best transformation for non-speech environmental sound recognition. A similar opinion was also articulated in [16,17]. These finding provide the essential motivation for us to use MFCC in extracting features for bathroom sound classification.

##### *Accuracy of Acoustic Event Detection/Classification*

Microphones that capture sounds with sufficient fidelity for automated processing are also much cheaper in comparison with other sensors. Acoustics-based behavioral understanding systems work from a distance: they do not constrain subjects by requiring them to wear special devices; a prerequisite for individuals with dementia. Computational auditory scene analysis (CASA)[18], the understanding of events through processing environmental sounds and human vocalizations, has become an increasing important

research area. There are however many classifier approaches in the literature, e.g., those based on Hidden Markov Models (HMMs), Artificial Neural Networks (ANN), Dynamic Time Warping (DTW), Gaussian Mixture Models (GMM), etc. From these options, we have chosen an approach based on HMM as the model has a proven track record for many sound classification applications [19]. Another advantage is that it can be easily implemented using the HMM Tool Kit (HTK) [20].

##### *Computing power of iPhone or Hand Held Devices*

The computing potentials of HHDs are increasing in a rapid manner with respect to memory and incorporation of full-fledged operating system but they are still inferior to personal computer comparing to the speed of data processing. Running of the core signal processing task requires high-end computing and therefore the processing should be done by means of external devices connected to the HHDs through the Bluetooth radio. Therefore in this case a light process to capture environmental sounds after some intervals will be running on the HHDs to be transmitted for activity detection.

##### *Privacy and Social Issues*

To manage their exposure, users should be able to disconnect from the virtual worlds at any time or be able to interrupt the sound capturing of their activities. The actual representation of users in the virtual world may be disclosed according to pre-configured user policies or users list. Additional privacy issues are related to the collection of environmental data by means of people carrying devices to different places and data collected

about them. We consider important privacy and security challenges related to people-centric sensing.

#### *Common sense knowledge base*

What this system try to create is a sense of presence that allows users to “see and feel” a remote place. From the analysis of captured signals the system gets aware of the objects available within the vicinity of the person carry the HHD. This object level information is mapped with a common sense knowledge base to infer the activity. For example, if the system detects that the sound cues represent frying pan, sink water, and chop board from consecutive input samples the developed common sense knowledge can infer a cooking activity. Moreover to minimize error in inference we incorporate temporal and ontology of daily life of an individual (e.g., student, salary man, old people etc.). For example, if system detects sound cues that initially inferred that the person is in the road, but according to the temporal information if it is at night 2:00 am on Monday and from the daily life ontology of a salary man (i.e., the user is a salary man in real world) the time frame indicates “resting”, the system may either show proactive behavior or ignore the initial inference.

#### *Scalability of the solution*

Our default approach is to run activity recognition algorithms on the mobile device to decrease communication costs and also to reduce the computational burden on the server. However, we recognize that signal processing based classification algorithms may be too computationally intensive for handheld devices and propose to run the classifier on a back-end server in this case. This may be particularly appropriate during the training phase of a classification model.

#### *B. Data Collection and Annotation*

Our collected sound samples are used to infer the location and activity being performed at that location by a person. Table 1 explains the list of activities and related locations of our interest. The sound clips are collected and annotated obliging such area of interest. For example, sound clips like, “Shower water falling” is annotated as a toilet event related to “taking shower” activity.

TABLE I  
SAMPLE SOUND CUES COLLECTED ACCORDING TO THIS LIST OF ACTIVITIES AND LOCATIONS OF INTEREST

| Location                 |                                  |            |                       |                   |                              |                               |                         |
|--------------------------|----------------------------------|------------|-----------------------|-------------------|------------------------------|-------------------------------|-------------------------|
| Indoor                   |                                  |            |                       |                   | Outdoor                      |                               |                         |
| At Home                  | At Office/<br>University         | At Kitchen | At Toilet             | At Gym            | On Road                      | In Train/<br>Train<br>Station | At Public<br>Place      |
| Activities               |                                  |            |                       |                   |                              |                               |                         |
| Watching TV              | Working<br>with PC               | Cooking    | Taking<br>Shower      | Doing<br>Exercise | Walking<br>along the<br>road | Waiting<br>for Train          | Eating at<br>Restaurant |
| Talking on Phone         | Attending<br>lecture/<br>meeting | Eating     | Urinating             |                   |                              | Traveling<br>by Train         | Playing in<br>Pachinko  |
| Talking with a<br>person | Talking on<br>Phone              |            | Washing<br>hand/ face |                   |                              |                               | Singing in<br>Karaoke   |
| Waking up from<br>sleep  | Talking<br>with a<br>person      |            |                       |                   |                              |                               | Shopping                |
| Getting ready            | Drinking<br>Tea/ Coffee          |            |                       |                   |                              |                               | Talking                 |
| Singing/ Music           |                                  |            |                       |                   |                              |                               | Hanging<br>around park  |
| Sleeping                 |                                  |            |                       |                   |                              |                               |                         |
| Drinking Tea/<br>Coffee  |                                  |            |                       |                   |                              |                               |                         |
| Vacuum Cleaning          |                                  |            |                       |                   |                              |                               |                         |
| Sitting Idle at<br>home  |                                  |            |                       |                   |                              |                               |                         |

### C. Ontology of Daily Life

In this initial prototyping three types of users are considered and based on empirical study the ontology of daily life (listed in Table 2) of each user type is considered along with common sense knowledge base to perform legitimate inference. For example, if the system

detects a sound cue related to eating activity during BN (before noon) time frame on a weekday and if the user is a service-holder, it doesn't log the event because the user is not happened to be present in a kitchen around that time. We also plan to provide the flexibility to personalize such ontology of daily life on the basis of each respective user.

TABLE II  
ONTOLOGY OF DAILY LIFE BASED ON THE PRESENCE AT PARTICULAR PLACES AT PARTICULAR TIME

| Role             | Weekday |                |                   |                |                   |                      |         | Weekend |            |                |                   |                      |                   |         |
|------------------|---------|----------------|-------------------|----------------|-------------------|----------------------|---------|---------|------------|----------------|-------------------|----------------------|-------------------|---------|
|                  | EM      | M              | BN                | AN             | E                 | N                    | LN      | EM      | M          | BN             | AN                | E                    | N                 | LN      |
| Graduate Student | H       | H, T           | H, U, K, T, R, Tr | U, G, T, P     | H, U, R, Tr, G, P | H, U, R, Tr, K, T, P | H, K, T | H       | H, T       | H, K, T, R, Tr | G, T, P           | H, R, Tr, K, T, G, P | H, R, Tr, K, T, P | H, K, T |
| Service-Holder   | H       | H, K, T, R, Tr | O, T              | O, T           | O, T, H, R, Tr, P | H, P, R, Tr, T       | H       | H       | H, K, T    | H, T, P, R, Tr | H, T, P, R, Tr, G | H, T, R, Tr, P       | H, P, R, Tr, T    | H       |
| Elderly People   | H, T    | H, T, R, K     | H, T, R, Tr, P    | H, T, R, Tr, P | H, T, R, Tr, K, P | H                    | H, T    | H, T    | H, T, R, K | H, T, R, Tr, P | H, T, R, Tr, P    | H, T, R, Tr, K, P    | H                 | H, T    |

H: Home; O: Office; U: University; K: Kitchen; T: Toilet; G: Gym; R: Road; Tr: Train; P: Public place

EM: Early Morning [03:00 – 05:00], M : Morning [05:00 – 8:00], BN: Before Noon [8:00 – 12:00], AN: After Noon [12:00 – 17:00], E : Evening [17:00 – 20:00], N : Night [20:00 – 01:00], LN: Late Night [01:00 – 03:00]

### III. SCHEMATIC SOLUTION

The system continuously listens to the environment and captures sound regularly at 20 seconds of interval for 10 seconds of duration. Each captured sound-clip is sent to a computer wirelessly where the clip is processed. Thus for a two-minute interval the system gets four sound clips. These four sound clips are considered to infer an activity at that moment. Each sound clip is processed by a HMM

network that outputs a particular sound class. Each sound class actually represents some real-word objects out of which interaction the sound maybe produced. For example, Let's assume that the system receives the following four sound clips (as shown in Fig. 2.) to process and the HMM classifies the four clips into the following four sound classes: "Cutting something on a Chop Board", "Water Falling on Sink", "Collision of Cooking pan and Spoon", and "Racking of Plates".

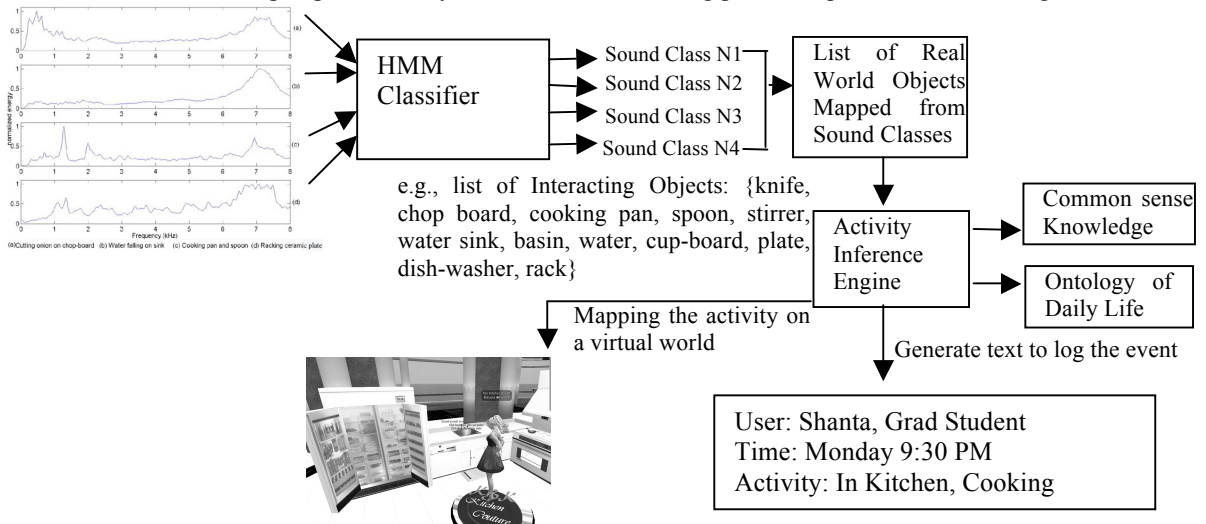


Fig. 2. A Schematic solution to infer "cooking" activity from sound cues

These four classes of sounds are pre-mapped as following,

- “Cutting something on Chop Board” → {Knife, Chop board}
- “Collision of Cooking pan and Spoon” → {Cooking pan, spoon, stirrer}
- “Water Falling on Sink” → {water sink, basin, water}
- “Racking of Plates” → {cup-board, plate, dishwasher, rack}

According to Fig. 2., the list of interacting objects is dealt with common sense knowledge. The common sense knowledge is extracted from ConceptNet [21] and outputs a probabilistic value representing a relationship between an activity and object. In the above example the objects yield a maximum probability of having a relationship with “cooking” activity and the near candidates are “eating”, “drinking tea/coffee”. From the ontology of daily life the system finds that it is likely to have “cooking” activity at the concerned time (Table 2). Therefore, the activity inference engine considers this event as a legitimate event and logs it by either animating this event on a virtual world or generating a text message.

#### IV. CONCLUSION

Software like Google Earth and Microsoft Virtual Earth map the real world with great accuracy in terms of geographical locations and local features. We believe that the next step is to enable a user to represent real world activities in this kind of virtual world whereby personal avatars will be able to reflect what the real persons are doing in the real world. This type of application is a source of fun for young generation while it has lots of potential regarding virtual shopping mall in e-commerce context, easy monitoring or life logging for elderly people for the caregivers etc.

#### REFERENCES

- [1] Bush, V., “As we may think”, *Atlantic Monthly*, 1945
- [2] B. Rhodes and T. Starner, “Remembrance Agent: A Continuously Running Automated Information Retrieval System,” Proc. 1st Int’l Conf. Practical App. of Intelligent Agents and Multi-Agent Technology, 1996, pp. 487-495.
- [3] B. Clarkson and A. Pentland, “Unsupervised Clustering of Ambulatory Audio and Video,” Proc. IEEE Int’l Conf. Acoustics, Speech, and SignalProcessing, IEEE CS Press, vol. 6, 1999, pp. 3037-3040
- [4] B. Clarkson, K. Mase, and A. Pentland, “The Familiar: A Living Diary and Companion,” Proc. ACM Conf. Computer-Human Interaction, ACM Press, pp. 271-272.
- [5] J. Gemmell et al., “MyLifeBits: Fulfilling the Memex Vision,” Proc. ACM Multimedia, ACM Press, 2002, pp. 235-238.
- [6] A. Fitzgibbon and E. Reiter, “‘Memories for Life’: Managing Information over a Human Lifetime,” UK Computing Research Committee Grand Challenge proposal, 2003.
- [7] S. Vemuri and W. Bender, “Next-Generation Personal Memory Aids,” *BT Technology J.*, vol. 22, no. 4, 2004.
- [8] M. Blum, A. Pentland, G. Troster, et al., “InSense: Internet-Based Life Logging”, *IEEE Multimeddia* vol 13, Issue 4, pp.40-48, 2006
- [9] C. Dickie et al., “Augmenting and Sharing Memory with eyeBlog,” Proc. 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences, ACM Press, 2004
- [10] J. Gemmell et al., “Passive Capture and Ensuing Issues for a Personal Lifetime Store,” Proc. Continuous Archival and Retrieval of Personal Experiences, ACM Press, 2004.
- [11] M. Raento, A. Oulasvirta, R. Petit, H. Toivonen, “ContextPhone: A Prototyping Platform for Context-aware Mobile Applications”, *IEEE Pervasive Computing*, vol 4, Issue 2, pp. 51–59, 2005.
- [12] K. Panu, M. Jani, K. Juha, K. Heikki, M. Esko-Juhani, “ContextPhone: Managing Context Information in Mobile Devices”, *IEEE Pervasive Computing*, vol 2, issue 3, pp. 42–51, 2003.
- [13] K. Kim, M. Jung, S. Cho, “KeyGraph-based chance discovery for mobile contents management system Source”, *Int’l Journal of Knowledge-based and Intelligent Engineering Systems archive*, Vol 11, Issue 5, pp.313-320, 2007
- [14] Linden Research Inc. SecondLife, <http://secondlife.com/>
- [15] M. Mirco, M. Emiliano, D. L. Nicholas, B. Shane, C. Tanzeem, T.C. Andrew, “The Second Life of a Sensor: Integrating Real-world Experience in Virtual Worlds using Mobile Phones”, Proc. Fifth Workshop on Embedded Networked Sensors (HotEmNets 2008), Virginia, 2008.
- [16] Okuno, H.G., Ogata, T., Komatani, K., Nakadai, K.: Computational Auditory Scene Analysis and Its Application to Robot Audition, International Conference on Informatics Research for Development of Knowledge Society Infrastructure (ICKS), (2004), 73–80
- [17] Eronen, A., Tuomi, J., Klapuri, A., Fagerlund, S., Sorsa, T., Lorho, G., Huopaniemi, J.: Audio-based Context Awareness - Acoustic Modeling and Perceptual Evaluation, Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03), vol. 5, (2003), V-529-532
- [18] Wang, D. Brown, G., 2006. Computational Auditory Scene Analysis: Principles, Algorithms and Applications, Wiley-IEEE Press.
- [19] Rabiner, L. R., Juang, B. H.: Fundamentals of Speech Recognition, PTR Prentice-Hall Inc., New Jersey, 1993
- [20] Young, S.: The HTK Book, User Manual, Cambridge University Engineering Department, 1995
- [21] Liu, H., and Singh, P., “ConceptNet: A Practical Commonsense Reasoning Toolkit”, *BT Technology Journal*, Vol 22, Issue 4, pp. 211-226, Kluwer Academic Publishers, 2004.

# A Novel Fuzzy Method to Traffic Light Control Based on Unidirectional Selective Cellular Automata for Urban Traffic

Moein Shakeri<sup>1</sup>, Hossein Deldari<sup>1</sup>, Alireza Rezvanian<sup>2</sup>, and Homa Foroughi<sup>1</sup>

<sup>1</sup> Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Khorasan, Iran

<sup>2</sup> Department of Computer Engineering, Azad University of Qazvin, Qazvin, Iran

**Abstract** — Vehicular travel which demands on the concurrent operations and parallel activities is increasing throughout the world. In this paper to control urban traffic, we study the optimization of traffic light controller in a city and present a fuzzy algorithm based on cellular automata. In existent system factors like priority of streets of intersection and width/length of streets are assumed equal. However, in real situations parameters like time during the day, density of vehicles of street and number of shopping centers have determinant effects on amount of traffic of streets. To overcome such limitations we propose a three leveled fuzzy system. At first level priority of street is computed based on fuzzy rules. At second level real velocity of vehicles is calculated. We use CA for simulating vehicles' density transmission. At third level by regarding priority of street and amount of density, decision for changing status of traffic light is done.

**Index Terms** — Cellular Automata Model, Fuzzy Control, Traffic Control, Traffic Light Control, Unidirectional Selective Cellular Automata, Urban Traffic.

## I. INTRODUCTION

Transportation research has the goal to optimize transportation flow of people and goods. As the number of road users constantly increases, and resources provided by current infrastructures are limited, intelligent control of traffic will become a very important issue in the future [6]. Optimal control of traffic lights using sophisticated sensors and intelligent optimization algorithms might therefore be very beneficial. Optimization of traffic light switching increases road capacity and traffic flow, and can prevent traffic congestions. In the recent years there were strong attempts to develop a theoretical framework of traffic science among the physics community. Consequentially, a nearly completed description of highway traffic [7, 11], e.g., the "Three Phase Traffic" theory, was developed. This describes the different traffic states occurring on highways as well as the transitions among them. Also the concepts for modeling vehicular traffic are well developed. Most of the models introduced in the recent years are formulated using the language of cellular automata (CA) [3,4,12,14-17]. Unfortunately, no comparable framework for the description of traffic states in city networks is present. In contrast to the highway networks, where individual highway segments can be

treated separated, the structure elements of city networks exert an immense influence onto the traffic dynamics [7].

In existent urban traffic systems priority of intersection streets and also width and length of streets are assumed equal; so they have no role in making decision for changing the status of the traffic light. To overcome such limitations, we propose a novel fuzzy system that momentarily computes priority of the street based on fuzzy rules and regarding to environmental factors. Furthermore, in the proposed system length of streets are not considered the same. Subsequently a cellular automaton is employed for modeling both density transmission and type of movement of vehicles. In fact each street forms a lattice and each member of this structure is regarded as a separate cell in cellular automata space.

The paper is organized as follows: in section II cellular automata is described concisely and then unidirectional selective cellular automata is introduced for further simulations. In section III we briefly review some existing traffic light controller systems. Our proposed system and its technical details are described in section IV. Simulation results are represented in section V and finally we draw conclusions in section VI.

## II. CELLULAR AUTOMATA

Cellular automata (CA) are a class of discrete dynamical systems [13], consisting of an array of nodes (cells) of some dimension,  $n$ . Each cell can be in one of  $k$  different states at a given tick of the clock. At each discrete tick of the clock, each cell may change its state, in a way determined by the transition rules of the particular CA [1,5]. The transition rules describe precisely how a given cell should change states, depending on its current state and the states of its neighbors. Let  $n$  be the dimension of the lattice,  $k$  the number of states,  $T$  the transition rule function,  $C_t(i_1, \dots, i_n)$  the state of the cell at position  $(i_1, \dots, i_n)$  at time  $t$ ,  $N_t(i_1, \dots, i_n)$  the values (given in a specific order) of the neighboring cells to this location at time  $t$ . Then the dynamics of the CA is completely specified by the initial states of all the cells,  $C_0$ , along with the recursion rule (Equation 1).

$$C_{t+1}(i_1, \dots, i_n) = T(N_t(i_1, \dots, i_n)) \quad (1)$$



### A. Unidirectional Selective Cellular Automata

General structure of CA is as follows:

Definition 1: A cellular automata is a 4-tuple  $CA = (Q, d, V, a)$  where:

1.  $Q$  is the nonempty state set that a cell can assume.
2. The dimension  $d$  of the cellular space, building a bounded subspace on  $Z^d$  ( $Z$  denoting the integers). Note that, when  $d = 1$ ; corresponds to the case of a linear array.
3. The neighborhood array  $(V_0, V_1, \dots, V_k), V_j \in Z; V_j \in [0, k]$ . Being  $x \in Z^d$  an automata location,  $V(x) = \{x + V_0, x + V_1, \dots, x + V_k\}$  describes the finite list of  $k + 1$  neighbors directly associated to it. In the case of a linear array,  $x = i \in Z$ ; the neighborhood array  $\{0, -1, +1\}$  index the central, the left and the right cells.
4. The transition function (or relation) defines the dynamical local state transformations to be applied to the cells. In standard automata without memory register the deterministic transitions are of the form  $: Q^{k+1} \rightarrow Q$ .

Based on urban traffic application, in this paper we change the basic characteristics of CA to define a new kind of it. Let us clarify our proposed CA visually. Fig.1 shows a street and each segment is a cellular space.

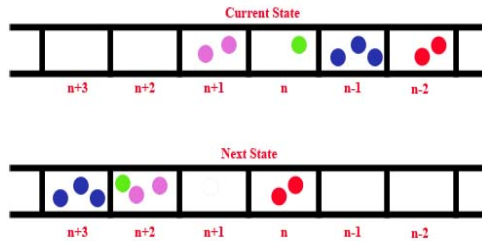


Fig. 1. Current and Next State of a Sample Street

a) While in classical CA, next state of each cell is defined according to state of its both left and right neighbors; in the proposed CA we take into account just the neighbors that are located in one side. For example in classical CA, when the neighborhood radius is 1, next state of  $n$ th cell is defined regarding to state of  $n$ th,  $n-1$ th and  $n+1$ th cells. Due to the fact that transmission of density is always in one direction in the street (in Fig.1 from right to left), one can say, only the right neighbors of each cell are effective in defining next state of it. Because of this new feature, proposed CA is named “Unidirectional”.

b) In the definition of CA, the neighbors of each cell are directly associated to it. However, in the proposed CA, neighbors of cell are selected based on a density transmission function at each step. This function is described in more details in section IV. It means that neighbors of  $n$ th cell are not necessarily  $n-1$ th and  $n+1$ th cells, but also they are selected via a predefined function. Due to this fact, the proposed CA is called “Selective”.

Combination of a and b features results in a new kind of cellular automata so-called Unidirectional Selective Cellular Automata (USCA). We use USCA at the second level of the proposed system toward simulating vehicles' density transmission function.

### III. RELATED WORK

Traffic dynamics bear resemblance with, for example, the dynamics of fluids and those of sand in a pipe. There are two common approaches for modeling traffic [6]; macroscopic and microscopic models.

- **Macroscopic Traffic Models:** Macroscopic traffic models are based on gas-kinetic models and use equations relating traffic density to velocity. These equations can be extended with terms for build-up and relaxation of pressure to account for phenomena like stop-and-go traffic and spontaneous congestions.

- **Microscopic Traffic Models:** In contrast to macroscopic models, microscopic traffic models offer a way of simulating various driver behaviors. A microscopic model consists of an infrastructure that is occupied by a set of vehicles. Each vehicle interacts with its environment according to its own rules.

Existent traffic prevention approaches can be categorized as follows [10]:

- **Global Strategies:** The considered global strategies are a “synchronized strategy”, a “green wave strategy” and a “random strategy”. In the case of the synchronized strategy all traffic lights switch synchronously to green (red) for the east (north) bound vehicles and vice versa. In the case of a green wave strategy, i.e., adjacent traffic lights switch with a defined offset. Additionally, an appropriate offset has to be determined for the green wave strategy; this is equal to the free flow travel time for the depicted case

- **Adaptive Strategies:** In the following three different adaptive strategies are presented. The first investigated adaptive strategy is the “switching based on the queue length”. Here a traffic signal switches if the length of a vehicle queue in front of a red light trespasses a certain value. Further investigated adaptive strategies are the “switching based on waiting time” and the “switching in analogy to a neural network”. In the first case a traffic light switches to red if the green phase is not used by a vehicle for a certain time. In the case of switching in analogy to a neural network the traffic lights act like integrate-and-fire neurons. After the switching process (fire) the potential is reset to zero again. The switching in analogy to a neural network strategy leads to similar results like the switching based on the waiting time [7, 8].

#### IV. THREE LEVELED FUZZY PROPOSED SYSTEM

In the proposed models utilize adjusting traffic light for urban traffic control, priority of streets of intersection and also width and length of streets are assumed equal. whereas in real world urban traffic, parameters like time during the entire day, density of vehicles of the street, number of shopping centers, offices that have plenty of clients; have determinant effects on amount of traffic of streets. Furthermore, the priority of each street is defined by considering these environmental factors. Taking into account these parameters we propose a fuzzy system that priority of each street is computed at each moment, also this system doesn't set any constraint on width and length of streets

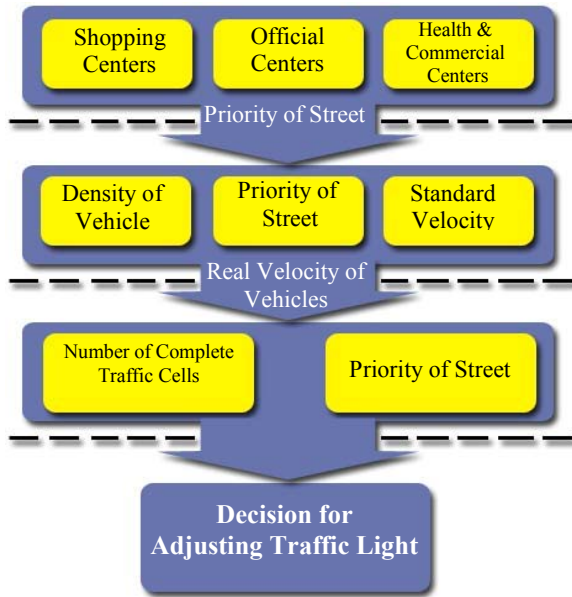


Fig. 2. Proposed System Flow

The proposed system has three levels; at the first level priority of each street is computed momentarily. At the second level real velocity of vehicles of every street is calculated at specific moments. Then cellular automata model is applied to specify density transmission and also type of movement of vehicles in the cellular spaces of each street. And finally at the third level, decision is done for changing status of traffic light.

##### A. First Level of the Proposed System

In this paper, pausing coefficient is defined as a significant factor to determine priority of each street. Streets with larger quantity of stopped or paused vehicles have more priority. Pausing coefficient is totally distinctive among three following categories:

##### • Health and Commercial Centers:

In these centers, density of clients is computed momentarily. Since number of clients directly affects on

number of paused or stopped vehicles in the street, priority of street is dependant to amount of density of clients with a linear function. So fuzzy sets of health and commercial centers are defined as Fig. 3. Note that M1, M2, M3, M4, M5 fuzzy sets are related to low, high, medium, high and low densities respectively.

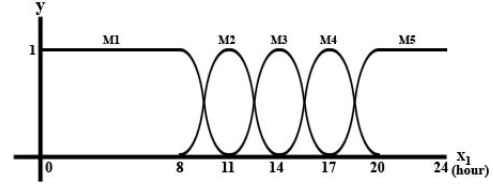


Fig. 3. Fuzzy Sets of Density of Clients of Health and Commercial Centers

##### • Official Centers:

In these centers density of clients at each moment could affect directly on priority of the street. Fuzzy sets of official centers are determined as Fig. 4. L1, L2, L3 fuzzy sets are related to low, high and low densities respectively.

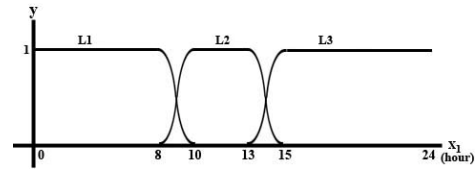


Fig. 4. Fuzzy Sets of Density of Clients of Official Centers

##### • Shopping Centers:

Like two previous centers amount of clients of these centers, has determinant effect on priority of the street, fuzzy sets of these centers can be considered as Fig. 5. Notice that K1, K2, K3, K4, K5 fuzzy sets are related to low, high, medium, high and low densities respectively.

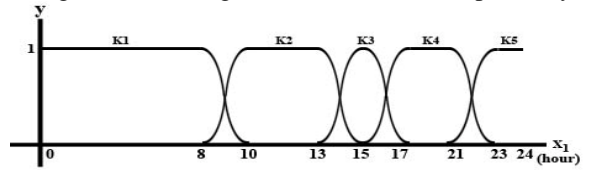


Fig. 5. Fuzzy Sets of Density of Clients of Shopping Centers

##### 1) Fuzzy Rules:

By definition of the input and output fuzzy sets, a total of 24 IF-THEN rules are defined as follows [2]:

$$\text{IF } (x_1 \text{ is } M1) \text{ and } (x_1 \text{ is } L1) \text{ and } (x_1 \text{ is } K1) \text{ then } F1=[A1*(1-y11)]*[A2*(1-y21)]*[A3*(1-y31)] \quad (2)$$

$$\text{IF } (x_1 \text{ is } M1) \text{ and } (x_1 \text{ is } L1) \text{ and } (x_1 \text{ is } K2) \text{ then } F2=[A1*(1-y11)]*[A2*(1-y21)]*[A3*y32] \quad (3)$$

$$\text{IF } (x_1 \text{ is } M1) \text{ and } (x_1 \text{ is } L2) \text{ and } (x_1 \text{ is } K1) \text{ then } F3=[A1*(1-y11)]*[A2*y22]*[A3*(1-y31)] \quad (4)$$

$$\text{IF } (x_1 \text{ is } M1) \text{ and } (x_1 \text{ is } L2) \text{ and } (x_1 \text{ is } K2) \text{ then } F4=[A1*(1-y11)]*[A2*y22]*[A3*y32] \quad (5)$$

$$\text{IF } (x1 \text{ is } M2) \text{ and } (x1 \text{ is } L1) \text{ and } (x1 \text{ is } K1) \text{ then} \\ F5=[A1*y12]*[A2*(1-y21)]*[A3*(1-y31)] \quad (6)$$

$$\text{IF } (x1 \text{ is } M2) \text{ and } (x1 \text{ is } L1) \text{ and } (x1 \text{ is } K2) \text{ then} \\ F6=[A1*y12]*[A2*(1-y21)]*[A3*y32] \quad (7)$$

$$\text{IF } (x1 \text{ is } M2) \text{ and } (x1 \text{ is } L1) \text{ and } (x1 \text{ is } K3) \text{ then} \\ F7=[A1*y12]*[A2*(1-y21)]*[A3*(1-|1-y33|)/2] \quad (8)$$

$$\text{IF } (x1 \text{ is } M2) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K1) \text{ then} \\ F8=[A1*y12]*[A2*y22]*[A3*(1-y31)] \quad (9)$$

$$\text{IF } (x1 \text{ is } M2) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K2) \text{ then} \\ F9=[A1*y12]*[A2*y22]*[A3*y32] \quad (10)$$

$$\text{IF } (x1 \text{ is } M2) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K3) \text{ then} \\ F10=[A1*y12]*[A2*y22]*[A3*(1-|1-y33|)/2] \quad (11)$$

$$\text{IF } (x1 \text{ is } M3) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K2) \text{ then} \\ F11=[A1*(1-|1-y13|)/2]*[A2*y22]*[A3*y32] \quad (12)$$

$$\text{IF } (x1 \text{ is } M3) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K3) \text{ then} \\ F12=[A1*(1-|1-y13|)/2]*[A2*y22]*[A3*(1-|1-y33|)/2] \quad (13)$$

$$\text{IF } (x1 \text{ is } M3) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K4) \text{ then} \\ F13=[A1*(1-|1-y13|)/2]*[A2*y22]*[A3*y34] \quad (14)$$

$$\text{IF } (x1 \text{ is } M3) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K2) \text{ then} \\ F14=[A1*(1-|1-y13|)/2]*[A2*(1-y23)]*[A3*y32] \quad (15)$$

$$\text{IF } (x1 \text{ is } M3) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K3) \text{ then} \\ F15=[A1*(1-|1-y13|)/2]*[A2*(1-y23)]*[A3*(1-|1-y33|)/2] \quad (16)$$

$$\text{IF } (x1 \text{ is } M3) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K4) \text{ then} \\ F16=[A1*(1-|1-y13|)/2]*[A2*(1-y23)]*[A3*y34] \quad (17)$$

$$\text{IF } (x1 \text{ is } M4) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K2) \text{ then} \\ F17=[A1*y14]*[A2*y22]*[A3*y32] \quad (18)$$

$$\text{IF } (x1 \text{ is } M4) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K3) \text{ then} \\ F18=[A1*y14]*[A2*y22]*[A3*(1-|1-y33|)/2] \quad (19)$$

$$\text{IF } (x1 \text{ is } M4) \text{ and } (x1 \text{ is } L2) \text{ and } (x1 \text{ is } K4) \text{ then} \\ F19=[A1*y14]*[A2*y22]*[A3*y34] \quad (20)$$

$$\text{IF } (x1 \text{ is } M4) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K2) \text{ then} \\ F20=[A1*y14]*[A2*(1-y23)]*[A3*y32] \quad (21)$$

$$\text{IF } (x1 \text{ is } M4) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K3) \text{ then} \\ F21=[A1*y14]*[A2*(1-y23)]*[A3*(1-|1-y33|)/2] \quad (22)$$

$$\text{IF } (x1 \text{ is } M4) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K4) \text{ then} \\ F22=[A1*y14]*[A2*(1-y23)]*[A3*y34] \quad (23)$$

$$\text{IF } (x1 \text{ is } M5) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K4) \text{ then} \\ F23=[A1*(1-y15)]*[A2*(1-y23)]*[A3*y34] \quad (24)$$

$$\text{IF } (x1 \text{ is } M5) \text{ and } (x1 \text{ is } L3) \text{ and } (x1 \text{ is } K5) \text{ then} \\ F24=[A1*(1-y15)]*[A2*(1-y23)]*[A3*(1-y35)] \quad (25)$$

In the above rules, x1 is input of the system and shows daytime, y1j,y2j,y3j demonstrate membership functions of Mj,Lj,Kj fuzzy sets respectively. Note that A1,A2,A3 are constant coefficients for Health/Commercial, Official and Shopping centers respectively.

## 2) Defuzzification

In this system defuzzification procedure is performed by using equation (26). Output of defuzzifier is considered as output of first level of system and shows priority of the street.

$$F = \frac{\sum_{i=1}^{24} Fi}{n} \quad (26)$$

In the above equation, Fi is output of ith rule (i=1...24), n is number of rules that have nonzero output and F is defuzzifier output (output of first level of system). Keep in mind that F is used as one of the inputs of second level of the system.

## B. Second Level of the Proposed System

In this step real velocity of vehicles of every street is calculated at specific moments and then type of transmission of density of vehicles is evaluated. Real velocity of vehicles is obtained by regarding three factors:

- Priority of Street: That's calculated dynamically based on fuzzy rules at the first level and on specific moments.
- Density of Vehicles: Since low-priority streets have low pausing coefficients, most of vehicles would be paused behind intersections. So if adjustment of the traffic light is done only based on priority of the street, vehicles may pause behind intersections for a long duration. To prevent this problem, number of existent vehicles in the street is also regarded as one of the inputs.
- Standard Velocity of Vehicles of Street: Since number of vehicles behind intersections, is a determinant factor for specifying distance between location of paused vehicle and traffic light, we used cellular automata for segmentation of street from intersection to end of street. So type of movement of vehicles from entrance time till pausing time behind intersection is achieved by using transition rules of cellular automata. One may say since standard velocity of vehicles (Standard velocity is constant for all vehicles of each street and computed based on width and length of street) must be specific while applying transitions rules so it's considered as third factor. So velocity of each vehicle is computed as follows:

$$V=(1-\gamma)(1-\alpha)*V_N+\epsilon \quad (27)$$

In the above equation,  $\gamma$  is priority of the street,  $V_N$  is standard velocity of the street,  $\alpha$  is a constant value (5km/h in this paper) and  $\epsilon$  is density of the street that's obtained from equation 28:

$$\alpha = \frac{N}{W * L * A} \quad (28)$$

Where W and L are width and length of the street respectively, A is the average occupied area by each vehicle, N is the number of available vehicles in the street (N is calculated as follows: W\*L is multiplied by summation of densities of cellular spaces). Equation 28 shows that more density or more priority of street, leads in

less velocity of vehicles. Based on this velocity, transition rules of cellular automata are determined.

#### 1) Cellular Automata Rules

Since velocity and displacement have a linear dependency ( $x=v*t$ ), we used a linear transition function. Because of extra computations, discrete time method is applied instead of continuous time approach and system is updated at specific seconds. In this paper, velocities of vehicles are divided in intervals of length 5 and length of each cell is considered as 10 meters. So at every 7.2" time interval, the vehicle moves 10 meters in a fixed 5km length(Updating progress is also done every 7.2 seconds). Based on these comments, we used specific transition function for obtaining cellular rules of density transmission in the street: In the proposed unidirectional selective cellular automata model, width of cell is regarded as width of the street and length of cell is set to a constant number (10). Note that transmission of density of vehicles means transmission of vehicles from one cell to another one, type of this transmission is specified as Fig. 6.

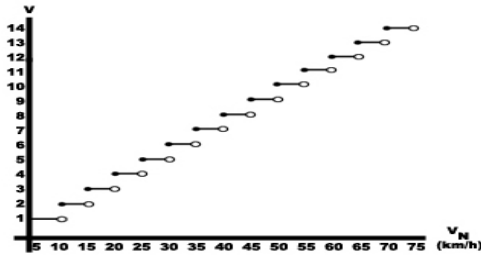


Fig. 6. Density Transmission Function Based on USCA Model

Maximum density of each cell is 1 and it is related to number of available vehicles at each cell. Maximum number of vehicles of each cell is computed as follows:

$$N = \frac{w * 10}{15} \quad (29)$$

In the above equation, W is width of the street, 10 is length of the street and 15 shows estimated area allocated for a vehicle. If number of available vehicles of each cell exceeds the maximum value (Obtained using equation 30), the surplus is transferred to previous cell. Note that in this paper all of the streets are bidirectional.

#### C. Third Level of the Proposed System

In this paper two factors are considered for adjusting traffic light. Priority of the street and number of complete traffic cells are inputs of third level of the proposed system. Since density and traffic in high-priority streets have more undesirable effects, priority of the street must be considered for adjusting traffic light. Furthermore,

number of paused vehicles behind intersections is the main factor for determining the time that status of traffic light is changed. This factor is illustrated by traffic cells and regarded as input of system.

In this proposed system two conversion functions are used for priority and number of traffic cells to determine time duration that green light is active.

Multiplication of outputs of two functions represents the time duration that traffic light is green for each street:

$$G = T_c * \gamma \quad (30)$$

Where G shows the time interval that green light is active,  $T_c$  is number of traffic cells and  $\gamma$  is priority of the street. G is computed for all of streets of intersection. Pay attention that these calculations are done concurrently and independently, eventually street with the highest G enables the green light and uses G seconds from it. To prevent consecutive variations of status of the traffic light equation 32 is employed:

$$\text{If } G_{\max} \geq 2 * G_{\text{green\_light}} \text{ THEN} \quad (31) \\ \text{Change Traffic light}$$

Where  $G_{\max}$  is the maximum value of G in every updating progress (Among streets of an intersection) and  $G_{\text{green\_light}}$  is the updated value of G for the street with green light. Indeed Change Traffic Light instruction activates green light for the street that has  $G_{\max}$ . In the proposed approach, at each moment the light green is active only for one of the streets.

## V. SIMULATION RESULTS

The proposed method was simulated with specific inputs and more desirable results were obtained in comparison with based adaptive method.

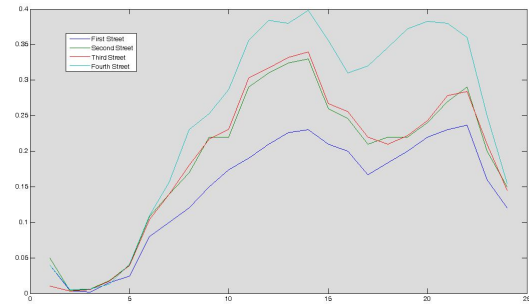


Fig. 7. Result of Fuzzy Proposed Approach for Representing Complete Traffic Cells Considering 1st Street as High-priority and 4th Street as Low- priority Street

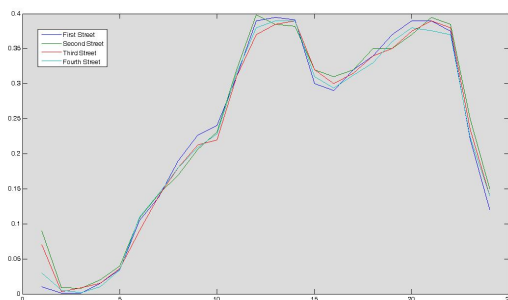


Fig. 8. Result of NN-based Approach for Representing Complete Traffic Cells Considering 1st Street as High-priority and 4th Street as Low- priority Street

when priorities of streets are different from each other and density of vehicles of streets of intersection streets is not the same, results of NN-based method don't change at all (Fig.8). Due to the fact that NN-based method doesn't consider priority of streets, so density of high-priority streets would increase and this matter leads to traffic increment throughout the street. Whereas in the proposed three leveled fuzzy approach, changing priority of streets and density of each street causes to significant changes in traffic light scheduling and also traffic cells. (Fig.7). This change causes to the street with more density and higher priority is selected to adjust the traffic light and eventually the traffic load and street density will be balanced.

Results of our proposed system and NN-based approach are shown in Table I:

Table I  
COMPARISON BETWEEN FUZZY PROPOSED SYSTEM  
AND NN-BASED METHOD

| Streets       | Fuzzy Proposed Method |                 |                 |                 |
|---------------|-----------------------|-----------------|-----------------|-----------------|
|               | 1 <sup>st</sup>       | 2 <sup>nd</sup> | 3 <sup>rd</sup> | 4 <sup>th</sup> |
| Traffic Cells | 2                     | 2               | 3               | 4               |
| Density       | 0.5834                | 0.5741          | 0.5208          | 0.4917          |
| Streets       | NN-based Method       |                 |                 |                 |
|               | 1 <sup>st</sup>       | 2 <sup>nd</sup> | 3 <sup>rd</sup> | 4 <sup>th</sup> |
| Traffic Cells | 2                     | 2               | 3               | 3               |
| Density       | 0.9571                | 0.7840          | 0.6425          | 0.5403          |

Although inconsiderable density has been added to 4th street, but density decrement is completely sensible in 1st, 2nd and 3rd streets. While existent systems only consider specific limited segment behind the intersection for adjusting the traffic light, our proposed system can control density of vehicles and traffic load throughout the street, also it prevent heavy traffics noticeably.

## VI. CONCLUSIONS

In this paper a three leveled fuzzy system was proposed for urban traffic control by adjusting the traffic light. Using this approach results in traffic decrement and also steady traffic dispreading all over the city. In the proposed system priority and length of streets are considered variable. Also real velocity of vehicles is computed based on fuzzy rules in different hours during the day. Considering mentioned factors leads to traffic load balancing throughout the city. Another dominant note in this paper is that priority of streets and also the maximum allowed velocity of the street are obtained by fuzzy sets with regard to urbanism principals. This matter also reduces traffic significantly.

## REFERENCES

- [1] Stephen Wolfram, "A New Kind of Science", 2002
- [2] Li-X Wang, "A Course In Fuzzy Systems and Control", 1996
- [3] M. Schreckenberg, R. Barlovic, W. Knospe, H. Klupfel, "Statistical Physics of Cellular Automata Models for Traffic Flow", Phd thesis, Germany, 2004.
- [4] S. Wolfram (ed.), "Theory and applications of CA", World Scientific, Singapore, 1986
- [5] Ron Breukelaar and Thomas Back, "Evolving Transition Rules for Multi Dimensional Cellular Automata", Springer-Verlag Berlin Heidelberg 2004
- [6] M. Wiering, J. van Veenen, J. Vreeken, A. Koopman, "Intelligent Traffic Light Control", Institute of Information and Computing Sciences, Utrecht University, 2004
- [7] R. Barlovic, T. Huisinga, A. Schadschneider, M. Schreckenberg, "Adaptive Traffic Light Control in the ChSch Model for City Traffic", Duisburg University, 2005
- [8] E. Brockfeld, R. Barlovic, A. Schadschneider, M. Schreckenberg, "Optimizing Traffic Light in a cellular automaton model for City Traffic", Duisburg University, PPHYSICAL RREVIEW E, VOLUME 64, 056132, 2001.
- [9] J. Neal Richter, David Peak, "Fuzzy Evolutionary Cellular Automata", Proceedings of ANNIE, 2002.
- [10] B.S. Kerner, "Empirical macroscopic features of spatial – temporal traffic patterns at highway bottlenecks", Physical Review, Volume 65, 046138, Stuttgart, Germany, 2002.
- [11] Traffic and Granular Flow, '97, '99, '01, Springer, (1996, 1998, 2000, 2003).
- [12] B.S. Kerner, "Traffic behavior near an off ramp in the cellular automaton traffic model", Phys. Rev. E 65, 046138 (2002).
- [13] R. Esperioueta, "Cellular Automata Dynamics", Math Department, Bakersfield College, 1997.
- [14] K. Nagel and M. Schreckenberg, "A Cellular Automaton Model for Freeway Traffic", J. Physique I 2, 2221 (1992).
- [15] N.H. Gartner and N.H.M. Wilson (eds), Transportation and Traffic Theory, (New York: Elsevier, (1987).
- [16] M. Fukui and Y. Ishibashi, J. Phys. Soc., "Evolution of Traffic Jam in Traffic Flow Model" Jpn. 62 (1993).
- [17] P. Wagner, K. Nagel, and D.E. Wolf, Physica A 234, 687 (1997).

# Bangla Vowel Sign Recognition by Extracting the Fuzzy Features

Mohammad Sarwar Kamal, Mohammed Moshiul Hoque, Md. Monjur Ul Hasan and Mohammad Shamsul Arefin

Department of Computer Science & Engineering, Chittagong University of Engineering & Technology, Chittagong-4349, Bangladesh

**Abstract** — This paper presents a fuzzy based Bangla vowel sign recognition system. Vowel and consonant are the main criteria to express any language. Exceptionally, Bangla has some vowel sign which are used to express Bangla word. Therefore, to recognize any Bangla word, at first it is necessary to recognize which vowel sign is used to express that word. Due to its low computational requirement and very high accuracy, fuzzy logic system is probably the most efficient method available for on-line vowel sign recognition. The most important task for this implementation is the building of the rule base using fuzzy logic that would describe the recognized vowel sign, which is more complicated as different people write the same vowel sign in completely different ways. In this paper, we propose a system that uses automatically generated fuzzy membership value that describes the handwritten Bangla vowel sign written by different individuals.

**Index Terms** — Vowel Sign, Geometric Features, Positional Features, Segmentation.

## I. INTRODUCTION

With the increasing the interest of computer applications, modern society needs the handwritten text into computable readable form. Since, exceptionally Bangla has some vowel sign, which are used to express Bangla word, so, to recognize any Bangla word; at first it is necessary to recognize which vowel sign is used to express that word. Generally, the objective of handwritten character recognition is the recognition of data that describe handwritten object. On line handwritten recognition deals with a time ordered sequence of data. It is relatively simple to write a computer program within a single document. The problem is complicated as different people write the same character in completely different ways. So, any system designed for the recognition of handwritten character must be extremely robust in term of being able to ignore differences in writing style. Although many works on character recognition have been done, works on Bangla character recognition are few. But, there is no task on recognition of Bangla vowel sign. Many methodologies have been developed for the recognition of handwritten characters. Some of the most commonly used techniques are Neural network [1], genetic algorithm [2], Bayesian inference [3], stochastic HMM [4], Fuzzy logic [5], Bangla handwritten character recognition [6], Handwritten alphanumeric character recognition [7]. Due

to the wide variety of different writing styles, all the proposed techniques have fall in some problem. The objectives of this paper are to segment the Bangla vowel sign by calculating angle difference and calculate the geometric features of each segment for recognizing the vowel sign.

The rest of the paper is organized as follows. Section II illustrates about fuzzy set and membership function. Section III describes proposed model for Bangla vowel sign recognition. Section IV depicts the experimental result and performance of proposed system. Finally section V concludes the paper.

## II. FUZZY SET AND MEMBERSHIP FUNCTION

The object and its membership value express fuzzy set. If a number between 0 and 1 indicates the degree of membership to a set then it is called fuzzy set [8]. A membership function maps every element of the universe of discourse  $X$  to the interval  $[0, 1]$  and is denoted by  $\mu_A(x)$ . We formally write this mapping as

$$\mu_A(x) : X \rightarrow [0, 1]$$

The notation of above equation is read as follows: there exist a function  $\mu_A(x)$  mapping every element of the set  $X$  to the set of interval  $[0, 1]$ .

### A. Different Types of Membership Function

There are different types of membership functions are in used. These are  $\Pi$ - shaped function, bell shaped function, triangular shaped function, trapezoidal shaped function etc. In this paper, we use the most popular membership function is triangular membership function which is represented by  $\Delta$  which have three parameters. This function defined as:

$$\Delta(x; b, c) = \begin{cases} 1 - 2 * |(x - c) \div c| & \text{if } (c - b \div 2) \leq x \leq (c + b \div 2) \\ 0 & \text{otherwise} \end{cases}$$

The triangular membership function is depicted in Fig.1.



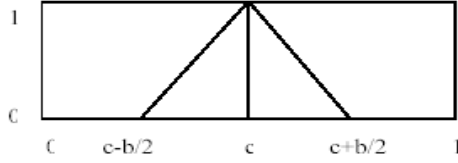


Fig. 1. Triangular Membership Function

Here  $c$  is the central point and  $b$  is the bandwidth of the triangular membership function and  $x$  is the determining variable.

### III. PROPOSED MODEL FOR BANGLA VOWEL SIGN RECOGNITION

The basic methodology of the online Bangla vowel sign recognition system is depicted in Fig. 2.

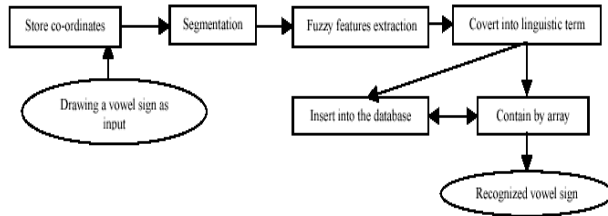


Fig. 2. Proposed architecture for Vowel Sign recognition

Here, the input of the model is drawn vowel sign that is passes through the different phases of the proposed model and generate output of the model, which is recognized standard format of vowel sign.

#### A. Segmentation Process

In segmentation phase, every vowel sign is divided into a number of segments. The pseudo code for the segmentation of vowel sign is depicted below:

```

Begin segmentation:
Step-1: Suppose a variable no_points. Set no_points = 0.
Step-2: If Pen down, initialize new segment
Step-3: Store co-ordinates of initialized segment, Set
 no_points = no_points + 1
Step-4: Get next co-ordinates
Step-5: If Pen up, store co-ordinates and Go to Step-1
ELSE
If abrupt change in direction and no_points > 4
Set no_points = 0 and Go to Step-2
ELSE
Go to Step-3
End segmentation

```

According to above algorithm we can divide the bangla

vowel sign 'ঠ', into two segments for the sample depicted in Fig. 3.

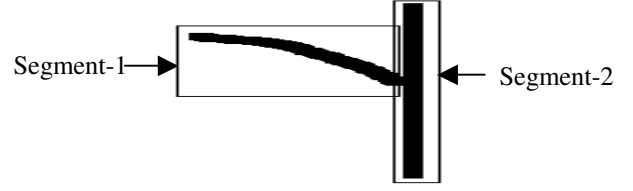


Fig. 3. Segmentation of Bangla vowel sign 'ঠ',

The segmentation method is based on the movement and angle differences between first four point connected line and consecutive second four point connected line. If the angle difference between two lines is more than 90 degrees, then first line is recognized a new segment. When electronic pen or mouse is up, we recognized also a new segment.

#### B. Fuzzy Features

Fuzzy feature are divided into three categories: global features, positional features and geometric features. Global features are the common for whole vowel sign. These global features are necessary for calculating the both positional features and geometric features.

##### B1. Global Features

Global features belong to whole vowel sign. Some global features are given below:

$$\min X^{seg(n)} = \min(x_i) \quad \max X^{seg(n)} = \max(x_i)$$

$$\min Y^{seg(n)} = \min(y_i) \quad \max Y^{seg(n)} = \max(y_i)$$

The global features of segment 1 of vowel sign 'ঠ', which is depicted in Fig. 3 can be describe by the Fig. 4.

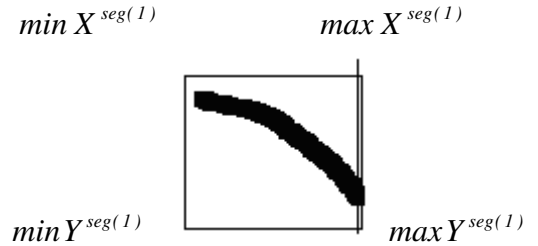


Fig. 4. Global feature of segment 1 of vowel sign 'ঠ'

##### B2. Positional Features

The universe of discourse is divided into two linguistic variables which are Vertical position (VP) and Horizontal Positional (HP). The Vertical Position is divided into three linguistic terms: {Top, Middle, Bottom} and the

Horizontal position is divided into three linguistic terms: {Left, Center, Right). Then the relative position of the identified segment is expressed by the following equations:

$$\mu_{HP} = \frac{(centreX^{seg(n)} - \min X^{seg(n)})}{(\max X^{seg(n)} - \min X^{seg(n)})} \quad (3)$$

$$\mu_{VP} = \frac{(centreY^{seg(n)} - \min Y^{seg(n)})}{(\max Y^{seg(n)} - \min Y^{seg(n)})} \quad (4)$$

### B3. Geometric Features

The geometric features are divided into two main categories, which are straight line and arc. The membership function for straightness and arc ness of the every segment are calculated by the following equation.

$$\mu_{straightness} = [d_{P(0)P(N)} / \sum_{K=1}^N d_{P(K)P(K+1)}] \quad (5)$$

Here,  $d_{P(k)P(K+1)}$  stands for the straight-line distance between point  $k$  and point  $(k+1)$  on the  $n^{th}$  segment.  $N$  depicts the number of element in the segment.

$$\mu_{Arcness} = 1 - \mu_{straightness} \quad (6)$$

If  $\mu_{straightness}$  is greater than 0.6 then the segment is straight line and otherwise it is an arc. Some geometric features are given below:

For vertical line:

$$\mu_{VL} = \max(\Delta(x, 90, 45), \Delta(x, 90, 270)) \quad (7)$$

For horizontal line:

$$\mu_{HL} = \max(\Delta(x, 90, 0), \Delta(x, 90, 180), \Delta(x, 90, 360)) \quad (8)$$

For positive slant

$$\mu_{PS} = \max(\Delta(x, 90, 45), \Delta(x, 90, 225)) \quad (9)$$

For negative slant

$$\mu_{NS} = \max(\Delta(x, 90, 135), \Delta(x, 90, 315))$$

The equation of the A-like curve is

$$\mu_{AL} = \min(1, \sum I_y / n) \quad (10)$$

Where  $I_y = 1$  if  $y > (y_s + y_e) / 2$  ; Otherwise 0

The equation of the U-like curve is

$$\mu_{UL} = \min(1, \sum I_y / n) \quad (11)$$

Where  $I_y = 1$  if  $y < (y_s + y_e) / 2$  ; Otherwise 0

The equation of the C-like curve is

$$\mu_{CL} = \min(1, \sum I_x / n) \quad (12)$$

Where  $I_x = 1$  if  $x < (x_s + x_e) / 2$  ; Otherwise 0

The equation of the D-like curve is

$$\mu_{DL} = \min(1, \sum I_x / n) \quad (13)$$

Where  $I_x = 1$  if  $x > (x_s + x_e) / 2$  ; Otherwise 0

The equation of the O-like curve is

$$\mu_{OL} = (\sum_{K=1}^N d_{P(K)P(K+1)}) / (2 * 3.1416 * r) \quad (14)$$

Where  $r$  is the radius of the curve.  $x_s, x_e$  is the start point and end point of segment respectively in the X axis.  $y_s, y_e$  is the start point and end point of segment respectively in the Y axis. Different types of curve and corresponding equation are given in the Table I.

TABLE I  
DIFFERENT TYPES OF GEOMETRIC FUZZY  
FEATURES

| Shape | Name            | Function                                                                                     |
|-------|-----------------|----------------------------------------------------------------------------------------------|
|       | Vertical line   | $\mu_{VL} = \max(\Delta(x, 90, 45), \Delta(x, 90, 270))$                                     |
| —     | Horizontal line | $\mu_{HL} = \max(\Delta(x, 90, 0), \Delta(x, 90, 180), \Delta(x, 90, 360))$                  |
| /     | Positive slant  | $\mu_{PS} = \max(\Delta(x, 90, 45), \Delta(x, 90, 225))$                                     |
| \     | Negative slant  | $\mu_{NS} = \max(\Delta(x, 90, 135), \Delta(x, 90, 315))$                                    |
| ∩     | A-Like Curve    | $\mu_{AL} = \min(1, \sum I_y / n)$ Where $I_y = 1$ if $y > (y_s + y_e) / 2$ ;<br>Otherwise 0 |
| ⊂     | C-Like Curve    | $\mu_{CL} = \min(1, \sum I_x / n)$ Where $I_x = 1$ if $x < (x_s + x_e) / 2$ ;<br>Otherwise 0 |
| ⊃     | D-Like Curve    | $\mu_{DL} = \min(1, \sum I_x / n)$ Where $I_x = 1$ if $x > (x_s + x_e) / 2$ ;<br>Otherwise 0 |

After calculating all fuzzy features the next step is to map these values into Linguistic terms according to predefined range. The range of value of membership function for evaluating angle difference and different linguistic term are shown in Fig. 5.

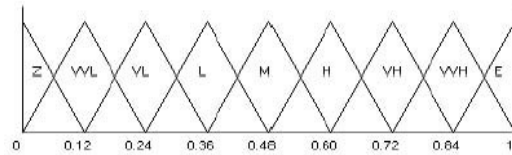


Fig. 5. Membership function for evaluating angle difference

### C. Creation of Fuzzy Rule Base

The rule base is depends on the number of segment contained in the given vowel sign.



There are a lot of rule in this works for vowel sign recognition. Each rule is unique for each vowel sign. The general format of fuzzy rule base for each vowel sign which are stored in the database is the features of a segment is combined with logical AND operator and the features of different segments is combined with logical OR operator.

#### D. Modes of Vowel Sign Recognition

There are two modes of operation for online handwritten Bangla vowel sign recognition technique. These are the Learning mode and the Recognition mode that are described in subsequent section.

##### D1. Learning Mode

In learning mode, user draws the image / vowel signs in the drawing pad. Now, we describe the learning mode by step by step.

**Step 1:** The segmentation process is explained by the algorithm described in section A

**Step 2:** Then fuzzy features extractions are occurred according to the equations described in section B.

**Step 3:** The important linguistic term is is described in TABLE II. In this table  $\mu_{ARC}$ ,  $\mu_{VF}$ ,  $\mu_{DL}$  and  $\mu_{CL}$  are stand for arcness, vertical position, D-like and C-like respectively. For sample 1 the value of  $\mu_{ARC}$ ,  $\mu_{VF}$ ,  $\mu_{DL}$  and  $\mu_{CL}$  are M, NB, VH and VL which are standing for 'medium', 'nearly bottom', 'very high' and 'very low' respectively.

TABLE II  
IMPORTANT LINGUISTICS TERM FOR ANY FEATURES

| SAMPLE  | $\mu_{ARC}$   | $\mu_{VF}$     | $\mu_{DL}$     | $\mu_{CL}$     |
|---------|---------------|----------------|----------------|----------------|
| 1       | M             | NB             | VH             | VL             |
| 2       | H             | BC             | VH             | VL             |
| 3       | M             | BC             | VH             | VL             |
| 4       | M             | B              | VH             | VL             |
| 5       | M             | BC             | VH             | VL             |
| 6       | VH            | BC             | VH             | VVL            |
| 7       | L             | C              | VH             | L              |
| 8       | M             | BC             | VVH            | VL             |
| Total=8 | Important = M | Important = BC | Important = VH | Important = VL |

**Step 4:** Then the store mechanism for fuzzy features into databases is happened according to created fuzzy rule base that are described in section C.

##### D2. Recognition Modes

The system architecture of recognition or output mode is described in below step by step.

**Step 1:** Recognition mode segmentation is same as the learning mode segmentation.

**Step 2:** Recognition mode feature extraction is same as the learning mode extraction.

**Step 3:** Store mechanism of data into array are described by the following source code:

```

If uol(i) <= 0 Then
 uolc(i) = "Z"
ElseIf uol(i) > 0 And uol(i) <= 0.125 Then
 uolc(i) = "VVL"
ElseIf uol(i) > 0.125 And uol(i) <= 0.25 Then
 uolc(i) = "VL"
End If
End If

```

**Step 4:** In comparison stage, array data are compared with the rule base data.

**Step 5:** Highest value is calculated from step 4 as percentage.

The vowel sign, which contains the highest recognition rate, is the recognized vowel sign.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

In the experimentation, we have collected twenty samples for each vowel sign from ten individuals. In Bangla, there are ten vowel signs. Therefore, the total numbers of samples are 2000. We got the very high recognition accuracy for the proposed system, which is more than 90%.

#### A. Segmentation Results

To find out attractive fuzzy features it is necessary to divide each vowel sign into a number of meaningful segments. In this paper, each vowel sign can be divided into maximum fifteen segments. But maximum accuracy is found in different specific number of segments for each different vowel sign. The segmentation result of Bangla vowel sign 'ঢ' is depicted in Fig. 6.

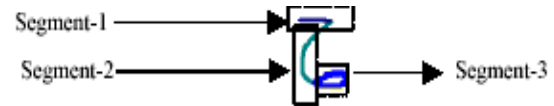


Fig. 6. Segmentation result of Bangla vowel sign 'ঢ'

#### B. Positional Features Extraction

The positional features of Bangla vowel sign 'ঢ', 'ঢ়' and 'ঢ়্' are shown in the TABLE III. This table describes the  $\mu_{HP}$  and  $\mu_{VP}$ , which are standing for horizontal and vertical position respectively. For vowel sign 'ঢ', the horizontal position of segment - 1 and segment - 2 is LC stand for

‘left corner’ and the vertical position of segment – 1 and segment – 2 are B stand for ‘Bottom ’ and C stand for ‘corner’ respectively.

TABLE III  
A SMALL FRAGMENT OF POSITIONAL  
FEATURES OF BANGLA VOWEL SIGN

| Sample of vowel sign | No. of Segment | Segment serial | μ <sub>up</sub> | μ <sub>vr</sub> |
|----------------------|----------------|----------------|-----------------|-----------------|
| ৳                    | 2              | 1              | LC              | B               |
|                      |                | 2              | LC              | C               |
| ৳                    | 4              | 1              | L               | BC              |
|                      |                | 2              | LC              | C               |
|                      |                | 3              | RC              | BC              |
|                      |                | 4              | R               | C               |
| ৳                    | 4              | 1              | L               | BC              |
|                      |                | 2              | LC              | C               |
|                      |                | 3              | RC              | BC              |
|                      |                | 4              | R               | C               |

### C. Geometric Features

The straightness features of Bangla vowel sign ৳, ৳ and ৳ are shown in TABLE IV. This table describes the  $\mu_{ST}$ ,  $\mu_{VL}$ ,  $\mu_{HL}$ ,  $\mu_{PS}$ ,  $\mu_{NS}$ ,  $\mu_{HLEN}$ ,  $\mu_{VLEN}$  and  $\mu_{SLEN}$  which are standing for straightness, vertical line, horizontal line, positive slanted, negative slanted, horizontal length, vertical length and slanted length respectively. For vowel sign ‘৳’,  $\mu_{ST}$ ,  $\mu_{VL}$ ,  $\mu_{HL}$ ,  $\mu_{PS}$ ,  $\mu_{NS}$ ,  $\mu_{HLEN}$ ,  $\mu_{VLEN}$  and  $\mu_{SLEN}$  of segment – 1 are E, M, E, H, VL, VL, L, L which are standing for ‘Excellent’, ‘Medium’, ‘High’, ‘Very low’ and Low’ respectively. The ‘-’ means that there is no straightness features because the selected segment contains the arcness fuzzy features.

TABLE IV  
STRAIGHTNESS FEATURES OF BANGLA VOWEL  
SIGN

| Sample of vowel sign | No. of Segment | Segment serial | μ <sub>ST</sub> | μ <sub>VL</sub> | μ <sub>HL</sub> | μ <sub>PS</sub> | μ <sub>NS</sub> | μ <sub>HLEN</sub> | μ <sub>VLEN</sub> | μ <sub>SLEN</sub> |
|----------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------------|-------------------|-------------------|
| ৳                    | 2              | 1              | E               | M               | E               | H               | VL              | VL                | L                 | L                 |
|                      |                | 2              | VVH             | H               | H               | E               | M               | VVL               | VL                | VL                |
| ৳                    | 4              | 1              | H               | VH              | VVH             | L               | M               | VL                | VVL               | E                 |
|                      |                | 2              | E               | VL              | Z               | VL              | VVL             | L                 | M                 | VH                |
|                      |                | 3              | H               | VH              | VVH             | L               | M               | VL                | VVL               | E                 |
|                      |                | 4              | Z               | -               | -               | -               | -               | -                 | -                 | -                 |
| ৳                    | 5              | 1              | H               | VH              | VVH             | L               | M               | VL                | VVL               | E                 |
|                      |                | 2              | H               | VH              | VVH             | L               | M               | VL                | VVL               | E                 |
|                      |                | 3              | Z               | -               | -               | -               | -               | -                 | -                 | -                 |
|                      |                | 4              | E               | VL              | Z               | VL              | VVL             | L                 | M                 | VH                |
|                      |                | 5              | H               | VH              | VVH             | L               | M               | VL                | VVL               | E                 |

Similarly, we can describe the arcness fuzzy features by tabular form.

### C. Recognition Accuracy of Bangla Vowel Sign

The recognition accuracy Bangla vowel sign ৳, ৳

and ৳ are shown in TABLE V. This table contains the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> highest percentage value for each vowel sign that are collect from different individuals. The original sign is selected as highest percentage of accuracy. Such as for ‘৳’, which consists three segments, 95% adjust with ‘৳’, 90% adjust with ‘৳’ and 80% adjust with ‘৳’. So, highest value of percentage of accuracy is 95%, which is adjust with ‘৳’. So, the recognized vowel sign is ‘৳’.

TABLE V  
RECOGNITION ACCURACY OF DIFFERENT  
BANGLA VOWEL SIGN

| Vowel Sign | No. of Segment | 1 <sup>st</sup> Recognized | 2 <sup>nd</sup> Recognized | 3 <sup>rd</sup> Recognized | Original sign |
|------------|----------------|----------------------------|----------------------------|----------------------------|---------------|
| ৳          | 3              | ৳ 95 %                     | ৳ 90 %                     | ৳ 80%                      | ৳             |
| ৳          | 4              | ৳ 80 %                     | ৳ 75 %                     | ৳ 67.5%                    | ৳             |
| ৳          | 9              | ৳ 71.6 %                   | ৳ 68.8 %                   | ৳ 67.2%                    | ৳             |
| ৳          | 7              | ৳ 53.5 %                   | ৳ 52.8 %                   | ৳ 52.1%                    | ৳             |
| ৳          | 6              | ৳ 85 %                     | ৳ 77 %                     | ৳ 67%                      | ৳             |
| ৳          | 7              | ৳ 70 %                     | ৳ 64 %                     | ৳ 57%                      | ৳             |

### E. System Performance

The performance of the system is described by the curve of number of segments versus percentage of accuracy for each individual sign, which are shown in Fig. 7. The highest, lowest and average percentage value of accuracy for each vowel sign with respect to number of segments has been mentioned in this figure. Such as for ‘৳’, for number of segment 1, the percentage of accuracy is 85%, for segment 2, the percentage of accuracy is 95%, for segment 3, the percentage of accuracy is 78%. Here, for ‘৳’, it is shows that the highest value of percentage of accuracy is found for segment 2 which is 95%, it is called potential point.

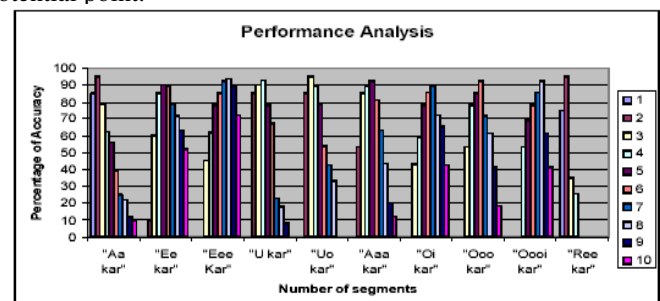


Fig. 7. Curve of number of segments versus percentage of accuracy

Now, we draw the curve for number of segments versus highest percentage of accuracy to find out the average system performance. The curve is depicted in Fig. 8. This figure describes the potential point at which segment, the percentage of accuracy is high. The numbers of segments along X axis are 2, 3, 8, 4, 3, 5, 7, 6, 8, and 2 which are standing for  $\text{ত}$ ,  $\text{ফ}$ ,  $\text{শ}$ ,  $\text{অ}$ ,  $\text{এ}$ ,  $\text{ক}$ ,  $\text{খ}$ ,  $\text{গ}$ ,  $\text{ট}$ , and  $\text{ঙ}$  respectively. The values along Y-axis are the corresponding highest percentage of accuracy.

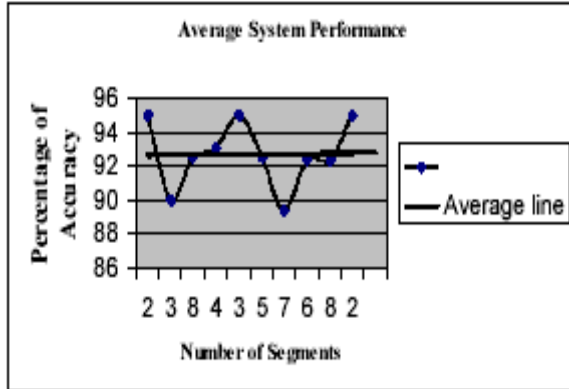


Fig. 8. Average system performance

#### V. CONCLUSION

With the increasing the interest of computer applications, modern society needs the handwritten text into computable readable form. So, handwritten recognition is a very interesting input method. The goal of this system is to develop an idea and to make software in order to recognize Bangla vowel sign regardless of various writing style. The recognition rate and speed are very good than others. This proposed system could be used to recognize the vehicle nameplate in Bangladesh. A cache system may introduce in further work. If Unicode can

improve it without Bangla software, then it will do with very easily. If the learning mode and recognition mode can be combined into a single mode, then this work performs better. That is the performance-improved dramatically.

#### REFERENCES

- [1] I. Guyon, "Application of Neural Networks to Character Recognition", in character and handwriting Recognition: Expanding frontiers, Ed: P.S.P Wang, World Scientific, pp.353-382, 1991.
- [2] G. Kim and S. Kim, "Feature Selection using Genetic Algorithm for Handwritten Character Recognition", *Proceeding of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002.
- [3] K. W. Cheung and D. Y. Yeung, "Bidirectional Deformable Matching with Application to Handwritten Character Extraction", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, No. 8, 2002
- [4] C. Bahlmann and H. Burkhardt, "Measuring HMM Similarity with the Bayes Probability of error and its Application to Online Handwriting Character Recognition", *6th International Conference on Document Analysis Recognition (ICDAR)*, pp. 406-411, 2001
- [5] A. Malaviya and L. Peters, "Extracting Meaningful Handwriting Features with Fuzzy Aggregation Method", *3rd International Conference on Document Analysis and Recognition (ICDAR)*, Montreal Canada, 1995.
- [6] M. M. Islam, S. M. F. Rahman, and M. M. Hoque, "Online Bengali Handwritten Recognition with Automatically Generated Fuzzy Linguistics Rules", in *proceeding of 8th International Conference on Computer and Information Technology (ICCIT)*, Dhaka, Bangladesh, pp.673-678, 2005.
- [7] R. Ranawana, V. Palade and G. E. M. D. C. Bandara, "Automatic Fuzzy Rule Base Generation for On-line Handwritten Alphanumeric Character Recognition", Oxford University Computing Laboratory, Oxford, OXI 3QD, United Kingdom
- [8] L. H. Tsoukalas and R. E. Uhrig, "Fuzzy and Neural Approaches in Engineering", pp-13-15.

# Active Noise Control for Industrial Applications

Syed Foysol Islam<sup>1</sup>, A.H.M. Saiful Islam<sup>2</sup>, M. M. Rajib Billah<sup>3</sup>

Electronics and Telecommunication Engineering Department<sup>1</sup>, Computer Science and Engineering Department<sup>2</sup>, Electronics and Telecommunication Engineering Department<sup>3</sup>  
University of Development Alternative, Dhanmondi, Dhaka-1205, Bangladesh  
Email: rumi\_cst@yahoo.com<sup>1</sup>, tamim\_saif@yahoo.com<sup>2</sup>, rajib\_ece2003@yahoo.com<sup>3</sup>

**Abstract** — In this paper we will show the construction of acoustic active noise control in ear defenders by the help of Filtered-x Least Mean Squares structure and using a low power fixed point digital signal processor. This system is implemented for 20dB to 30dB attenuation of broad band noise and ca 60dB for sinusoidal interference. This is well suited for industrial applications.

**Index Terms** — Attenuation, Adaptive filters, Digital Signal Processors, Digital filters, Microphones

## I. INTRODUCTION

In any workplace i.e. construction sites or factories sound and noise are emitted and that noise is so much unhealthy for human. This noise might cause of giant damage to the hearing and impair the overall health of personnel. So they need protection. Traditional ear defenders can protect from upper range of frequencies but for lower range (below 500Hz) it is inadequate. In that kind of ear defenders the passive protection is used [1]. But for higher protection we need active methods. So, to make it better we are using Active Noise Control (ANC) method within the ear cup so that the sound pressure level is minimized.

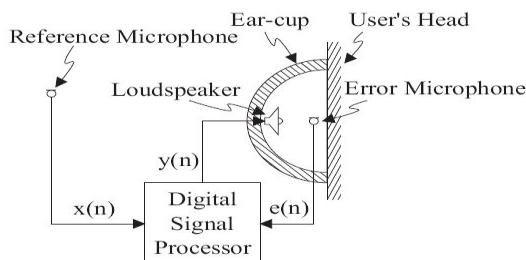


Fig.1. Basic structure of ANC in Ear defender.

Analog techniques are highly cost effective and have less efficiency [4]. Now this days reconfigurable digital solution is used which is based on Digital Signal Processor. Now we are using FXLMS algorithm to make the system more effective.

This contribution presents and evaluates a fixed point ANC implementation in ear defenders using low power Digital Signal Processor.

## II. ACTIVE CONTROL OF NOISE USING FXLMS

The basic setup of ear defender given in Fig. 1. shows the primary design of ANC system. In this section we introduce some necessary terms and the details of the setup.

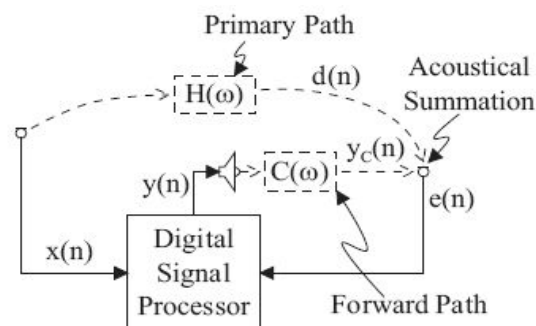


Fig. 2. ANC in ear-muffs.

We see a transfer function transferred from reference microphone to error microphone through the primary path  $H(w)$  in Fig. 2. There is another path denoted as forward path,  $C(w)$ , through which another transfer function travel from the input of the loud speaker to the error microphone [6, 7].

So, there are two acoustic signals in the microphone, noise signal and anti-noise signal and the model is a linear summation of these two signals. But path from the loud speaker to the external reference microphone can be neglected due to the characteristic of the ear-muffs [5, 8].

The implementation of FXLMS is shown in Fig. 3.

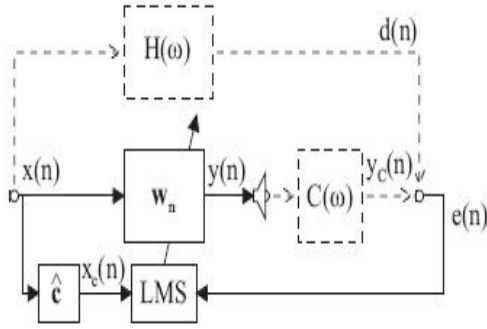


Fig. 3. Model of ANC using FXLMS.

Here a forward path estimate is,  $\hat{C}(w)$ , with an impulse response,  $\hat{c}(n)$ , which recompense for the delays creates by the forward path [6,8]. We can have the following expressions by the derivation of the FXLMS structure.

$$y(n) = \sum_{k=0}^{K-1} w_n(k)x(n-k) = \mathbf{w}_n^T \mathbf{X}_n \quad (1)$$

$$x_c(n) = \sum_{m=0}^{M-1} \hat{c}(m)x(n-m) = \hat{\mathbf{C}}^T \mathbf{X}'_n \quad (2)$$

Here,

Adaptive filter coefficient,

$$\mathbf{w}_n = (w_n(0), \dots, w_n(K-1))^T \quad (3)$$

Forward path estimate,

$$\hat{\mathbf{C}} = (\hat{c}(0), \dots, \hat{c}(M-1))^T \quad (4)$$

The size of  $\mathbf{x}_n$  and  $\mathbf{x}'_n$  is K and M respectively and they contain present and former reference signal values.

The adaptive filter coefficient update is,

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n) \mathbf{X}_{c,n} \quad (5)$$

Here,  $\mu$  = Algorithm step size.

$$\mathbf{X}_{c,n} = (x_c(n), \dots, x_c(n-(K-1)))^T \quad (6)$$

Because of the interaction of the adaptive filter,  $\mathbf{w}_n$  and the forward path estimate,  $\hat{\mathbf{C}}$  the robustness to the forward path estimate error is high [8].

### III. FIXED POINT FXLMS IMPLEMENTATION

To implement the FXLMS structure we will use a fixed point processor [3, 8]. We are using the processor which uses 16 bit for arithmetic calculation and can extend capability up to 48 bit.

Digital signal processing unit consist of two vital parts [3]:

- i. Digital Filter.
- ii. Adaptive Algorithm.

During the architecture we may face some problems. The secondary path introduces delay that causes instability to the standard LMS algorithm since the error signal is not correctly “aligned” in time with the reference signal [4, 5]. In Fig. 4. the utilization of LMS update algorithm has been shown with a 2-ADC/1-DAC configuration.

Filtering the reference signal with an estimate of the secondary path resolves the instability and results in the FXLMS algorithm. The estimate of the secondary path is performed off-line prior to the adaptation.

Feedback can be resolved by introducing an extra canceling filter in parallel to the feedback path and employing adaptive Infinite Impulse Response (IIR) filter [2, 7].

Now, adaptation is a sensitive issue. If the filter co-efficient update is stopped then an undesirable result can occur. To minimize this problem adaptive filter co-efficient are stored in an 32 bit format [8],

$$\mathbf{W}_n = 2^{16} \mathbf{W}_n^{hi} + \mathbf{W}_n^{lo} \quad (7)$$

Where,  $W_n^{hi} = 16$  bit and  $W_n^{lo} = 16$  bit.

Though the whole 32 bits are used in adaptation phase in the filtering phase only 16 bits are used. It can be expressed [8] as,

$$y(n) = \left\lfloor 2^{-8} \sum_{k=0}^{K-1} W_n^{hi} x(n-k) \right\rfloor = \left\lfloor 2^{-8} W_n^{hi^T} X_n \right\rfloor \quad (8)$$

Floor operator " $\lfloor \rfloor$ " only allow the floor values of a function. By using floor operator the filter co-efficient can be expressed as  $2^{-8} W_n^{hi}$ .

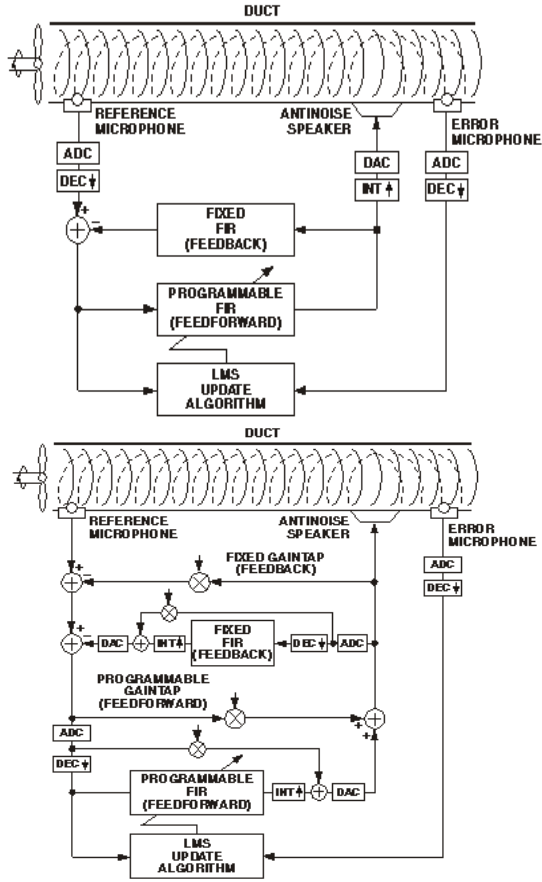


Fig. 4. A 2-ADC/1-DAC configuration with no high-speed gain taps.

The adaptive filter co-efficient update is

$$W_{n+1} = W_n + \mu_q e(n) x_{c,n} \quad (9)$$

Here  $\mu_q = \left\lfloor 2^{-8} \mu \right\rfloor$  is the adaptive filter step size but the actual size is

$$\mu = 2^{-8} \mu_q \quad (10)$$

#### IV. EXPERIMENTAL SETUP

To design experimental circuit we need:

- i. Printed Circuit Board (PCB).
- ii. Micro Controller.
- iii. Low pass Filter.

Printed Circuit Board (PCB):

For designing the experimental circuit of the fixed point FXLMS a PCB is manufactured. Block diagram of the PCB is shown in Fig. 5. The PCB can support two input and output channels.

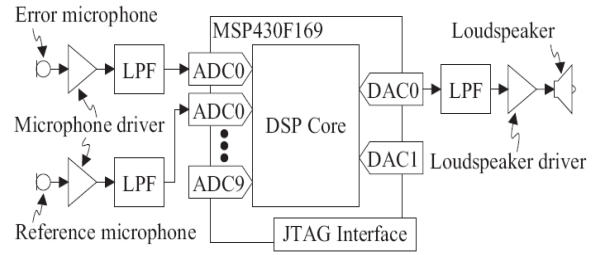


Fig. 5. Block Diagram of the PDB.

Micro Controller:

We are using MSP430F169 micro processor [3] which is situated on the PCB (Fig. 6.).

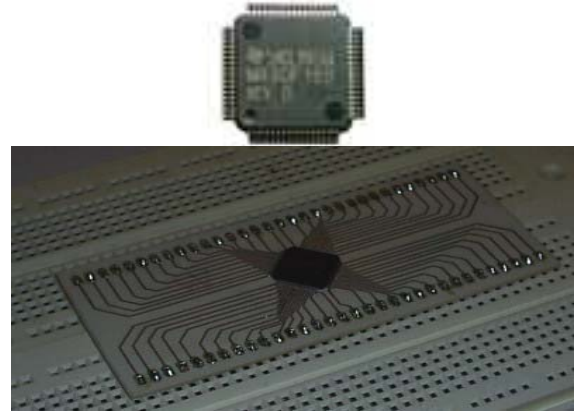


Fig. 6. MSP430F169 (1<sup>st</sup> image), Processor on board (2<sup>nd</sup> image).

The maximum clock speed of the processor is 8Mz [3, 7]. Average execution rate is 2 MIPS. The micro controller has eight Analog-to-Digital Converter (ADC) and dual Digital-to-Analog



Converter (DAC) on board and both are of 12 bits.

Low pass Filter:

Properly configured 10<sup>th</sup> order low pass filter LTC1569-6 is situated on board prior to ADC and after DAC [8]. It is for aliasing reduction and signal re-construction.

The range of input output signal is  $\pm 1.25V$ . Externally there are micro amplifiers and anti noise loud speakers [6]. The positions in the ear-cup are shown in Fig. 7.

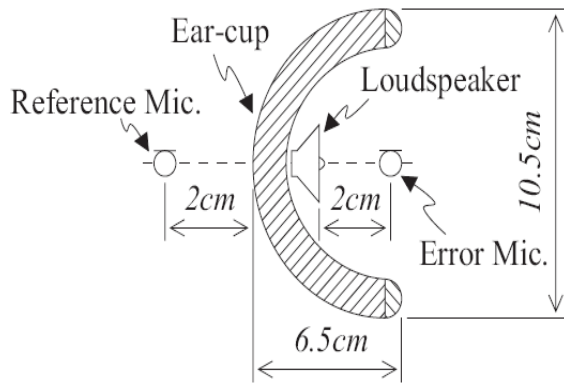


Fig. 7. Ear-cup

We can see the passive attenuation performance where attenuation  $> 20\text{dB}$  and frequencies  $> 500\text{Hz}$ . in Fig. 8.

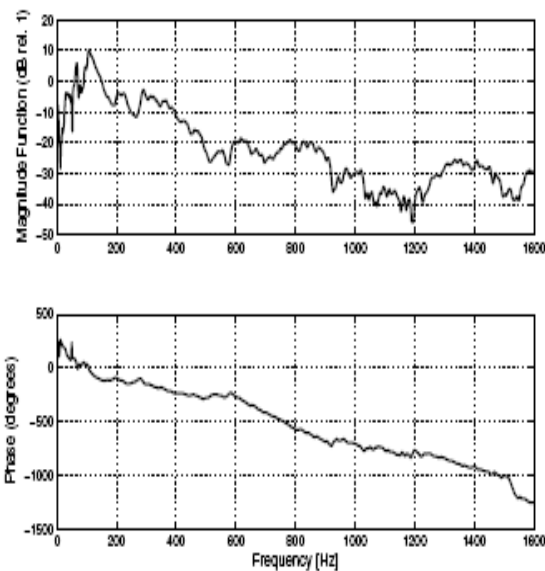


Fig. 8. Passive Attenuation Performance.

Hence, the sampling frequency,  $F_s = 100\text{Hz}$  and the maximum number of coefficients for the adaptive filter and forward path estimate is 40.

## V. RESULT

The basic idea of Active noise control to generate a signal (secondary noise) is equal to a disturbance signal (primary noise) in amplitude and frequency, but has opposite phase. Combination of these signals results in cancellation of the primary (unwanted) noise. For example,

$$W_{opt}(\omega)C(\omega) = -H(\omega) \quad (11)$$

where  $W_{opt}(\omega)$  is the Fourier transform of the adaptive filter coefficients after convergence. In Fig. 9. we can see the filter chain which is presented after convergence.

We can also see the power spectral density [8] of the error microphone signal with and without the control algorithm applied in a broad band noise case in Fig. 10. Corresponding power spectrum of the spectra is shown in Fig.11. for suppression of tonal interference (sinusoidal at 200 Hz).

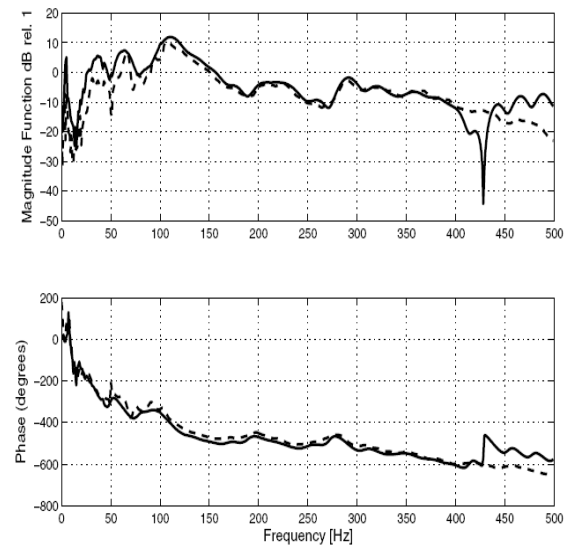


Fig. 9. Filter chain of the converged adaptive filter.

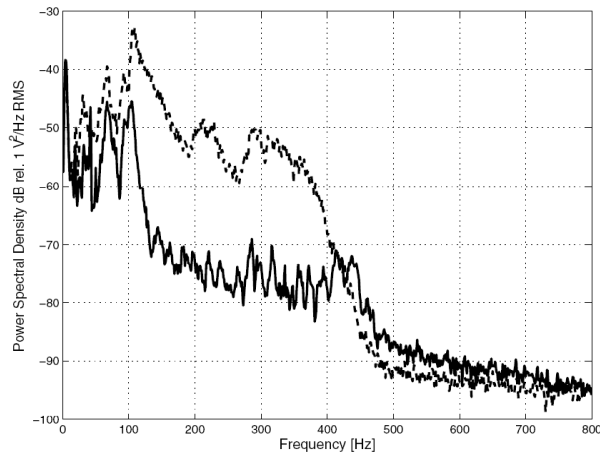


Fig. 10. Power spectral density of the error signal without (dashed) and with (solid) active noise control.

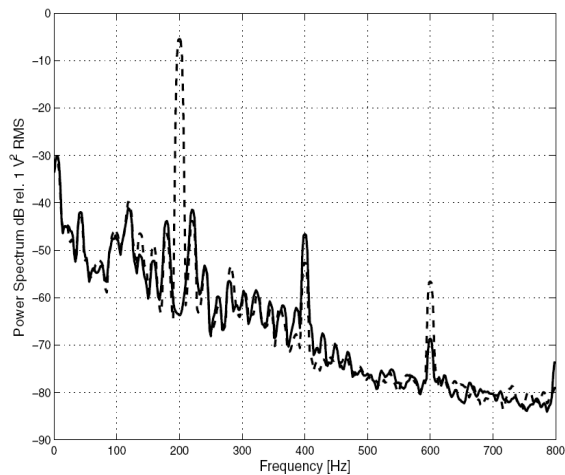


Fig. 11. Power spectrum of the error signal without (dashed) and with (solid) active noise control. Tonal interference sinusoidal at 200Hz.

#### REFERENCES

- [1] <http://en.wikipedia.org>, 25<sup>th</sup> February, 2008, 11 p.m.
- [2] <http://www.ait.gr:80/>, 20<sup>th</sup> March, 2008, 10:07 a.m.
- [3] <http://www.analog.com>, 20<sup>th</sup> March, 2008, 10:10 a.m.
- [4] <http://publik.tuwien.ac>, 25<sup>th</sup> March, 2008, 11 p.m.
- [5] <http://scholar.lib.vt.edu>, 9<sup>th</sup> April, 2008, 07:10 p.m.
- [6] <http://www.erci.edu>, 4<sup>th</sup> May, 2008, 10a.m.
- [7] <https://tagteamdbserver.mathworks.com>, 17<sup>th</sup> June, 2008, 09 a.m.
- [8] Benny Sallberg, Lars Hakansson and Ingvar Claesson – “Active Noise Control for Hearing Protection Using a Low power Fixed Point Digital Signal Processor”.



# A Lossless Image Compression Technique Using Generic Peano Pattern Mask Tree

Mohammad Kabir Hossain<sup>1</sup>, Shams M Imam<sup>1</sup>, Khondker Shajadul Hasan<sup>1</sup>, William Perrizo<sup>2</sup>

<sup>1</sup>Computer Science & Engineering, North South University, Dhaka, Bangladesh.

<sup>2</sup>Computer Science, North Dakota State University, Fargo, ND, USA

Email: mkhossain@northsouth.edu, shams.mahmood@gmail.com, shajadul@northsouth.edu, william.perrizo@northsouth.edu

**Abstract** — Digital Image processing has become ubiquitous in our daily life and the demands to produce and process images are ever increasing. Large amounts of space are required to store these images. Image compression techniques are in high demand as they allow reduction in this storage space. The basis for image compression, as is for most other compression techniques, is to remove redundant and unimportant data. Lossless image compression techniques retain the original information in compact form and do not introduce any errors when decompressed. In this paper, we discuss such a lossless technique using a data structure that we name “Generic Peano Pattern Mask Tree”. It is an improvement over a previously discussed Lossless Image compression technique – “Peano Pattern Mask Tree”. Both these structures are based on the data structure – Peano Mask Tree.

**Index terms** — Lossless Image Compression, Peano Tree, Peano Mask Tree, Peano Pattern Mask Tree, Generic Peano Pattern Mask Tree, Data mining Ready.

## I. INTRODUCTION

For many applications lossless image compression is extremely important to be able to store more images in a given amount of disk or memory space at the same time to retain the original image. Such applications include medical images, legal papers or image scans made for archival purposes, aerial photography from satellites etc. In these cases a perfect reproduction of the original image is necessary and any sort of degradation cannot be tolerated [2, 3].

There are basically two types of Image Compression methods: *Lossy* and *Lossless* [4]. Lossy compression involves the loss of some information that makes it capable of achieving much higher compression at the cost of accuracy. When the information from the compressed data using lossy technique is reconstructed, it introduces compression artifacts. In lossless compression schemes, the

reconstructed image, after compression, is numerically identical to the original image. A perfect reproduction of the original image can be achieved using it.

In this paper we are proposing a method of lossless image compression technique which is an improvement over a previous work done by Shams Mahmood et. al. [1] and Fazle Rabbi et. al. [2]. In the next sections of the paper we are giving a short description of Peano Count Tree, Peano Mask Tree and Peano Pattern Mask Tree. Also we will discuss how these data structures can be used in lossless image compression. After that we will introduce the improved version of Peano Pattern Mask Tree which we call ‘Generic Peano Pattern Mask Tree’ and its application in lossless image compression.

## II. PEANO TREES AND ITS VARIATIONS

### A. The Peano Count Tree

Peano Count Tree (P-tree) is a lossless, compressed, and data-mining ready data structure. The P-tree is a quadrant-based tree representation of the original spatial data [5, 6]. The general concept of P-tree is to recursively divide the entire spatial data into quadrants and to record the count of 1-bits for each quadrant, thus forming a quadrant count tree. Using P-tree structure, all the count information can be calculated quickly. For example, a P-Tree for the given  $8 \times 8$  image of single bit in figure 1 is explained below. Corresponding P-Tree is shown in figure 2.

```

1 1 0 0 1 0 1 1
0 1 0 0 1 0 0 1
1 1 1 1 1 0 1 0
0 1 1 1 1 0 1 0
1 1 1 1 1 1 1 1
1 1 1 1 0 1 1 1
1 1 1 1 1 1 1 0
1 1 1 1 0 1 1 0

```

Fig. 1. 8-bit by 8-bit image.

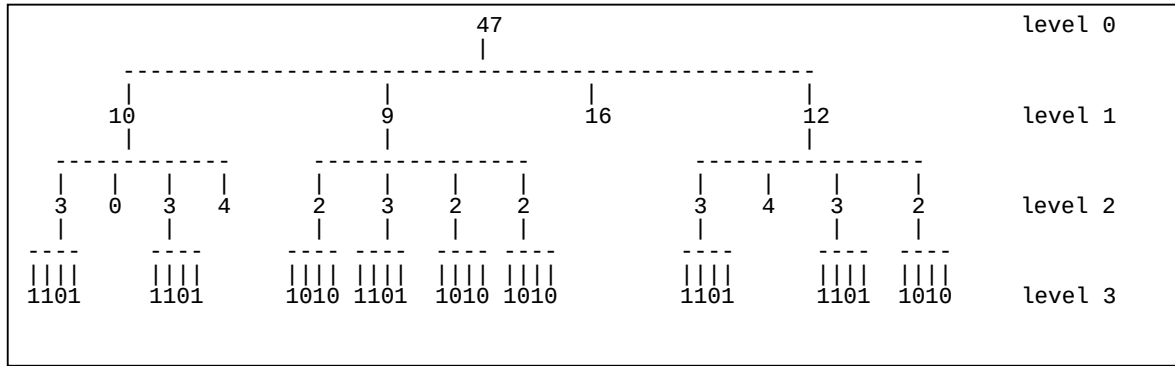


Fig. 2. P-Tree for data in figure 1.

In this example, 47 is the number of 1's in the entire image which is stored in root node which is labeled as level 0. The numbers 10, 9, 16, and 12 found at the next level (level 1) are the 1-bit counts for the four major quadrants in raster order, or Z order (upper left, upper right, lower left, and lower right). This process of storing the count of the number of ones for each of the major quadrants is repeated. When quadrants are composed entirely of 1-bits (called pure-1 quadrants), sub-trees are not needed, and these branches terminate. Similarly, quadrants composed entirely of 0-bits (called pure-0 quadrants), also cause termination. A quadrant which is neither pure-1 nor pure-0 is called mixed quadrant. This pattern is continued recursively using the Peano-ordering, or Z-ordering (recursive raster ordering), of the four sub-quadrants at each new level. Eventually, every branch terminates (since, at the "leaf" level, all quadrants are pure either pure-1 or pure-0).

#### B. The Peano Mask Tree

A variation of the P-tree data structure is the Peano Mask Tree (PM-Tree) in which masks, rather than counts, are used. In a PM-tree, we use three valued logic to represent pure-1, pure-0, and mixed quadrants. (A 1 denotes pure-1; 0 denotes pure-0; and M denotes mixed.) The PM-Tree for the previous

example is given in figure 3. We can easily construct the original P-tree from its PM-tree by calculating the counts from the leaves to the root in a bottom-up fashion.

This PM-Tree data structure was used for compression in [2]. It is obvious that this method will work excellent when neighborhood pixels share the same bit values with high probability. This is usually the case in higher-order bits for pixels in image data. However, neighboring lower order bits have different values due to precision difference. The PM-Tree structure does not provide good compression, if any at all, for such lower order bits. This happens as quadrants that are mixed are introduced and recursive definition of data quadrants goes on till pure-1 or pure-0 quadrants are one bit long. The above mentioned paper thus chose to use only the higher four bits to construct the PM Tree. However in instances where these higher order bits lack a high degree of correlation a similar problem to the one just described might occur. A PM-Tree has three states; at least two bits are required to save the information for each node of the tree. Two bits are able to provide us with four combinations; hence, one combination remains unexploited in the PM-Tree method. This lacking was addressed in [1] which proposed a new data structure, The Peano Pattern Mask Tree.

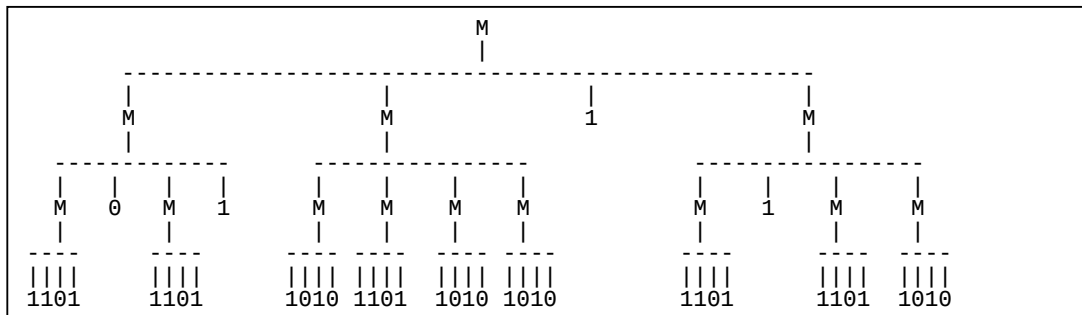


Fig. 3. PM-Tree for data in figure 1.

### C. The Peano Pattern Mask Tree

The Peano Pattern Mask Tree (PPM-Tree) is another variation of the P-Tree. It shares many features of the PM-Tree and addresses some unexplored area of the PM-Tree [1]. However, instead of using three-value logic, the PPM-Tree uses four-value logic. We name this fourth logic as P, for pattern. This pattern is not fixed like pure-1 or pure-0, but rather assumes dynamic values depending on the tree constructed. In fact, the pattern state is a special case of a mixed state. An example will help us better understand the method. To construct a PPM Tree for any data set, we first need to construct a PM Tree for the same data set. Once the PM Tree is created, we search through the leaf level quadrants to look for the most frequently occurring four-bit combination. In the example for PM Tree shown in figure 3, we see that “1101” is the most frequently occurring quadrant at the leaf level. We make this our Pattern and change the original PM Tree. All M nodes having the Pattern Parent property are transformed into a P node and all children removed. This search is carried out in a bottom-up manner where nodes at the lowest levels are transformed before any node at a higher level is transformed. As a result of this, we finally get the PPM Tree as shown in figure 4.

### D. Opportunities Unexplored in PPM-Tree

As mentioned earlier lower order bits for pixels in images share the same bit values with a low probability. As a result quadrants in the PM-Tree and PPM-Tree can usually be 1-bit long, which usually means no compression at all. Also due to the low probability of sharing bit values in neighborhood pixel, pure-0 and pure-1 quadrants are also rare. Hence the decision to stick with pure-0 and pure-1 as two of the four states do not help in compressing, i.e.

reducing the number of levels, the tree structure. We intend to address this lacking in the PPM-Tree structure and propose a new data structure, The Generic Peano Pattern Mask Tree.

### III. THE GENERIC PEANO PATTERN MASK TREE

The Generic Peano Pattern Mask Tree (GPPM-Tree) is another variation of the P-Tree. It shares many features of the PPM-Tree and addresses some unexplored areas. Like the PPM Tree, the GPPM uses four-value logic. Like the PPM-Tree there is one state reserved for the mixed (M) type. However unlike the PPM-Tree there are no fixed patterns, like the pure-0 and pure-1 in the PPM-Tree. Hence there are three states available for dynamic patterns which we name pattern-1 (P1), pattern-2 (P2), and pattern-3 (P3). P1, P2 and P3 are determined dynamically during the tree is created to result in maximum compression of the tree.

The GPPM-Tree improves on the PPM-Tree by being able to choose three of the available sixteen patterns. In the PPM-Tree we were able to choose only one of available fourteen patterns, “0000” and “1111” being reserved as pure-0 and pure-1 states respectively. It can be clearly seen that the GPPM-Tree structure is a more “generic” version of the PPM-Tree as none of the available sixteen patterns are given any special privilege prior to the construction of the tree. An example will help us understand better the GPPM-Tree structure. Consider the 8-bit by 8-bit image of Figure 1. When broken down to the leaf levels in the PM-Tree structure we get the tree in figure 5. We then perform a frequency distribution of the bit patterns in the leaf nodes. (Result in Table I.)

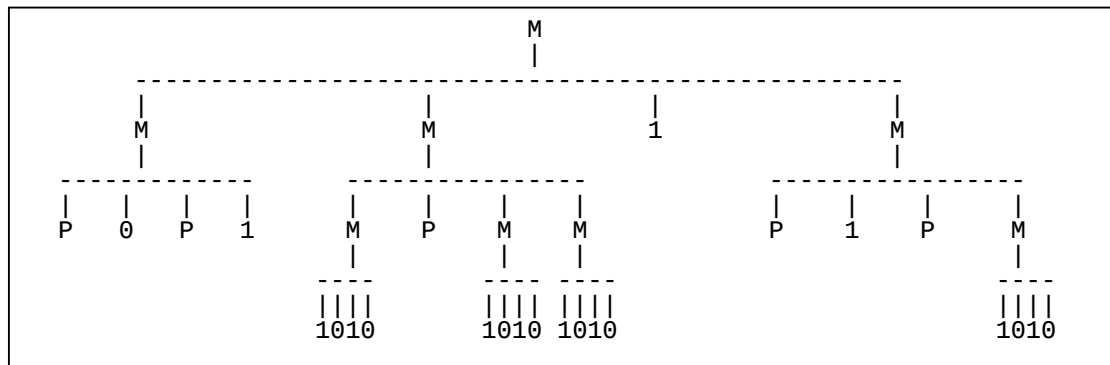


Fig. 4. PPM-Tree for data in figure 1.



an  $n$  pixel image. In our proposed scheme, we first construct  $24 \times n \times n$  dimensional arrays using each of the 24 bits of each pixel. We then create a GPPM tree for each of these arrays. If two of the patterns for the GPPM-Tree turn out to be “0000” and “1111”, then we store the tree as a PPM-Tree avoiding the overhead of storing the two additional patterns, else we store the tree as a GPPM-Tree. Next we decide if storing this GPPM-Tree/PPM-Tree for the corresponding data set from which it was generated requires lesser space than the raw data. If so then the tree is stored, else the original raw data is stored. We require a header file to keep track of which sets are stored in raw or one of the two compressed forms. Since we have three forms of storing the data we need two bits to store this information. The header file format is described below in table 2:

TABLE II  
PROPOSED FILE FORMAT

| Description                                                                                                        | Size in bytes |
|--------------------------------------------------------------------------------------------------------------------|---------------|
| Length of original image as a power of 2                                                                           | 1             |
| Bits representing which segment are stored in raw or compressed form.<br>[ $24 \times 2$ bits = 48 bits = 6 bytes] | 6             |
| Data for Red Component (8 Segments)                                                                                | Variable size |
| Data for Green Component (8 Segments)                                                                              | Variable size |
| Data for Blue Component (8 Segments)                                                                               | Variable size |

While storing the GPPM-Tree, we store the patterns in the first twelve bits (represented in bold in the figure 7) and then store the nodes of the tree in breadth-first order. Since we do this, we do not require storing a count of the number of nodes in the tree. While decoding the tree, decoding stops when all nodes are non-mixed nodes or nodes with quadrant size of 1 bit are reached. We can summarize the results of storing the raw bit pattern in Table 3.

|                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>1111 1101 1010</b> 11 11 11 00 11 01 11 01 00<br>10 01 10 10 01 00 01 10 0000<br><br>The first three 4-bit (bolded text) segments are the three chosen patterns. The next series of 2-bits (normal text) are representations for each of the non-1-bit length nodes/quadrants. The last series of 4-bits (emphasized text) are the patterns of the 1-bit length quadrants. |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Fig. 7. Bit Representation of the GPPM-Tree of figure 6.

## V. EXPERIMENTAL RESULTS

In this paper we only compared the performance of our method to the compression scheme using PPM-

Tree which is mentioned in [1]. The superiority of PPM-Tree over PM-Tree is shown in [2]. A program was written for implementing the image compression scheme using GPPM-Tree. The result is summarized in Table IV. The images used in this experiment can be found at:

[http://picasaweb.google.com/shams.mahmood/AdvancedPPTImages?authkey=K7eg9oaO\\_p4](http://picasaweb.google.com/shams.mahmood/AdvancedPPTImages?authkey=K7eg9oaO_p4)

TABLE III  
SPACE REQUIRED IN STORING THE BIT PATTERN OF  
FIGURE 1 IN VARIOUS FORMATS

| Techniques | Number of bit requirements                                           |
|------------|----------------------------------------------------------------------|
| Raw data   | <b>64</b> bits                                                       |
| GPPM-Tree  | $(3 \times 4) + (17 \times 2) + (1 \times 4)$ bits = <b>50</b> bits. |
| PPM-Tree   | $(1 \times 4) + (17 \times 2) + (4 \times 4)$ bits = <b>54</b> bits  |
| PM-Tree    | $(17 \times 2) + (9 \times 4)$ bits = <b>70</b> bits.                |

The experimental results show that our proposed algorithm performs better than method in [1]. We see significant compression improvements have been achieved in certain cases (Jupiter\_Aurora and Bangladesh). We consider this as a quite significant improvement as the PPM-Tree structure was previously thought to be a mature and optimal evolution of the P-Tree structure. We also proof with the help of our results that the GPPM-Tree is a further improvement for storing data using P-Trees

TABLE IV  
EXPERIMENTAL RESULTS

| Image Name     | Image Dimension | Size in KB | Size using PPM-Tree | Size using proposed GPPM-Tree |
|----------------|-----------------|------------|---------------------|-------------------------------|
| Circles        | 128×128         | 48.0       | 28.15               | 27.90                         |
| Jupiter Aurora | 512 x 512       | 768.0      | 381.73              | 376.72                        |
| Bangladesh     | 512 x 512       | 768.0      | 60.56               | 59.54                         |
| Bright         | 512 x 512       | 768.0      | 398.24              | 398.07                        |
| Dell-wallpaper | 512 x 512       | 768.0      | 453.462             | 453.458                       |

## VI. CONCLUSION

P-tree has been successfully applied in data mining applications ranging from Classification and Clustering with K-Nearest - Neighbor, to Classification with Decision Tree Induction, to Association Rule Mining [7-10]. The P-Tree based decision tree induction method is significantly faster than existing classification methods [10, 11]. P-Tree

data structure allows computing the Bayesian probability values efficiently. Bayesian classification with P-trees has been used successfully on remotely sensed image data to predict yield in precision agriculture [7]. Experimental results showed that using p-tree techniques in an efficient association rule-mining algorithm, P-ARM has significant improvement compared to FP-growth and A-priori algorithms. [7]. Basic P-Trees can be constructed from the compressed file trivially after reconstruction of the GPPM Tree. Once the basic P-Trees have been created, we get data mining ready structure that facilitates efficient data mining tasks [1]. Thus the proposed scheme, the GPPM-Tree structure, is also a data mining ready structure providing it an upper hand over other compression techniques when used in data mining applications.

#### REFERENCES

- [1] Shams Mahmood, Rezaul Hoque, Mohammad Kabir Hossain, William Perrizo, "A New Technique of Lossless Image Compression using PPM-Tree", in the Proceedings of 8<sup>th</sup> International Conference of Computer and Information Technology (ICCIT) 2005, Dhaka, Bangladesh.
- [2] Fazle Rabbi, Mohammad Kabir Hossain, William Perrizo, "Lossless Image Compression using P-Tree," in the Proceedings of 6<sup>th</sup> International Conference of Computer and Information Technology (ICCIT), 2003, Dhaka, Bangladesh.
- [3] Erickson B J, Manduca A, Persons K R, et. al. "Evaluation of irreversible compression of digitized posterior-anterior chest radiographs," *J Digit Imaging* 1997; 10(3): 97-102.
- [4] B. C. Vemuri, S. Sahni, F. Chen, C. Kapoor, C. Leonard, and J. Fitzsimmons "Lossless image compression," Available at <http://citeseer.nj.nec.com/559352.html>
- [5] H. Samet, "The Quadtree and related hierarchical data structure," *ACM Computing Survey*, 16, 2, 1984.
- [6] W. Perrizo, Peano count tree lab notes, Technical report NDSU-CSOR-TR-01-1, Fargo, ND, 2001.
- [7] Mohammad Kabir Hossain, et. al. "Bayesian Classification for Spatial Data Using P-tree," Proceedings of IEEE INMIC 2004, December 24-26, 2004 in Lahore, Pakistan.
- [8] Mohammad Kabir Hossain and W. Perrizo, "Automatic fingerprint identification system using p-tree," in the Proceedings of 5<sup>th</sup> International Conference of Computer and Information Technology (ICCIT) 2002, Dhaka, Bangladesh.
- [9] William Perrizo, William Jockheck, Amal Perera, "Multimedia Data Mining using P-Trees," Available at [http://www-staff.it.uts.edu.au/~simeon/mdm\\_kdd2002/abstracts/14.html](http://www-staff.it.uts.edu.au/~simeon/mdm_kdd2002/abstracts/14.html)
- [10] Quang Ding, Qin Ding and William Perrizo. "Decision tree classification of spatial data streams using peano count trees," Proceedings of ACM Symposium on Applied Computing (SAC '02), Madrid, Spain, March 2002, pp. 413-417.
- [11] Mohammad Kabir Hossain, et. al. "Automatic Face Recognition System using P-tree and K-Nearest Neighbor Classifier", in the Proceedings of 7<sup>th</sup> International Conference of Computer and Information Technology (ICCIT) 2004, Dhaka, Bangladesh.

# A Combined Local-Global Optical Flow Approach for Cranial Ultrasonogram Image Sequence Analysis

Mohiuddin Ahmad, Taohida Taslima, Lutfunnesa Lata, and Masayoshi Yamada\*

Dept. of Electrical and Electronic Engineering, Khulna University of Engineering and Technology,  
Khulna – 9203, Bangladesh

\*Dept. of Electronics and Information Science, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, Japan

**Abstract** – Tissue-motions can represent the strength of artery pulsation of a newborn baby for pediatrics diagnosis. In this paper, we propose a novel method to estimate the tissue-motion quantitatively by combining local and global motion estimation methods, namely combined local-global (CLG) optical flow technique in cranial ultrasonogram of newborn babies. Tissue motions are estimated in different coronal sections successfully. Further, we analyze the time variant tissue-motion by using discrete Fourier transform. We also observed the pulsation in the time variant motion images and strong pulsation is observed in the harmonic frequency of tissue-motion that has a relation to the heartbeat frequency of a newborn baby.

**Index Terms** — Tissue-motion, artery pulsation, newborn baby head, CLG optical flow, cranial ultrasonogram.

## I. INTRODUCTION

Ultrasonogram is routinely used at the various sites of medical diagnosis because of its convenience, easy to use, low cost in compared with magnetic resonance imaging (MRI) and X-ray computed tomography (CT). Cranial ultrasound scans are essential part of the routine investigation during neonatal intensive care. The scan ultrasound image sequences are not only use for real-time diagnosis but also recorded for future diagnosis. Pediatricians use an ultrasonogram to find the cranial abnormalities in a newborn baby, since they can observe the brain tissue through the anterior fontanel [2, 5]. A key point of diagnosis is whether there is the artery pulsation accompanied with blood flow or not [5, 6]. However, there are a few attempts were taken to analyze the tissue-motion from the moving images of ultrasonogram for pediatrics diagnosis. In order to detect artery pulsation from a series of noisy ultrasound echo images of newborn baby head for pediatric diagnosis, a digital image processing system was developed by using the algorithm based on Fourier transform [6]. In order to observe artery pulsation in neonatal cranium at the site of diagnosis, a real time processing system was developed for continuous

display and detection of artery pulsation from moving images of neonatal cranial ultrasonogram [10]. The entire algorithms used above were mainly based on the time sequence variation of pixel value (intensity) in a series of ultrasound echo images. However, the pixel value can easily vary. Therefore, the above methods were very sensitive to the brightness of the original ultrasonogram movie. In this paper, we calculate the tissue-motion due to artery pulsation in cranial ultrasound image sequences by using CLG optical flow technique, which is not sensitive to image brightness change. This is very helpful for quantitative characterization of artery pulsation. Moreover, the time variant tissue-motions are analyzed by using Fourier transform, emphasizing that the tissue-motion at a particular frequency has a relation with heartbeat frequency of newborn baby.

The rest of the paper is organized as follows: Section II gives a brief introduction of brain tissue-motion and artery pulsation. Section III describes the CLG optic flow technique for estimating tissue-motion. In section IV, the proposed analysis of tissue-motion by Fourier transform is illustrated. Some experimental results of our method are illustrated in section V. Finally, we conclude this paper in section VI.

## II. TISSUE-MOTION AND ARTERY PULSATION

The velocity experience by brain tissue due to blood flow is known as *tissue-motion velocity*. The nature of the tissue motion velocity is periodic due to periodic nature of heartbeat and it varies time to time, baby to baby, and it depends on physical conditions. Therefore, tissue motion contains many harmonic components of the pulsation frequency. The tissue-motion velocities repeat themselves due to the periodic nature of heartbeat. Since the heartbeat also varies due to physical condition of newborn babies, the tissue motion velocities vary from time to time, baby to baby.

*Artery pulsation* is the periodic beating of arteries as blood pumps through arteries, and pulsation strength

represents the rating of blood flow to the brain-tissue through essential arteries in the brain. The tissue motion velocity is associated with blood flow. Therefore, pulsation strength can consider as the magnitude of tissue motion velocity, and therefore, a relationship exists between motion and heartbeat interval. Therefore, pulsation strength is lower or higher depends upon the blood flow rate or flow velocity.

### III. ULTRASOUND IMAGE TISSUE-MOTION ESTIMATION

It is well known that all motion estimation techniques are application-dependant and B-mode echocardiography motion estimation itself is a special frame work which is completely different from well-known ground truth sequences. Echocardiography signal has a particular speckled texture that is derived from the primary tissue microstructure and motion detection is based on this primary pattern displacements. Our contribution in this paper is implementing a combined form of optical flow based on two previous famous motion estimations methods (Lucas-Kanade and Horn-Schunck) using image moments instead of primary image pixels. To do that, appropriate window sizes and moment orders and sizes were obtained.

The cranial ultrasound image sequence of a newborn baby is taken as noisy, speckle-like gray scale image including non-rigid tissues such as cerebrums and arteries with local motion due to artery pulsation of blood flow [2]. Therefore, the tissue-motion may vary with time and place. For our analysis, we assume that artery in the cranial ultrasound is locally rigid and keeps its structure during instantaneous motion. Several methods [1, 3, 8] are available for estimating object motion. In order to calculate tissue-motion as the instantaneous optical flow over a local region, we use a combined local-global approach.

We use the optical flow to estimate local motion, because it can precisely establish the motion at any pixel. We calculate the optical flow velocity of any pixel within the region of interest by using the CLG optical flow technique [7]. The CLG optic flow method has complementary advantages over either the global [4] or local [9] method alone, due to its robustness against noise and the combination of local-global optical flow.

#### A. Global Motion Estimation (Horn and Schunck)

The general differential formulation of optical flow computation consists of assuming that the intensity of a particular point in a moving pattern does not change with time and we can write,  $p_x u + p_y v + p_t = 0$  [3]. In global optimization, a functional is defined where a smoothness constraint is used to regularize the solution of a partial differential equation and the smoothness

constraints is weighted with a positive constant. The functional is defined as:

$$\iint \left( (p_x u + p_y v + p_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2) \right) dx dy \quad (1)$$

where,  $p$  is pixel intensity with gradients toward spatial and temporal dimensions of the image, and  $u$  and  $v$  is the motion vectors toward  $x$  and  $y$  directions and  $\alpha$  is a weighting parameter [1, 3]. However primary global methods have visually attractive motion fields but lack the discontinuity preservation property.

#### B. Local Motion Estimation (Lucas-Kanade)

To overcome the sub-dimensionality of the optical flow formulation in Equation (2), Lucas-Kanade [9] proposed equation (3) to solve the problem on small local windows where the motion vector is assumed to be constant. We ought to minimize this equation:

$$\sum W^2 (p_x v_x + p_y v_y + p_t)^2 \quad (2)$$

where,  $p$  is pixel density and  $W$  is a weighted window and  $u$  and  $v$  are the motion vectors towards  $x$  and  $y$  directions [9].

#### C. Combined Local Global (CLG) Motion Estimation

As discussed before, to solve the problem of sensitivity to noise, Bruhn et al. [7] proposed a method to combine local [9] with global [3]. To give a combined local-global method, let us reformulate the previous formulae: We consider spatio-temporal CLG motion,  $w = [u, v, 1]^T$  at a pixel  $x = (x, y)$  in time  $t$ , velocity gradient:

$$\nabla w = [\nabla u, \nabla v]^T, \text{ intensity gradient } \nabla_3 p = (p_x, p_y, p_t)^T,$$

and motion tensor :  $J_p(\nabla_3 p) = K_p * (\nabla_3 p \nabla_3 p^T)$ .  $K_p$  is the smoothing kernel (spatial or spatial-temporal). It is evident that Lucas-Kanade minimizes the quadratic form:

$$\begin{aligned} E_{LK} &= K_p (p_x u + p_y v + p_t)^2 \\ &= \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cdot \begin{bmatrix} p_x^2 & p_x p_y & p_x p_t \\ p_x p_y & p_y^2 & p_y p_t \\ p_x p_t & p_y p_t & p_t^2 \end{bmatrix} \cdot \begin{bmatrix} u & v & 1 \end{bmatrix} \quad (3) \\ &= w^T J_p(\nabla_3 p) w \end{aligned}$$

where,  $K_p$  is the local kernel (usually a Gaussian),  $E_{LK}$  is the energy function of Lucas-Kanade and  $p_x, p_y, p_t$  denote the image intensity gradient with respect to Cartesian and temporal dimensions. We have a similar formulation for Horn and Schunck with an iterative recursion:



$$E_{HS}(w) = \int_{\Omega} (w^T J_0(\nabla_3 p)w + \alpha |\nabla w|^2) dx dy \quad (4)$$

where,  $E$  is the energy function of HS (Horn-Schunck),  $J$  is the motion tensor and  $w$  is the motion vector [10]. Then we can rewrite:

$$\nabla_3 p = (p_x, p_y, p_t)^T = J_p(\nabla_3 p) = K_p * (\nabla_3 p \nabla_3 p^T) \quad (5)$$

By replacing the structure tensor  $J_0(\nabla_3 p)$  instead of matrix  $J_p(\nabla_3 p)$ , we will have:

$$\Delta u - \frac{1}{\alpha} (p_x^2 u + p_x p_y v + p_t p_x) = 0 \quad (6)$$

$$\Delta v - \frac{1}{\alpha} (p_x p_y u + p_y^2 v + p_t p_y) = 0 \quad (7)$$

Its minimization induces below equation which satisfies Euler-Lagrange formula:

$$\Delta u - \frac{1}{\alpha} (K_p(p_x^2)u + K_p(p_x p_y)v + K_p(p_x p_t)) = 0 \quad (8)$$

$$\Delta v - \frac{1}{\alpha} (K_p(p_x p_y)u + K_p(p_y^2)v + K_p(p_y p_t)) = 0 \quad (9)$$

Now, the spatiotemporal combined local-global approach is defined as:

$$E_{CLG}(w) = \int_{\Omega} (w^T J_p(\nabla_3 p)w + \alpha |\nabla w|^2) dx dy \quad (10)$$

Combining the temporal extended variant of both [4] and [12] method, we obtain a spatio-temporal version of the combined local-global functional which is given by,

$$E_{CLG}(w) = \int_{video} (w^T J_p(\nabla_3 p)w + \alpha |\nabla w|^2) dx dy dt \quad (11)$$

Therefore, based on Lagrange equations, we have:

$$0 = \sum_{j \in N(i)} \frac{u_j - u_i}{h^2} - \frac{1}{\alpha} (J_{11i} u_i + J_{12i} v_i + J_{13i}) \quad (12)$$

$$0 = \sum_{j \in N(i)} \frac{v_j - v_i}{h^2} - \frac{1}{\alpha} (J_{21i} u_i + J_{22i} v_i + J_{23i}) \quad (13)$$

for  $i = 1, 2, \dots, N$  and  $N(i)$  denote the set of neighborhood of pixel  $i$ . After a simple calculation, the iterative formulations in (14) and (15) are resulted.

$$u_i^{k+1} = (1 - \omega) u_i^k + \omega \frac{\sum_{j \in N^-(i)} u_j^{k+1} + \sum_{j \in N^+(i)} u_j^k - \frac{h^2}{\alpha} (J_{12i} v_i^k + J_{13i})}{|N(i)| + \frac{h^2}{\alpha} J_{11i}} \quad (14)$$

$$v_i^{k+1} = (1 - \omega) v_i^k + \omega \frac{\sum_{j \in N^-(i)} v_j^{k+1} + \sum_{j \in N^+(i)} v_j^k - \frac{h^2}{\alpha} (J_{21i} v_i^k + J_{23i})}{|N(i)| + \frac{h^2}{\alpha} J_{22i}} \quad (15)$$

where,  $N^-(i) = \{j \in N(i) | j < i\}$ ,  $N^+(i) = \{j \in N(i) | j > i\}$ , and  $|N(i)|$  is the number of neighborhood pixel region for implementing Gauss-Seidel iteration. Here,  $i$  and  $j$  indices determine the backward and forward Gauss-Seidel iterative pixels and  $k$  defines the iteration number [7]. From motion components, the absolute value of tissue-motion,  $V(x, y, t) = \sqrt{v_x^2 + v_y^2}$  and the phase angle,  $\phi = \tan^{-1}(v_y / v_x)$  is determined.

#### IV. DFT ANALYSIS OF MOTIONS

The discrete Fourier transform is meaningful to inspect magnitude versus frequency plot for changes the discrete tissue motion versus time at a particular frequency than to observe the tissue-motion with time variation. Since time variant tissue-motions contain different frequency components, so the analysis by Fourier transform is helpful to detect such frequency in which pulsation occurs. The tissue-motion at any pixel position  $V(x, y, t)$  can be stated as  $V(k\delta t)$  at any position  $(x, y)$  with sampling interval,  $\delta t$  where  $k$  is the sample index  $[0 \leq k \leq n-1]$ . Now, the discrete Fourier transform,  $V(f)$  of  $n$ -number of samples, time variant tissue-motion as one-dimensional function,  $V(k\delta t)$  is defined as [4]:

$$V(f) = \sum V(k\delta t) e^{-\frac{j 2 \pi f k}{n}} \quad (16)$$

where,  $f$  = frequency index. At a particular frequency  $V(f)$  has two parts; i.e. one is real part that contains cosine term,  $R(f) = \Re[V(f)]$  and the other is imaginary term,  $I(f) = \Im[V(f)]$  and therefore the power spectral density of motion,  $P(f)$  is given by:

$$P(f) = |V(f)|^2 = R^2(f) + I^2(f) \quad (17)$$

Pulsation frequency,  $f_p$  at which the maximum spectral density occurs, evaluates among different frequency components by the following relationship:

$$f_p = \max_{1 \leq f \leq n/2} \{P_f(x, y)\} \quad (18)$$

where,  $P_f$  is the power spectrum at any position in the harmonic image.

## V. EXPERIMENTAL RESULTS AND DISCUSSION

The original cranial ultrasound images of the newborn baby head were taken in B-mode by using an ultrasound probe of 5 MHz and recorded on video tapes at the site of pediatricians, which is shown in Fig. 1.

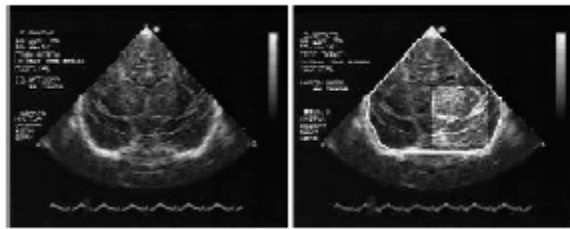
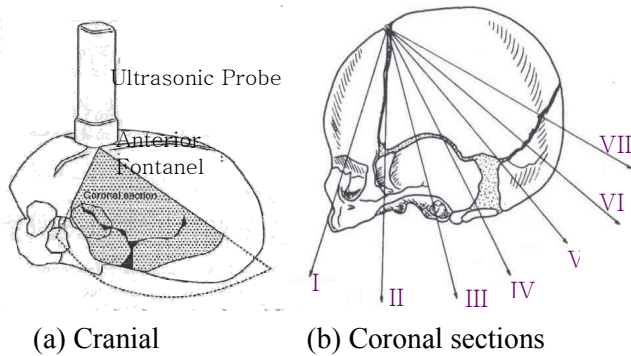


Fig. 1. Cranial Ultrasonogram of newborn baby head with an image of coronal section II.

Then a series of ultrasound echo images (32 frames,  $640 \times 480$  pixels/frame, 8 bits/pixel, 33 ms/frame) was captured by a PCI-based PC with video digitizer. Figure 1 (c) shows a typical image in the anterior coronal section (section II) in the cranial ultrasonogram. We also show the region of interest (ROI) by the white border line in Fig. 1 (d). The detailed scheme of essential tissue in this section is shown in Fig. 2. In Fig. 1 (d), the necessary tissues and organs are shown within a white boundary line. Outside it, there are cranial bones and echo artifacts. The coronal image shows the symmetrical property with left and right parts.

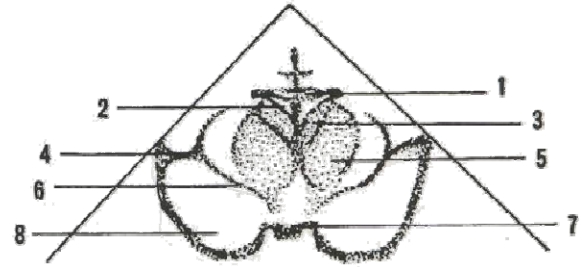


Fig. 2. The essential tissue-scheme in the anterior coronal section (section II). 1: Corpus callosum, 2: Lateral ventricle, 3: Head of caudate nucleus, 4: Sylvian fissure, 5: Lenticular nucleus, 6: Middle cerebral artery, 7: Body of sphenoid bone, 8: Temporal lobe.

Therefore, we can use a portion of the image [square block with enhanced brightness in Fig. 1(d)] in a sequence of our analysis. As preprocessing steps, we have checked several critical conditions of the image sequence, such as (i) sway of the ultrasound probe, (ii) sway of the head of baby, (iii) discontinuity of the image frames, and (iv) completely dark or bright of any image frame in the sequence.

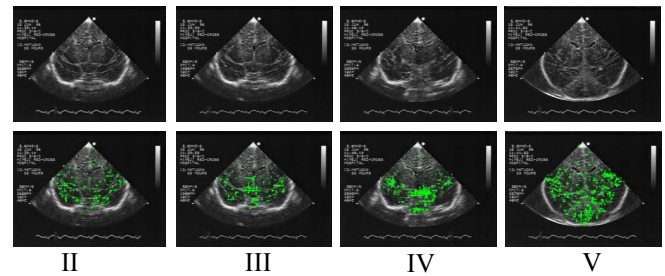


Fig. 3. Different coronal section images with tissue-motion overlapping on the images. (II, III, IV, and V mean coronal sections). Green color vectors show the motion.

We calculate the tissue-motion velocity at each pixel by using a combined local global method for all the coronal sections in consecutive two frames of cranial ultrasound echo images of the newborn baby head by using the spatial and temporal gradients. The ultrasound images as well as the tissue-motions overlapping the images are shown in Fig. 3. We also show the tissue-motion vector of the ROI and the precise vector diagram of the specified region of Fig. 1 is shown in Fig. 4. In the arterial region, the motions have significant value.

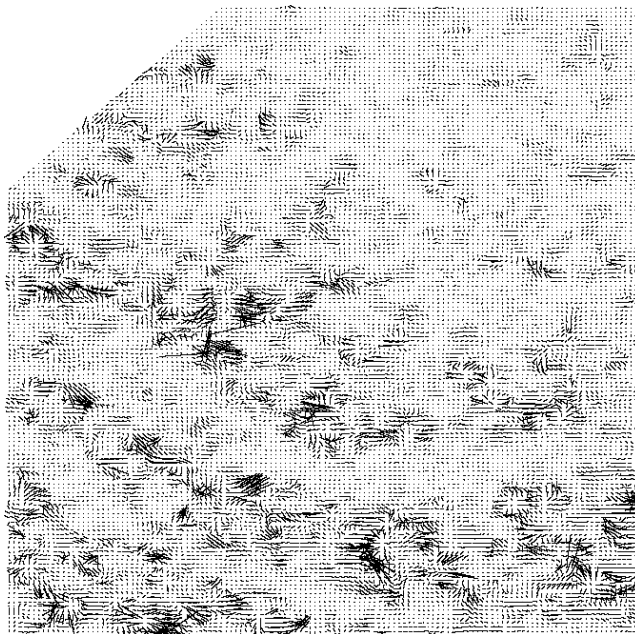


Fig. 4. Two-dimensional maps of tissue-motion vectors in the anterior coronal section of cranial ultrasonogram of newborn baby head.

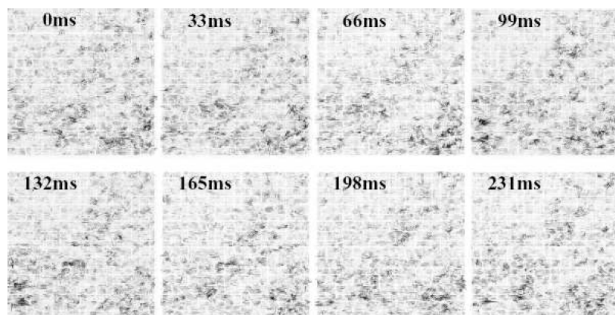


Fig. 5. Time-variant tissue-motion velocities

For computing the frequency components of tissue-motion, we use fast Fourier transform [8]. We do not consider the motion appeared outside cranial bone. In addition, no motion is found in the region of cranial bone.

For simplicity, we have used a portion of the image in the sequence, since it contains the essential tissues and organs such as middle cerebral artery, sylvian fissure and temporal lobe, is shown as enhanced in Fig.1.

Figure 5 shows the two-dimensional maps of time-variant tissue-motion vector of the portion of cranial ultrasonogram. The magnitude of tissue-motion vectors are normalized so that the distance between two neighbor pixels in the map corresponds to 0.5 pixels/frame. The tissue-motion vectors vary with place and its magnitude becomes up to a few pixel/frame at several region within

the cranial bone. The middle cerebral artery is one of the essential arteries in the brain and there are fine arteries around sylvian fissure and corpus callosum. The spatial distribution of large motion vectors has found in the above regions. On the other hand, motion vectors become zero on the cranial bone and at such tissues as lenticular nucleus which is also shown in Fig. 4. Furthermore, the maximum vectors are aligned to the same directions at many regions of the images.

The motion vectors do not vary only place but also with time, which is shown in Fig. 5. Pulsation regions were found with strong pulsation and the magnitude of motion is high in the region of middle cerebral artery and near region. Thus, it is very useful to observe the periodical variation of motion as well as the magnitude of tissue motion over time in different area where different organs are involved.

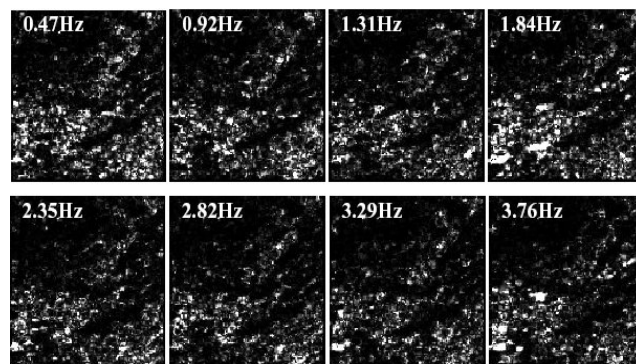


Fig. 6. Power spectrum of tissue-motion at different frequency.

The heart rate of the typical baby is 140 per minute observed at the time of image sequence capturing. Therefore, the heartbeat interval of this typical baby is 429ms. The corresponding pulsation frequency is 2.33Hz. By using sampling interval 33 ms and the number of samples 64, the frequency resolution is obtained at 0.47 Hz. Therefore, fundamental frequency of the heartbeat occurs in between 4th and 5th harmonic frequency of tissue-motion. This is illustrated in Fig. 6, where the power spectrum of tissue-motion at different frequency is shown.

We show only a few images out of a large number of images. The strong pulsation is found in the image that corresponds to the frequency of 1.84Hz and 2.35 Hz. Therefore, strong pulsation is found in a frequency, which is not the fundamental frequency of heartbeat frequency but the harmonic frequency of pulsation frequency.

This fact represents that pulsation is not sinusoidal, and the period of Fourier transform is not equal to the integer time of the period of heartbeats. The normalized power spectral density is shown in Figure 6 at different

frequency. Here 8 frequency components are shown. It is clear that tissue-motion contains several harmonic pulsation frequency components. From Fig. 7, it is clear that the strong pulsation occurs at 4th harmonic frequency of heartbeat frequency and the tendency of the neighbor frequency is high enough. We have tested several sections, using both 32 and 64 samples, several cranial ultrasonogram images of newborn babies, and find the similar results. The babies exhibit different pulsation, since their physical conditions were different. The plot in Fig. 9 shows the normalized power spectrum of some babies at different pulsation frequency.

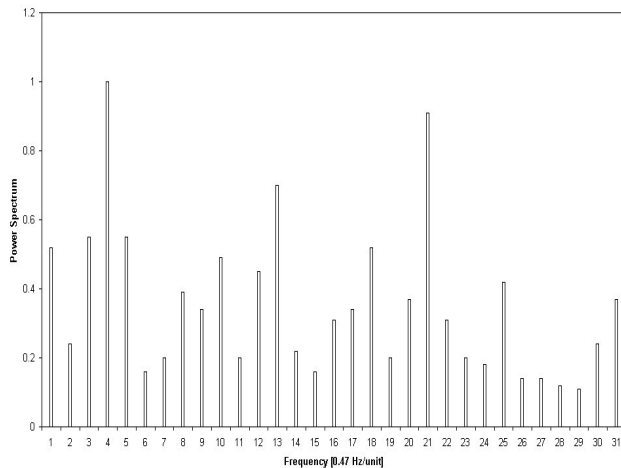


Fig. 7. Histogram showing the power spectral density at different frequency.

## VI. CONCLUSIONS

We successfully computed the tissue-motion by using a combined local-global optical flow method for characterizing artery pulsation quantitatively without losing the benefits of local and global method. Since tissue-motion is not affected by spatial variation of image brightness, therefore the strength of artery pulsation can be quantitatively estimated by using the magnitude of tissue-motion velocity. From frequency analysis, it is found that pulsation is associated with blood flow is periodically changed with heartbeat is verified by using Fourier transform of time variant motion. Strong pulsation is found at a harmonic frequency, which is related to heartbeat frequency, and the tissue-motions are not sinusoidal. Our future work includes precise and automatic detection and estimation of artery pulsation from noisy cranial ultrasound image sequences.

## REFERENCES

- [1] M. Ahmad, M. S. Yusuf, M. Z. Chowdhury, and M. Yamada, "Global optimization of optical flow technique on the computation of tissue-motion in neonatal cranial ultrasonogram," *Journal of Electrical & Electronic Engg, Istanbul University*, vol. 4, no. 1, pp. 1007-1016, 2004.
- [2] M. Fukazawa, H. Kubo, Y. Kitsunezuka, and M. Yamada, "Motion analysis of artery pulsation in neonatal cranial ultrasonogram," *SPIE*, vol. 3661, 1999.
- [3] Horn, B. K. P., and B.G. Schunck, "Determining Optical Flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.
- [4] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing*, Addison-Wesley Publishing Company, 1993.
- [5] Y. Kitsunezuka, and H. Nakamura, "Syonika-Rinsyo," no. 38, pp. 297-308 [in Japanese], 1985.
- [6] M. Yamada, M., M. Fukuzawa, Y. Kitsunezuka, J. Kishida, N. Nakamori, H. Kanamori, T. Sakurai and S. Kodama, "Pulsation detection from noisy ultrasound echo moving images of newborn baby head using Fourier transform," *Jpn. J. Appl. Phys.*, vol. 34, p. 2854-2856, 1995.
- [7] A. Bruhn, J. Weikert, and C. Schnorr, "Lucas-Kanade meets Horn and Schunck: combining local and global optic flow methods," *Int'l J. Computer Vision*, vol. 61, no. 3, pp. 211-231, 2005.
- [8] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int'l J. Computer Vision*, vol. 12, no. 1, pp. 43-77, February 1994.
- [9] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *In Proc. Seventh Int'l Joint Conf. on Artificial Intelligence*, pp. 674-679, Vancouver, Canada, 1981.
- [10] M. Fukuzawa, Y. Kitsunezuka and M. Yamada, "A real-time processing system for pulsation detection in neonatal cranial ultrasonogram," *Jpn. J. Appl. Physics*, vol. 37, pp. 1175-1178, 1998.

# Distortion Measurement Using Arc-Length-Parameterisation within a Vertex-Based Shape Coding Framework

Ferdous Ahmed Sohel and Mohammed Bennamoun

School of Computer Science and Software Engineering  
The University of Western Australia, Crawley, WA 6009, Australia.

Emails: [Ferdous.Sohel@csse.uwa.edu.au](mailto:Ferdous.Sohel@csse.uwa.edu.au); [m.bennamoun@csse.uwa.edu.au](mailto:m.bennamoun@csse.uwa.edu.au)

**Abstract** — Existing vertex-based shape coding algorithms use a number of different distortion measurement techniques including the *shortest absolute distance* (SAD), *distortion band* (DB) or *tolerance band* (TB), *accurate distortion measurement technique for shape coding* (ADMSC) and *distortion measurement based on chord-length-parameterisation* (DMCLP). Among these techniques DMCLP is computationally the fastest. It however, is a relaxed measurement technique and does not fully utilise the admissible distortion within the rate-distortion optimisation. This paper presents a novel *distortion measurement technique based on arc-length-parameterisation* (DMALP) within the vertex-based shape coding framework. Experimental results vindicate its superior performance over the existing geometric distortion measurement techniques.

**Index Terms** — Image coding, image/video coding and transmission, video coding.

## I. INTRODUCTION

Despite facilitating increasingly effective retrieval, manipulation and interactive editing functionality for both natural and synthetic video sequences, object-oriented video coding using shape information remains a challenging research topic [1-4]. The universal pursuit for efficiency in existing communication technologies means that applications such as video-on-demand, video streaming over the Internet, real-time applications and mobile video transmissions for handheld devices as well as real-time identification and authentication systems will derive significant benefits from strategies enabling fast shape coding.

The rigorous review of shape coding algorithms presented in [1] draws the conclusion that the classical vertex-based shape coding framework is optimal in an *operational rate-distortion* (ORD) sense. With both polygonal and quadratic B-splines (BS) based strategies being deployed, it has become the core for several other shape coding algorithms [2-5]. The broad aim of these algorithms is that for some prescribed distortion, a shape contour is optimally encoded in terms of the number of bits, by selecting a set of *control points* (CP) that incurs the lowest bit-rate and vice versa. Distortion measures thus play a vital role in these algorithms, as evidenced in [1], either the *shortest absolute distance* (SAD) or

*distortion band* (DB) measure is applied. A generalisation of DB is the *tolerance band* (TB) which supports a variable *admissible distortion* (AD) [2, 4], though it inherits the same performance characteristics of the DB including the propensity for trivial solutions. Furthermore, while both SAD and DB guarantee all points on the approximating curve are bounded by the peak distortion, they do not ensure that each boundary point maintains the AD particularly at shape corners. This anomaly has been solved by the new *accurate distortion measurement technique for shape coding* (ADMSC) [5]. Since distortion measurement is seamlessly integrated into the ORD optimal shape coding framework, the overall computational complexity of these algorithms is highly dependent on the distortion measurement overhead. In order to ensure computationally efficient encoding, it is essential to employ a fast technique. For both polygon and BS-based encoding, the DB and TB incur  $O(N_B^2)$  computational complexity, where  $N_B$  is the number of boundary points. Additionally, the computational impost for DB and TB proportionally increases with AD as the width of the respective admissible bands widens. In contrast, the SAD and ADMSC measures require  $O(N_B)$  time for polygonal encoding, but still mandate  $O(N_B^2)$  time for BS-based encoding. BS encoding is mostly used since it affords a twofold superiority over its polygonal counterpart, in firstly requiring a lower bit-rate since higher degree curves are used and secondly curve smoothness, due to the inherent parametric continuity of BS. To address these issues a new distortion measurement technique based on chord-length-parameterisation (DMCLP) [6] was proposed. DMCLP incurs only  $O(N_B)$  computational cost for both polygon and BS-based shape coding.

Since DMCLP is based on *chord-length-parameterisation* (CLP), it determines a value of  $u$  for each contour point, which is then employed to obtain the approximating BS (or polygon) point for that contour point. The location of these approximating points are not only dependent on  $u$ , but also dependent on the distance between the consecutive CP. Concentration of these points is greater in the area close to a smaller control polygon edge than a large one. Consequently, DMCLP over-

calculates (the calculated distortion is more than the actual distortion) the distortion and can reject a candidate CP set on the ground that it did not maintain AD, despite being maintained the opposite. Therefore, it can lead to an unnecessarily high bit-rate requirement. To overcome this problem this paper presents a novel *distortion measurement technique that utilises the arc-length-parameterisation* (DMALP). In fact, *arc-length-parameterisation* (ALP) produces an equal unit distance between the consecutive BS-points, which resembles the one unit distance between consecutive contour points on a one pixel neighbourhood. In this technique the values of  $u$  are calculated based on a modified CLP and then combined with the weights obtained from the ALP [7]. Experimental results confirm the superior performance of DMALP compared with the DMCLP and the others. It should also be noted that alike DMCLP, DMALP incurs only  $O(N_B)$  computational time for both polygon- and BS-based shape coding.

The rest of this paper is organised as follows: Section II provides a short overview of the existing distortion measurement techniques within the vertex-based shape coding framework. Section III details the new DMALP technique for both BS and polygon based encoding. Section IV provides a thorough discussion on the empirical results and performance comparison with DMALP embedded into the original ORD framework. Some concluding remarks are given in Section V.

## II. EXISTING DISTORTION MEASUREMENT TECHNIQUES WITHIN THE VERTEX-BASED ORD OPTIMAL SHAPE CODING FRAMEWORK

Existing ORD optimal shape coding algorithms seek to determine and encode a set of CP representing a particular shape. Let the contour  $B = \{b_0, b_1, \dots, b_{N_B-1}\}$  be an ordered set of points.  $S$  is an ordered set of CP used to approximate  $B$  with  $S \subseteq C$ , where  $C$  is the ordered set of vertices in the admissible CP band. The original ORD framework detailed in [1-3], is valid for both polygon and BS-based models and determines an optimal  $S$  for  $B$  within RD constraints. As the choice of the distortion metric clearly has a significant impact upon both computational speed and quality of this framework, the existing distortion measurement techniques will now be briefly examined.

*The shortest absolute distance:* This is the most straightforward and widely used geometric shape distortion measurement technique [1]. With SAD, distortion for any arbitrary shape point is defined as the perpendicular distance of that point from the corresponding line segment (or its extension) of the approximated shape polygon. Though it is simple to

compute, this metric however does not always accurately reflect the true distortion since the perpendicular distance may not always correspond to the actual distance for all shape points. For instance, when the shortest absolute distance from the extended part of a line segment is considered to be the minimum distance between a shape point and the line, it fails to represent the actual distance [5]. Apart from this, while the overall computational complexity of SAD is  $O(N_B)$  for polygonal encoding, it requires  $O(N_B^2)$  cost for BS-based coding.

*Tolerance band:* TB works as follows [2, 4]: draw a circle around each  $b_r$  of radius  $T[r]$  which is the AD at  $b_r$  and is determined from the prescribed admissible bounds ( $T_{\max}$  and  $T_{\min}$ ); TB consists of the set of all points that lie inside the circles; check the distortion, if all points on a candidate edge (curve) lie inside the TB, it is considered that the candidate edge (curve) maintains the distortion criterion. Thus, every point on the candidate edge (curve) is required to be checked to see whether it belongs to the TB points set. The number of points compatible with the TB-grid on a candidate edge (curve) is of  $O(N_B)$  while TB itself comprises a set of  $O(N_B)$  points, so the entire checking process for a candidate curve necessitates  $O(N_B^2)$  computational cost in the worst case. However, as the distortion of each individual associated boundary point is not tested to see if it lies within TB, this inherits the same accuracy measurement problems as DB [5].

*Accurate distortion measurement technique for shape coding:* ADMSC [5] considers all possible relative positions of a point with respect to an edge, it always measures the distortion accurately. Hence, the inaccuracy problems associated with both SAD and TB techniques are solved with this ADMSC. In polygonal encoding, the edge-distortion for all associated boundary points are calculated from the candidate edge and checked against the corresponding AD which takes  $O(N_B)$  computational time. BS encoding however, requires  $O(N_B^2)$  cost to check the distortion because a BS curve is in fact a concatenation of piecewise polygon-edges and so for each boundary point associated with a candidate curve, the individual distortion has to be measured from all edges that form the curve. The minimum edge-distortion is then taken as the distortion for that particular boundary point and checked with the corresponding AD.

*Distortion measurement using chord-length-parameterisation:* The philosophy in DMCLP [6] is that if there is an associated approximated point to each contour point, distortion can be measured as the Euclidean distance between these two points. Therefore, for each contour point it determines a value of BS control parameter  $u$  and thus an approximating BS point.



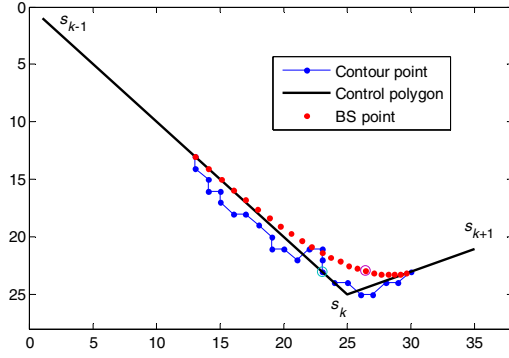


Figure 1: Illustration of the problem with the DMCLP.

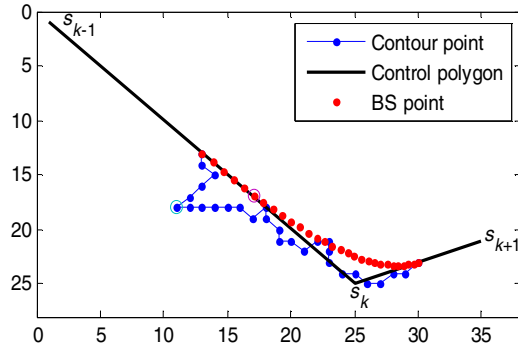
It is computationally the most efficient amongst these distortion measurement techniques and incurs  $O(N_B)$  time for both polygon and BS-based encoding. However, DMCLP over-calculates the distortion. The location of BS points does not only depend on  $u$ , but also on the distance between consecutive CP. The concentration of these points is greater in the area close to the smaller control polygon edge than the large edge. For instance, as in Figure 1, the concentration of BS points is higher in the part close to the edge  $s_k s_{k+1}$  than that of  $s_k s_{k-1}$ . It shows the 17<sup>th</sup> point of both the contour and the approximating BS curve encircled. It is clearly visible that despite the distortion at the 17<sup>th</sup> contour point with respect to the entire approximating BS curve is much lower than the point to point distance obtained by DMCLP. Therefore, despite producing a smaller distortion, DMCLP can reject this CP set and hence lead to an unnecessarily high bit-rate requirement. To overcome this problem the following section presents a novel distortion measurement technique based on ALP.

### III. DISTORTION MEASUREMENT TECHNIQUE USING ARC-LENGTH-PARAMETERISATION

The philosophy behind the ALP is that it define a set of weights for the parametric curve coefficients so that the generated curve points obtain a ‘unit speed’ [7]. It alternatively means the distance between the consecutive curve points is ‘equal’. It resembles the fact that the contour points also form a chain in a one pixel neighbourhood. So if the number of BS points, i.e., the number of steps in  $u$  is equal to the number of contour points, then a BS point would lie within a close area of the corresponding contour point. The overall process is performed in two modules. In the first module the value of  $u$  is determined using a modified CLP. For an arbitrary curve segment having start and end indices  $k_1$  and  $k_2$  respectively of the associated boundary points in  $B$ , the values of  $u$  can be determined as:

$$u_r = \begin{cases} 0 & \text{if } r = k_1 \\ \frac{|b_{k_1} b_{k_1+1}| + |b_{k_1+1} b_{k_1+2}| + \dots + |b_{r-1} b_r|}{|b_{k_1} b_{k_1+1}| + |b_{k_1+1} b_{k_1+2}| + \dots + |b_{k_2-1} b_{k_2}|} & \text{otherwise} \end{cases} \quad (1)$$

where  $u_r$  is associated to  $b_r$ . In DMCLP  $|b_{r-1} b_r|$  is the Euclidean distance between  $b_{r-1}$  and  $b_r$ . However, this approach has a limitation. In this form  $u$  monotonically increases with the value of  $r$ . But in many cases, despite the fact that there is an increment in the contour point index, there is no real advancement of the position of the contour point with respect to the baseline (the line joining the two end control points). Consequently, though the contour point is not moved, the corresponding BS point is moved, and hence can lead to the over-calculation of distortion. This scenario can be explained with the example in Figure 2. It shows that for the 3<sup>rd</sup> contour point the BS point is quite close to it. But although for the next few contour points there is a little advancement in the direction of baseline, the BS points have moved away. For example, the 6<sup>th</sup> contour point and its corresponding BS point are encircled and the distance between them is quite high. To overcome this limitation, it is important to use the projection of the distance between  $b_{r-1}, b_r$  on the baseline, rather than using the geometric distance. If the angle is  $\beta$ , then the projection would be  $\text{distance}(b_{r-1} b_r) \cdot \cos(\beta)$ . Since, 1 pixel neighbourhood is considered for the contour chain, if the angle between the baseline and the global coordinate is also converted to the same pixel grid, depending on the angle the values of



$|b_{r-1} b_r|$  would be either 0 or 1.

Figure 2: Illustration of the problem with the CLP.

In the second module, ALP is performed to obtain the weights  $w$  so that an equal distance between the CP is achieved. As detailed in [7], for a quadratic curve ALP can be obtained as follows:

$$Q_6 = \left( |\Delta s_{k-1}|^2 + 6 \cdot |\Delta s_k|^2 + 3 \cdot |\Delta s_{k-1}| \cdot |\Delta s_k| \right) \quad \text{and} \\ -Q_0 = \left( 6 \cdot |\Delta s_{k-1}|^2 + |\Delta s_k|^2 + 3 \cdot |\Delta s_{k-1}| \cdot |\Delta s_k| \right)$$

where  $|\Delta s_k| = s_{k+1} - s_k$ .

$$\text{So } \alpha_{\min} = \left( 1 + \sqrt{-Q_6/Q_0} \right)^{-1}$$

Hence, the optimal parameterisation coefficients are

$$w_j = (1 - \alpha_{\min})^j \cdot (\alpha_{\min})^{2-j} \quad \text{for } j = 0, 1, 2. \quad (2)$$

The combination of these weights and the respective value of  $u$ , will be employed to calculate the approximating point corresponding to a contour point. The distortion is then calculated using point to point distance between them. The overall DMALP technique is formalised in Algorithm 1.

**Algorithm 1:** The proposed distortion measurement based on arc-length-parameterisation (DMALP).

Inputs:  $B$  - boundary,  $s_{k-1}, s_k, s_{k+1}$  - control points,  $T$  - AD,  $k_1, k_2$  - indices of the associated boundary points.  
Output: *Flag* - 1 means distortion criteria upheld; 0 means not upheld.

- 
1.  $Flag = 1$ ;  $chord\_len_{k_1} = 0$ ;
  2. FOR  $\forall r, k_1 < r \leq k_2$
  3.      $chord\_len_r = chord\_len_{r-1} + |b_{r-1}b_r|$  using (1);
  4. FOR  $\forall j, 0 \leq j \leq 2$
  5.     Determine  $w_j$  using (2);
  6. FOR  $\forall r, k_1 \leq r \leq k_2$
  7.      $u_r = \frac{chord\_len_r}{chord\_len_{k_2}}$ ;
  8.     Calculate the approximating point  $q$  using  $u_r$  and  $w$ ;
  9.      $dist = |qb_r|$ ;
  10. IF  $dist > T[r]$   $Flag = 0$ ;
- 

It will now be proven in Lemma 1 that DMALP maintains AD for all boundary points.

**Lemma 1:** If  $d[r]$  and  $T[r]$  are respectively the generated and AD for an arbitrary boundary point  $b_r$ , the DMALP technique always upholds the AD so that  $d[r] \leq T[r]$ .

**Proof:** By embedding DMCLP into the vertex-based ORD optimal shape coding framework, every boundary point will have a respective point on the encoded shape. If the distortion defined in Step 9 of Algorithm 1 at  $b_r$  is  $dist_r$  and distortion (shortest distance) for the same point with respect to the final approximated shape is  $d[r]$ , then  $d[r] \leq dist_r$ . In the distortion test (Step 10), for any curve segment to be included in the encoded shape,  $dist_r \leq T[r]$  must hold true, so that  $d[r] \leq T[r]$ .

**Computational complexity analysis:** In terms of the order of complexity of Algorithm 1, Steps 2-3 calculate

the cumulative chord-length of the boundary at a cost, in the worst case of  $O(N_B)$ , while Steps 4-5 determine the reparameterisation  $O(N)$  time, where  $N$  is the degree of the curve. For quadratic BS and polygon based approximation the value of  $N$  is 2 and 1 respectively. Steps 6-10 firstly determine the value of  $u$  and then generate the approximating point before checking to see whether the AD is maintained which also incurs  $O(N_B)$ . The overall complexity for Algorithm 1 in the worst case is therefore,  $O(N_B)$ , which is equal to the DMCLP and one degree less than all other existing distortion measurement techniques for BS based encoding. The main difference between DMALP and DMCLP is the calculation of reparameterisation coefficients in Steps 4-5 and their use in Step 8.

#### IV. RESULTS AND ANALYSIS

To analyse the performance of DMALP compared with the existing distortion measurement techniques, the ORD shape coding algorithms including the SW strategy were implemented in Matlab (R2007a) and applied to a number of test shape sequences. The experiment platform was 2.67GHz Pentium-4 microprocessor, with four Gigabyte RAM under Windows XP.

Figure 3(a) plots the rate vs distortion results for BS-based encoding upon the popular *Kids* video sequence (SIF, 300 frames) while Figure 3(b) provides a zoom-in of the section highlighted by the rectangle in (a). It should be noted that  $D_n$  is MPEG-4 standard distortion metric. From the plot it is apparent that DMALP produced superior rate-distortion (RD) results than the DMCLP. This is due to the ALP of the control parameter. DMALP also produced better RD results compared with TB/DB and SAD. ADMSC always produces the best rate-distortion result because it always calculates the actual distortion. However, when the computational time complexity is the concern, as shown in Figure 4, ADMSC is outperformed by both DMCLP and DMALP, which warrants the theoretical fact that ADMSC mandates  $O(N_B^2)$  time while both DMALP and DMCLP take only  $O(N_B)$  time. The slightly higher time requirement in DMALP compared with DMCLP is due to the additional computation required for reparameterisation of the weight coefficients. In terms of both bit-rate and computational time, DMALP outperforms TB and SAD.

Table I summarises the numerical results produced by the different distortion measurement techniques couple with polygon-based framework upon various test sequences for  $T_{\max} = 2$ ,  $T_{\min} = 1$  pixels. It shows that the RD performance of DMALP is consistently better than DMCLP. These results also exhibit that DMALP,



TABLE I

AVERAGE BIT-RATE (BITS PER FRAME) REQUIREMENTS (WITH THE OBTAINED DISTORTION IN PARENTHESIS WHENEVER IT IS DIFFERENT FROM THE ADMISSIBLE PEAK) FOR THE VARIOUS TEST SEQUENCES WITH  $T_{\max} = 2$ ,  $T_{\min} = 1$  PIXELS USING VARIOUS COMBINATIONS OF POLYGON-BASED ALGORITHMS.

| Algorithms→      | SAD       | TB        | ADMSC    | DMCLP    | DMALP    |
|------------------|-----------|-----------|----------|----------|----------|
| Video sequence↓  | Bit-rate  | Bit-rate  | Bit-rate | Bit-rate | Bit-rate |
| MissAmerica.qcif | 343 (3.0) | 338 (3.0) | 348      | 355      | 350      |
| Akiyo.qcif       | 312 (2.8) | 310 (3.0) | 313      | 320      | 314      |
| Bream.qcif       | 421 (3.0) | 415 (3.0) | 421      | 430      | 422      |
| Kids.sif         | 1592      | 1593      | 1593     | 1600     | 1595     |
| Stefan.sif       | 580       | 582       | 585      | 589      | 587      |
| Kids.sdtv        | 4500      | 4499      | 4507     | 4512     | 4510     |
| Stefan.sdtv      | 1080      | 1080      | 1085     | 1092     | 1090     |

ADMSC and DMCLP always maintain the AD which is not the case for either SAD or TB.

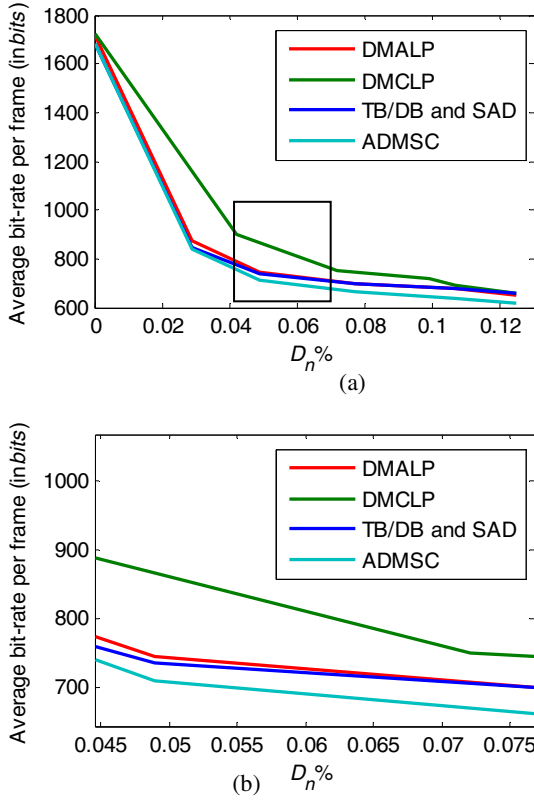


Figure 3: Performance comparison – (a) rate-distortion and (b) zoom-in.

A cursory look may reveal that SAD and TB required fewer bits than the others. But it is not due to the efficacy of SAD or TB, it is rather due to the fact that they fail to maintain the AD (see Sections I and II, and [5] for details) and thereby saving a few bits at the cost of erroneous and

large distortion. Since the computational complexity of the distortion measurement process has a direct impact on that of the shape coding framework, a further series of experiments was conducted to compare the computational time requirements incurred by algorithm combinations for various AD pairs. Table II summarises the total *central processing unit* (CPU) time for various algorithm implementations.

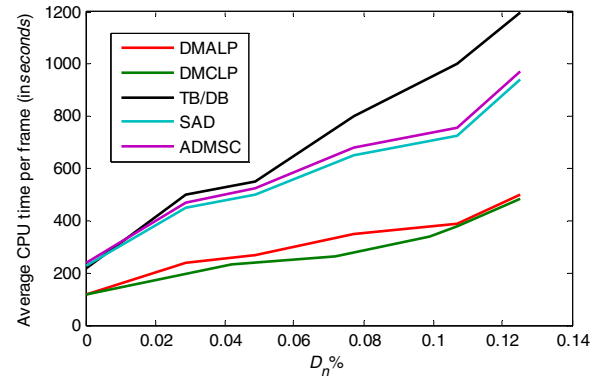


Figure 4: Time-distortion performance comparison.

From the results in Table II, it is readily apparent that those algorithms using DMALP as their distortion measure were computationally faster compared with the respective SAD, TB and ADMSC counterparts and comparable to DMCLP, to underscore the theoretical time complexity analysis presented in Sections I and II. For instance, with  $T_{\max} = 3$ ,  $T_{\min} = 1$  pixels, *B-spline-SAD*, *B-spline-TB*, *B-spline-ADMSC*, *B-spline-DMCLP* and *B-spline-DMALP* respectively required 565 , 620.3 , 578.8, 312.5 and

313 seconds, so vindicating that while SAD, TB and ADMSC cost  $O(N_B^2)$  for BS-based encoding, the

TABLE II

COMPUTATIONAL TIME (SECONDS) REQUIRED FOR THE NECK REGION OF THE 31TH FRAME OF THE MISSAMERICA.QCIF BY DIFFERENT ORD OPTIMAL SHAPE CODING ALGORITHMS FOR VARIOUS AD PAIRS ( $T_{\max}$ ,  $T_{\min}$  IN PIXELS) (DISTORTION VALUES PRODUCED USING SAD/TB ALSO SHOWN IN PARENTHESES).

| Admissible distortion → | $T_{\max} = 1$ ,<br>$T_{\min} = 1$ | $T_{\max} = 2$ ,<br>$T_{\min} = 1$ | $T_{\max} = 2$ ,<br>$T_{\min} = 2$ | $T_{\max} = 3$ ,<br>$T_{\min} = 1$ | $T_{\max} = 3$ ,<br>$T_{\min} = 2$ |
|-------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Algorithms ↓            | Time                               | Time                               | Time                               | Time                               | Time                               |
| <i>Polygon-SAD</i>      | 1.59 (1.42)                        | 1.80 (2.24)                        | 1.90 (2.23)                        | 2.0                                | 2.0                                |
| <i>Polygon-TB</i>       | 4.26 (2.24)                        | 6.03(2.24)                         | 7.73 (4.0)                         | 11.35 (5.0)                        | 12.66 (5.0)                        |
| <i>Polygon-ADMSC</i>    | 1.63                               | 1.89                               | 2.01                               | 2.15                               | 2.25                               |
| <i>Polygon-DMCLP</i>    | 1.61                               | 1.83                               | 1.92                               | 1.97                               | 2.02                               |
| <i>Polygon-DMALP</i>    | 1.62                               | 1.83                               | 1.93                               | 1.98                               | 2.03                               |
| <i>B-spline-SAD</i>     | 120 (2.0)                          | 550 (2.45)                         | 560 (5.65)                         | 565 (7.0)                          | 570 (8.0)                          |
| <i>B-spline-TB</i>      | 90.60 (2.0)                        | 510.50 (3.6)                       | 545.50 (2.8)                       | 620.30 (6.0)                       | 680.40 (6.0)                       |
| <i>B-spline-ADMSC</i>   | 554.20                             | 575.00                             | 582.10                             | 587.80                             | 591.60                             |
| <i>B-spline-DMCLP</i>   | 270.20                             | 290.30                             | 297.00                             | 312.50                             | 314.30                             |
| <i>B-spline-DMALP</i>   | 271.05                             | 291.20                             | 297.50                             | 313.00                             | 315.20                             |

overhead for DMALP is only  $O(N_B)$ , which is the same as DMCLP. The slightly higher time requirement for DMALP compared to DMCLP is due to the reason of reparameterisation discussed in Section III. For polygon-based encoding, DMALP algorithms also required less time than their ADMSC based counterparts, while as expected TB algorithms consistently incurred a significantly greater time overhead because SAD, DMALP and ADMSC all are of order  $O(N_B)$  compared with  $O(N_B^2)$  for TB.

## V. CONCLUSION

This paper has presented a novel *distortion measurement* technique based on *arc-length-parameterisation* (DMALP) which can be seamlessly embedded into the classical vertex-based rate-distortion optimal shape coding framework to improve the overall RD performance without increasing the computational time complexity order. It has been proven that DMALP always maintains the admissible distortion. Experimental results have shown that among the existing distortion measurement techniques DMALP is the best next to DMCLP in terms of computational time requirement and has outperformed DMCLP in terms of RD with a much larger margin while it is comparable with the others.

## ACKNOWLEDGEMENT

The authors acknowledge this work has been supported by an ARC discovery project (DP0771294) and a University of Western Australia Postdoctoral Fellowship.

## REFERENCES

- [1] A. K. Katsaggelos, L. P. Kondi, F. W. Meier, J. Ostermann, and G. M. Schuster, "MPEG-4 and rate-distortion-based shape-coding techniques," *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1126-1154, June 1998.
- [2] L. P. Kondi, G. Melnikov, and A. K. Katsaggelos, "Joint optimal object shape estimation and encoding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 4, pp. 528-533, 2004.
- [3] G. M. Schuster and A. K. Katsaggelos, *Rate-distortion based video compression: optimal video frame compression and object boundary encoding*. Kluwer Academic Publishers, 1997.
- [4] F. A. Sohel, L. S. Dooley, and G. C. Karmakar, "New dynamic enhancements to the vertex-based rate-distortion optimal shape coding framework," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 10, pp. 1408-1413, 2007.
- [5] F. A. Sohel, L. S. Dooley, and G. C. Karmakar, "Accurate distortion measurement for generic shape coding," *Pattern Recognition Letters*, vol. 27, no. 2, pp. 133-142, 2006.
- [6] F. A. Sohel, G. C. Karmakar, and L. S. Dooley, "Fast distortion measurement using chord-length parameterisation within the vertex-based rate-distortion optimal shape coding framework," *IEEE Signal Processing Letters*, vol. 14, no. 2, pp. 121-124, 2007.
- [7] R. T. Farouki, "Optimal parameterizations," *Computer Aided Geometric Design*, vol. 14, pp. 153-168, 1997.

# Time-Frequency Representation of Audio Signals Using Hilbert Spectrum with Effective Frequency Scaling

Md. Khademul Islam Molla<sup>2</sup>, Mostafa Al Masum Shaikh<sup>1</sup> and Keikichi Hirose<sup>1</sup>

<sup>1</sup>Dept. of Information and Communication Engineering, The University of Tokyo, Tokyo, Japan

<sup>2</sup>Dept. of Computer Science and Engineering, The University of Rajshahi, Bangladesh

Email: {molla, almasum, hirose}@gavo.t.u-tokyo.ac.jp

**Abstract** – The efficiency of Hilbert spectrum (HS) in time-frequency representation (TFR) of audio signals is investigated in this paper. HS is derived by applying empirical mode decomposition (EMD), a newly developed data adaptive method for nonlinear and non-stationary signal analysis together with Hilbert transform. EMD represents any time domain signal as a sum of a finite number of bases called intrinsic mode functions (IMFs). The instantaneous frequency responses of the IMFs derived through Hilbert transform are arranged to obtain the TFR of the analyzing signal yielding the HS. A new frequency scaling method is introduced here for proper interpretation of the energy spectra in HS. The performance of HS is compared with well known and widely used short-time Fourier transform (STFT) technique for TFR. The experimental results show that HS based method performs better than STFT in time-frequency representation of the audio signals

**Index Terms** — Empirical mode decomposition (EMD), Hilbert-Huang transformation (HHT), instantaneous frequency, time-frequency representation (TFR).

## I. INTRODUCTION

The time-frequency distribution of a signal provides information about how the spectral content of the signal evolves with time, thus providing an ideal tool to dissect, analyze and interpret non-stationary signals. The need for a time-frequency representation (TFR) is stemmed from the inadequacy of either time domain or frequency domain analysis to fully describe the nature of non-stationary signals. The TFR of audio signal has many potential applications including blind signal separation, speech recognition, signal characterization and speech enhancement. A variety of methods for obtaining the energy density as a function of time and frequency simultaneously have been devised, most notably the short time Fourier transform (STFT), the wavelet transform (WT) and the Wigner-Ville distribution (WVD). The TFR is also used to clarify, and provide additional information about the behaviour of the non-stationary signals.

The short-time Fourier transform (STFT) localizes the signal by modulating it with a window function, before performing the Fourier transform to obtain the frequency content of the signal in the region of the window. Choosing the window to have roughly the same shape as the analyzing signals improves the performance although

these benefits are usually slight however [1]. Another popular method of TFR is the wavelet transform (WT) which expand the time domain signal in terms of wavelet functions which are localized in both time and frequency [2]. The wavelet function makes use of a dilation factor and a translation factor. It is a critical decision to select the proper wavelet basis function for decomposition based of the characteristics of the analyzing signals. For specific applications, the basic wavelet function can be modified according to the special need, but the form has to be given before the analysis [3]. Although, wavelet transform is widely used in TFR, it does not carry the exact phase information. The proper reconstruction of the signal after some further processing in time-frequency space is rather difficult without the phase information.

In this paper, the performance of Hilbert spectrum (HS) is explored as a technique for TFR of audio signals. Any signal is represented as a sum of bases called intrinsic mode function (IMF) obtained by applying EMD, a data adaptive method to analyze non-stationary and nonlinear signals [4]. The instantaneous frequency and amplitude of each IMF is derived by Hilbert transformation. Then the instantaneous responses of the IMFs are arranged to construct the HS i.e. time-frequency space corresponding to the time domain signals. A new frequency scaling method is introduced in deriving the HS using the instantaneous frequency vector. The basis functions of the EMD are directly derived from data. Each basis function has frequency response as a function of time preserving the non-stationarity of the original signal. The STFT assumes any signal as piecewise stationary and uses suitable window function to produce the short time spectral characteristics of the signal. Whereas, most of the real world signals including speech, audio, radar signals are non-stationary in nature. The Fourier transform (FT) based technique treat the signal as a sum of predefined harmonics. It performs better when the analyzing signal consists of harmonics and poorly performed otherwise. It is noted that the voiced speech signal contains some harmonics. When the application domain becomes broad (e.g. audio signals other than voiced speech) and the signal characteristics are unknown, a robust and adaptive TF representation method is the candidate for efficient analysis. The TFR using HS is based on EMD method is fully data adaptive and the signal decomposition is

performed without any pre-consideration of being harmonic or stationary. The segmental SNR and spectral distortion are used to measure the performance of TF representation. The performance of HS is compared with STFT and the experimental results show the superiority of Hilbert spectrum.

## II. EMD AND HILBERT SPECTRUM

Empirical mode decomposition (EMD) represents an time-domain signal into a finite set of AM-FM oscillator components which are the bases of the decomposition. The key benefit of using EMD is that it is an automatic decomposition and fully data adaptive. The principle of the EMD technique is to decompose a signal  $x(n)$  into a sum of band-limited functions,  $g_b(n)$ , called intrinsic mode functions (IMFs). Each IMF must satisfy two properties: (i) the number of extrema and the number of zero crossings are either equal or differ by one; (ii) the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. A special “sifting” process is employed to extract all of the IMFs [4, 5]. This sifting process is described as follows.

Firstly, the upper- and lower-envelopes of the signals  $x(n)$ , as well as their mean value  $\mu_1(n)$ , are calculated respectively. The first step of the sifting process is to calculate the difference  $h_1(n) = x(n) - \mu_1(n)$ . However,  $h_1(n)$  rarely satisfies the two IMF properties and is not taken as the first IMF of the signals straightway. Therefore, the sifting usually has to be implemented for more times, where the “difference” obtained in the previous sifting is taken as “signals” in present sifting. If after  $(k+1)^{\text{th}}$  sifting, corresponding difference  $h_k(n) = h_{(k-1)}(n) - \mu_k(n)$  satisfies the IMF properties, then it can be taken as the first IMF component, denoted by  $g_1(n)$ , that is,  $g_1(n) = h_k(n)$ . In practice, to determine whether or not  $h_k(n)$  well satisfies the IMF properties, we usually use so-called standard deviation ( $\delta$ ) criterion, that is, to check if the following inequality holds [4]:

$$\delta_k = \sum_{n=0}^{N-1} \left[ \frac{|h_{(k-1)}(n) - h_k(n)|^2}{h_{(k-1)}^2(n)} \right] \leq \sigma, \quad (1)$$

where  $N$  is the frame length and  $0.2 \leq \sigma \leq 0.3$ . Next, taking rest data  $r_1(n) = x(n) - g_1(n)$  as “new” signals and implementing the sifting process on it, we can obtain the second IMF  $g_2(n)$ . This procedure should be repeatedly used for  $B$  (total number of IMF components) times until the last residue  $r_B(n)$  becomes a monotonic function. When the decomposition procedure finished, the signals then can be expressed as

$$x(n) = \sum_{b=1}^B g_b(n) + r_B(n), \quad (2)$$

where  $g_1(n), g_2(n), \dots, g_B(n)$  are all of the IMFs included in the signals and  $r_B(n)$  is a negligible residue.

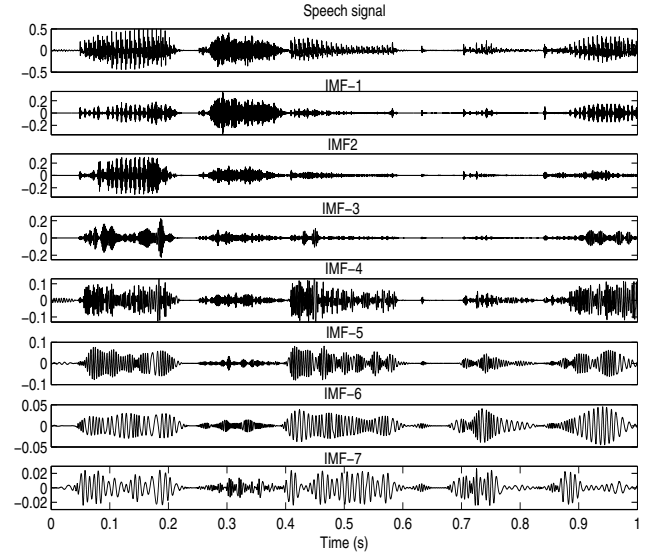


Fig 1. The illustration of EMD. A noisy speech signal at 10 dB SNR and its first 7 IMFs out of 15.

A noisy speech signal and some selected IMF components are shown in Fig 1. It can be observed that higher order IMFs contain lower frequency oscillations than that of lower order IMFs. This is reasonable, since sifting process is based on the idea of subtracting the component with the longest period from the data till an IMF is obtained. Therefore the first IMF will have the highest oscillating components; the components with the highest frequencies.

### A. Instantaneous Frequency

Instantaneous frequency (IF) represents signal's frequency at any time instance and it is defined as the rate of change of the phase angle at the instant of the “analytic” version of the signal. Every IMF is a real valued signal. The discrete Hilbert transform (HT) is used to compute the analytic signal for an IMF. The analytic version of the  $b^{\text{th}}$  IMF  $g_b(t)$  is defined as:

$$z_b(t) = g_b(t) + jH_d[g_b(t)] = a_b(t)e^{j\theta_b(t)} \quad (3)$$

where  $a_b(t)$  and  $\theta_b(t)$  are instantaneous amplitude and phase respectively of the  $b^{\text{th}}$  IMF and  $H_d[\cdot]$  denotes the discrete Hilbert transform.

The analytic signal is advantageous in determining the instantaneous quantities such as energy, phase and frequency. The discrete-time IF of  $b^{\text{th}}$  IMF is then given as the derivative of the phase  $\theta_b(t)$ -calculated at  $t$  i.e.

$$f_b(t) = \frac{d\tilde{\theta}_b(t)}{dt} \quad (4)$$

where  $\tilde{\theta}_b(t)$  represents the unwrapped version of instantaneous phase  $\theta_b(t)$ . The derivative in Eq. (4), is evaluated at discrete instant of  $t$ . It should be noted that such derivative introduces the abrupt fluctuations of IF and hence nonlinear smoothing is required. Here, the moving average smoothing filtering is used to remove such fluctuations. The filtering scheme improves the effectiveness of computing IF using discrete derivative. The instantaneous frequencies of the IMFs shown in Fig. 1 are illustrated in Fig 2.

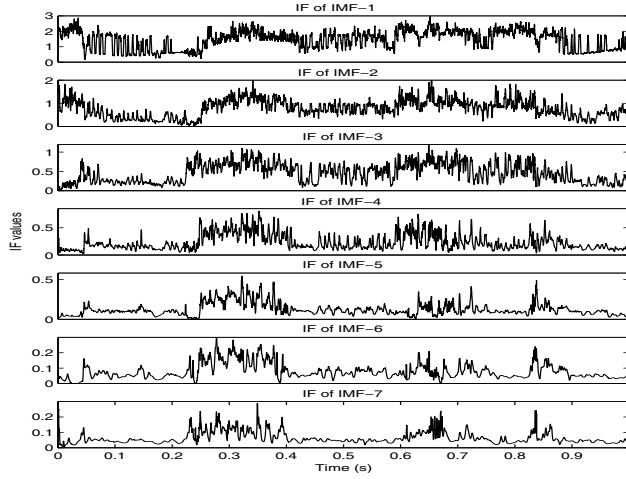


Fig 2. Instantaneous frequencies of the seven IMF components shown in Fig 1.

### B. Frequency Scaling

The EMD is based on the simple assumption that any data consists of different simple intrinsic modes of oscillations. Each intrinsic mode, linear or nonlinear, represents a simple oscillatory mode as a counterpart to the simple harmonic function, but it is much more general. Instead of constant amplitude and frequency, as in a simple harmonic component, the IMF can have a variable amplitude and frequency as functions of time. The derived instantaneous frequencies are also a function of time. The method of deriving the IMF components is fully data adaptive; it does not employ any mathematical formulation. As EMD based IF does not have any specific scaling function. For any further processing e.g. to construct the Hilbert spectrum based time-frequency representation, it is required to normalize the IF vector by multiplying the elements of the vector with  $1/IF_{max}$  where  $IF_{max} = \text{Max}(f_1, f_2, \dots, f_b, \dots, f_B)$ .

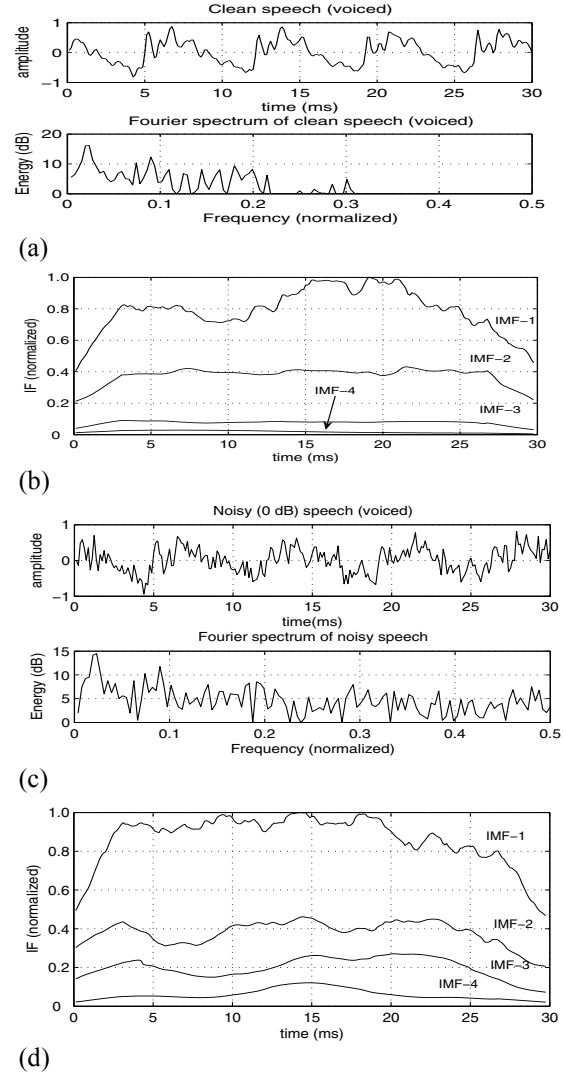


Fig 3. Clean and noisy (0 dB) voiced speech signals and its frequency responses; (a) clean speech and its Fourier spectra, (b) IF response of the first four IMFs of clean speech, (c) noisy speech and its Fourier spectra, and (d) IF response of the first four IMFs of noisy speech.

The normalized IF vector does not make any noticeable difference for the signals of different spectral characteristics. Two signals of distinguishable spectra and their IF vectors are shown in Fig. 3. Observing the IF components, it is difficult to realize the real spectral components contained by the analyzing signals. A frequency scaling method is introduced here by concatenating white Gaussian noise with the analyzing signal. Let  $\mathcal{V}(n)$  be the uniformly distributed Gaussian noise (uGn). The Fourier spectra of uGn containing noticeable energy spectrum over the entire frequency range, and the spectra of its first four IMFs are shown in Fig 4.

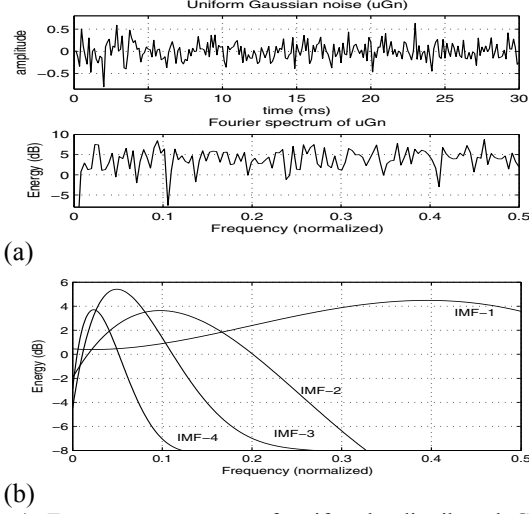


Fig 4. Frequency response of uniformly distributed Gaussian noise (uGn) and its IMFs; (a) Fourier spectra of uGn, (b) frequency response of its first four IMFs.

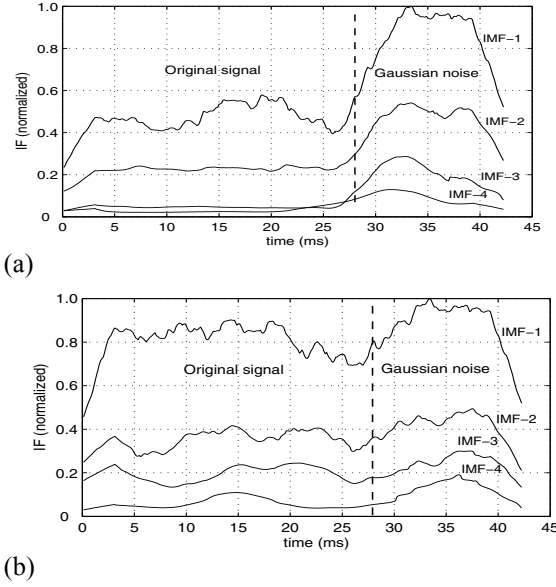


Fig 5. Normalized IF response of the modified analyzing signals (MAS), (a) IFs of the first four IMFs of MAS corresponding to clean speech of Fig. 3(a); (b) IFs of the first four IMFs of MAS corresponding to the noisy speech signal as shown in Fig. 3(c).

If  $x(n)$  is the analyzing signal, the modified analyzing signal (MAS)  $y(n)$  is designed as:

$$y(n) = \begin{cases} x(n) & ; 1 \leq n \leq N_x \\ v(n) & ; N_x < n \leq N_x + N_v \end{cases} \quad (5)$$

where  $N_x$  and  $N_v$  are the length of  $v(n)$  and  $x(n)$  respectively and in this experiment  $N_v=100$ . It is already reported in [5] that the EMD works as dyadic filter-bank for fractional Gaussian noise (fGn).

The EMD is effectively a dyadic filter [6]. The normalized IF vectors of MAS corresponding to the signal shown in Fig. 3 are illustrated in Fig. 5. The left side of the vertical dashed line corresponds the scaled IF response of the original signal and the right side represents the IF response of the uGn which is discarded during any further processing of IF information. It is observed that the IF vector analyzing signal  $x(n)$  is properly scaled by the IF information of the concatenated Gn. The EMD technique does not include any mathematical formulation of IF scaling, and hence the proposed scaling method is useful for any further of the obtained IF information.

### C. Hilbert Spectrum

Hilbert spectrum represents the distribution of the signal energy as a function of time and frequency. It is also designated as Hilbert amplitude spectrum  $H(k,t)$  or simply Hilbert spectrum(HS). This process first normalizes the IF values. The bin spacing of the HS is  $0.5/F_c$ , where  $F_c$  is the number of desired frequency channels. The overall HS is expressed as the superposition of the individual IMFs' HSs defined as:  $H(k,t) = \sum_{b=1}^B H^{(b)}(k,t)$  for  $b=1, 2, \dots, B$  [7]. Hence, each element  $H(k,t)$  of the overall HS is defined as the weighted sum of the instantaneous amplitudes of all the IMFs at  $k^{\text{th}}$  frequency bin.

$$H(k,t) = \sum_{b=1}^B a_b(t) w_b^{(k)}(t) \quad (6)$$

where the weight factor  $w_b^{(k)}(t)$  takes 1 if  $f_b(t)$  falls within  $k^{\text{th}}$  band, otherwise is 0.

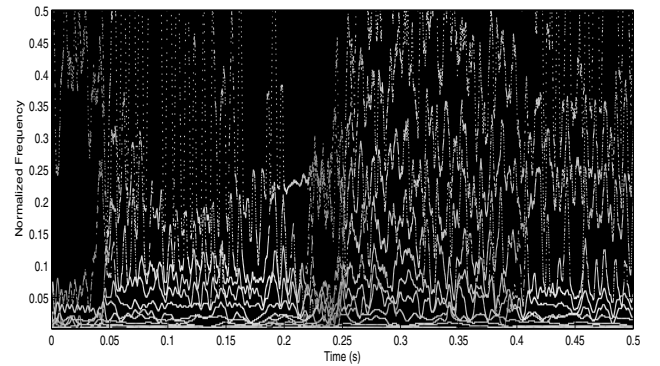


Fig 6. Hilbert spectrum (HS) of speech signals.

After computing the elements over the frequency bins,  $H$  represents the instantaneous signal spectrum in time-frequency space as a 2D table.

It is noted that the time resolution of  $H$  is equal to the sampling rate and the frequency resolution can be chosen up to Nyquist limit. Fig. 6 represents the Hilbert spectrum (in dB) of the voiced speech signal shown in Fig. 1 using 128 frequency bins.

### C. Performance of Inverse Transformation

In many speech processing applications including audio source separation, it is often required to reconstruct the time domain signal after performing the processing in time-frequency domain. The perfect reconstruction of the time domain signal is possible if the time-frequency transformation technique supports proper inverse transformation with minimum required parameters. The experiments demonstrate that EMD as well as Hilbert spectrum is a reversible representation. One can also reconstruct the time domain signal from the Hilbert spectrum. During the Hilbert transform the real part of the signal remains unchanged [8]. The time domain signal  $s_r(t)$  is easily be reconstructed by filtering out the imaginary part from the HS and summing over frequency bins as:

$$s_r(t) = \sum_k H(k, t) \cdot \cos[\phi(k, t)] + r_B(t) \quad (7)$$

where  $\phi(k, t)$  is the phase matrix of Hilbert spectrum and  $r_B(t)$  is the final residue of EMD. The phase matrix is saved during the construction of Hilbert spectrum to be used in re-synthesis. The signal  $s_r(t)$  reconstructed from the Hilbert spectrum is shown in Fig. 7. The reconstruction error (of the order  $1 \times 10^{-15}$ ) is considered as negligible in reality.

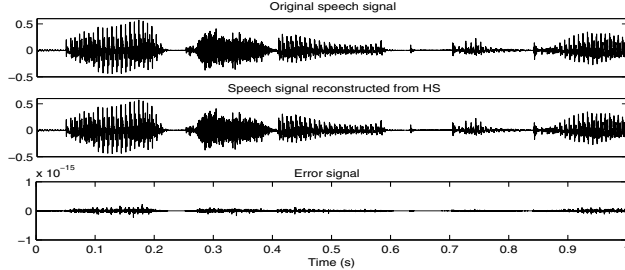


Fig 7. The speech signal (top), the reconstructed signal from HS (middle), the error between the original and reconstructed signal (lower one).

On the other hand, it is difficult to reconstruct the original time domain signal from the STFT representation, especially when the window function is used in an overlapping fashion. The inverse transformation becomes more troublesome when the tapered window is used instead of square one. It is also required to know the parameters (e.g. number of FFT points, window function and its length, amount of overlap) what are used in STFT for proper inverse transformation. Whereas, no parameter is required to get back the time domain signal from the corresponding HS.

The segmental SNR (SSNR) is a good measure of distortion in time domain between original and

reconstructed signals [9]. The SSNR of the  $i^{\text{th}}$  frame is defined as

$$SSNR(i) = 10 \log_{10} \left[ \frac{\sum_{t=0}^{T-1} |s^{(i)}(t)|^2}{\sum_{t=0}^{T-1} |s^{(i)}(t) - \hat{s}^{(i)}(t)|^2} \right], \quad (8)$$

where  $s^{(i)}(t)$  and  $\hat{s}^{(i)}(t)$  are the original and reconstructed signals respectively of the  $i^{\text{th}}$  frame. Then the average SSNR (avgSSNR) is expressed as

$$\text{avgSSNR} = \frac{1}{I} \sum_{i=1}^I SSNR(i), \quad (9)$$

where  $I$  is the total number of speech frames. The measure of SSNR is a direct time domain method to illustrate the difference between two signals. This error measure is quite well understood and mathematically tractable. Its main characteristic is that it puts great emphasis on large errors and little emphasis on small errors.

The measure of error between two signals can also be performed in the frequency domain and can be interpreted as a goodness of fit between a given signal spectrum and a model spectrum that approximates it. For determining the quality of the reconstructed power spectrum envelop, its segmental spectral distortion (SSD) is calculated over the power spectrum over the frequency plane as an objective measure. The SSD for the  $i^{\text{th}}$  frame is defined as in [10],

$$SSD(i) = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} \left( 10 \log_{10} \frac{S^{(i)}(k)}{\hat{S}^{(i)}(k)} \right)^2}, \quad (10)$$

where,  $S^{(i)}(k)$  and  $\hat{S}^{(i)}(k)$  are the original and reconstructed power spectrum for the  $i^{\text{th}}$  speech frame. Then the average SSD (avgSSD) is expressed as:

$$\text{avgSSD} = \frac{1}{I} \sum_{i=1}^I SSD(i), \quad (11)$$

where  $I$  is the total number of speech frames. The reduction of SSD indicates the spectral envelop estimate's level of accuracy. The minimization of the spectral distortion, we expect a better fit of  $\hat{S}^{(i)}(k)$  to  $S^{(i)}(k)$ .

### III. EXPERIMENTAL RESULTS AND DISCUSSION

To explore the performance of Hilbert spectrum (HS) in time-frequency representation of audio signals in terms of average segmental SNR average segmental spectral distortion, the experiments are conducted on the speech (male and female speech) signals from TIMIT database. All the speech signals are with 16 kHz sampling rate and 16 bits amplitude resolution. Each speech signal is transformed to time-frequency domain using HS and

STFT based methods. The simulation results are compared with the STFT based time-frequency representation.

In many applications of TF representation including audio source separation, it is required to reconstruct the time domain signal even after performing some processing in TF domain. The reconstruction from spectrogram outputs some spurious error comparing with the original signal. The signals reconstructed from HS are with very high SNR (greater than 300dB) which is treated as noise less. In the context of reconstruction error, the HS can be proposed as the reversible technique between time and TF domain. The EMD parameters i.e. number of IMFs, and the reconstruction error in terms of average SSNR (avgSSNR) and average SSD (avgSSD) with HS and STFT based TFR of some typical audio signals are presented in Table 1. It is observed that the HS offers better accuracy of inverse transformation to get back the time domain signals. It is noted that same parameters are used for STFT and its inverse transformation.

Table 1: Performance comparison between HS and STFT

| Audio signals | IMF | avgSSNR (dB) |       | avgSSD (dB) |       |
|---------------|-----|--------------|-------|-------------|-------|
|               |     | HS           | STFT  | HS          | STFT  |
| MS_1(25s)     | 16  | 315.50       | 30.92 | 1.01e-13    | 1.853 |
| MS_2 (28s)    | 17  | 314.82       | 29.84 | 2.02e-13    | 1.213 |
| FS_1 (30s)    | 16  | 325.41       | 31.32 | 1.52e-13    | 0.945 |
| FS_2 (25s)    | 15  | 330.94       | 30.71 | 1.71e-13    | 2.014 |
| Jazz (40s)    | 19  | 311.08       | 28.89 | 2.13e-13    | 2.776 |
| Flute (20s)   | 18  | 318.63       | 31.47 | 1.03e-13    | 2.012 |

The use of STFT is faced with the problem of resolution in both time and frequency axes at a time. Some researches argued that TF domain based audio signal separation technique has better separation efficiency with higher resolution in TF representation [11]. Narrow windows give good time resolution, but poor frequency resolution. Wide windows give good frequency resolution, but poor time resolution; furthermore, wide windows may violate the condition of stationarity. All the parameters required in STFT are signal dependent. The most difficulty is occurred when the signal characteristics are fully unknown. Whereas, in HS only one parameter, the number of desired frequency bins is involved. Moreover, STFT includes additional cross-spectral energy due to window overlapping. The use of HS is capable to overcome all of the mentioned limitations of the STFT based time-frequency representation.

#### IV. CONCLUSIONS

We have presented the effectiveness of Hilbert spectrum (HS) in time-frequency representation of audio

signals. The experimental results provide the evidence that Hilbert spectrum based TFR performs better. The specialty of Hilbert spectrum is that the time resolution can be as precise as the sampling period and the frequency resolution depends on the choice up to Nyquist limit. Hence it serves the effective resolution of TFR yielding better performance in different signal processing applications. Another major problem of STFT is that it includes a noticeable amount of cross-spectral energy terms during the use of window function with a overlapping fashion, whereas, HS does not make use of window function. Hence, HS performs better than STFT in time-frequency representation of audios signals.

#### REFERENCES

- [1] L. A. Ronald and W. M. Duncan, "Signal Analysis: Time, Frequency, Scale and Structure", *Wiley-Interscience Publisher*, 2004.
- [2] A. Mertins, "Signal Analysis: Wavelets, Filter Banks, Time-Frequency Transform and Applications", *Wiley Publisher*, 1999.
- [3] W. C. Lang and K. Forinash, "Time-frequency analysis with the continuous wavelet transform", *American Journal of Physics*, Vol. 66, No. 9, pp: 794-797, Sep, 1998.
- [4] N. E. Huang et al., "The Empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proc. Roy. Soc. Lond. A*, Vol. 454, pp: 903-995, 1998.
- [5] P. Flandrin, G. Rilling, P. Goncalves, "Emperical Mode Decomposition as a filter bank" *IEEE Sig. Proc. Letter*, Vol. 11, No. 2, pp: 112-114, Feb 2004.
- [6] B. Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition", *Proc. R. Soc. Lond. A*, 460, pp: 1597-1611, 2004.
- [7] M. K. I Molla, K. Hirose and N. Minematsu, "Separation of mixed audio signals by decomposing Hilbert spectrum with modified EMD", *IEICE Trans. Fundamentals*, Vol. E89-A, No. 3, pp:727-734, 2006
- [8] M. Cooke, "Modeling Auditory processing and Organisation", *Cambridge University Press*, 1993.
- [9] M. K. I Molla and K. Hirose, "Multiband linear prediction of speech signals with adaptive order using empirical mode decomposition", *Journal of Signal Processing*, Vol. 11, No. 6, pp: 471-479, Nov., 2007.
- [10] M. A. Ramirez and M. Minami 'Split-order linear prediction for segmentation and harmonic spectral modeling", *IEEE Signal Processing Letters*, Vol. 13, No. 4, pp. 244-247, April, 2006.
- [11] M. K. I. Molla, K. Hirose and N. Minematsu, "Audio source separation from their mixture using empirical mode decomposition with independent subspace analysis", *Proc. of ICSLP'04*, October, 2004.



# Server Consolidation Using OpenVZ: Performance Evaluation

Mohuiddin Ahmed, Showayb Zahda, Majed Abbas  
Computer Science Department  
International Islamic University Malaysia

**Abstract**—Virtualization is abstracting computer resources. It is the ability to run multiple operating systems simultaneously on a single physical machine. In this paper we evaluate the performance of the operating system level virtualization software (OpenVZ). The evaluation will be based on Web hosting, in other words, we evaluate the performance of Web servers on two different machines (dubbed as Baru and Lama) and we observe the behavior of OpenVZ and its scalability for Web servers on different architectures (32 bits and 64 bits). The study evaluates the CPU, Memory and Throughputs under different loading conditions.

**Index Terms** — OpenVZ, performance, server consolidation, virtualization.

## I. INTRODUCTION

“Server consolidation is an approach to the efficient usage of computer server resources in order to reduce the total number of servers or server locations that an organization requires”[1]. Simply put, server consolidation is the combination of several servers (i.e. Web Server, Mail Server, Java Server, etc) into one physical machine, in order to, save cost and utilize the computer resources.

One way to achieve server consolidation is to use virtualization. Virtualization as a word does not exist in most English dictionaries. However, the closest word to it in Oxford Concise Dictionary is “virtual” which means “Computing: not physically existing as such but made by software to appear to do so” [2]. However, virtual is not a new term in the computer jargons, it was used before such as virtual memory, virtual reality, etc.

Virtualization in our context according to [3] is defined as “the abstraction of computer resources” and “abstraction is the process of separating hardware functionality from the underlying hardware.” In other words, virtualization is sharing, and partitioning the resources of computer hardware amongst more than one running operating systems on a single physical machine simultaneously. Each one of the running operating systems is called a Virtual Machine (VM). Virtualization offers the ability to run multiple operating systems (Linux, Windows, MacOS X) on one physical machine at the same time.

Currently, virtualization is grabbing the attention of the data centers as well as the small and medium enterprises due to its enormous advantages. This led processor manufacturer mainly Intel and AMD to develop processors to support virtualization i.e. Intel VT and AMD-V respectively.

Indeed, virtualization brings a huge economical value to the organizations. Mainly, it reduces the Total Cost of Ownership (TCO) by consolidating a bunch of physical servers into one physical server. For example, 10 Web servers can use one physical machine instead of using 10 physical machines. This reduces a lot of cost -beside the cost of the machines- such as electricity bills, space, air conditioning, etc. In addition, organizations can couple virtualization with open source software solutions (e.g. OpenVZ, Xen) to achieve more reduction in TCO and not to bother anymore about licensing issues. This is handy when multiple virtual machines or virtual machine backups are kept online for redundancy, restoration, or disaster recovery purposes [4].

Virtualization technology can be implemented in several techniques namely, hardware emulation, full virtualization, paravirtualization, operating system level virtualization, application virtualization, and desktop virtualization. However, the discussion about the differences amongst them is beyond this paper. But, a brief discussion about operating system level virtualization will be provided in OpenVZ section.

Besides cost savings, virtualization in general has great advantages for the organizations, due to the fact that the virtualization layer isolates and separates the VMs. None of the VMs is aware about the existence of other VMs on the same machine. They seem to each other as if they were running on different physical machines. The isolation provides higher availability and security for the organization. If one of the VMs crashes or goes down, it has no effects on other VMs. Moreover, system administrators can fire up a VM in less than a minute. As for security, running different services on different VMs improves the security. For instance, if an intruder manages to hack one server, only that particular server goes down and others will remain running as if nothing happened.

Live migration of VMs is yet another great advantage of virtualization. Assume one of the VMs is exhausting the hardware resources, then the administrator can migrate the server into another physical machine in less than one minute. Or if the physical machine should go down for upgrade or maintenance, the VMs can be migrated to another machine. Therefore, migration boosts the availability of the running servers.

Software developers and testers use virtualization to run several platforms to develop and test the products on (i.e. Linux, Windows, MacOS X). It reduces the time of installing the platforms as well as increasing the number of platforms

---

[1] The corresponding author.

that can be used for testing and development.

Virtualization is indeed a very helpful technology for education purposes. Now, universities and educational institutions no longer need to buy extra machines to cater their student's need of running several operating systems. Moreover, each student can have a root access which allows them to have full authority over the system.

Virtualization has disadvantages too; a big damage to all the running servers might occur if the power of the host machine goes down. Or, if the administrator is upgrading the system hardware, he/she has to shutdown the machine. However, by using migration, the system administrators can easily migrate the VMs to another physical machine in less than one second. The rest of the paper is organized as follows: in section II OpenVZ is briefly described, in section III we describe the methodology of the performance evaluation, section IV describes the test bed and its components, in section V we analyze the test results followed by concluding remarks.

## II. OPENVZ

OpenVZ is a free and open source virtualization solution. It is the basis of the commercial virtualization solution Virtuozzo from Parallel (formerly known as SWsoft). OpenVZ implements the operating system level virtualization technique which virtualizes servers on the operating system (kernel) layer [5]. In other words, OS-level Virtualization adds a new layer on top of the operating system. This layer manages all the issues related to OpenVZ containers and resource management.

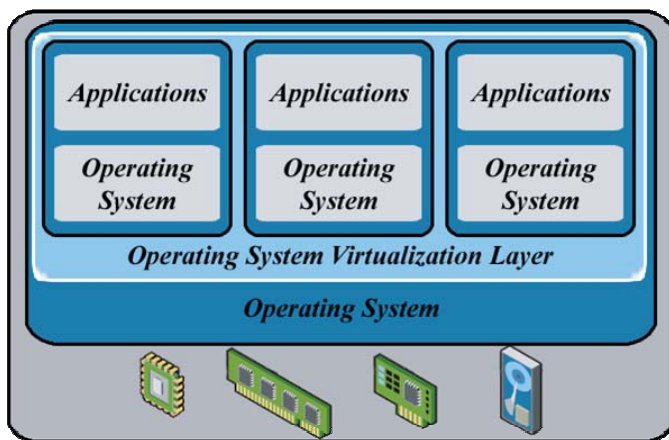


Fig. 1. Operating System level Virtualization [5]

OpenVZ has its own terminologies regarding virtualization technology. The physical machine that is running OpenVZ is called a Host System or Hardware Node. A running instance of the operating system that partitions and shares the resources of the hardware node is called a container, Virtual Environment (VE), and/or Virtual Private Server (VPS) [6]. In this paper we will refer to the running instance as a container.

Operating system-level virtualization shares the kernel amongst all running containers. The kernel (OpenVZ kernel creates the virtualization layer) bears the duty of managing and partitioning the resources of the computer, and isolating

the containers from each other. However, sharing one kernel introduces a drawback. In case the kernel crashes all the containers are down.

The performance of OpenVZ is closed to the native performance of the machine. The overhead only consumes about 1-3% of the resources. This indeed makes OpenVZ a good choice for server consolidation especially for web hosting [7].

OpenVZ is based on Linux operating system and it only runs in Linux distributions. Moreover, OpenVZ can also be tested without installation using OpenVZ LiveCD<sup>1</sup>.

It takes a few minutes to setup one container using OpenVZ. This is because of the usage of Operating System Templates. An OS-template is a set of packages from some Linux distribution used to populate a VE. With OpenVZ, different distributions can co-exist on the same hardware box, so multiple OS templates are available. [6]

Furthermore, changing the configuration of any container can be done on the fly. There is no need to restart any container in order to give or take from it some resources.

OpenVZ comes with a command line control utilities that allow the administrators to create, start, stop, destroy and migrate a container. The utilities also allow him/her to set the quota, manage the resources and execute commands on the containers from outside the container itself. For instance, "vzsplit" is a command that creates a configuration file which splits the resources of the computer equally amongst a number of containers the user specifies (vzsplit -n 4).

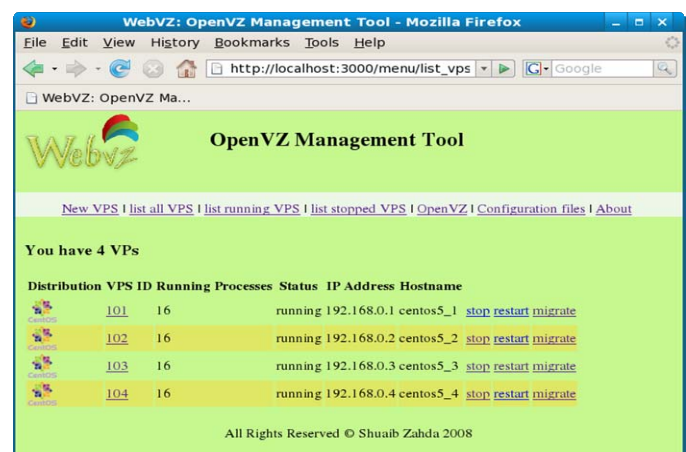


Fig. 2. WebVZ Management Tool

WebVZ is a web management tool for OpenVZ and was developed by the authors and available in [8] under GPL GNU. WebVZ allows system administrators to control the containers locally and remotely. Fig 2 shows a screenshot of WebVZ. Resource allocation among the containers is done on the basis of demand. If there is need to migrate a container to another server due to resource hogging, this can easily be attained by migration to a different machine which is lightly

<sup>1</sup> Live CD: is an operating system that runs without the need for installation on the hard disk and it does not have any effect on the existing data. Once the CD is ejected, everything goes back to its original state

loaded. Fig 3 shows the output of the command “vzlist” (list all containers)

```
[root@localhost ~]#
[root@localhost ~]# vzlist
VEID NPROC STATUS IP_ADDR HOSTNAME
101 16 running 192.168.0.1 centos5_1
102 16 running 192.168.0.2 centos5_2
103 16 running 192.168.0.3 centos5_3
104 16 running 192.168.0.4 centos5_4
105 16 running 192.168.0.5 centos5_5
106 16 running 192.168.0.6 centos5_6
107 16 running 192.168.0.7 centos5_7
108 16 running 192.168.0.8 centos5_8
109 16 running 192.168.0.9 centos5_9
110 16 running 192.168.0.10 centos5_10
111 16 running 192.168.0.11 centos5_11
112 16 running 192.168.0.12 centos5_12
113 9 running 192.168.0.13 centos5_13
114 9 running 192.168.0.14 centos5_14
115 9 running 192.168.0.15 centos5_15
116 9 running 192.168.0.16 centos5_16
117 8 running 192.168.0.17 centos5_17
118 7 running 192.168.0.18 centos5_18
119 7 running 192.168.0.19 centos5_19
120 7 running 192.168.0.20 centos5_20
[root@localhost ~]#
```

Fig 3 vzlist output

### III. METHODOLOGY

In order to test the migration between containers we injected http queries from a server and then used a wrapper to increase the number of requests until the system chokes or crashes.

In order to evaluate the performance of the containers on OpenVZ kernel, first we evaluated the performance of both machines (we named them Baru and Lama respectively) as a single unit without running OpenVZ kernel so as to know how many requests per second each machine can handle and to find out the relationships between the containers and the Host System. Second, we evaluated the containers by scaling the number of containers from 4 to 8, and finally to 12 on the two machines. All the containers receive the same amount of requests per second concurrently. In other words, each container for example is receiving 100 connections per second at the same time.

In order to measure the performance of each running Web server we used a free and open source HTTP performance measurement tool from HP called “httpperf” version 0.9.0. According to the manual page of httpperf “httpperf speaks the HTTP protocol both in its HTTP/1.0 and HTTP/1.1 flavors and offers a variety of workload generators. While running, it keeps track of a number of performance metrics that are summarized in the form of statistics that are printed at the end of a test run. The most basic operation of httpperf is to generate a fixed number of HTTP GET requests and to measure how many replies (responses) came back from the server and at what rate the responses arrived.”

However, httpperf is limited to one test at a time which generates a fixed number of HTTP GET. In order to run httpperf several times and increase the number of HTTP GETs each time, another tool called “autobench” can be used which automates web server benchmarking using httpperf. According to autobench version 2.1.1 manual page “autobench runs httpperf against the specified host or hosts, ramping up the number of requested connections, and logging the results in TSV or CSV format files.”

In order to capture the reading of the CPU and the RAM we used “top -b -d 1” to read the values every second. While the throughput of each test was reported by “httpperf” for plotting and display the result, “bench2graph” is a tool comes with autobench, was used to plot the output of the tests.

This paragraph explains how autobench runs httpperf. (E.g. *autobench --single\_host --host1 192.168.0.1 --low\_rate 50 --high\_rate 100 --rate\_step 10 --timeout 7 --file ve1.tsv*). This command will run “httpperf” 5 times. First time httpperf initiates 50 connections per second until the total number of connections reaches 5000 connections (the default) which is the total number of connections that are initiated during one test, and each connection contains 10 requests (the default). In the second test httpperf initiates 60 connections per second until the total number of connections reaches 5000 and so on, each time the number of connections per second increases according to the specified rate step value. The output of httpperf is shown in Fig 4 which is extracted and tabulated by autobench as shown in Fig 5.

```
httpperf --timeout=7 --client=0/1 --server=192.168.0.1 --port=80 --uri=/ --rate=50 --send-buffer=4096 --recv-buffer=16384 --num-conns=5000 --num-calls=10
Maximum connect burst length: 1

Total: connections 5000 requests 50000 replies 50000 test-duration 99.985 s

Connection rate: 50.0 conn/s (20.0 ms/conn, <=2 concurrent connections)
Connection time [ms]: min 4.7 avg 5.6 max 30.2 median 5.5 stddev 1.6
Connection time [ms]: connect 0.3
Connection length [replies/conn]: 10.000

Request rate: 500.1 req/s (2.0 ms/req)
Request size [B]: 64.0

Reply rate [replies/s]: min 500.0 avg 500.0 max 500.0 stddev 0.0 (19 samples)
Reply time [ms]: response 0.5 transfer 0.0
Reply size [B]: header 250.0 content 325.0 footer 0.0 (total 575.0)
Reply status: 1xxx=0 2xx=50000 3xx=0 4xx=0 5xx=0

CPU time [s]: user 5.58 system 94.37 (user 5.6% system 94.4% total 100.0%)
Net I/O: 312.1 KB/s (2.6*10^6 bps)

Errors: total 0 client-time 0 socket-time 0 connrefused 0 connreset 0
Errors: fd-unavail 0 addrunavail 0 ftab-full 0 other 0
```

Fig 4 httpperf output sample

```
dem_req_rate req_rate_192.168.0.1 con_rate_192.168.0.1 min_rep_rate_192.168.0.1
avg_rep_rate_192.168.0.1 max_rep_rate_192.168.0.1 stddev_rep_rate_192.168.0.1
resp_time_192.168.0.1 net_io_192.168.0.1 errors_192.168.0.1
500 500.1 50.0500.0 500.0 500.0 0.0 0.5 312.1 0
600 600.1 60.0600.0 600.0 600.0 0.0 0.5 374.5 0
700 700.1 70.0700.0 700.0 700.1 0.0 0.5 436.9 0
800 800.1 80.0800.0 800.0 800.1 0.0 1.0 499.3 0
900 900.1 90.0899.5 900.1 900.7 0.3 0.5 561.7 0
1000 1000.1 100.0 998.3 999.9 1000.7 0.7 0.6 624.1 0
```

Fig. 5. autobench output sample

### IV. TESTBED

The tests were conducted on two different servers.

First server’s specifications (dubbed as Baru) were Intel Core 2 Duo 1.83 GHz, 2 GB of RAM, and 100 Megabit network interface (64 bit machine). And second server’s specifications (dubbed as Lama) were Intel Pentium 4 2.6 GHz, 1 GB of RAM, and 100 Megabit network interface (32 bits machine). Both machines hosted OpenVZ on Centos 5. The tests were conducted from two clients both running Ubuntu 7.10. Fig 6 portrays the architecture of the test bed.

The tests were conducted based on the instructions and practices of [9]. We conducted the tests in isolated network environment.

The following configuration of “autobench” was done to measure the performance. When we benchmarked the whole machine, the total number of connections to be sent to the machine was set to 20,000. The number of requests in each connection was set to 10. The client time out was set to 7 seconds. However, when we benchmarked the containers, for each container, 5,000 connections were sent in each test. And the rest of the configuration parameters were the same as above.

OpenVZ kernel version 2.6.18, and Centos 5 templates from OpenVZ website were used and Apache web server version 2.2.3 was used as the Web server. The http queries were generated by using Httperf version 0.9.0 and autobench 2.1.2. We conducted various tests by scaling the number of containers from 4 to 8 and finally to 12.

The configurations of Apache were left to the default except the “KeepAlive Off” which we changed to “On” in order to allow multiple requests in one connection.

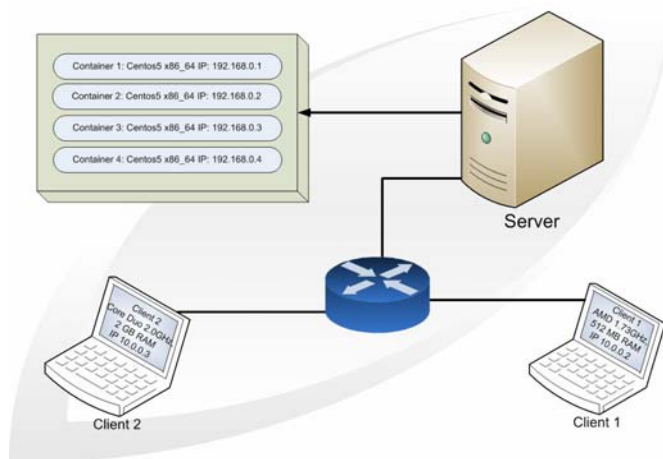


Fig. 6. The Test Bed architecture

The configuration files of all the containers were the same and a sample is shown in Fig 7.

```
ONBOOT="yes"
UBC parameters (in form of barrier:limit)
KMEMSIZE="11055923:11377049"
LOCKEDPAGES="256:256"
PRIVVMPAGES="65536:69632"
SHMPAGES="21504:21504"
NUMPROC="240:240"
PHYSPPAGES="0:2147483647"
VMGUARPPAGES="33792:2147483647"
OOMGUARPPAGES="26112:2147483647"
NUMTCPSOCK="360:360"
NUMFLOCK="188:206"
NUMPTY="16:16"
NUMSIGINFO="256:256"
TCPSNDBUF="1720320:2703360"
TCPRCVBUF="1720320:2703360"
OTHERSOCKBUF="1126080:2097152"
DGRAMRCVBUF="262144:262144"
NUMOTHERSOCK="360:360"
DCACHESIZE="3409920:3624960"
NUMFILE="9312:9312"
AVNUMPROC="180:180"
NUMIPTENT="128:128"
Disk quota parameters (in form of softlimit:hardlimit)
DISKSPACE="1048576:1153024"
DISKINODES="200000:220000"
QUOTATIME="0"
CPU fair sheduler parameter
CPUUNITS="1000"
```

Fig. 7. Configuration of a typical container

## V. DISCUSSION

### A. OpenVZ

There are general observations about OpenVZ that we would like to talk about in this section.

It is easy to fire up a new container using OpenVZ. It would take about one minute. Moreover, a container occupies only a small amount of memory. However, this depends on the size of the distribution as well as the running services. In our case, we only fired up the web server and the default services that run on the Centos 5 template from OpenVZ Website. We found that each VE takes around 25 MB of RAM and about 400 MB of hard disk. Accordingly only 20 containers can be running web servers using 450 MB of RAM and about 8 GB of hard disk.

We ran 55 containers on Baru and 23 on Lama at the same time. This almost filled the memory of each machine respectively. Moreover, we ran the containers with no application overhead within itself.

### B. The host machine (Baru)

After running “autobench” (from 1000 to 7000 requests per second), the machine managed to handle about 6000 requests per second with no error. This amount of concurrent requests has saturated the CPUs of the machine. However, the throughput was not saturated due to the size of the files (~80 Kilo bytes).

Figure 8 shows the results of “autobench” which uses httperf to benchmark the machines. In Fig 8 the X-axis represents the number of requests per second that httperf is supposed to send each test, and Y-axis represents different values that changes based on the figure’s keys. In the following paragraph, we will describe what each line means.

Sent request per second is the number of requests that was sent to the machine each second. Average reply rate per second is the average number of replies from the server per second. Response time is the reply time of the request in milliseconds. Throughput is the size of the bandwidth that was used during each test in Kilo bit per second. And finally we observed the percentage of errors for each test as packet loss.

As for Fig 8 we see that the first test sent 1000 requests per second and the replies were all intact. In the next test which was 1500 requests per second also it was fine. Autobench conducted 14 tests starting from 1000 requests per second and ending at 7000 requests per second. Each test autobench increases the number of requests per second by 500. So, we see that until 6000 requests per second everything was fine and no errors occurred. After 6000 requests per second we notice there is a drop in the number of requests sent to the server which is a bit steep. And so did every thing. The drop occurred due to the CPUs saturation which was utilized 100%. Fig 9 shows the CPU usage in relation with the time. X-axis represents the time line of the test. We captured the reading of the CPU each second for the whole duration of the tests. Y-axis shows the percentage of the idle CPU. Why the idle? Because the percentage of the idle CPU show the percentage of the free portion of CPU, we can see whether the whole system including the web server is utilizing the CPU or not.



Fig 9 shows the CPU idle percentage. We observe that at the beginning of the test the usage of the CPU was about 10%. This value keeps fluctuating each second and keeps going down until it reaches 0% which means that the CPU is fully used. In this test, this point was reached at approximately at 210 seconds and error occurrences were increasing from this point and onward (i.e. 6500 requests per second and beyond).

Fig 10 shows the usage of the RAM of the whole duration of this test.

The results indicate that Baru machine can handle about 6000 requests per second and the CPU was saturated. However, in reality, usually benchmarking is done up to 80% of the CPU usage.

### C. The host machine (Lama)

Since we are testing the performance of OpenVZ on two different machines, section A presented the results of the first machine (Baru) whereas this section presents the second machine (Lama). In this section we will not present the graphs and we will only mention the results.

After running “autobench”, the machine managed to handle about 3300 requests per second with no error. This amount of concurrent requests has saturated the CPU of the machine. However, the throughput was not saturated due to the size of the files (~80 Kilo bytes).

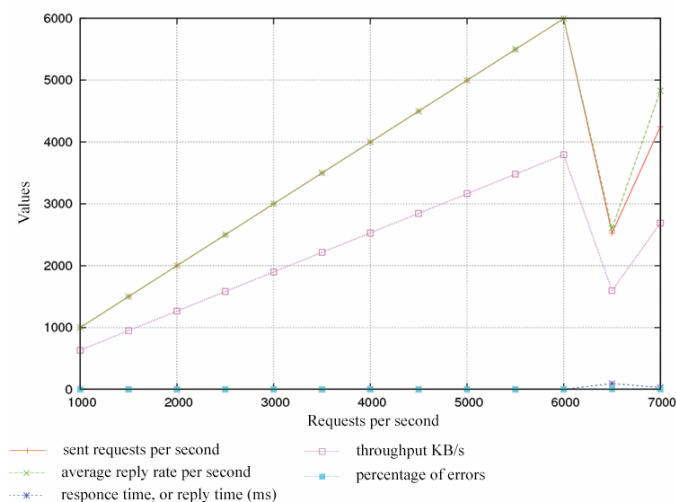


Fig. 8. The performance of Baru.

### D. OpenVZ on Baru and Lama

In this section we will describe the results of running OpenVZ on the two machines. The results are tabulated but brief discussion is provided here about the result for clarification.

As mentioned before we scaled the number of containers from 4 to 8 and finally to 12 and we conducted the tests on them so as to find the relationships between the machine and OpenVZ. Surprisingly, the two machines have had the same relationship with OpenVZ. The results are described first and later we show the relationship between the machine and OpenVZ.

#### Running 4 containers on Baru:

We set up Baru with 4 identical containers and we ran autobench against them concurrently.

In other words, all the containers receive the same load at the same time. Fig 11 shows the results of autobench on one of the containers (the other containers have almost an identical graphs so we present one of them here). As the graph shows autobench starts from 200 requests and ends at 2000 requests. However, the errors started to occur at about 1600 requests per second. This means that the machine is receiving about 6400 requests per second at the same time. Note that this number is about the same as the performance of machine Baru in section A. Surprisingly, the CPU usage (Fig 12) was almost the same as in section A and is saturated at the same the error occurrences. As for the RAM (Fig 13) it has also the same usage as in section A.

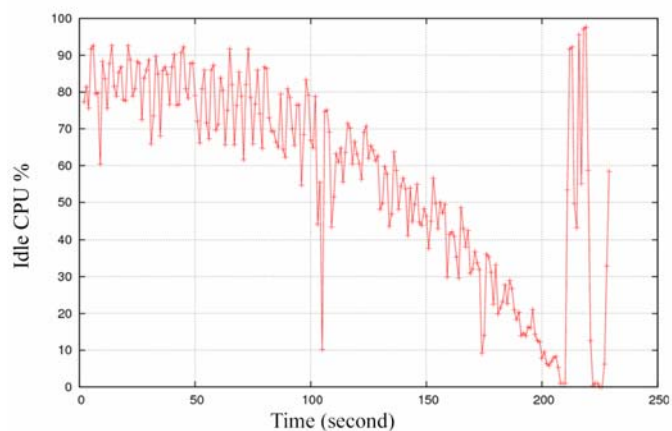


Fig. 9. CPU Usage of Baru

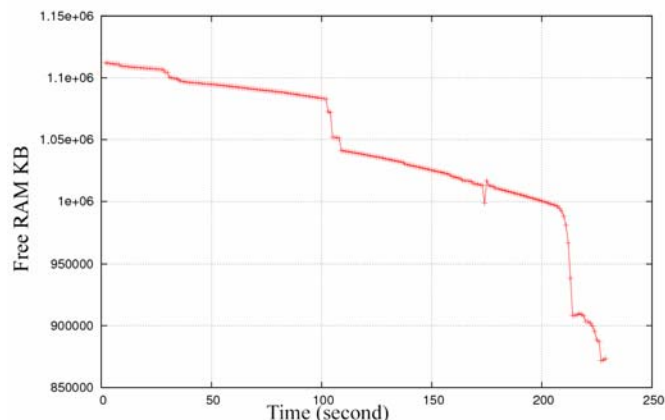


Fig. 10. RAM usage of Baru

#### Running 8, and 12 containers on Baru:

The results of 8 and 12 containers were almost the same as when we were using four containers. In 8 containers, each container handled about 750 requests per second, and in 12 containers, each container handles 500 requests per second with no errors. The graphs for the CPU and RAM were the same as Baru machine and 4 containers. Table 1 shows the relationship between the performance of Baru and 4, 8, and 12 containers. We can see from the table that OpenVZ has nearly 0% overhead. This means that we do not lose any hardware resources when we are running OpenVZ.

TABLE 1  
Performance of Baru and its containers

| Machine       | Number of servers | Requests per second | Total requests per second |
|---------------|-------------------|---------------------|---------------------------|
| Baru          | 1                 | ~6000               | ~6000                     |
| 4 containers  | 4                 | ~1600               | ~6400                     |
| 8 containers  | 8                 | ~750                | ~6000                     |
| 12 containers | 12                | ~500                | ~6000                     |

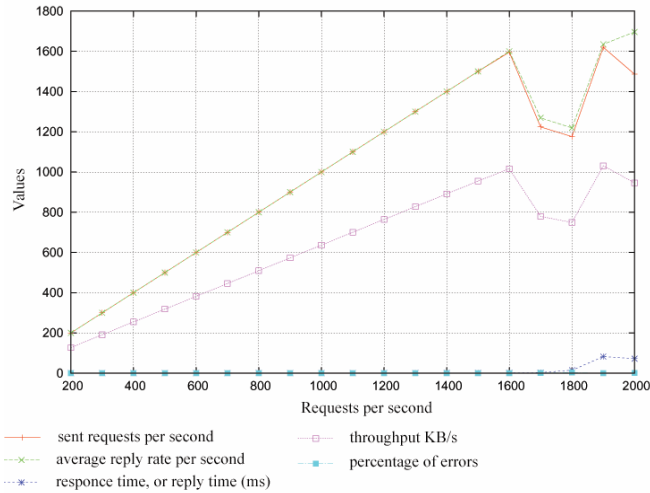


Fig. 11. Performance of one of the containers

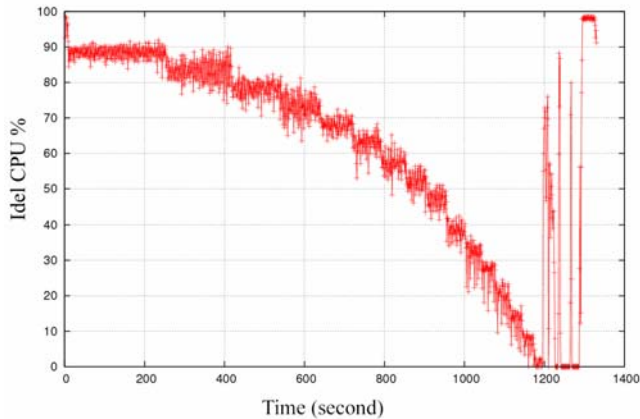


Fig. 12. CPU usage of 4 containers

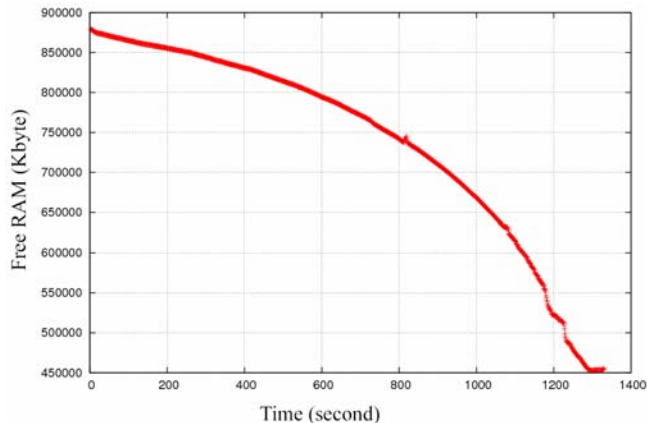


Fig. 13. RAM usage of the 4 containers

As for Lama we mentioned before that it handled about 3300 requests per second. So, when we ran 4 containers on it we found that 700 requests per seconds were handled by each container which means  $700 * 4 = 3200$  requests per second which is about the same number of requests that were handled by Lama. Note that we did not manage to run 12 containers on Lama due to RAM limitations.

TABLE 2  
Performance of Lama and its containers

| Machine      | Number of servers | Requests per second | Total requests per second |
|--------------|-------------------|---------------------|---------------------------|
| Lama         | 1                 | ~3300               | ~3300                     |
| 4 containers | 4                 | ~700                | ~3200                     |
| 8 containers | 8                 | ~400                | ~3200                     |

## VI.CONCLUSION

We found that when using OpenVZ, almost no overhead on the computer resources and it shares the resources of the computer fairly amongst the running containers. Moreover, we found that both machines (Baru and Lama) can handle about 6000 and 3300 requests per second respectively. And so do containers of OpenVZ when we run all of them concurrently on the both Baru and Lama. This means that OpenVZ performance does not change with the change of the architecture (32 bits and 64 bits). We can safely say that if we want to run 15 Web servers using OpenVZ containers, each VE can handle 400 requests per second at the same time (this is on Baru). The WebVZ is a very useful tool for administrating OpenVZ containers both locally and remotely.

## REFERENCES

- [1] Information Technology @ Johns Hopkins-Glossary of Technical Terms and Acronyms. Available <http://it.jhu.edu/glossary/pqrs.html>
- [2] Concise Oxford English Dictionary (Eleventh Edition). Computer Software.
- [3] Wolf C., Halter E. (2005). Virtualization from the desktop to the enterprise. Apress, 2005. p.1.
- [4] Shields G., The Shortcut Guide to Selecting the Right Virtualization Solution, electronic book. Available: <http://nexus.realtimepublishers.com/SGSRVS.htm>
- [5] OS Virtualization. Available: <http://www.parallels.com/en/products/virtuozzo/os/>
- [6] OpenVZ Wiki. Available: <http://wiki.openvz.org/Category:Definitions>
- [7] Padala P., Xiaoyun Zhu, Zhikui Wang, Singhal S., Shin K. Performance Evaluation of Virtualization Technologies for Server Consolidation. HP Laboratories Palo Alto, April 11, 2007. Available: <http://www.hpl.hp.com/techreports/2007/HPL-2007-59.pdf>
- [8] WebVZ. Available: <http://webvz.sourceforge.net>
- [9] The Linux HTTP Benchmarking HOWTO. Available: <http://www.xenoclast.org/doc/benchmark/HTTP-benchmarking-HOWTO/>

# A New Distributed Evolutionary Computation Technique for Solving Large Number of Equations

Moslema Jahan, M. M. A Hashem, and Gazi Abdullah Shahriar

Department of Computer Science and Engineering  
Khulna University of Engineering & Technology (KUET)  
Khulna-9203, Bangladesh

semi\_cse2k3@yahoo.com, mma\_hashem@hotmail.com, rana\_kuet@yahoo.com

**Abstract** — Evolutionary algorithm is more effective to gain optimal solution to solve equations than traditional methods. It also provides quick solution for solving large number of equations having huge parameters with less expense in distributed manner. This paper presents a new distributed evolutionary computation technique, which decomposes decision vectors into smaller components and achieves optimal solution in short time. In this technique, A Jacobi-based Time Variant Adaptive (JBTV) Hybrid Evolutionary Algorithm is distributed. Moreover, a new selection method named Best All Selection (BAS) is introduced for selecting best individuals. Experimental results show that optimal solution is achieved for different kinds of problems having huge parameters and considerably speedup is formed in proposed distributed system.

**Index Terms** — Master-Slave architecture, Linear Equations, Evolutionary Algorithms, Hybrid Algorithm, Speedup.

## I. INTRODUCTION

In recent years, application of evolutionary algorithms is increasing in a greater extent. Evolutionary algorithms (EAs) are stochastic search methods that have been applied successfully in many search, optimization and machine learning problems. Successful use of evolutionary algorithm for solving linear equations is applied in [1], [2], [3]. However, it is often very difficult to estimate the optimal relaxation factor, which is the key parameter of the successive over relaxation (SOR) method. Optimal solution is achieved quickly as relaxation factors are adapted automatically in evolutionary algorithm. Equation solving abilities was extended in [2], [3], [4] by using time variant parameter. The invention of hybrid evolutionary algorithm [5], [6] brought a greater benefit to have the solutions of linear equations within very short time. The hybrid algorithm combining the jacobi-based successive relaxation (SR) method with evolutionary computation techniques was implemented in C++ framework and the related speedup was calculated in [4]. In this case more computation time will be needed to solve the problems with huge parameters. Solving large number of equations with huge parameters needs high computational cost in case of single processor. One approach to overcome this kind of limitation is by formulating the problem into distributed computing structure.

Parallel and distributed systems are two of intrinsic natures of EAs. The main parallel achievements in the algorithmic families include the evolutionary computation, parallel models, and parallel implementations and are discussed in [7]. A distributed cooperative coevolutionary algorithm is developed in [8] which is beneficial for solving complex problems. A distributed technique [9] is proposed for parallelizing fitness evaluation time. Fitness evaluation time is high but other operations of evolutionary algorithm takes more time. This paper proposes a distributed technique in which mutation and adaptation times are distributed. So, a high quality speedup is achieved.

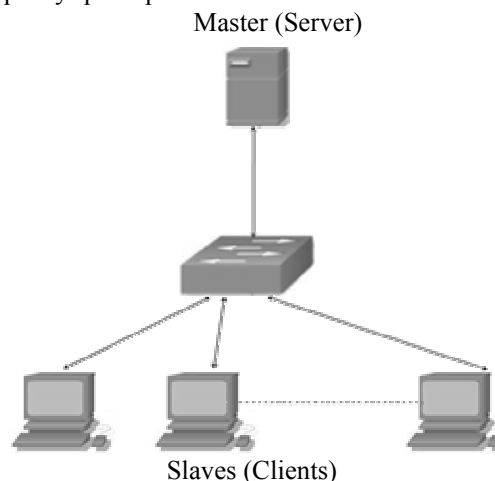


Fig. 1. Distributed Master-Slave Architecture.

Basic master-slave architecture is shown in Fig. 1 which is used in proposed distributed technique follows a server-client paradigm where connections are closed after each request. The basic approach of this system is to divide a large problem into many subproblems and to evolve subproblems separately. These subproblems are then combined and actual solution is searched. This process is continued until reaching the desired position.

The remainder of this paper is structured as follows: Section II mention Jacobi method of solving linear equations. Distributed model and timing calculation are discussed in section III. Section IV examines results of the experiment of various problems and provides a comprehensive comparison of single and distributed

system. Finally, Section V presents our concluding remarks.

## II. JACOBI METHOD OF SOLVING LINEAR EQUATIONS

Consider the following linear equations:

$$Ax = b \text{ or } Ax - b = 0 \quad (1)$$

Where  $A \in \mathbb{R}^n \times \mathbb{R}^n$  and  $x, b \in \mathbb{R}^n$

Let,  $A$  is  $n \times n$  matrix, where  $a_{ii}$  is the diagonal elements,  $a_{ij}$  is other elements of the  $A$  matrix and  $b_i$  is elements of  $b$  matrix. For solving linear equations Jacobi method is used [10].

Let,

Linear equations is  $Ax = b$  and  $|A| \neq 0$

Assuming without loss of generality that none of the diagonal entries is of zero; otherwise interchange it rows. Then

$$(D + U + L)x = b, \text{ Where } A = (D + U + L)$$

$$\text{or } Dx = b - (U + L)x$$

$$\text{or } x = D^{-1}b - D^{-1}(U + L)x$$

$$\text{or } x = H_j x + V_j$$

Where  $H_j = D^{-1}(-L - U)$  and  $V_j = D^{-1}b$

The linear equation can be written as

$$\sum_{j=1}^n a_{ij}x_j = b_j, (i = 1, 2, \dots, n) \quad (2)$$

In Jacobi method by using SR technique [11] is given by

$$x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}} \left( b_i - \sum_{j=1}^n a_{ij}x_j^{(k)} \right), (i = 1, 2, \dots, n) \quad (3)$$

In matrix form matrix-vector equation is

$$x^{(k+1)} = H_\omega x^{(k)} + V_\omega \quad (4)$$

Where  $H_\omega$  called Jacobi iteration matrix, and  $V_\omega$  are given successively by

$$H_\omega = D^{-1} \{ (1 - \omega)I - \omega(L + U) \} \quad (5)$$

And

$$V_\omega = \omega D^{-1}b \quad (6)$$

Iterative process is continued using equation (3) until the satisfactory result is achieved. Based on this method, different hybrid evolutionary algorithms are developed [1]

[2]. Problems of different categories will be generated by varying  $a_{ii}$ ,  $a_{ij}$  and  $b_i$ 's.

## III. PROPOSED MODEL OF DISTRIBUTED EVOLUTIONARY ALGORITHMS

In reality, the computation cost for evolutionary optimization in terms of time and hardware increases as the size of the problem increases. One approach to overcome such a limitation is to exploit the inherent parallel nature of EA by formulating the problem into a distributed computing structure suitable for parallel processing, i.e., to divide a task into subtasks and to solve the subtasks simultaneously using multiple processors.

### A. Distributed Technique

The availability of powerful networked computers presents a wealth of computing resources to solve problems with large computational effort. The distributed technique which is considered here, large problems are decomposed into smaller subtasks that are mapped into the computers available in a distributed system. The Distributed Hybrid Evolutionary Algorithm (DHEA) is as follows:

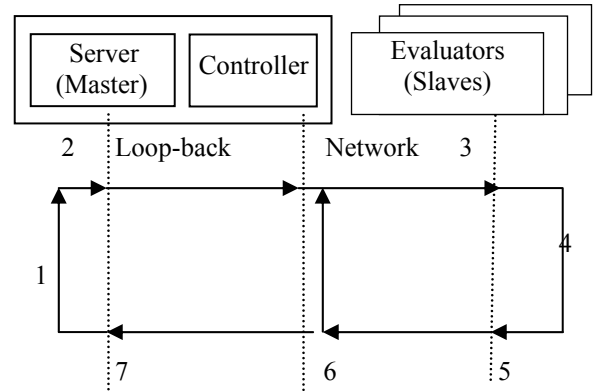


Fig. 2. Architecture of proposed model.

In Fig. 2; all steps of distributed hybrid algorithm are expressed using numbering system in each position. Step 1, 2, 6, 7 are completed in master processor and steps 3, 4, 5 steps occurred in slave processors.

1) *Initialization*: In step1, for a specific problem master (server) initializes with population and relaxation factor. Initial population  $X^{(0)} = \{x_1^{(0)}, x_2^{(0)}, \dots, x_N^{(0)}\}$  is generated randomly using normal distribution at master. Here  $N$  is the population size. Let  $k \leftarrow 0$  where  $k$  the generation counter is. This population is considered as parent. Relaxation factors according to individuals are also initialized at this position.

$$\omega^{(0)} = \{\omega_1^{(0)}, \omega_2^{(0)}, \dots, \omega_N^{(0)}\} \quad (7)$$



2) *Recombination*: In step 2, Recombination operation is performed on parent at master and population  $X^{(k+c)} = R(X^{(k)})^t$  is obtained.

Where  $R = (r_{ij})_{N \times N}$  and  $\sum_{j=1}^n r_{ij} = 1$  and  $r_{ij} \geq 0$  for  $1 \leq i \leq N$   
Matrix  $R$  is a stochastic matrix.

Population generated after recombination operation is distributed among slaves. Then mutation, fitness evaluation and adaptation operations are performed on that distributed subpopulations.

3) *Mutation*: In step 3, after completing recombination operation Mutation operation is performed on that subpopulation in slave processors and mutated subpopulation  $X^{(k+m)}$  is obtained. For each individual  $x_i^{(k+c)}$  ( $1 \leq i \leq N$ ) in subpopulation  $x^{(k+c)}$  produces an offspring using following equation

$$x_i^{(k+m)} = H_{\omega} x_i^{(k+c)} + V_{\omega}, (i = 1, 2, \dots, n)$$

Where  $H_{\omega}$  is called Jacobi iteration matrix.

Consider, number of slaves is  $M$  and number of individuals is  $N$ , then Number of individuals in each subpopulation,

$$N_{sub} = \frac{N}{M}$$

Mutation on subpopulation is

$$\sum_{i=1}^{N_{sub}} x_i^{(k+m)} = H_{\omega} x_i^{(k+c)} + V_{\omega} \quad (8)$$

4) *Fitness Evaluation*: After completing mutation operation, fitness evaluation is performed in step 4. Error function is

$$E = \sum_{i=1}^{N_{sub}} E_i \quad (9)$$

Where  $E = \sum_{j=1}^{N_{sub}} |a_{ij}(x_j) - b_i|, i = 1, 2, \dots, N_{sub}$

The fitness evaluation of an individual measures the accuracy of an individual for a particular problem and calculates the error rate which is used for selecting best individuals.

5) *Adaptation*: In step 5, adaptation is performed on mutated offspring according to fitness evaluation. Consider two individuals  $X$  and  $Y$ , corresponding errors  $\|e_x\|$  and  $\|e_y\|$  and relaxation factors  $\omega_x$  and  $\omega_y$ .  
If  $\|e_x\| > \|e_y\|$  then  $\omega_x$  is adapted to  $\omega_y$ .

$$\omega_x^m = (0.5 + p_x)(\omega_x + \omega_y)$$

Similarly if  $\|e_y\| > \|e_x\|$  then  $\omega_y$  is adapted to  $\omega_x$  and if  $\|e_x\| = \|e_y\|$  then no adaptation is performed.

In step 6, controller checks whether all offspring from all slaves have reached or not. If all offspring have reached then controller sends them to server otherwise looping is continued.

After reaching all offspring in server, selection operation is performed and best individuals are selected.

6) *Selection*: In BAS method, best individuals are selected among parent and all mutated offspring. At position 7, all mutated offspring from all slaves are combined and selection mechanism is performed on all parent and all offspring.

Mathematically,

$$Selection(i) = err_{\min} \{x_1, x_2, x_3, \dots, x_{2N-i+1}\} \quad (10)$$

$$Selection = \sum_{i=1}^N Selection(i) \quad (11)$$

In case of selection, direct competition is occurred among individuals and best individuals are selected according to error rate. If error is less or equal to threshold value then selected individuals are the solution for that problem. Otherwise this process is continuing until error is less than threshold value.

There is no direct communication among the subpopulations. All communications are performed between the subpopulation and the central server. Such communications take place in every generation before reaching the accurate result and communication increases linearly with respect to the number of computers.

#### B. Computation Time Analysis of Single Processor and Distributed Processors

All mathematics, comparison and assignment statements in the computer are executed in a particular time. For a specific optimization problem, recombination time, mutation time, fitness evaluation time, adaptation time and selection time can be denoted by  $T_r, T_m, T_f, T_a$  and  $T_s$  respectively. Total time in single processor is indicated by  $T_{single}$ . Time in single processor is represented as follows,

$$T_{single} = T_r + T_m + T_f + T_a + T_s \quad (12)$$

In case of distributed system, time is calculated in master and slaves separately. Then total time is calculated by combining master and slave time.

1) *Time in master processor*: For a particular problem, Recombination operation is performed on initial population and time  $T_r$  is calculated. Then the server distributes individuals among all slaves connected to it. Number of subpopulation is  $p$  which is distributed in slaves. In this case, Marshalling and transmission time is considered. Marshalling time is the time to construct data stream from object and transmission time is the time to place data stream in channel. Marshalling time of  $i^{th}$  individual is  $T_{mi}$  and Transmission time with  $i^{th}$  individual is  $T_{transi}$ .

If Marshalling and transmission time with  $p$  subpopulation =  $p[T_{mi} + T_{transi}]$

2) *Time in slave processors*: Mutation time, fitness evaluation time, adaptation time with  $p$  subpopulation is  $T_m$ ,  $T_f$  and  $T_a$  respectively.

After completing adaptation operation, slaves send mutated offspring to server. Unmarshalling time  $T_{umi}$  is considered here. Unmarshalling time is the time to create object from data stream. After reaching all offspring, selection operation is performed on mutated offspring and selection time  $T_s$  is obtained in master. The processing time in each slave is not equal because of communication delay. In experimental calculation maximum time is considered.

In BAS technique, selection mechanism is performed in master processor.

3) *Time in Master processor for selection*: Selection time with  $2N$  subpopulation is  $T_s$ ,  $N$  individuals for parent and  $N$  individuals for offspring so total is  $2N$  subpopulation. There is an important thing that maximum time is considered for calculating speedup among all slaves. Consider  $m$  is the number of slaves. Total time for distributed processors is as follows,

$$T_{Distributed} = T_r + p[T_{mi} + T_{transi} + \max_{1 \dots m}(T_m + T_f + T_a)] + pT_{umi} + T_s \quad (13)$$

Where  $T_{distributed}$  represents total distributed time,  $T_r$ ,  $T_{mi}$ ,  $T_{transi}$ ,  $T_m$ ,  $T_f$ ,  $T_a$ ,  $T_{umi}$ ,  $T_s$  represents recombination time, marshalling time, transmission time, mutation time, fitness evaluation time, adaptation time, unmarshalling time and selection time, respectively.

Speedup for distributed technique,

$$Speedup = T_{Single} / T_{Distributed} \quad (14)$$

Percentage of improvement, % = Speedup/Number of computer (15)

Computation time in distributed processors will be less than single processor. So, speedup will be gradually rising with increasing number of individuals in a population.

### C. Achieving Optimal Solution

Generally, problems can be solved using iterative methods or other conventional numerical methods but accuracy, effectiveness, less computation time etc. are the properties desirable to all of us. By adding evolutionary algorithm with conventional method, it is possible to get optimal solution because of its excellent capacity in case of changing the value of relaxation factor  $\omega$ . Accuracy is coming but a high amount time is needed for solving large number of equations having huge parameters. This situation will be removed using our proposed distributed technique in which large number of equations with huge parameters is solved within a short time. Table I sustains the efficiency of proposed distributed technique.

## IV. EXPERIMENTAL RESULTS

Using the proposed mathematical model we can see realistic experimental results. Consider a cluster with 20 homogeneous computers and 100 Mbits/sec Ethernet switch. Assume a population of 80 individuals and dimension 170 is evolved. This system is implemented in Java on a personal computer with Intel Pentium IV and 512MB RAM. Various results are displayed by considering different aspects. In our experiment, individual values are generated randomly. Here, random values are generated using normal distribution. -15 to 15 range is used for choosing random values. As normal distribution is used, result converges with few generations. Random values are generated with 80 individuals and 170 dimensions. Here, testing problem is  $a_{ii} = 20.0$ ,  $a_{ij} \in (0, 1)$ ,  $b_i = 10 * i$ ; that means P1 problem in Table I All graphs show the results of testing problem. Table I represents a few problems based on (1) and reflects performance and efficiency of distributed system.

In BAS method, marshalling and unmarshalling time is considered. When number of individuals are 30 and number of parameters are 100 for 15 number of computers in a cluster, then marshalling and transmission time is  $3.7 \times 10^{-3}$  sec, mutation time is  $7.6 \times 10^{-3}$  sec, fitness evaluation time is  $4.5 \times 10^{-3}$  sec, adaptation time is  $3.5 \times 10^{-3}$  sec and unmarshalling time is  $8.9 \times 10^{-4}$  sec. Number of individuals is 40 and number of parameters is 100 for 5 and 10 number of computers in a cluster.

Table I summarizes experimental results of various problems. In this experiment, number of computers in a

TABLE I  
EXPERIMENTAL RESULTS OF TIME FOR DIFFERENT PROBLEMS IN SINGLE AND DISTRIBUTED PROCESSOR

| Problem Number | Problem                                                                  | Individual Number | Dimension number | Time in single processor (seconds) | Time in distributed processor (seconds) | Error              |
|----------------|--------------------------------------------------------------------------|-------------------|------------------|------------------------------------|-----------------------------------------|--------------------|
| P1             | $a_{ii} = 20.0; a_{ij} \in (0, 1); b_i = 10 * i;$                        | 40                | 170              | $4.02 \times 10^{-1}$              | $6.88 \times 10^{-2}$                   | $9 \times 10^{-4}$ |
| P2             | $a_{ii} = 20n; a_{ij} = j; b_i = i;$                                     | 40                | 120              | $3.68 \times 10^{-1}$              | $6.75 \times 10^{-2}$                   | $9 \times 10^{-9}$ |
| P3             | $a_{ii} = 2i^2; a_{ij} = j; b_i = i;$                                    | 40                | 120              | $3.38 \times 10^{-1}$              | $6.16 \times 10^{-2}$                   | $9 \times 10^{-9}$ |
| P4             | $a_{ii} \in (-100, 100);$<br>$a_{ij} \in (-10, 10); b_i \in (-100, 100)$ | 20                | 50               | $1.63 \times 10^{-1}$              | $1.79 \times 10^{-2}$                   | $9 \times 10^{-2}$ |
| P4             | $a_{ii} \in (-100, 100);$<br>$a_{ij} \in (-10, 10); b_i \in (-100, 100)$ | 20                | 80               | $1.64 \times 10^{-1}$              | $2.73 \times 10^{-2}$                   | $9 \times 10^{-2}$ |
| P4             | $a_{ii} \in (-100, 100);$<br>$a_{ij} \in (-10, 10); b_i \in (-100, 100)$ | 20                | 100              | No solution                        | $3.57 \times 10^{-2}$                   | $9 \times 10^{-2}$ |
| P5             | $a_{ii} = 5; a_{ij} = \{1, 0\};$<br>$b_i = 2;$                           | 20                | 50               | $1.53 \times 10^{-1}$              | $1.62 \times 10^{-2}$                   | $9 \times 10^{-2}$ |
| P6             | $a_{ii} \in (-70, 70);$<br>$a_{ij} \in (0, 7); b_i \in (0, 70);$         | 20                | 100              | $1.62 \times 10^{-1}$              | $1.60 \times 10^{-2}$                   | $9 \times 10^{-1}$ |
| P7             | $a_{ii} = 70; a_{ij} \in (-10, 10);$<br>$b_i \in (-70, 70);$             | 40                | 10               | $1.85 \times 10^0$                 | $2.13 \times 10^{-2}$                   | $9 \times 10^{-4}$ |
| P8             | $a_{ii} \in (2, 10);$<br>$a_{ij} \in (-2, 2); b_i \in (-5, 5)$           | 40                | 10               | $7.26 \times 10^{-1}$              | $4.30 \times 10^{-2}$                   | $9 \times 10^{-4}$ |

It is observed that P4 problem is tested in three cases and there is no solution in the last case for single processor but solution is occurred for distributed system because of computation complexity with huge parameters.

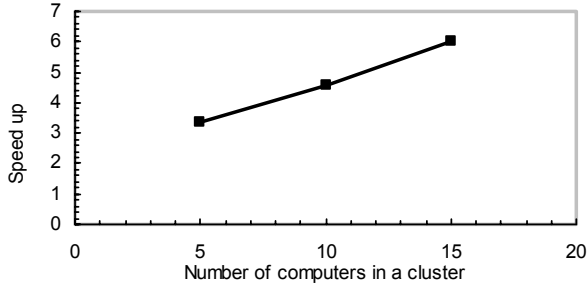


Fig. 3. Results for measuring speed according to number of computers in a cluster.

In Fig. 3. Using (14) speedup is shown that when number of slave computers is 5, 10 and 15 then, speedup is 3.36, 4.57 and 6, respectively. When number of slave computers is 5, 10 and 15 then percentage of improvement is sequentially 67.2 %, 45.7 % and 40 % is calculated based on (15). Fig. 4. Shows the time required for performing genetic operation in each slave. Here, number of slave computers in a cluster is 15 and individual is 30 and parameters are 100. There is no load balancing system so different slave computers need different time. Intuitively, workload balancing for a distributed system can be difficult because the working environment in a network is often complex and uncertain.

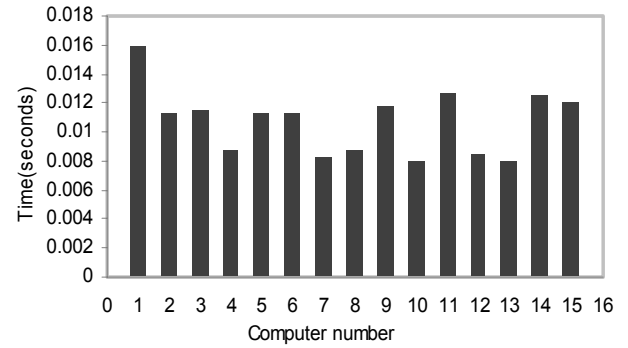


Fig. 4. Results for measuring time (seconds) according to computer number.

From Fig. 4, we can see that computer number 1 needs the highest computation time among all slaves. For calculating total distributed time, this maximum value is considered.

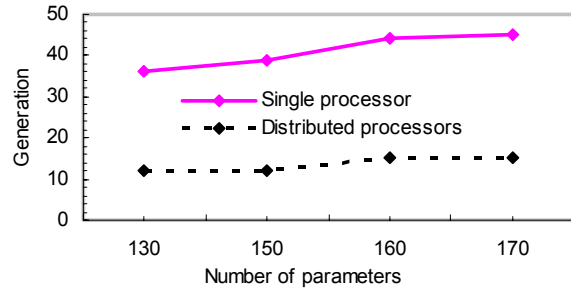


Fig. 5. Results for measuring generation number according to the number of parameters.

In Fig 5, generation number is considered for single and distributed processors according to parameter number. Here, number of slave computers in a cluster is 20 and individual is 80 and parameters are 130,150,160,170. To reach convergence generation number for single processor is 24, 27, 29, 30 and generation number for distributed processors is 12, 12, 15, and 15 according to number of parameters.

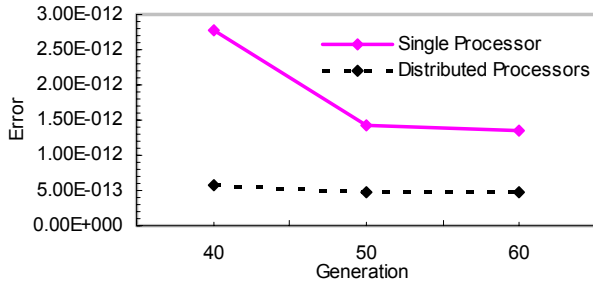


Fig. 6. Results for measuring error according to generation.

In Fig 6, error is considered according to generation. In this case, number of parameters is 100 and individual number is 40. Generation numbers are 40, 50 and 60.

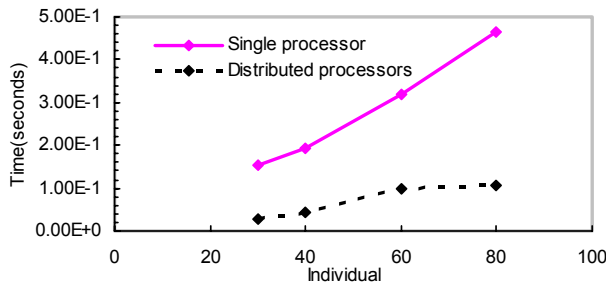


Fig. 7. Results for measuring Time in distributed processors and single processor according to Individual.

Fig. 7 shows comparison of time between single and distributed processors. When individual is 40 and 80 then number of slave computers in a cluster is 10 and 20. When individual is 30 and 60 then number of slave computers in a cluster is 15. In this case, number of parameters is always 100.

## V. CONCLUDING REMARKS

It is easy to solve linear equations with less number of parameters in single processor but high computational cost is needed for large number of parameters. This cost is dramatically reduced using distributed system. This paper

introduced a new distributed algorithm for solving large number of equations with huge parameters. It also introduced a new selection method called BAS for selecting offspring. In this method, best individuals are selected and computation load is distributed in each slaves and mutation, fitness evaluation and adaptation operations are performed on distributed load. As distributed technique is used, computation time is reduced. Different problems are tested and high quality speedup is achieved in distributed system.

## REFERENCES

- [1] Jun He, Jiyou Xu, and Xin Yao, "Solving Equations by Hybrid Evolutionary Computation Techniques", *IEEE Trans. Evolutionary Computations.*, vol. 4, no. 3, pp. 229-235, 2000.
- [2] A.R.M. Jalal Uddin Jamali, Md. Bazlar Rahman and M.M.A. Hashem, "Solving Linear Equations by Classical Jacobi-SR Based Hybrid Evolutionary Algorithm with Time-Variant Adaptation Technique," *Ganit: Journal of Bangladesh Mathematical Society*, vol. 24, pp. 67-79, 2005.
- [3] A.R.M. Jalal Uddin Jamali, M.M.A. Hashem and Md. Bazlar Rahman, "An approach to Solve Linear Equations Using a Time-Variant Adaptation Based Hbrid Evolutionary Algorithm," *Jahangirnagar University Journal of Science*, vol. 27, pp. 277-289, 2004.
- [4] Jamali, A.R.M.J.U, "Study of Solving Linear Equations by Hybrid Evolutionary Computation Techniques," *M.Phil Thesis*, Khulna University of Engineering & Technology, Khulna, Bangladesh, 2004.
- [5] Hashem, M.M.A., "Global Optimization Through a New Class of Evolutionary Algorithm," *Ph.D. dissertation*, Diss. No. 19, Saga University, Japan, pp. 1-25, 1999.
- [6] K. Watanabe and M.M.A. Hashem, "Evolutionary Computations: New Algorithms and their Applications to Evolutionary Robots," vol. 3, Springer-Verlag, Berlin, 2004.
- [7] E. Alba and M. Tomassini, "Parallelism and Evolutionary Algorithms," *IEEE Trans. Evolutionary Computations.*, vol. 6, no. 5, pp. 443-462, 2002.
- [8] K. C. Tan, Y. J. Yang and C. K. Goh, "A Distributed Cooperative Coevolutionary Algorithm for Multiobjective Optimization" *IEEE Trans on Evolutionary Computation*, vol. 10, no. 5, 2006.
- [9] M. Dubreuil, C. Gagné, and M. Parizeau, "Analysis of a Master-Slave Architecture for Distributed Evolutionary Computations", *IEEE Trans. On Systems, Man and Cybernetics-Part B*, vol. 36, no. 1, pp. 229-235, 2006.
- [10] Gerald, C. F. and P.O. Wheatley, "Applied Numerical Analysis (5<sup>th</sup> edition)", Addison-Wesley, New York, 1994.
- [11] Engell-Müllges, G.E. and F. Uhlig, "Numerical Algorithms with C", Springer-Verlog, Heidelberg, 1999

# Achieving Guaranteed Performance on Computational Grid

Ajanta De Sarkar, Dibyajyoti Ghosh, Rupam Mukhopadhyay and Nandini Mukherjee

Department of Computer Science and Engineering, Jadavpur University, Kolkata 700 032, West Bengal, India.

**Abstract:** For the dynamic nature of the Grid, the applications need to adapt to the changing resource status and usage scenarios. This is required to ensure that the performances of the components do not suffer. Objective of this paper is to present an adaptive execution scheme for achieving guaranteed performance as has been agreed upon through service level agreements (SLAs). The scheme has been implemented based on the notion of performance properties and by deploying a set of autonomous agents within an integrated performance-based resource management framework.

**Index Terms:** Adaptive Execution, Performance Properties, Grid

## I. INTRODUCTION

In a computational Grid environment, an application is executed as a set of jobs (components) [12], which are allocated to different resources in a heterogeneous and dynamic resource pool. The major objective is to achieve the performance of an application as desired by a client and as specified in the *service level agreements* set up between the client and the resource owners. However, Grid being a dynamic and complex environment, it is difficult to predict the appropriate set of resources, which will satisfactorily deliver the required performance. Often performance models are used to estimate the capabilities of the resource set and an application is mapped onto the resource set based on this model. But, such models generally fail because of the frequently changing resource load and resource usage scenario and the volatility of the resource pool, which resources may join and leave frequently. Thus, the initial mapping of the application components onto a set of selected resources may need to be reviewed and changed during run-time so that performances of any of the components do not suffer. Applications running on a Grid environment should therefore adapt themselves to certain changes in the environment so that any kind of performance degradation of any application component can be overcome.

This paper presents an adaptive execution scheme for achieving *guaranteed performance* of components (hereafter referred as jobs) of an application as has been agreed upon through service level agreements (SLAs). The scheme has been implemented by deploying a set of autonomous agents within an integrated multi-agent framework for performance-based resource management [12]. Within the framework, all the agents cooperate to

accomplish tasks like *resource brokering*, *performance monitoring and analysis* and *adaptive execution* of the jobs. A hierarchical agent framework is developed for performance analysis of these jobs at different levels [6]. The paper mainly deals with the adaptive execution strategies and algorithms within the multi-agent framework. Results of the adaptive execution scheme using local tuning techniques are implemented in this paper and adaptive execution scheme using migration technique is presented in [12].

The remaining part of this paper is organized as follows. Related work is discussed in Section 2. Section 3 presents the adaptive execution scheme. The concepts related to performance properties and their relevance in the current work are discussed in Section 4 and notion of severity along with its computation techniques are discussed in Section 5. Algorithms for maintaining quality of services and the agent interactions are described in Section 6. Section 7 and Section 8 give details of the experimental setup and results of the experiments that have been carried out to demonstrate the adaptive execution scheme and its effectiveness. Section 9 concludes with a direction of the future work.

## II. RELATED WORK

Adaptive execution has received a lot of attention from the Grid researchers. Grads project [2],[3] presents a program execution framework, which supports adaptive reallocation if performance degrades because of changes in the availability of Grid resources. It also implements rescheduling methods on the basis of resource selection model and monitors performance contracts using simple stop/migrate/restart strategy. The performance contracts specify the expected performance of modules as a function of available resources in Grid. In [5], a job-monitoring technique and an analyzer tool for automatic generation of the required job behaviour description are proposed. This work gathers resource access information, produces job behaviour description after the analysis. Then the job behaviour description is used to choose the scheduling algorithms of jobs in Grid. A low-cost rescheduling policy is proposed in [14]. With this policy, after the initial mapping, some jobs are selectively rescheduled on the basis of their run-time performance analysis. It considers rescheduling at a few, carefully selected points during execution. In this work, the change in the resource pool is not considered. In [17],

HEFT-based adaptive rescheduling algorithm is presented with collaboration between workflow planner and executor. In this approach, the executor will notify the planner about any run time event, through which planner gets information about resource unavailability or discovery of new resources. This research adapts the resource pool change to achieve better performance and reschedule the remaining jobs if necessary.

The OpenMP based adaptive parallelism is presented in [15]. In this system, task parallel OpenMP applications can execute on network of workstations (NOW) with a variable of nodes. Adaptive parallelism allows all processes to complete their current task, as processes do not have a compute-relevant private process state. The problem of resource selection and adaptation in grid environments is discussed in [16]. Adaptation of resources is done on the basis of degree of parallelism. In this work the adaptation coordinator periodically collects performance statistics and computes the weighted average efficiency. It compares the weighted average efficiency with certain threshold to add or remove processors from the cluster. The system, however, does not support run-time job migration.

Previous research works mainly concentrate on run-time job migration. On the other hand, the work presented in this paper supports adaptive execution either by application of local tuning techniques or through run-time job migration depending upon the specific situation. Our research focuses on an integrated multi-agent-based framework, in which agents (including mobile agents) are deployed to coordinate the adaptive execution of multiple concurrent jobs. This is also a novelty of our system. In order to detect the performance problem, we have used the notion of performance properties and their severities. Moreover, the algorithms and strategies presented in this work not only aims at maintenance of performance guarantee, but also attempts to keep the resource utilization cost as minimal as possible. Thus, any over-provisioning is controlled by iteratively measuring the severity of performance properties and taking appropriate action immediately (sometimes even decreasing the allocated resources). In the next section, we discuss the basic adaptive execution scheme.

### III. ADAPTIVE EXECUTION

In Grid environment, resources in a dynamic resource pool are owned by different resource-owners and are administered under multiple administrative domains. Status and load of the resources change frequently and the jobs need to adapt to the changes so that desired performance can be achieved. In this section, we present an adaptive execution scheme using local tuning techniques that is implemented within our multi-agent framework. However, adaptive execution scheme using

run-time job migration technique is discussed and implemented in [12].

#### A. Initial Allocation

Whenever, a client submits a batch of jobs (components of one or more applications), the client also specifies resource requirements for each of them. Along with the requirements, a job metadata, which describes the job, is also submitted. On the basis of the job requirements and the resource capabilities, a resource selection algorithm is run to select an initial set of resources to which the batch of jobs is allocated. For each job, an SLA is established between the selected resource provider and the client. The SLA provides detailed information about the requirements of the job and availability of resources, while job metadata contains information regarding significant parts (such as loops) of the job [7]. The SLA also contains the *expected completion time* ( $T_{est}$ ), which provides the basis of the performance guarantee for the job. Consequently, the next objective of the multi-agent framework is to maintain this performance guarantee by adaptively executing the job.

In general, a job may adapt to its current run-time environment depending on its own resource requirement and the present resource usage scenario. Additionally, adaptive execution may also be provided by automatic migration of the job in any of the following cases (but not limited to them) (i) performance degradation, (ii) changes in job requirements, (iii) resource failure, (iv) any kind of fault in the execution environment. Our multi-agent framework enables adaptive execution by considering the above two situations. It firstly applies local tuning techniques, i.e. attempts to improve the underlying execution environment as and when necessary and secondly supports automatic migration of the job to another appropriate resource provider when resource requirement cannot be fulfilled on the current resource provider.

#### B. Performance Analysis

Following the SLA establishment, each job is submitted onto the selected resource provider and a JobExecutionManager Agent (JEM Agent), which is a mobile agent, becomes associated with it and is deployed on the same host along with the job. A Node-level Analysis Agent (NA) [6] [7] sitting on the resource provider monitors execution of all jobs submitted to that resource provider and analyses their performances. Performance analysis for every job is done at periodic intervals. The NA instructs the JEM Agent to suspend the job after executing certain portion of it, gathers all relevant performance data at this point and analyses the data. In case the NA identifies a performance problem or predicts (explained in the next section) that the resource

provider may fail to maintain the performance guarantee, it invokes a Tuning Agent (TA).

### C. Tuning Actions

The TA decides either to locally tune the job or to migrate the job to a different resource provider. The decision regarding the next course of action (local tuning or migration) is taken on the basis of the nature of the performance problem and the resource availability and current status. Thus, the job may execute using more number of threads on more number of processors or may be tuned to reduce the load imbalance in case sufficient resources are available on the same resource provider. However, if resources on the same resource provider are overloaded or a resource provider with higher capability is required to achieve the guaranteed performance, then the job is migrated. The JEM Agent then selects (explained later) a new suitable resource provider and reschedules the job to this newly selected resource provider and resumes its execution from the point at which it has been suspended. The JEM Agent, being a mobile agent, moves along with the job, controls execution of the job and performs resource-brokering task when rescheduling is required.

## IV. PERFORMANCE PROPERTIES

One of the major differences of our work with other similar research works presented in literature is that it *depends on the notion of performance properties*. Performance properties have been introduced in [8] and thoroughly discussed in [9]. A performance property characterizes specific performance behaviour of a program, such as load imbalance, cache misses and communication. Every performance property can be checked by a set of conditions and is associated with a severity figure, which shows how important the performance property is in the context of the performance of an application. When the severity figure crosses a pre-defined threshold value, the property becomes a performance problem and the performance property with highest severity value is considered to be a performance bottleneck.

In our system, the resource provider defines a set of performance properties which are relevant for the machine architecture and the programming paradigm used for executing jobs. At the time of execution of a job these performance properties are checked at certain *measuring points* and for each, the severity value is computed. The severity values are computed on the basis of performance property specifications (define the condition, confidence and severity for each property), which are defined and stored on the resource provider. During the execution of a job, performance data are collected at certain intervals (at measuring point) and on the basis of this data and the performance property specifications, the severities of specified performance

properties are evaluated and a *Performance Details Report* is generated [6]. In order to evaluate the severity of certain performance property, some parameters from the SLA are also required (next section).

Two performance properties are used for the current work. First is the *InadequateResources* property that identifies the problem related to execution of a portion of job sequentially (or on less number of processors) when additional number of processors is required to achieve performance guarantee.<sup>1</sup> Second is the *LoadImbalance* property. Severity computation procedures for these properties are discussed in the next section.

## V. QOS AND SEVERITY COMPUTATION

In a computational Grid environment, *quality of services* (QoS) is primarily defined in terms of the execution performance of an application or the components of the application submitted by a client. QoS desired by a client and guaranteed by a resource provider is specified in an SLA. At the time of initial allocation of a job onto a resource provider, the client first sends *Task Service Level Agreement* (TSLA) to the resource broker component. The TSLA contains two parts; in the first part the requirements for a job (such as preferred start-time, expected completion time, number of processors) are specified and in the second part information about every significant region (loop) in the program (such as nesting level, loop bounds) are specified in the form of job metadata. A *Resource Service Level Agreement* (RSLA) containing information about the capabilities of a resource provider is sent by the resource provider to the resource broker component. When a specific resource provider is selected for a job [13], the SLA is finalized in the form of a *Binding Service Level Agreement* (BSLA) which is a tripartite agreement between the client, the resource provider and the brokering component of the system. BSLA contains information about the requirements of the job and availability of resources as committed by the resource provider.

The *expected completion time* ( $T_{ec}$ ) for a job is generally based on some pre-execution analysis of the job or on historical data gathered after any previous execution of it. Severities of the performance properties in some cases are computed using the information stored in the JobMetaData of the BSLA. For every job, a JobMetaData is submitted by the client that contains information about its significant regions. In this work only loops are considered as significant regions. Severity

---

<sup>1</sup> A property called *Non-parallelized code* has been used in [9] in order to capture the overhead due to sequential part of the program. However, the property defined by us not only captures the problem due to sequential part of a job, but also it detects performance problems when a job is running parallelly using less number of threads than it requires to use in order to maintain the guaranteed performance.

values are computed during the execution of these significant regions only. Whenever, a significant region starts executing, the NA introduces certain measuring points within the region by communicating a factor  $f$  to the JEM Agent which controls the execution of the job. The factor  $f$  indicates the distance between two measuring points (initially distance of the first measuring point from the start of the significant region). At each point, performance data regarding execution of the job are collected and sent to the NA for computation of severity values. Occasionally, NA may revise the factor  $f$  depending on the requirements.

In case of Inadequate Resources Property, the severity figure for a specific loop  $L_i$  is given by,

$$\text{sev\_resr}(L_i) = [T_{\text{fact}}^f(L_i) / T_{\text{fect}}^f(L_i)] \quad (1)$$

where, expected completion time of  $f$  portion of a loop ( $L_i$ ) is given by  $T_{\text{fect}}^f(L_i)$ , and actual completion time of  $f$  portion of the same loop ( $L_i$ ) (as measured at the time of execution) is given by  $T_{\text{fact}}^f(L_i)$ .  $T_{\text{fect}}^f(L_i)$  is computed using the information in the JobMetaData. Thus,

$$T_{\text{fect}}^f(L_i) = [L_{\text{frac}}^i * T_{\text{ect}} * f] \quad (2)$$

$L_{\text{frac}}^i$  is the proportionate execution time of the loop ( $L_i$ ) during its current invocation with respect to the total execution time of the job (we consider this to have been collected from historical data or through profiling and to be part of BSLA).

In case of Load Imbalance Property, execution times for every thread are measured. Thus, the severity figure for a loop  $L_i$  is given by

$$\text{sev\_load}(L_i) = [(T_{\text{max}}^f(L_i) - T_{\text{avg}}^f(L_i)) * 100] / T_{\text{max}}^f(L_i) \quad (3)$$

where,  $T_{\text{max}}^f(L_i)$  is the maximum time spent in a thread to execute  $f$  portion of the loop and  $T_{\text{avg}}^f(L_i)$  is the average taken over the times spent in all the threads to execute  $f$  portion of the loop.

Thus, when the above severity values are greater than a threshold (here greater than 1), the corresponding performance problem is detected. Such problems indicate that the QoS cannot be maintained, or in other words guaranteed performance cannot be achieved. Therefore, the agents involved with the process take some action so that the performance problem can be overcome. Next section describes the operations and interactions of the agents in order to explain how QoS is maintained in the system.

## VI. QOS MAINTENANCE

In this section, we focus on the services of the agents and their interactions from three different perspectives: (i) run-time performance analysis of jobs, (ii) run-time local tuning of jobs and (iii) run-time job migration to other resources.

### A. Performance Analysis

NAs are always active in each resource provider in the Grid and are responsible for run-time performance analysis of all the jobs running using the local resources. When execution of a job starts, the JEM Agent associated with it sends 'Registration' message to the NA. NA, in response sends a message back to the JEM Agent indicating the fraction of each significant region to be executed before collection of any performance data (the factor  $f$ ). JEM Agent collects performance data and sends the data to the NA and suspends the job. NA computes severities of different performance properties (such as  $\text{sev\_resr}(L_i)$  and  $\text{sev\_load}(L_i)$ ) the particular significant region. If any of the properties is detected as severe, NA invokes TA and sends 'Warning' message to the JEM Agent. In case, none of the measured properties are severe, NA sends 'Resume' message to the JEM Agent for resumption of the job. The 'Resume' message also carries a revised value of  $f$  that indicates the next portion of the significant region to be executed before arriving at a measuring point. NA continues this iterative performance analysis till the end of the job.

### B. Local Tuning

Like NAs, TAs are always active in each local resource provider in the Grid. A TA is invoked by NA and receives the severity message from NA containing the job\_id, performance property which has been identified as a performance problem and the severity value. TA checks the resource availability for tuning for the *InadequateResources* property on the local resource provider and if resources are available it sends a message indicating the required number of resources to the JEM Agent. If the resources are not available it sends an appropriate message to the JEM Agent. In case of *LoadImbalance*, local scheduling strategy to different threads may be changed. The TA instructs the JEM Agent to change the strategy, and the JEM Agent changes it and executes the remaining job with the new strategy. Currently, TA uses a simple algorithm because it handles only two performance properties. However, in a complicated scenario a better algorithm and a proper use of a knowledge base may be introduced.

### C. Job Migration

The major decision to be taken at the time of migration of the job is where the job is to be migrated. In the hierarchical agent framework, a higher-level agent, called Grid Site Agent (GSA) [7], is used to monitor all the grid resources located at a particular geographical site. A ResourceProvider Agent (RPA) [12], on the other hand, is responsible for maintaining information and



establishing SLAs with the client. Each JEM Agent carries a *ResourceProviderList*, which contains all the Resource Providers who can meet the requirements of the job (in ascending order of resource utilization cost [13]). During the execution of a job, if the TA indicates that migration is required, it sends the *ResourceProviderList* to the GSA. GSA checks resource availability and current load of each resource using Ganglia information provider [10]. It sends back a *RevisedResourceProviderList*, which contains only the resource providers who can currently fulfill the requirements of the job in order to achieve the guaranteed performance. Because Grid is a dynamic environment, this step helps to take into account any change in the resource pool. The JEM Agent selects a Resource Provider from the list, sends a BSLA to the corresponding remote RPA. If the RPA accepts the BSLA, job is migrated to the particular resource provider along with the JEM Agent; otherwise, the next Resource Provider from the list is contacted.

## VII. EXPERIMENTAL SETUP

The above adaptive execution scheme has been tested on a local Grid test bed that has been set up with heterogeneous nodes running in Linux. The Grid test bed is built with Globus Toolkit 4.0 (GT4) [11]. Ganglia Information Provider [10] is installed in each node of the Grid test bed. The multi-agent system (which includes the hierarchical analysis agent framework) is implemented on top of GT4. The agents (NA, JEM Agent, TA, RPA and GSA) are deployed on these nodes. These agents interact with different components of Grid. The agent framework is developed using Jade technology [1].

We have used Java codes with JOMP [4] directives as test codes, and executed them in the above environment. Next section demonstrates the results of running two test codes, one is LU factorization, and other is matrix multiplication. Experiments are primarily carried out on a HP ProLiant ML570 G4 with 4 Intel core2 Duo processors (referred as *HPserver*) and HP NetServer LH 6000 with 2 processors (referred as *Netserver*) in our local Grid test bed.

## VIII. RESULTS

We have experimented with periodic and iterative performance analysis and tuning techniques for both *InadequateResources* and *LoadImbalance* performance properties according to a given priority. As discussed in Section 6, the NA first checks the *InadequateResources* Property and then the *LoadImbalance* Property periodically. Each job starts executing with a single thread (on a single processor) and *static* scheduling strategy. Whenever, a performance problem is detected, a tuning action is taken. Fig. 1 and Fig. 2 compare the

expected completion time ( $ect$ , provided by the client), projected completion time ( $P_{act}$ , obtained after extrapolating the measured runtime after executing a portion of the job) and actual completion time ( $f_{act}$ , time needed to complete the job) for LU Factorization and Matrix Multiplication respectively.

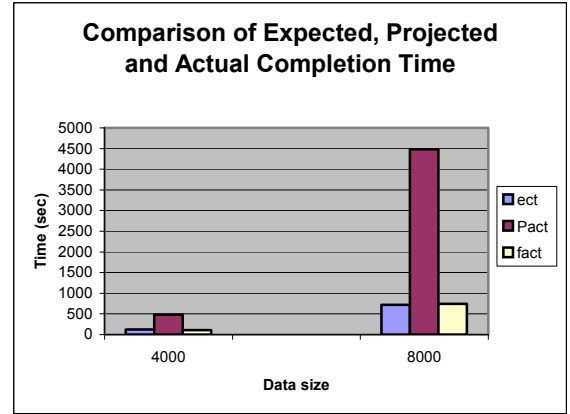


Fig. 1: Result of the experiments with LU Factorization job

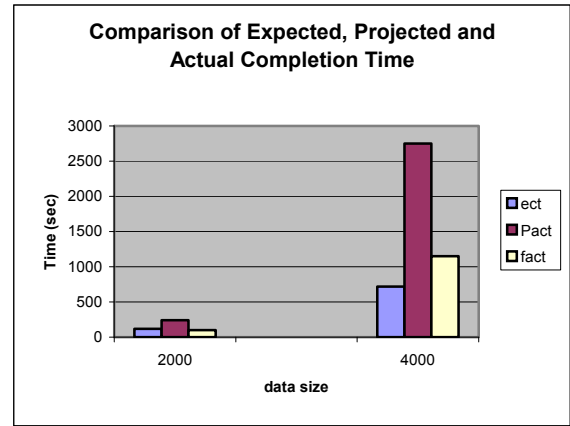


Fig. 2: Result of the experiment for Matrix Multiplication job

*Discussion* - The results demonstrate that using local tuning techniques we could achieve guaranteed performance within some tolerance level. If the tuning techniques were not applied, execution time of the job would be far more than what desired by the client (shown by  $P_{act}$ ). But tuning techniques have been applied (increasing number of processors and changing scheduling strategies) repeatedly on the jobs in order to eliminate the observed performance problems and therefore the actual completion time is close to (usually less than) the expected completion time. However we may not always achieve 100% guaranteed performance. The reason is that our algorithm supports run-time allocation of processors to the job, hence number of processors are increased and decreased dynamically (based on the severity values) and often in the last portion of the job, the guaranteed performance is slightly violated. In future, we will test our algorithm with several other test codes for guaranteed performance and

also attempt to add more sophisticated techniques. A major strength of our techniques is that, because of dynamic allocation, resource utilization cost is optimized.

## IX. CONCLUSION

An adaptive execution scheme based on a Hierarchical Analysis Agent framework is presented in this paper. The scheme supports performance analysis of components of applications executing parallelly on different resource providers and dynamically improving their execution performances. The scheme is based on the notion of performance properties, and supports both local tuning and migration of jobs depending on resource availability and the current resource usage scenarios. Some experiments have also been carried out which demonstrate the effectiveness of the framework. It has also been observed, that the agent control overheads are negligible even when multiple jobs are submitted concurrently onto the same resource provider. The work presented in this paper can be extended further to include more complicated situations and to exploit sophisticated knowledge base for appropriate tuning of the jobs.

## REFERENCES

- [1] Bellifemine, F., Poggi, A. and Rimassa, G., "Developing multi agent systems with a FIPA-compliant agent framework", in *Software - Practice & Experience*. V31: 103-128.
- [2] Berman F., et al., "New Grid Scheduling and Rescheduling Methods in the Grads Project", *International Journal of Parallel Programming*, 33(2), pp:209-229, 2005.
- [3] Berman F., et al., "Toward a Framework for Preparing and Executing Adaptive Grid Programs", *Proceedings of the 16<sup>th</sup> International Parallel and Distributed Processing Symposium*, pp:322, 2002.
- [4] Bull J.M., M.E. Kambities, "JOMP - an OpenMP -like interface for java", *Proceedings of the ACM2000 Java Grande Conference*, pp. 44-53, June 2000.
- [5] Csaba L., Lorincz, T.Kozsik, Ulbert A., Horvath Z.: "A Method for Job Scheduling in Grid based on Job Execution Status" in *Multiagent and Grid Systems*, Volume 1, Issue 3, 2005, pp: 197-208.
- [6] De Sarkar A., D. Ghosh, R. Mukhopadhyay and N. Mukherjee, "Implementation of a Grid Performance Analysis and Tuning Framework Using Jade Technology" has been accepted in *The 2008 International Conference on Grid Computing and Applications (GCA '08: July 14-17, 2008)* to be held in Las Vegas, USA.
- [7] De Sarkar, A., S. Kundu and N. Mukherjee, "A Hierarchical Agent framework for Tuning Application Performance in Grid Environment" in the proceedings of the 2<sup>nd</sup> IEEE Asia-Pacific Service Computing Conference, IEEE APSCC 2007, pp: 296-303, Tsukuba, Japan, December 11-14, 2007.
- [8] Fahringer T., M. Gerndt, G. Riley, J. Traff, "Knowledge Specification for Automatic Performance Analysis, APART Technical Report", Workpackage 2, Identification and Formalization of Knowledge, Technical Report FZJ-ZAM-IB-9918, D-52425 Julich, Nov. 1999.
- [9] Fahringer, T., Gerndt, M., Riley, G.D., and Traiff, J. L., "Formalizing OpenMP Performance Properties with ASL", In *Proceedings of the Third International Symposium on High Performance Computing* (October 16-18, 2000), ISHPC, pp. 428-439.
- [10] Ganglia Information Provider – available in <http://ganglia.sourceforge.net>
- [11] Globus Toolkit 4.0 – available in [www.globus.org/toolkit](http://www.globus.org/toolkit).
- [12] Roy S. and N. Mukherjee, "Utilizing Jini Features to Implement a Multiagent Framework for Performance-based Resource Allocation in Grid Environment", In *Proceedings of International Conference on Grid Computing and Applications (GCA'06)*. The 2006 World Congress in Computer Science, Computer Engineering, and Applied Computing, pp. 52-60, Las Vegas, June 26-29, 2006.
- [13] Roy S., M. Sarkar and N. Mukherjee, "Optimizing Resource Allocation for Multiple Concurrent Jobs in Grid Environment", in the *Proceedings of Third International Workshop on scheduling and Resource Management for Parallel and Distributed systems, SRMPDS '07*, Hsinchu, Taiwan, December 5-7, 2007.
- [14] Sakellariou R., H. Zhao, "A Low-cost Rescheduling Policy for Efficient Mapping of Workflows on Grid Systems", *Scientific Programming*, 12(4), pp:253-262, 2004.
- [15] Scherer, A., T. Gross, W. Zwaenepoel, "An Evaluation of Adaptive Execution of OpenMP Task Parallel Programs" in the proceedings of *Languages, Compilers, and Runtimes for Scalable Computing*, May 2000.
- [16] Wrzesinska G., J. Maassen, H. E Bal, "Self-adaptive Applications on the Grid" accepted in *ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PpoPP '07)*, San Jose, CA, USA, 14-17 March 2007.
- [17] Yu, Z., W. Shi, "An Adaptive Rescheduling Strategy for Grid Workflow Applications", in the *Proceedings of IPDPS*, pp: 1-8, 2007.

# On Hot-Spot Traffic Pattern of TESH Network

M.M. Hafizur Rahman\*, Yasushi Inoguchi<sup>†</sup>, Yukinori Sato<sup>†</sup>, Yasuyuki Miura<sup>††</sup>, Susumu Horiguchi<sup>‡</sup>

\*Dept. of CSE, KUET, Khulna-9203, Bangladesh

E-mail: hafiz90305@gmail.com

<sup>†</sup>Center for Information Science, JAIST, Nomi-Shi, Ishikawa 923-1292, Japan

E-mail: inoguchi@jaist.ac.jp, yukinori@jaist.ac.jp

<sup>††</sup>Shonan Institute of Technology, Fujisawa, Kanagawa 251-8511, Japan

E-mail: miu@info.shonan-it.ac.jp

<sup>‡</sup>GSIS, Tohoku University, Aoba 6-3-09, Aramaki, Sendai 980-8579, Japan

E-mail: susumu@ecei.tohoku.ac.jp

**Abstract**— In this paper, we present a deadlock-free routing algorithm for the Tori connected mESH (TESH) network using 2 virtual channels – 2 being the minimum number for dimension-order routing – and evaluate the network's dynamic communication performance under the hot-spot traffic pattern, using the proposed routing algorithm. We evaluate the dynamic communication performance of TESH, mesh, and torus networks by computer simulation. It is shown that the dynamic communication performance of the TESH network is better than that of the mesh and torus networks. It is also shown that the relative difference in maximum throughput between torus and TESH networks decreases with the increase of hot spot traffic.

**Index Terms**— TESH network, deadlock-free routing, hot-spot traffic pattern, dynamic communication performance.

## I. INTRODUCTION

Interconnection networks are the key elements for building massively parallel computers. A TESH network [1] is an interconnection network for large-scale 3D multicomputers. It consists of multiple basic modules (BMs) which are 2D-mesh networks. The BMs are hierarchically interconnected by a 2D-torus to build higher level networks.

While the details of TESH networks are available in the literature [1–4], its major features can be summarized as follows: it is hierarchical, thus allowing exploitation of computation locality as well as easy expansion up to hundreds or thousands of nodes, it permits efficient VLSI realization, it is designed for fault tolerance by making use of redundancy for defect circumvention, it is well suited for 3-D stacked implementation, and it provides good dynamic communication performance under uniform traffic pattern. In this paper, we are specifically interested to investigate the effect of hot-spot traffic pattern in the TESH network.

Deterministic, dimension-order routing is popular in multicomputers because it has minimal hardware requirements and allows the design of simple and fast routers.

Wormhole routing [5] has become the dominant switching technique because it makes latency independent of the message distance. Virtual channels [6, 7] were originally introduced to solve the problem of deadlock in wormhole-routed networks. Since the hardware cost increases as the number of virtual channels increases, the unconstrained use of virtual channels is not cost-effective in parallel computers. In this study, we use a deadlock-free routing algorithm for the TESH network using 2 virtual channels – 2 being the minimum number for dimension-order routing.

The dynamic communication performance of the TESH network with a dimension-order routing algorithm under uniform traffic pattern using 2 and 4 virtual channels and hot-spot traffic pattern using 4 virtual channels was evaluated [4], and it is proved to be better than that of conventional mesh network, but not that of torus network. However, the details analysis of the dynamic communication performance on the TESH network under the hot-spot traffic patterns has not yet been studied. The main objective of this paper is to investigate the impact of hot-spot traffic pattern on the TESH network.

The remainder of the paper is organized as follows. In Section II, we briefly describe the basic structure of the TESH network. The dynamic routing algorithm is proposed in Section III and its freedom from deadlock is also proved using minimum number of virtual channels. The dynamic communication performance of the TESH network under hot-spot traffic pattern is discussed in Section IV. Finally, Section V presents the conclusion.

## II. INTERCONNECTION OF THE TESH NETWORK

### A. Architecture of TESH Network

The *Tori-connected mESH (TESH) Network* is a hierarchical interconnection network consisting of Basic Modules (BM) that are hierarchically interconnected to form a higher level network. The BM of the TESH network

is a 2D-mesh network of size  $(2^m \times 2^m)$ . In this paper, BM refers to a Level-1 network. Successively higher level networks are built by recursively interconnecting immediately lower level subnetworks in a 2D-torus network. A higher-level network is built using a 2D-toroidal connection among  $(2^{2m})$  immediate lower level subnetworks. A TESH( $m, L, q$ ), which by definition is constructed using  $2^m \times 2^m$  basic modules, has  $L$  levels of hierarchy. The last parameter  $q$  is the inter-level connectivity.

If  $m = 2$ , the size of the BM is  $(4 \times 4)$ , and in this paper, we focus attention on  $m = 2$  i.e.,  $(4 \times 4)$  BMs. A BM of  $(4 \times 4)$  is shown in Figure 1. As seen in the figure, the BM has some free ports in the periphery for higher level interconnection. All ports of the interior Processing Elements (PEs) are used for intra-BM connections. All free ports of the exterior PEs, either one or two, are used for inter-BM connections to form higher level networks.

In principle,  $m$  could be any positive integer value. However, if  $m = 1$ , then the network degenerates to a hypercube network. Hypercube is not a suitable network, because its node degree increases along with the increase of network size. If  $m = 2$ , then it is considered the most interesting case, because it has better granularity than the large BMs. In the rest of this paper we consider  $m = 2$ , therefore, we focus on a class of TESH( $2, L, q$ ) networks.

A  $2^m \times 2^m$  BM has  $2^{m+2}$  free ports at the contours for higher level interconnection. For each higher level interconnection, a BM uses  $4 \times (2^q) = 2^{q+2}$  of its free links,  $2(2^q)$  free links for vertical interconnections and  $2(2^q)$  free links for horizontal interconnections. Here,  $q \in \{0, 1, \dots, m\}$ , is the inter-level connectivity.  $q = 0$  leads to minimal inter-level connectivity, while  $q = m$  leads to maximum inter-level connectivity.

With  $q = 0$ , for example,  $L_{max} = (2^{2-0} + 1) = 5$ . Level-5 is the highest possible level for a  $(4 \times 4)$  BM. The limitation of having a maximum level of hierarchy is not a serious constraint. For the case of  $(4 \times 4)$  BM with  $q = 0$ , a network built with the highest level, Level-5, consists of 1 million PEs. Successive higher level networks are built recursively by interconnecting the immediately lower level sub-networks. A higher-level network having  $(2^m \times 2^m)$  BM is built using a  $(2^{2m})$  subnetworks using a 2D-toroidal connection. For example, considering  $(m = 2)$  a Level-2 subnetwork, can be formed by interconnecting  $2^{2 \times 2} = 16$  BMs. Similarly, a Level-3 network can be formed by interconnecting 16 Level-2 subnetworks, and so on. This phenomena is illustrated in Figure 2, a Level-2 TESH network can be formed by interconnecting 16 BMs as a  $(4 \times 4)$  2D-torus network. Each BM is connected to its logically adjacent BMs. To avoid clutter, the wraparound links of the BMs are not shown.

### B. Addressing of Nodes

Base-4 numbers are used for convenience of address representation. As seen in Figure 1, nodes in the BM

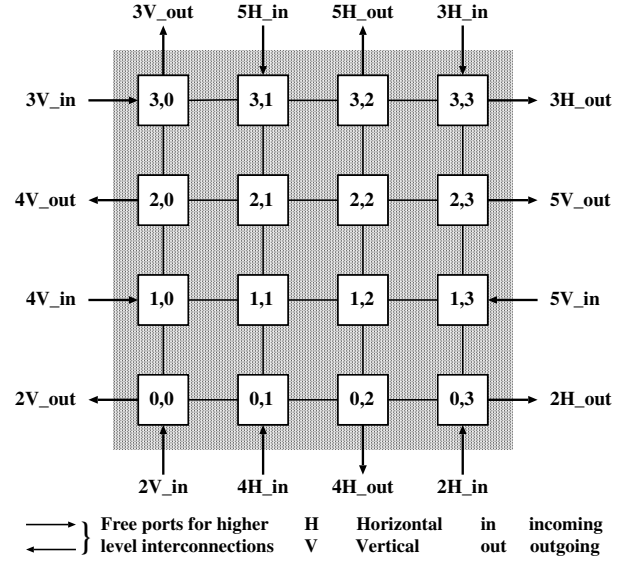


Fig. 1. Basic module of the TESH network

are addressed by two digits, the first representing the row index and the next representing the column index. More generally, in a Level- $L$  TESH network, the node address is represented by:

$$\begin{aligned}
 A &= A^L A^{L-1} A^{L-2} \dots A^2 A^1 \\
 &= a_{n-1} a_{n-2} a_{n-3} \dots a_2 a_1 a_0 \\
 &= a_{2L-1} a_{2L-2} a_{2L-3} a_{2L-4} \dots a_3 a_2 a_1 a_0 \\
 &= (a_{2L-1} a_{2L-2}) (a_{2L-3} a_{2L-4}) \dots \\
 &\quad \dots (a_3 a_2) (a_1 a_0)
 \end{aligned} \tag{1}$$

Here, the total number of digits is  $n = 2L$ , where  $L$  is the level number.  $A^L$  is the address of level  $L$  in row-major scheme, and  $(a_{2L-1} a_{2L-2})$  is the co-ordinate position of Level- $(L-1)$  for Level- $L$  network. Pairs of digits run from group number 1 for Level-1, i.e., the BM, to group number  $L$  for the  $L$ -th level. Specifically,  $i$ -th group  $(a_{2i-1} a_{2i-2})$  indicates the location of a Level- $(i-1)$  subnetwork within the  $i$ -th group to which the node belongs;  $1 \leq i \leq L$ . In a two-level network the address becomes  $A = (a_4 a_3) (a_1 a_0)$ . The first pair of digits  $(a_4 a_3)$  identifies the BM to which the node belongs, and the last pair of digits  $(a_1 a_0)$  identifies the node within that basic module.

## III. ROUTING ALGORITHM FOR TESH NETWORK

### A. Dynamic Routing Algorithm

Routing of messages in the TESH network is performed from top to bottom. That is, it is first done at the highest level network; then, after the packet reaches its highest level sub-destination, routing continues within the subnetwork to the next lower level sub-destination. This process is repeated until the packet arrives at its final destination.

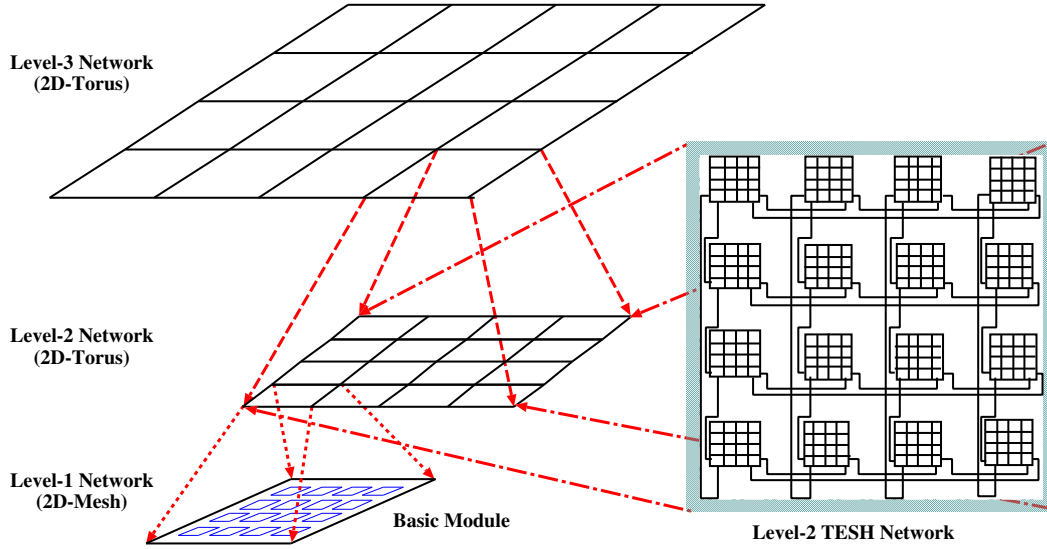


Fig. 2. Interconnection of a TESH network

When a packet is generated at a source node, the node checks its destination. If the packet's destination is the current BM, the routing is performed within the BM only. If the packet is addressed to another BM, the source node sends the packet to the outlet node which connects the BM to the level at which the routing is performed.

We have considered the dimension order routing algorithm for the TESH network. We use the following strategy: at each level, vertical routing is performed first. Once the packet reaches the correct row, then horizontal routing is performed. Routing in the TESH network is strictly defined by the source node address and the destination node address. Let a source node address be  $s = (s_{2L-1}, s_{2L-2}), (s_{2L-3}, s_{2L-4}), \dots, (s_3, s_2), (s_1, s_0)$ , a destination node address be  $d = (d_{2L-1}, d_{2L-2}), (d_{2L-3}, d_{2L-4}), \dots, (d_3, d_2), (d_1, d_0)$ , and a routing tag be  $t = (t_{2L-1}, t_{2L-2}), (t_{2L-3}, t_{2L-4}), \dots, (t_1, t_0)$ , where  $t_i = d_i - s_i$ . Figure 3 shows the routing algorithm for the TESH network.

### B. Deadlock-Free Routing

A deadlock-free routing algorithm can be constructed for an arbitrary interconnection networks by introducing virtual channels. In this section, we investigate the number of virtual channels required to make the routing algorithm deadlock-free for the TESH network. We also present a proof that the TESH network is deadlock-free by these number of virtual channels. To prove the proposed routing algorithm for the TESH network is deadlock free, we divide the routing path into three phases, as follows:

- *Phase 1*: Intra-BM transfer path from source node to the outlet node of the BM.
- *Phase 2*: Higher level transfer path.

sub-phase 2.i.1 : Intra-BM transfer to the outlet PE of Level  $(L - i)$  through the  $y$ -link.

sub-phase 2.i.2 : Inter-BM transfer of Level  $(L - i)$  through the  $y$ -link.

sub-phase 2.i.3 : Intra-BM transfer to the outlet PE of Level  $(L - i)$  through the  $x$ -link.

sub-phase 2.i.4 : Inter-BM transfer of Level  $(L - i)$  through the  $x$ -link.

- *Phase 3*: Intra-BM transfer path from the outlet of the inter-BM transfer path to the destination PE.

The proposed routing algorithm enforces some routing restrictions to avoid deadlocks [6]. Since dimension-order routing is used in the TESH network, routing of message first performed in the vertical direction and then in the horizontal direction. The interconnection of the BM is a mesh network and the higher level network is a torus connection. The number of virtual channels required to make the routing algorithm deadlock-free for the TESH network is determined using Lemma 1 and 2.

*Lemma 1*: If the message is routed in the  $y \rightarrow x$  direction in a 2D-mesh network, then the network is deadlock free with 1 virtual channels. [9]

*Lemma 2*: If the message is routed in the  $y \rightarrow x$  direction in a 2D-torus network, then the network is deadlock free with 2 virtual channels. [9]

*Theorem 1*: A TESH network with 2 virtual channels is deadlock free.

*Proof*: In phase-1 and phase-3 routing, packets are routed in the source-BM and destination-BM, respectively. The BM of the TESH network is a 2D-mesh network. According to Lemma 1, the number of necessary virtual channels for phase-1 and phase-3 is 1. The routing of the message in source-BM and destination-BM is carried out separately. The virtual channels required in phase-1

Routing Algorithm for a Level-L TESH:

```

Routing(s,d);
source,s={s2L-1,s2L-2,...,s0}; destination,d={d2L-1,d2L-2,...,d0};
tag;t2L-1,t2L-2,...,t0; group;g;

for i = 2L-1:2;
 if (di-si+2m) mod 2m <= 2m/2 then
 routedir = plus; ti= (di-si+2m) mod 2m;
 else routedir = minus; ti= 2m- (di-si+2m) mod 2m; endif;

 g = get_group_number(s,d,routedir);

 while(ti!= 0) do
 if i is even number then
 outlet_nodex= outlet_x(g,i/2+1,H,routedir);
 outlet_nodey= outlet_y(g,i/2+1,H,routedir);endif;
 if i is odd number then
 outlet_nodex= outlet_x(g,i/2+1,V,routedir);
 outlet_nodey= outlet_y(g,i/2+1,V,routedir);endif;
 BM_routing(outlet_nodex, outlet_nodey);
 } (a)

 if routedir = plus then send packet to next BM;
 else send packet to previous BM; endif; } (b)

 ti= ti- 1;
endwhile;
endfor;

BM_routing(d1,d0); → (c)
end.

BM_routing(dx, dy);
source,sx,sy; destination,dx,dy;
tag;tx,ty;

tx= dx- sx;
ty= dy- sy;
while(ty!= 0) do
 if ty> 0 then move packet to upper node; ty= ty- 1; endif;
 if ty< 0 then move packet to lower node; ty= ty+ 1; endif;
endwhile;
while(tx!= 0) do
 if tx> 0 move packet to right node; tx= tx- 1; endif;
 if tx< 0 move packet to left node; tx= tx+ 1; endif;
endwhile;
end.

```

Fig. 3. Routing algorithm of the TESH network

and *phase-3* can share each other. Intra-BM links between inter-BM links are used in sub-phases 2.i.1 and 2.i.3. Thus, sub-phases 2.i.1 and 2.i.3 utilize channels over intra-BM links, sharing the channels of either *phase-1* or *phase-3*. The free links of the BM are used in inter-BM routing, i.e., sub-phases 2.i.2 and 2.i.4, and these links form a 2D-torus network for the higher level network. According to Lemma 2, the number of necessary virtual channels for this 2D-torus network is also 2. According to the structure of the TESH network, the inter-BM links other than the wrap-around links in the end-to-end node of the higher level networks used in sub-phases 2.i.2 and 2.i.4 routes messages like mesh network. Thus, the routing of messages in this part can share the channels of either *phase-1* or *phase-3*. And the wrap-around links uses the another links of the 2D-torus in the higher level networks.

Therefore, the total number of necessary virtual channels for the whole network is 2.

#### IV. DYNAMIC COMMUNICATION PERFORMANCE

The overall performance of a massively parallel computer system is affected by the performance of the in-

terconnection network, as well as by the performance of the nodes. Continuing advances in VLSI/NoC technologies promise to deliver more power to individual nodes. On the other hand, low performance of the communication network will severely limit the speed of the entire system. Therefore, the success of massively parallel computers is highly dependent on the efficiency of their underlying interconnection networks.

##### A. Performance Metrics

The dynamic communication performance of a massively parallel computer systems is characterized by *message latency* and *network throughput*. Message latency refers to the time elapsed from the instant when the first flit is injected to the network from the source, to the instant when the last flit of the message is received at the destination. In wormhole routing, it is the average value of the time elapsed between injection of the header flit into the network from the source, and reception of the last unit of the data flit at the destination. Network throughput refers to the maximum amount of information delivered per unit of time through the network. For the network to have good performance, low latency and high throughput must be achieved. When the dynamic communication performance of various networks are compared by computer simulation and wormhole routing is used, latency is measured in simulator clock cycles and throughput is measured in flits per node and per clock cycle.

##### B. Hot-Spot Traffic Pattern

Traffic patterns are pairs of nodes that communicate. In an interconnection network, sources and destinations for messages form the traffic pattern. Traffic characteristics such as message length, message arrival time, and destination distribution have significant performance implications. Message destination distributions vary a great deal depending on the network topology and the application's mapping onto different nodes. Depending on the characteristics of an application, some nodes may communicate with each other more frequently than with others. Consequently, non-uniform traffic patterns are frequent, and cause uneven usage of traffic resources, significantly degrading the dynamic communication performance of the network.

When a hot spot occurs, a particular communication link experiences a much greater number of requests than the rest of the links – more than it can service. In a remarkably short period of time, the entire network may become congested. Hot spots are particularly insidious because they may result from the cumulative effects of very small traffic imbalances. Hot spots often occur because of the burst nature of program communication and data requirements. For generating hot spot traffic we used a model proposed by Pfister and Norton [10]. According to this model, each node first generates a random number. If that number is less than a predefined threshold, the message will be

sent to the hot-spot node. Otherwise, the message will be sent to other nodes, with a uniform distribution. Here, in uniform distribution, message destinations are chosen randomly with equal probability among the nodes.

To observe the effect of non-uniformity, in this paper, we are specifically interested in how the hot-spot traffic pattern affects the dynamic communication performance of the TESH network.

### C. Simulation Environment

The total number of nodes in a TESH is  $N = 2^{2mL}$ . If  $m = 2$  and  $L = 3$  then the total number of nodes of the TESH is 4096. That is, Level-3 TESH network has 4096 nodes.  $64 \times 64$  mesh network and  $64 \times 64$  torus network have 4096 nodes. We have evaluated the dynamic communication performance of TESH, mesh, and torus networks with 4096-nodes.

To evaluate dynamic communication performance, we have developed a wormhole routing simulator. In our simulation, we use a dimension-order routing algorithm and hot-spot traffic pattern. The dimension-order routing algorithm, which is exceedingly simple, provides the only route for the source-destination pair. Two virtual channels per physical channel are simulated, and the virtual channels are arbitrated by a round robin algorithm. For all of the simulation results, the packet size is 16 flits. Two flits are used as the header flit.

In the evaluation of dynamic communication performance, flocks of messages are sent in the network to compete for the output channels. For each simulation run, we have considered that the message generation rate is constant and the same for all nodes. Flits are transmitted at 20,000 cycles; in each clock cycle, one flit is transferred from the input buffer to the output buffer, or from output to input if the corresponding buffer in the next node is empty. Therefore, transferring data between two nodes takes 2 clock cycles.

### D. Dynamic Communication Performance Evaluation

In real application, it may happen that there are some packets (hot-spot packets) which remain in the network, and request rates are very high. Here, the simulations were carried out under the condition that, for TESH network,  $\text{Node}(n_5, n_4)(n_3, n_2)(n_1, n_0)$  are source nodes and  $\text{Node}(n_5, n_4)(0, 0)(0, 0)$  are hot-spot nodes with the hot-spot flit generation probability  $P_h = 0.02, 0.05, 0.10$ , and  $0.20$ , i.e., the hot-spot percentages are assumed to be 2%, 5%, 10%, and 20% for all networks. For mesh and torus networks, we divide the networks into  $4 \times 4$  matrix. From each part, the node which is closest to the center is assumed to be the hot-spot node, with same hot-spot percentages as listed above.

Figure 4 depicts the message latency versus network throughput curves for the hot-spot traffic pattern with different hot-spot percentage. Figure 4(a) represents the

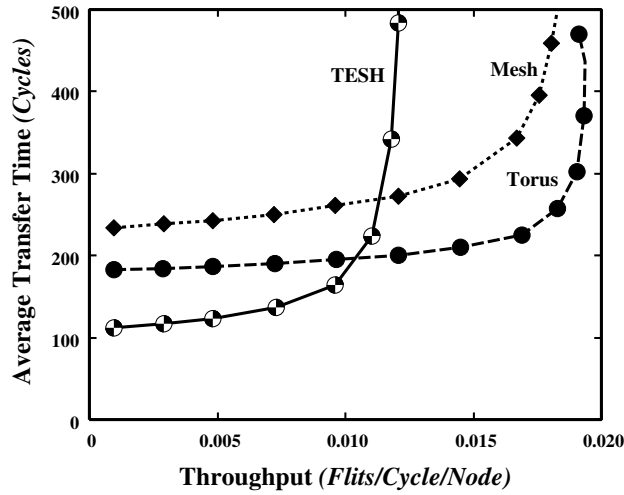
result of simulations using 2% hot-spot traffic. It is shown that the average transfer time of the TESH network is far lower than that of the mesh and torus networks. The maximum throughput of the TESH network is lower than that of those networks. Figure 4(b) represents the result of simulations using 5% hot-spot traffic. It is shown that the average transfer time of the TESH network is far lower than that of the mesh and torus networks. The maximum throughput of the TESH network is still lower than that of the mesh and torus networks. However, the relative difference in maximum throughput between TESH and mesh and torus networks is decreases with the increase of hot-spot traffic from 2% to 5%. Figure 4(c) represents the result of simulations using 10% hot-spot traffic. It is also shown that the average transfer time of the TESH network is far lower than that of the mesh and torus networks. The maximum throughput of the TESH network is higher than that of torus network and almost equal to that of mesh network. Figure 4(d) represents the result of simulations using 20% hot-spot traffic. As usual, it is shown that the average transfer time of the TESH network is far lower than that of the mesh and torus networks. The maximum throughput of the TESH network is far higher than that of the mesh and torus networks. Therefore, with the increase of hot-spot traffic percentage in hot-spot traffic pattern, the hierarchical TESH network yields better dynamic communication performance than that of the conventional mesh and torus networks.

One interesting point to be noted here is that the relative difference in maximum throughput between torus and the hierarchical TESH network decreases with the increase of hot-spot traffic, and it is shown in Figure 4(d) that with 20% hot spot traffic the maximum throughput of the TESH networks are higher than that of the torus network.

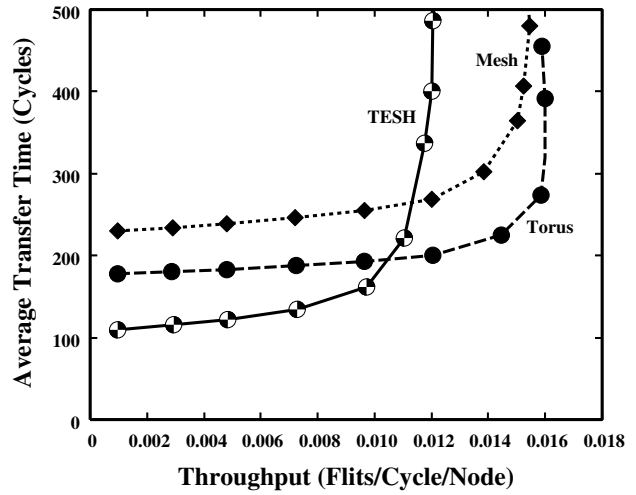
## V. CONCLUSION

A deadlock free routing algorithm using dimension-order routing with a minimum number of virtual channels has been proposed for the TESH network. It has been proved that 2 virtual channels per physical channel are sufficient for the deadlock free routing algorithm of the TESH network; 2 is also the minimum number of virtual channels for dimension-order routing. By using the routing algorithm described in this paper and the hot-spot traffic pattern, we have evaluated the dynamic communication performance of the TESH network as well as that of several other commonly used networks. The average transfer time of TESH network is lower than that of the mesh and torus networks. Maximum throughput of the TESH network is also higher than that of those networks with the increase of hot-spot traffic. A comparison of dynamic communication performance reveals that the TESH network outperforms the mesh and torus networks because it yields low latency and high throughput, which are indispensable for high performance massively parallel computers.

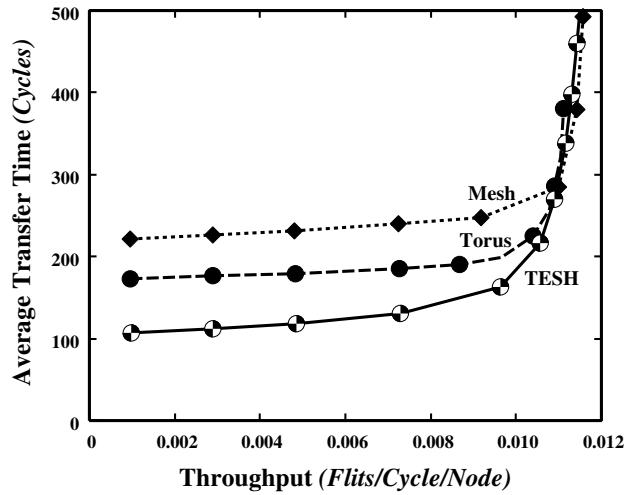




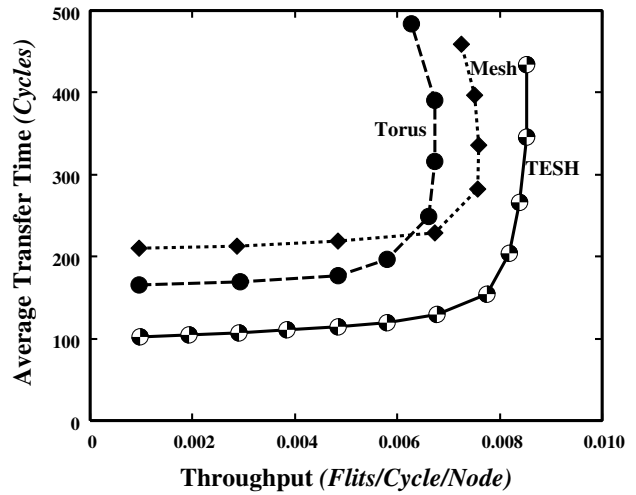
(a) 2% Hot-spot traffic



(b) 5% Hot-spot traffic



(c) 10% Hot-spot traffic



(d) 20% Hot-spot traffic

Fig. 4. Dynamic communication performance of dimension-order routing with hot spot traffic pattern on various networks: 4096 nodes, 2 VCs, 16 flits, and  $q = 0$

The performance degradation with the presence of large number of long length links is removed by replacing the long length electronic links by optical links. The performance evaluation of this opto-electronic TESH network is kept in mind as future works.

#### REFERENCES

- [1] V.K. Jain, T. Ghirmai, and S. Horiguchi, "TESH: A new hierarchical interconnection network for massively parallel computing", *IEICE Trans. on Inf. & Syst.*, vol.E80-D, no.9, pp.837-846, 1997.
- [2] V.K. Jain and S. Horiguchi, "VLSI Considerations for TESH: A New Hierarchical Interconnection Network for 3-D Integration", *IEEE Trans on VLSI Systems*, vol.6, no. 3, pp. 346-353, 1998.
- [3] M. Maziarz, and V. K. Jain, "Automatic reconfiguration and yield of the TESH multicomputer network", *IEEE Trans. on Computers*, pp. 963-972, 2002.
- [4] Y. Miura, "Wormhole Routing for Hierarchical Interconnection Networks", *Ph.D. Dissertation*, School of Information Science, JAIST, 2002.
- [5] L.M. Ni and P.K. McKinley, "A Survey of Wormhole Routing Techniques in Direct Networks", *IEEE Computer*, vol.26, no.2, pp.62-76, 1993.
- [6] W.J. Dally and C.L. Seitz, "Deadlock Free Message Routing in Multiprocessor Interconnection Networks", *IEEE Trans. on Computers*, vol.36, no.5, pp. 547-553, 1987.
- [7] W.J.Dally, "Virtual-Channel Flow Control", *IEEE Trans. on Parallel and Distributed System*, vol.3, no.2, pp.194-205, 1992.
- [8] Y.Miura and S.Horiguchi, "A Deadlock-Free Routing for Hierarchical Interconnection Network: TESH", *Proc. of the 4th Int'l. Conf. on HPC Asia*, pp.128-133, 2000.
- [9] W.J.Dally, "Performance Analysis of  $k$ -ary  $n$ -cube Interconnection Networks", *IEEE Trans. on Computers*, vol. 39, no. 6, (1990) 775-785.
- [10] G.F. Pfister and V.A. Norton, "Hot Spot Contention and Combining in Multistage Interconnection Networks", *IEEE Trans. on Computers*, vol. 34, no. 10, pp. 943-948, 1985.



# Dynamic Communication Performance of a TESH Network under the Nonuniform Traffic Patterns

M.M. Hafizur Rahman\*, Yasushi Inoguchi<sup>†</sup>, Yukinori Sato<sup>†</sup>, Yasuyuki Miura<sup>††</sup>, Susumu Horiguchi<sup>‡</sup>

\*Dept. of CSE, KUET, Khulna-9203, Bangladesh

E-mail: hafiz90305@gmail.com

<sup>†</sup>Center for Information Science, JAIST, Nomi-Shi, Ishikawa 923-1292, Japan

E-mail: inoguchi@jaist.ac.jp, yukinori@jaist.ac.jp

<sup>††</sup>Shonan Institute of Technology, Fujisawa, Kanagawa 251-8511, Japan

E-mail: miu@info.shonan-it.ac.jp

<sup>‡</sup>GSIS, Tohoku University, Aoba 6-3-09, Aramaki, Sendai 980-8579, Japan

E-mail: susumu@ecei.tohoku.ac.jp

**Abstract**—Interconnection networks play a crucial role in the performance of massively parallel computers. Hierarchical interconnection networks provide high performance at low cost by exploring the locality that exists in the communication patterns of massively parallel computer systems. The *Tori-connected mESH (TESH) Network* is a 2D-torus network of multiple basic modules, in which the basic modules are 2D-mesh networks that are hierarchically interconnected for higher level networks. In this paper, we present a deadlock-free routing algorithm for the TESH network using 2 virtual channels – 2 being the minimum number for dimension-order routing – and evaluate the network's dynamic communication performance under various nonuniform traffic patterns, using the proposed routing algorithm. We evaluate the dynamic communication performance of TESH, mesh, and torus networks by computer simulation. It is shown that the dynamic communication performance of the TESH network is better than that of the mesh and torus networks.

**Index Terms**—TESH network, deadlock-free routing, nonuniform traffic patterns, dynamic communication performance.

## I. INTRODUCTION

Interconnection networks are the key elements for building massively parallel computers. A TESH network [1] is an interconnection network for large-scale 3D-multicomputer systems. It consists of multiple basic modules (BMs) which are 2D-mesh networks. The BMs are hierarchically interconnected by a 2D-torus to build higher level networks.

While the details of TESH networks are available in the literature [1–4], its major features can be summarized as follows: it is hierarchical, thus allowing exploitation of computation locality as well as easy scalable up to thousands or tens of thousands of processors, it permits efficient VLSI realization, it is designed for fault tolerance by making use of redundancy for defect circumvention, it is well suited for 3-D stacked implementation, and

it provides good dynamic communication performance under uniform and hot-spot traffic patterns with 4 virtual channels [4]. In this paper, we are specifically interested to investigate the effect of nonuniform traffic patterns in the TESH network with minimum number of virtual channels.

Deterministic, dimension-order routing is popular in multicomputers because it has minimal hardware requirements and allows the design of simple and fast routers. Wormhole routing [5] has become the dominant switching technique because it makes latency independent of the message distance. Virtual channels [6, 7] were originally introduced to solve the problem of deadlock in wormhole-routed networks. Since the hardware cost increases as the number of virtual channels increases, the unconstrained use of virtual channels is not cost-effective in parallel computers. In this study, we use a deadlock-free routing algorithm for the TESH network using 2 virtual channels – 2 being the minimum number for dimension-order routing.

The dynamic communication performance of the TESH network with a dimension-order routing algorithm under uniform traffic pattern using 2 and 4 virtual channels and hot-spot traffic pattern using 4 virtual channels was evaluated [4], and it is proved to be better than that of conventional mesh network, but not that of torus network. However, the dynamic communication performance of the TESH network under the various nonuniform traffic patterns has not yet been evaluated. The main objective of this paper is to investigate the impact of various nonuniform traffic patterns on the TESH network.

The remainder of the paper is organized as follows. In Section II, we briefly describe the basic structure of the TESH network. The dynamic routing algorithm is proposed in Section III and its freedom from deadlock is also proved using minimum number of virtual channels. The dynamic communication performance of the TESH

network under the various nonuniform traffic patterns is discussed in Section IV. Finally, Section V presents the conclusion.

## II. INTERCONNECTION OF THE TESH NETWORK

### A. Architecture of TESH Network

The *Tori-connected mESH (TESH) Network* is a hierarchical interconnection network consisting of Basic Modules (BM) that are hierarchically interconnected to form a higher level network. The BM of the TESH network is a 2D-mesh network of size  $(2^m \times 2^m)$ . In this paper, unless specified otherwise, BM refers to a Level-1 network. Successively higher level networks are built by recursively interconnecting immediately lower level subnetworks in a 2D-torus network. A higher-level network is built using a 2D-toroidal connection among  $(2^{2m})$  immediate lower level subnetworks.

If  $m = 2$ , the size of the BM is  $(4 \times 4)$ , and in this paper, we focus attention on  $m = 2$  i.e.,  $(4 \times 4)$  BMs. A BM of  $(4 \times 4)$  is shown in Figure 1. As seen in the figure, the BM has some free ports in the periphery for higher level interconnection. All ports of the interior Processing Elements (PEs) are used for intra-BM connections. All free ports of the exterior PEs, either one or two, are used for inter-BM connections to form higher level networks.

In principle,  $m$  could be any positive integer value. However, if  $m = 1$ , then the network degenerates to a hypercube network. Hypercube is not a suitable network, because its node degree increases along with the increase of network size. If  $m = 2$ , then it is considered the most interesting case, because it has better granularity than the large BMs. In the rest of this paper we consider  $m = 2$ , therefore, we focus on a class of TESH(2,  $L$ ,  $q$ ) networks.

A  $2^m \times 2^m$  BM has  $2^{m+2}$  free ports at the contours for higher level interconnection. For each higher level interconnection, a BM uses  $4 \times (2^q) = 2^{q+2}$  of its free links,  $2(2^q)$  free links for vertical interconnections and  $2(2^q)$  free links for horizontal interconnections. Here,  $q \in \{0, 1, \dots, m\}$ , is the inter-level connectivity.  $q = 0$  leads to minimal inter-level connectivity, while  $q = m$  leads to maximum inter-level connectivity.

With  $q = 0$ , for example,  $L_{max} = (2^{2-0} + 1) = 5$ . Level-5 is the highest possible level for  $(4 \times 4)$  BM interconnection. The limitation of having a maximum level of hierarchy is not a serious constraint. For the case of  $(4 \times 4)$  BM with  $q = 0$ , a network built with the highest level, Level-5, consists of 1 million PEs. Successive higher level networks are built recursively by interconnecting the immediately lower level sub-networks. A higher-level network having  $(2^m \times 2^m)$  BM is built using a  $(2^{2m})$  subnetworks using a 2D-toroidal connection. For example, considering  $(m = 2)$  a Level-2 subnetwork, can be formed by interconnecting  $2^{2 \times 2} = 16$  BMs. Similarly, a Level-3 network can be formed by interconnecting 16 Level-

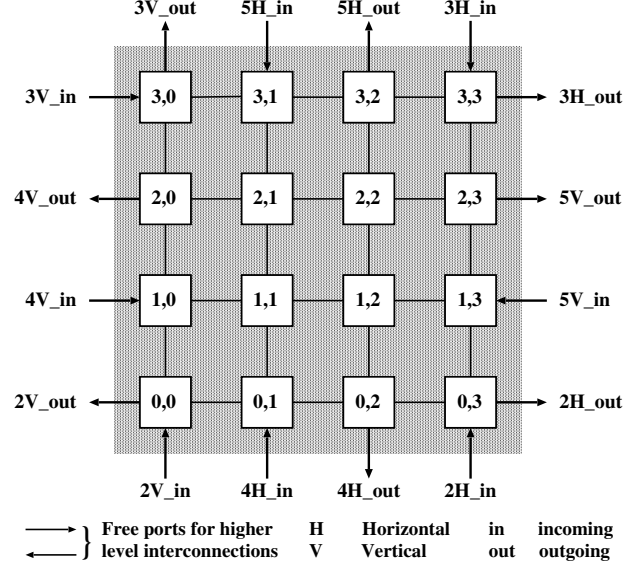


Fig. 1. Basic module of the TESH network

2 subnetworks, and so on. This phenomena is illustrated in Figure 2, a Level-2 TESH network can be formed by interconnecting 16 BMs as a  $(4 \times 4)$  2D-torus network. Each BM is connected to its logically adjacent BMs. To avoid clutter, the wraparound links of the BMs are not shown.

### B. Addressing of Nodes

Base-4 numbers are used for convenience of address representation. As seen in Figure 1, nodes in the BM are addressed by two digits, the first representing the row index and the next representing the column index. More generally, in a Level- $L$  TESH network, the node address is represented by:

$$\begin{aligned}
 A &= A^L A^{L-1} A^{L-2} \dots \dots A^2 A^1 \\
 &= a_{n-1} a_{n-2} a_{n-3} \dots \dots a_2 a_1 a_0 \\
 &= a_{2L-1} a_{2L-2} a_{2L-3} a_{2L-4} \dots \dots a_3 a_2 a_1 a_0 \\
 &= (a_{2L-1} a_{2L-2}) (a_{2L-3} a_{2L-4}) \dots \dots \\
 &\quad \dots \dots (a_3 a_2) (a_1 a_0)
 \end{aligned} \tag{1}$$

Here, the total number of digits is  $n = 2L$ , where  $L$  is the level number.  $A^L$  is the address of level  $L$  in row-major scheme, and  $(a_{2L-1} a_{2L-2})$  is the co-ordinate position of Level- $(L-1)$  for Level- $L$  network. Pairs of digits run from group number 1 for Level-1, i.e., the BM, to group number  $L$  for the  $L$ -th level. Specifically,  $i$ -th group  $(a_{2i-1} a_{2i-2})$  indicates the location of a Level- $(i-1)$  subnetwork within the  $i$ -th group to which the node belongs;  $1 \leq i \leq L$ . In a two-level network the address becomes  $A = (a_4 a_3) (a_1 a_0)$ . The first pair of digits  $(a_4 a_3)$  identifies the BM to which the node belongs, and the last pair of digits  $(a_1 a_0)$  identifies the node within that basic module.

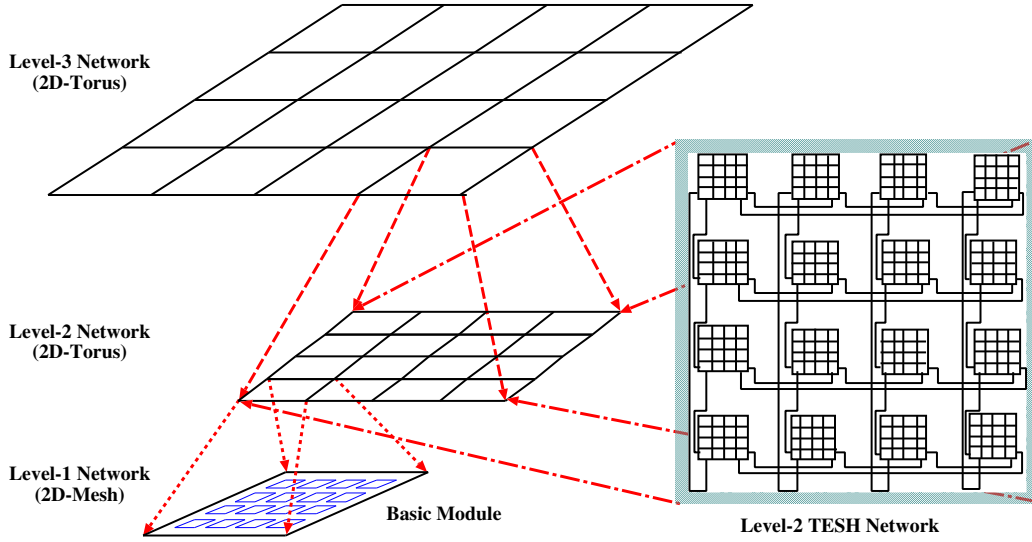


Fig. 2. Interconnection of a TESH network

### III. ROUTING ALGORITHM FOR TESH NETWORK

#### A. Dynamic Routing Algorithm

Routing of messages in the TESH network is performed from top to bottom. That is, it is first done at the highest level network; then, after the packet reaches its highest level sub-destination, routing continues within the subnetwork to the next lower level sub-destination. This process is repeated until the packet arrives at its final destination. When a packet is generated at a source node, the node checks its destination. If the packet's destination is the current BM, the routing is performed within the BM only. If the packet is addressed to another BM, the source node sends the packet to the outlet node which connects the BM to the level at which the routing is performed.

We have considered the dimension order routing algorithm for the TESH network. We use the following strategy: at each level, vertical routing is performed first. Once the packet reaches the correct row, then horizontal routing is performed. Routing in the TESH network is strictly defined by the source node address and the destination node address. Let a source node address be  $s = (s_{2L-1}, s_{2L-2}), (s_{2L-3}, s_{2L-4}), \dots, (s_3, s_2), (s_1, s_0)$ , a destination node address be  $d = (d_{2L-1}, d_{2L-2}), (d_{2L-3}, d_{2L-4}), \dots, (d_3, d_2), (d_1, d_0)$ , and a routing tag be  $t = (t_{2L-1}, t_{2L-2}), (t_{2L-3}, t_{2L-4}), \dots, (t_1, t_0)$ , where  $t_i = d_i - s_i$ . Figure 3 shows the routing algorithm for the TESH network.

#### B. Deadlock-Free Routing

A deadlock-free routing algorithm can be constructed for an arbitrary interconnection networks by introducing virtual channels. In this section, we investigate the number of virtual channels required to make the routing algorithm deadlock-free for the TESH network. We also present a

proof that the TESH network is deadlock-free by these number of virtual channels. To prove the proposed routing algorithm for the TESH network is deadlock free, we divide the routing path into three phases, as follows:

- *Phase 1*: Intra-BM transfer path from source node to the outlet node of the BM.
- *Phase 2*: Higher level transfer path.
  - sub-phase 2.i.1 : Intra-BM transfer to the outlet PE of Level  $(L - i)$  through the  $y$ -link.
  - sub-phase 2.i.2 : Inter-BM transfer of Level  $(L - i)$  through the  $y$ -link.
  - sub-phase 2.i.3 : Intra-BM transfer to the outlet PE of Level  $(L - i)$  through the  $x$ -link.
  - sub-phase 2.i.4 : Inter-BM transfer of Level  $(L - i)$  through the  $x$ -link.
- *Phase 3*: Intra-BM transfer path from the outlet of the inter-BM transfer path to the destination PE.

The proposed routing algorithm enforces some routing restrictions to avoid deadlocks [6]. Since dimension-order routing is used in the TESH network, routing of message first performed in the vertical direction and then in the horizontal direction. The interconnection of the BM is a mesh network and the higher level network is a torus connection. The number of virtual channels required to make the routing algorithm deadlock-free for the TESH network is determined using Lemma 1 and 2.

*Lemma 1*: If the message is routed in the  $y \rightarrow x$  direction in a 2D-mesh network, then the network is deadlock free with 1 virtual channels. [9]

*Lemma 2*: If the message is routed in the  $y \rightarrow x$  direction in a 2D-torus network, then the network is deadlock free with 2 virtual channels. [9]

*Theorem 1*: A TESH network with 2 virtual channels is deadlock free.

Routing Algorithm for a Level-L TESH:

```

Routing(s,d);
source,s={s2L-1,s2L-2,...,s0}; destination,d={d2L-1,d2L-2,...,d0};
tag;t2L-1,t2L-2,...,t0; group;g;

for i = 2L-1:2;
 if (di-si+2m) mod 2m <= 2m/2 then
 routedir = plus; ti= (di-si+2m) mod 2m;
 else routedir = minus; ti= 2m- (di-si+2m) mod 2m; endif;

 g = get_group_number(s,d,routedir);

 while(ti!= 0) do
 if i is even number then
 outlet_nodex= outlet_x(g,i/2+1,H,routedir);
 outlet_nodey= outlet_y(g,i/2+1,H,routedir);endif;
 if i is odd number then
 outlet_nodex= outlet_x(g,i/2+1,V,routedir);
 outlet_nodey= outlet_y(g,i/2+1,V,routedir);endif;
 BM_routing(outlet_nodex, outlet_nodey);
 } (a)

 if routedir = plus then send packet to next BM;
 else send packet to previous BM; endif; } (b)

 ti= ti- 1;
endwhile;
endfor;

BM_routing(d1,d0); → (c)
end.

BM_routing(dx, dy);
source,sx,sy; destination,dx,dy;
tag;tx,ty;

tx= dx- sx;
ty= dy- sy;
while(ty!= 0) do
 if ty> 0 then move packet to upper node; ty= ty- 1; endif;
 if ty< 0 then move packet to lower node; ty= ty+ 1; endif;
endwhile;
while(tx!= 0) do
 if tx> 0 move packet to right node; tx= tx- 1; endif;
 if tx< 0 move packet to left node; tx= tx+ 1; endif;
endwhile;
end.

```

Fig. 3. Routing algorithm of the TESH network

*Proof:* In phase-1 and phase-3 routing, packets are routed in the source-BM and destination-BM, respectively. The BM of the TESH network is a 2D-mesh network. According to Lemma 1, the number of necessary virtual channels for phase-1 and phase-3 is 1. The routing of the message in source-BM and destination-BM is carried out separately. The virtual channels required in *phase-1* and *phase-3* can share each other. Intra-BM links between inter-BM links are used in sub-phases 2.i.1 and 2.i.3. Thus, sub-phases 2.i.1 and 2.i.3 utilize channels over intra-BM links, sharing the channels of either phase-1 or phase-3. The free links of the BM are used in inter-BM routing, i.e., sub-phases 2.i.2 and 2.i.4, and these links form a 2D-torus network for the higher level network. According to Lemma 2, the number of necessary virtual channels for this 2D-torus network is also 2. According to the structure of the TESH network, the inter-BM links other than the wrap-around links in the end-to-end node of the higher level networks used in sub-phases 2.i.2 and 2.i.4 routes messages like mesh network. Thus, the routing of messages in this part can share the channels of either phase-1 or phase-3. And the wrap-around links uses the

another links of the 2D-torus in the higher level networks.

Therefore, the total number of necessary virtual channels for the whole network is 2.

#### IV. DYNAMIC COMMUNICATION PERFORMANCE

The overall performance of a massively parallel computer system is affected by the performance of the interconnection network, as well as by the performance of the nodes. Continuing advances in VLSI/NoC technologies promise to deliver more power to individual nodes. On the other hand, low performance of the communication network will severely limit the speed of the entire system. Therefore, the success of massively parallel computers is highly dependent on the efficiency of their underlying interconnection networks.

##### A. Performance Metrics

The dynamic communication performance of a massively parallel computer systems is characterized by *message latency* and *network throughput*. Message latency refers to the time elapsed from the instant when the first flit is injected to the network from the source, to the instant when the last flit of the message is received at the destination. In wormhole routing, it is the average value of the time elapsed between injection of the header flit into the network from the source, and reception of the last unit of the data flit at the destination. Network throughput refers to the maximum amount of information delivered per unit of time through the network. For the network to have good performance, low latency and high throughput must be achieved. When the dynamic communication performance of various networks are compared by computer simulation and wormhole routing is used, latency is measured in simulator clock cycles and throughput is measured in flits per node and per clock cycle.

##### B. Traffic Patterns

Traffic patterns are pairs of nodes that communicate. In an interconnection network, sources and destinations for messages form the traffic pattern. Traffic characteristics such as message length, message arrival time, and destination distribution have significant performance implications. Message destination distributions vary a great deal depending on the network topology and the application's mapping onto different nodes. Depending on the characteristics of an application, some nodes may communicate with each other more frequently than with others. Consequently, non-uniform traffic patterns are frequent, and cause uneven usage of traffic resources, significantly degrading the dynamic communication performance of the network.

When a hot spot occurs, a particular communication link experiences a much greater number of requests than the rest of the links – more than it can service. In a remarkably short period of time, the entire network may become congested. Hot spots are particularly insidious

because they may result from the cumulative effects of very small traffic imbalances. Hot spots often occur because of the burst nature of program communication and data requirements. Bit permutation and computation [11] is a class of non-uniform traffic patterns which are very common in scientific applications, where the source node sends messages to a predefined destination. Both dimension order routing and bit permutation & communication create significant congestion under dimension order routing in the network, and when congestion occurs, the network throughput decreases precipitously. To observe the effect of non-uniformity, in this paper, we are specifically interested in how the nonuniform traffic patterns affect the dynamic communication performance of the TESH network.

1) *Hot Spot traffic Pattern*: A hot spot is a node that is accessed more frequently than other nodes in the uniform traffic distribution. In hot-spot traffic pattern, each node generates a random number. If that number is less than a threshold, the message is sent to the hot spot node. Otherwise it is sent to other nodes with a uniform distribution. Here, in uniform distribution, message destinations are chosen randomly with equal probability among the nodes in the network.

2) *Complement Traffic*: The binary representation of the node address is  $b_{\beta-1}, b_{\beta-2} \dots b_1, b_0$ . In complement traffic, the node  $(b_{\beta-1}, b_{\beta-2} \dots b_1, b_0)$  communicates with the node  $(\overline{b_{\beta-1}}, \overline{b_{\beta-2}}, \dots \overline{b_2}, \overline{b_1}, \overline{b_0})$ .

3) *Perfect Shuffle Pattern*: The node with binary coordinates  $b_{\beta-1}, b_{\beta-2} \dots b_1, b_0$  communicates with the node  $(b_{\beta-2}, b_{\beta-3}, \dots b_0, a_{\beta-1})$ . That is, rotate left 1 bit.

### C. Simulation Environment

The total number of nodes in a TESH is  $N = 2^{2mL}$ . If  $m = 2$  and  $L = 3$  then the total number of nodes of the TESH is 4096. That is, Level-3 TESH network has 4096 nodes.  $64 \times 64$  mesh network and  $64 \times 64$  torus network have 4096 nodes. We have evaluated the dynamic communication performance of TESH, mesh, and torus networks with 4096-nodes.

To evaluate dynamic communication performance, we have developed a wormhole routing simulator. In our simulation, we use a dimension-order routing algorithm and various nonuniform traffic patterns. The dimension-order routing algorithm, which is exceedingly simple, provides the only route for the source-destination pair. Two virtual channels per physical channel are simulated, and the virtual channels are arbitrated by a round robin algorithm. For all of the simulation results, the packet size is 16 flits. Two flits are used as the header flit.

In the evaluation of dynamic communication performance, flocks of messages are sent in the network to compete for the output channels. For each simulation run, we have considered that the message generation rate is constant and the same for all nodes. Flits are transmitted at 20,000 cycles; in each clock cycle, one flit is transferred

from the input buffer to the output buffer, or from output to input if the corresponding buffer in the next node is empty. Therefore, transferring data between two nodes takes 2 clock cycles.

### D. Dynamic Communication Performance Evaluation

In real application, it may happen that there are some packets (hot-spot packets) which remain in the network, and request rates are very high. Here, the simulations were carried out under the condition that, for TESH network,  $\text{Node}(n_5, n_4)(n_3, n_2)(n_1, n_0)$  are source nodes and  $\text{Node}(n_5, n_4)(0, 0)(0, 0)$  are hot-spot nodes with the hot-spot flit generation probability  $P_h = 0.20$ , i.e., the hot-spot percentages are assumed to be 20% for all networks. For mesh and torus networks, we divide the networks into  $4 \times 4$  matrix. From each part, the node which is closest to the center is assumed to be the hot-spot node, with same hot-spot percentages as listed above.

Figure 4 depicts the message latency versus network throughput curves for various network models under hot-spot traffic pattern. It is shown in the Fig. 4 that the average transfer time of the TESH network is far lower than that of the mesh and torus networks. The maximum throughput of the TESH network is far higher than that of the mesh and torus networks. Therefore, with the hot spot traffic pattern, TESH network yields better dynamic communication performance than mesh and torus networks.

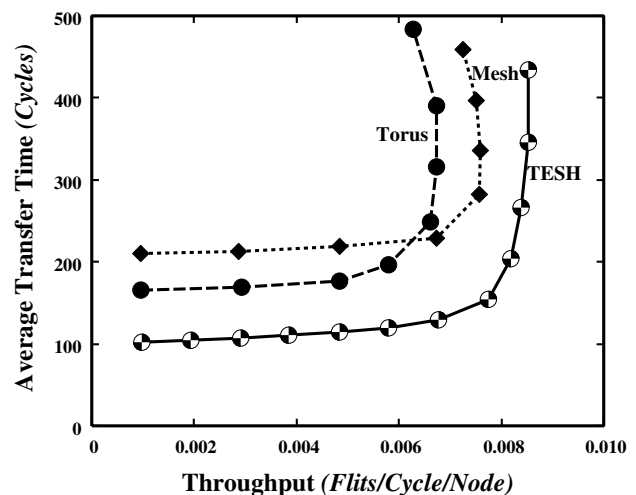


Fig. 4. Dynamic communication performance of dimension-order routing with hot spot traffic pattern on various networks: 4096 nodes, 2 VCs, 16 flits, and  $q = 0$

Figure 5 portrays the results of simulations for the various network models under complement traffic pattern. From Figure 5, it is seen that the average transfer time of the TESH network is far lower than that of the mesh network and significantly lower than that of torus network. The maximum throughput of the TESH network is far

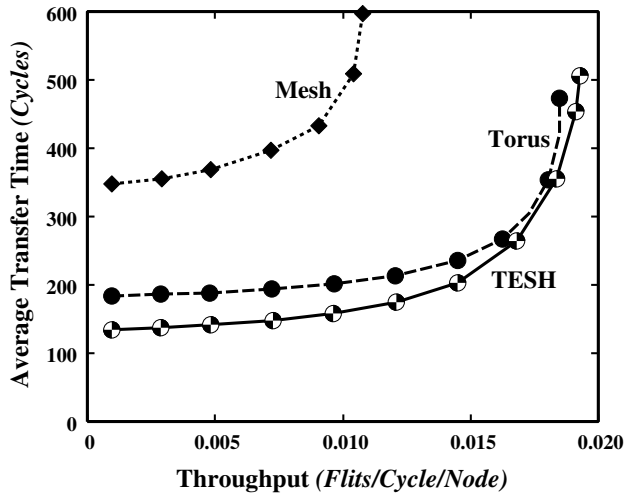


Fig. 5. Dynamic communication performance of dimension-order routing with hot spot traffic pattern on various networks: 4096 nodes, 2 VCs, 16 flits, and  $q = 0$

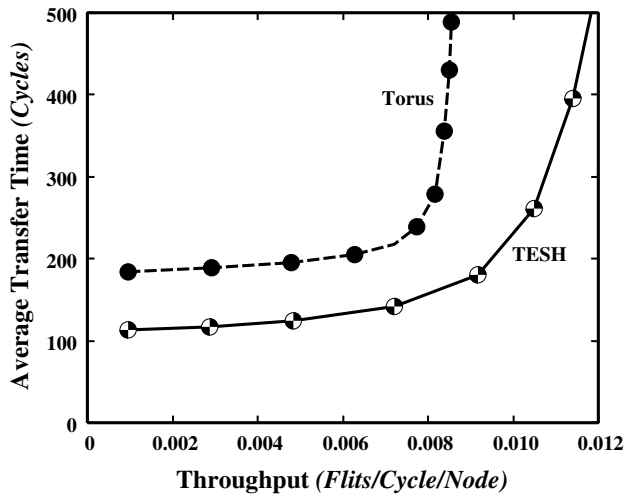


Fig. 6. Dynamic communication performance of dimension-order routing with hot spot traffic pattern on various networks: 4096 nodes, 2 VCs, 16 flits, and  $q = 0$

higher than that of mesh network and noticeably higher than that of torus network.

Figure 6 shows the comparison of dynamic communication performance between TESH and torus network under perfect shuffle traffic pattern. It is shown that the average transfer time of the TESH network is far lower than that of the torus network. And, the maximum throughput of the TESH network is far higher than that of the torus network. Therefore, TESH network yields better dynamic communication performance than that of torus network.

## V. CONCLUSION

A deadlock free routing algorithm using dimension-order routing with a minimum number of virtual channels

has been proposed for the TESH network. It has been proved that 2 virtual channels per physical channel are sufficient for the deadlock free routing algorithm of the TESH network; 2 is also the minimum number of virtual channels for dimension-order routing. By using the routing algorithm described in this paper and various nonuniform traffic pattern, we have evaluated the dynamic communication performance of the TESH network as well as that of several other commonly used networks. The average transfer time of TESH network is lower than that of the mesh and torus networks. Maximum throughput of the TESH network is also higher than that of those networks under various non uniform traffic patterns. A comparison of dynamic communication performance reveals that the TESH network outperforms the mesh and torus networks because it yields low latency and high throughput, which are indispensable for high performance massively parallel computers.

The performance degradation with the presence of large number of long length links is removed by replacing the long length electronic links by optical links. The dynamic communication performance evaluation of this opto-electronic (hybrid) TESH network is kept in mind as future works.

## REFERENCES

- [1] V.K. Jain, T. Ghirmai, and S. Horiguchi, "TESH: A new hierarchical interconnection network for massively parallel computing", *IEICE Trans. on Inf. & Syst.*, vol.E80-D, no.9, pp.837-846, 1997.
- [2] V.K. Jain and S. Horiguchi, "VLSI Considerations for TESH: A New Hierarchical Interconnection Network for 3-D Integration", *IEEE Trans on VLSI Systems*, vol.6, no. 3, pp. 346-353, 1998.
- [3] M. Maziarz, and V. K. Jain, "Automatic reconfiguration and yield of the TESH multicomputer network", *IEEE Trans. on Computers*, pp. 963-972, 2002.
- [4] Y. Miura, "Wormhole Routing for Hierarchical Interconnection Networks", *Ph.D. Dissertation*, School of Information Science, JAIST, 2002.
- [5] L.M. Ni and P.K. McKinley, "A Survey of Wormhole Routing Techniques in Direct Networks", *IEEE Computer*, vol.26, no.2, pp.62-76, 1993.
- [6] W.J. Dally and C.L. Seitz, "Deadlock Free Message Routing in Multiprocessor Interconnection Networks", *IEEE Trans. on Computers*, vol.36, no.5, pp. 547-553, 1987.
- [7] W.J.Dally, "Virtual-Channel Flow Control", *IEEE Trans. on Parallel and Distributed System*, vol.3, no.2, pp.194-205, 1992.
- [8] Y.Miura and S.Horiguchi, "A Deadlock-Free Routing for Hierarchical Interconnection Network: TESH", *Proc. of the 4th Int'l. Conf. on HPC Asia*, pp.128-133, 2000.
- [9] W.J.Dally, "Performance Analysis of  $k$ -ary  $n$ -cube Interconnection Networks", *IEEE Trans. on Computers*, vol. 39, no. 6, (1990) 775-785.
- [10] G.F. Pfister and V.A. Norton, "Hot Spot Contention and Combining in Multistage Interconnection Networks", *IEEE Trans. on Computers*, vol. 34, no. 10, pp. 943-948, 1985.
- [11] M. Grammatikakis, D. F. Hsu, M. Kratzel and J. F. Sibeyn, "Packet routing in fixed connection networks: a survey", *Journal of Parallel and Distributed Computing*, vol. 54, no. 2, pp. 77-132, 1998.
- [12] K. Bolding, M. Fulgham, and L. Synder, "The Case of Chaotic Adaptive Routing", *IEEE Trans. on Computers*, vol. 46, no. 12, pp. 1281-1292, 1997.
- [13] P.R. Miller, "Efficient Communications for Fine-Grain Distributed Computers", *Ph.D. Dissertation*, Southampton University, U.K., 1991.

# SAGLET- Secure Agent Communication Model

Elhum Nusrat, Abu Shohel Ahmed, Gazi Mushfiqur Rahman, Lafifa Jamal

Department of Computer Science and Engineering, University of Dhaka, Bangladesh

Email: elhum10@gmail.com,  
shohel\_csdu@yahoo.com, gazimushfiqurrahman@gmail.com

**Abstract** -- Mobile agent offers a new programming paradigm in which a program can move from one node to another node in a network. Since data as well as code move at the same time, security issue becomes more important than traditional RPC system. To overcome the security hurdles, this paper presents a new model based on existing Aglets architecture. A new service agent along with a policy file is incorporated with the existing Aglets architecture. This novel service agent which controls & co-ordinates all communication between the agent and the platform provides an extra layer of security by a set of API. This extra layer of security in Aglets will enhance the overall security of the platform.

**Index Terms**-- Aglets, mobile agent, SAGLET, security, service agent.

## I. INTRODUCTION

A *mobile agent* is a program, which represents a user in a computer network and is capable of migrating autonomously from node to node in the network to perform some computation on behalf of the user. Its tasks are determined by the agent application. A mobile agent can roam the network and can either come back to their 'home site' after completing its intended task. Harrison et al. [1] identified several advantages of the mobile agent paradigm. These are:

- reduce network usage
- increase asynchrony between clients and server
- add clients-specified functionality to server
- dynamically update server interfaces

Various advantages are provided by the mobile agent. Advantages like reduce network usage, increase asynchrony between clients and server are obvious for its design structure. Though the mobile agent system provides various advantages, it can also add significant problem in the field of security. Malicious agents, similar to viruses and Trojan horses, expose these hosts to the risk of system penetration. Similarly, the agents should be protected from the hosts they visit.

Aglets is a Java based platform which is one of the most popular medium for developing agents. At present, there is no encryption service in the Aglets Framework. This paper discusses issue involved protecting the aglets' **state** (the data objects the aglet carries) from tampering, eavesdropping and replay attacks and other malicious activities. To ensure security this paper proposes a new security model 'SAGLET' for Aglets

platform. The main component of SAGLET architecture is the service agent. All communication between agents and platform are done through a defined message passing standard. Service agent provides services to other visiting agent. Service agent contains modules for Authentication manager, Message communication manager and Resource Manager. Authentication manager authenticates the visiting agent. Message communication manager defines the standard for message communication between service agent and visitor agent. Resource manager allocates resource to agent according to the policy file. Public/Private key is implemented for authentication and data encryption. In addition for long message encryption, symmetric key based encryption is used. This overall approach of this proposed model will provide a new direction towards agent security.

## II. WHY AGENT SECURITY?

Though the mobile agent system provides various advantages, it can also add significant problem in the field of security. The use of mobile agents requires that each cooperating host in the distributed system provide a facility for executing them. Malicious agents, similar to viruses and Trojan horses, expose these hosts to the risk of system penetration. Four threat categories [2] are identified: (1) threats stemming from an agent attacking an agent platform, (2) an agent platform attacking an agent, (3) an agent attacking another agent on the agent platform, and (4) other entities attacking the agent system.

## III. SECURITY IN AGLETS PLATFORM

Aglets uses an organizational approach whereby all agent systems in a certain domain are deemed trustworthy, and evaluates the authenticity of the agent depending on the domain in which it has been roaming around. A user first authenticates him/her to the system, and the system then issues the credentials of the user's agent. The agent system then evaluates the authenticity of the credentials, to determine whether or not they were issued within the same domain. Host based authentication is used to identify the domain to which the communicating agents belongs. Following security features are provided by the Aglets Framework.

- a) Authentication of domains.
- b) Integrity checked communication between servers within a domain.
- c) Fine-grained authorization similar to the JDK1.2 security model.

Aglets servers are able to authenticate whether the contacting server belongs to a certain domain. All servers that belong to one domain share a secret key, and can authenticate the contacting server by means of that secret key using MAC (Message Authentication Code). In current Aglets system, agent code is only identified by its codebase. Agent code signing is not supported at all.

#### A. Limitation of Security in Aglets:

For secure agent execution, the agent system must provide the following security services:

1. Authentication of the Sender, the Manufacturer and the Owner of the Agent.
2. Authorization of the Agent (or Its Owner)
3. Secure Communication between Agent systems.
4. Non-repudiation and Auditing of data and entity.

Current Aglets platform only provides authentication within a domain by a shared key. Therefore, it cannot uniquely identify the originator of the agent. For the same reason visiting agent's identity cannot be determined. This lack of authentication also creates non-repudiation problem.

### IV. PROPOSED SECURITY MODEL– SAGLET

The proposed new model enhances security for Aglets platform by adding an extra layer of protection. No change to the Main Aglets framework source code was carried out during the project. We only provide an extension of the frame work based on the existing one. Security is implemented in three major areas.

**User Authentication** - Strong user authentication is the first line of defense for an unattended, lost, or stolen device. Multiple modes of authentication increase the work factor for an attacker. This model adds public/private key base authentication.

**Content Encryption** - Content encryption is the second line of defense for protecting sensitive information. This model uses both symmetric key algorithm Rijndael and public/private key algorithm for data encryption.

**Controls** - Policy rules are enforced in the model for all programs regardless of associated privileges, protect critical components from modification and limit access to security related information. In SAGLET model shows the way of converting non-secure Aglets platform to secure Aglets platform. An extra layer of security is developed by a set of APIs for secure Agent communication. The working model of the proposed SAGLET's agent communication is shown in Fig. 1.

#### A. Working principle

Working principle of the Origin node and Visiting node are described below:

**Origin site:** According to the proposed model origin node is the owner of the Agent. Originator will decide what information it needs to collect, where the agent should go and what resources it will acquire from remote sites. Originator node will fill the path vector of the agent with addresses where it wants to visit. This filled Itinerary vector is signed by the owner node to prevent the possible modification by visiting nodes. In addition, origin node defines the resource collection mechanism. This is done through a new Message passing standard proposed in this model. Agents pass messages according to their need and services are provided for the request with consultation of policy file. Besides, Origin site node defines security profile to identify itself to the visiting platform. Authentication of agent is implemented by Digital signature approach.

**Visitor site:** Visitor's site is the node where agent visits to collect information. In the proposed architecture all visitors' node contains a Service agent named *TesterAglet*. Service agent works as a combination of service provider, authentication manager, resource manager and access control manager. *TesterAglet* provides a set of message passing standards. The standards are defined and published in well known secure site, so service requestor agent can implement the standard message rule. The agent that needs service must communicate to the service agent through standard message communication rule which should be agreed by the communicating parties. The first role the service agent performs is authentication of the incoming agent. It checks the Itinerary object and uses the public key of the originator node of the agent to unveil the Itinerary objects authenticity. After unveiling the itinerary object, visitor can check whether it is in the route path or not.

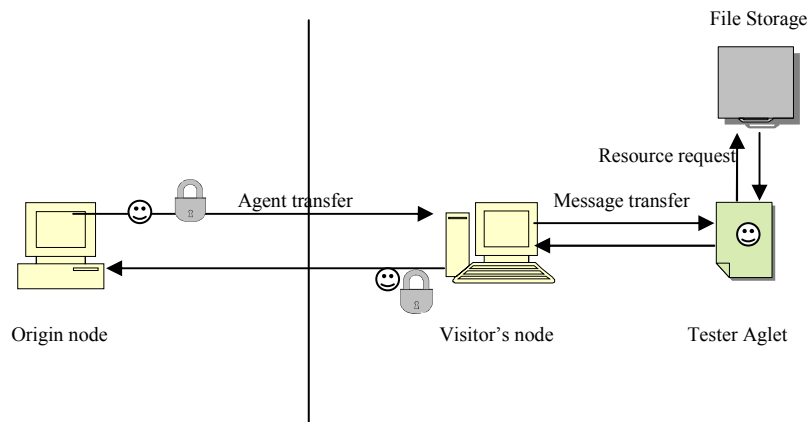
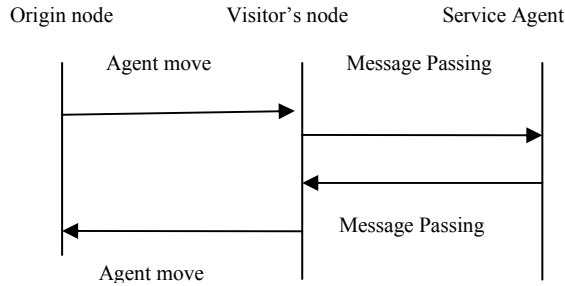


Fig. 1: Agent Data communication in SAGLET



Secondly, Service agent acts as a resource manager and access controller. For each originator node and signer Agent, Service agent contains a policy entry in the policy file. Access right and resources are allocated to the agent as defined in the policy file. Policy file is a hashed file maintained in each platform which contains policy entry for each agent in the system.



**Fig. 2: Agent Communication Model in SAGLET**

Authorization manager's job is to authenticate the visiting agent. It tests the sign of the visitor agent by consulting the security provider. Message communication manager defines the standard for message communication between service agent and visitor agent. Access control manager checks the specific access right for each signer. Resource manager handles the resource request from the visitor agent. Resources are allocated according to the access permission in the policy file.

To get resource access, Visitor node needs to communicate with service agent, authenticate itself and according to security policy the access is granted to the resources. If a visitor node has access right to resources(i.e. database, file system), it expresses its desire to the service agent, service agent do all the data collection from the home node and reply the data through message passing.

### B. SAGLET Architecture

SAGLET Architecture is based on existing Aglets [5] architecture. In addition, during implementation SAGLET uses Aglet API [5] and AgletSec API [4]. It works on application level. The main component of SAGLET architecture is service agent. Service agent contains modules for Authorization manager, Message communication manager and Privacy manager and Resource Manager. The design is modularized so that it can be enhanced later on with little effort.

### C. Message Communication Standards

Communication between service agents and visitor agent is done by message passing standard which is known to all the communicating parties. Some of the mostly used commands used by the Agent and Service provider are given below.

**msg: access-** This command is used by Service provider to check the signer of the agent.

**msg: accessDB-** This is used by visitor agent to access database resource of the visiting platform. Additional information is sent by argument setting.

**msg:accessFile:** This is used by visitor agent to access File resource of the visiting platform. Additional information is sent by argument setting.

### D. Prevention of known Attacks in proposed architecture

Now, let's see how the new model prevents known attacks on agents and agent platform.

**a) Replay attacks:** An attack in which a service is already authorized and completed is forged by another "duplicate request" in an attempt to repeat authorized commands. This model used two separate mechanisms to prevent Repay Attacks. First one is TTL (time to live): when an agent is dispatched from the Origin node, the originator defines the time duration – the time it will alive. After the time duration, the agent simply dies. So the service agent will deny any further request from this agent.

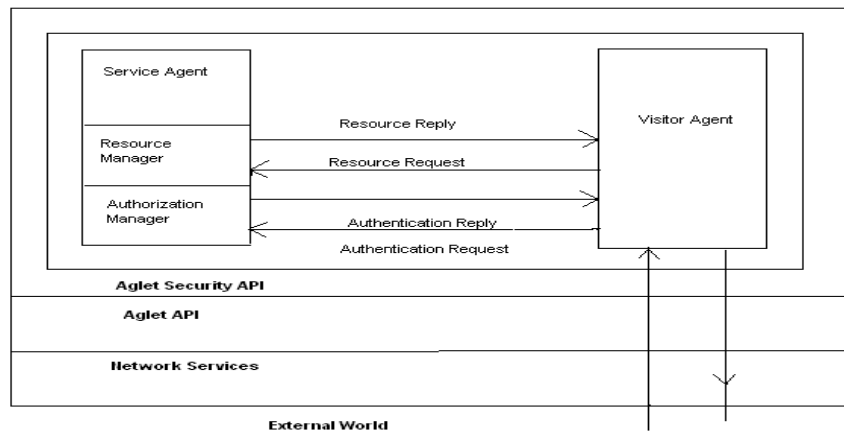
Second method is random number generation by origin node at the agent's creation moment. When an agent arrives at a visiting node, it must send the random number to the Service Agent. Then Service Agent stores this random number for a specific period defined in the TTL. So any replay attack within this time is successfully revealed through checking the random number duplication. Random number generation also prevents cut & paste attack in some cases.

**b) Eavesdropping and Wiretapping, Impersonation:** The term eavesdrops implies overhearing without expending any extra effort. An attacker can pick off the contents of a communication passing in the clear format. In this model, when information is inserted into the agent, visitor node encrypts the data by symmetric key. This symmetric key is encrypted by the agent originator node's public key, so that only the originator node can access the data. In an impersonation one host pretends to be another. Public/Private key use for object protection & identification removes impersonation problem. But if some agent can impersonate itself during key distribution process then the whole system will be compromised.

**c) Man in the Middle Attack:** In man-in-the-middle attack one entity intrudes between two other entities. Successful use of Public key cryptography will spoil the man-in-the-middle attack. Key distribution is the most sensitive part for preventing man-in-the-middle attack. If the attacker somehow passes his key instead of the intended one, then the intruder is successful in attack.

## V. BENEFITS OF PROPOSED SYSTEM OVER EXISTING SYSTEM

**Privacy:** Privacy means that sender and receiver expects confidentiality. Existing Aglets platform doesn't support any encryption API. In addition, Aglets designers are not concerned about privacy/confidentiality issues. The new proposed system provides a Security API. Using this API visiting platform encrypts the data by symmetric key during sending information



**Fig 3: SAGLET-Architecture Model**

from the visitor platform. This encryption mechanism ensures that privacy is maintained in communication system.

**Integrity:** **Integrity** implies that data must arrive at the receiver as they were sent by the sender. Data integrity is not supported directly by Aglets platform. Any agent platform or hostile agent can compromise the data it carried along the way. In our proposed system, itinerary object is signed by the originator. Agent or agent platform having the originator's public key can view the path vector. Only public key of the originator can open the itinerary object. If some middle party modifies the signed itinerary object, then sign of the object also changed. Thus it will destroy the integrity of the path vector. So, signing Itinerary object ensures integrity of the path vector. Now, when data comes back from the visiting platform to the originator node, data is encrypted. If some middle party modifies the Agent's data, the decryption key cannot open the object. So the use of encryption during agents' communication ensures integrity of data.

**Non-repudiation & Authentication:** The term **Non-repudiation** reveals the fact that a receiver must be able to prove that a received message came from a specific sender. As Code signing is not supported in Aglets, it has no way to prove that an agent comes from a specific sender. Signing any object provides non-repudiation. Signing is used in two phases, firstly, during the agent creation time; the path vector's authenticity is set. Secondly, when information comes back to the Origin site, data are encrypted by symmetric key. This key is signed by private key of the visiting platform user. As the private key of the visiting platform ensures the signer's identity, there is no way to deny his/her activity after signing. **Authentication** means to prove his/her identity. Using the above signing process the model supports authentication of both the originator and visitor nodes.

**Reduction of vulnerabilities:** Mobile agent moves from one platform to another platform. Agent moves from one platform, collects data and goes to another platform. In normal case, an agent accesses data directly from the visitor platform. It means malicious agent can maneuver the platform. In our proposed system, authentication means to prove his/her identity. All resource collection is done by service agent for the visitors. Thus it creates a sandbox, which gives enhanced security to the platform.

**Mobility:** **Mobility** is the most important feature of mobile agents. Some security models [6] proposed a set of constraints on mobility feature for security reasons. Our proposed system does not define constraint on mobility. Agents can come from any platform and free to go from that platform to another one.

**Isolation:** *Instead* of interacting with any of the user host platforms, the interacting agents migrate to a multi-agent computing service platform. All the agents that arrive at a platform act like an isolated agent, all communication is done through service agent.

## VI. PERFORMANCE OVER EXISTING SYSTEMS

In the existing Aglets platform, communication to the Agent platform is achieved by directly calling aglets API. So it requires less time to execute than proposed security model. As the new model has an extra layer of protection mechanism it certainly creates some overhead. On the other hand, SeMoA [7] uses jar signing of the code. SeMoA also includes too much security in file & folder level. Thus encryption and signing is performed many times to ensure security in SeMoA platform. That's why SeMoA is much slower than SAGLET platform. In case of another mobile agent system Concordia developed by Mitsubishi Electric, it implements a weak mobility mechanism requiring the agent to be serialized. There is also a project named TACOMA that does not rely on any specific language to guarantee the host integrity. It uses the underlying operating system mechanism to accomplish it. This characteristic increases flexibility but can also impact performance, depending on the operating system.[8]

However, our proposed system uses Digital Signature for signing itinerary object. Itinerary object only contains the URL's the agent wants to visit. This itinerary object is signed by originator platform. As we know that typically an itinerary object is not very large, signing it requires less time. When the object returns from the visitor's node to the home site, data are encrypted by symmetric key algorithm. Symmetric key encryption takes less time than asymmetric encryption. This key is signed by the visitor node. The signed key and the platform's IP address are merged together and encrypted by the public key of the originator to ensure authenticity & privacy of the key. As this key signing algorithm is applied to a small number of data, the required time is less than those algorithm which uses public/private key encryption on the whole data.

Performance enhancement can also be gained by choosing appropriate encryption and signing algorithm. For each environmental context, the new architecture can use different type of algorithm to enhance performance of the security model.

## VII. FUTURE SCOPE OF THE PROPOSED SYSTEM

1. Denial of service (DOS) attack is a major threat for agent & agent platform. Excessive message generation and too many agents in a platform can overwhelm the whole system. Therefore, in future models may contain DOS prevention. We provide a DOS prevention algorithm that is based on "Agent based Distributed Denial of Service Prevention" [9]. In existing implemented model of SAGLET, replay attack prevention technique meets some requirement of DOS prevention. Besides the model will incorporate the following strategy:

Service agent monitors message activities of the visitor agent. If visitor's agent sends too many messages or if it doesn't follow standard message rule then visitor's agent is simply discarded. Service agent also tracks all the agents in the platform, if it finds some agent is doing malicious activities, it is discarded. If platform is overwhelmed by too many agents then message is sent from service agent to the upper level router to discard agents from malicious hosts.

2. Proposed system should include key distribution mechanism in agent architecture. We need to define which key distribution model best suit the security model.

3. Enhancing message communication standard. We need to think more about formation of message standard that will lead to a security model which will serve the interest of all types of request.

4. Prevention of *cut & paste* attack.

5. Enhancing thread mechanism in agent architecture. Multiple requests can be serviced at a time if we enable thread mechanism in service agent.

6. Access control policy should be improved. Policy setting should be based on agent manufacturer, agent owner, context service provider, domain authority.

7. Security model need to include context level security [10]. When agents try to arrive from another platform, security policy should be enforced at that time.

## VIII. CONCLUSION

Agent reveals a new technological arena for Internet users. This emerging technology will replace the existing one if we can remove the flaw in it. Security is the main hurdles for implementing agent technology in large scale. To remedy this problem the paper discusses a modular architecture where service agent deals all communication between agent platform and agents. Service agent works as controls authentication and manages other resources. Policy file is used by service agent for resource management. Service agent also performs the message communication and resource allocation job. Authentication part is enhanced by using public/private key cryptography.

Combining this whole approach, this implementation will ensure security in agent communication.

## REFERENCES

- [1] Colin G. Harrison, David M. Chess, and Aaron Kershenbaum. "Mobile Agents Are they a good idea?" Technical report IBM Research Division T.J. Watson Research Center, March 1995.
- [2] NIST Special Publication 800-19, Mobile Agent Security, Wayne Jansen, Tom Karygiannis. National Institute of Standards and Technology.
- [3] Karnik, Neeran, "Security in Mobile Agent Systems", Computer Science and Engineering, University of Minnesota, 1998 (1/31).
- [4] Andrew phalidas. "Security in Aglets ", Dept. of Electronic System Engineering, University of Essex, UK.
- [5] Aglets Specification 1.1 Draft, September, 1998 . IBM corp.
- [6] Md. Nurul huda, Eiji kamioka, Shigeki Yamada. "Privacy protection in Multi-Agent based Applications", 8<sup>th</sup> ICCIT, 2005.
- [7] Ulrich Pinsdorf, Jan Peters, Mario Hoffmann, and Piklu Gupta. *Context-aware Services based on Secure Mobile Agents*. In Proceedings of 10th International Conference on Software, Telecommunications & Computer Networks (SoftCOM 2002), pp. 366-370, IEEE Communication Society, Ministry of Science and Technology Republic of Croatia and University of Split, R. Boskovic, HR-21000 Split, Croatia, October 2002. Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture. ISBN 953-6114-52-6.
- [8] D. Johansen, N.P. Sudmann, R.V. Renesse, "Performance Issues in TACOMA", in *3<sup>rd</sup> Workshop on Mobile Object Systems, 11<sup>th</sup> European Conference on Object-Oriented Programming*, Finland, Jun. 1997.
- [9] Asif Hossain khan, Abu Shohel Ahmad, Gazi Mushfiquur Rahman, Naimul Huda. "Agent based Distributed Denial of Service Prevention", 8<sup>th</sup> ICCIT, 2005.
- [10] Gunter Karjoth, Danny B. Lange, Mitsuru Oshima, "A Security Model for Aglets", IBM Research Division, Tokyo Reseach Laboratory.

# Side Channel Attack Prevention for AES Smart Card

Mohammad Zahidur Rahaman and Mohammad Akram Hossain  
rmzahid@juniv.edu, akram864@yahoo.com

**Abstract**—This paper describes an AES Smart Card implementation highly tamper resistant to side channel attacks. Smart Cards are gaining popularity in applications that require high security and store sensitive information. Modern Smart Cards, highly capable of complicated cryptology, provide a high assurance of tamper resistance and thus commonly used in payment application. Although advanced Smart Cards can not protect attackers from being defrauded by different side channel attacks. Small, embedded integrated circuits (ICs) such as Smart Cards are vulnerable to side-channel attacks (SCAs). We describe the development of differential power attacks and describe how to perform differential power kind of side-channel attack on an AES implementation, using simulated power traces. We also discuss the security prevention from such corresponding attacks, such as randomized masking techniques for software implementations.

**Index Terms**—AES, DPA, cryptanalysis, SCA, power consumption model, DPA prevention.

## I. INTRODUCTION

Most attacks in the past have centered on the mathematics of the algorithm including differential[1], [2] and linear cryptanalysis[1]. It has long been assumed that if the algorithm itself is secure then any implementations of it also secure. The truth is that a hardware implementation of an encryption scheme is weakness to several types of attacks where important information is leaked during cryptographic operations. These attacks, typically referred to as side channel[4] attacks include timing attacks, simple power analysis and differential power analysis. All of the above mentioned side channel attacks work because there is a correlation between what is measured and the internal operations of the device, which relate to the secret key [3]. Here provides a brief introduction to differential power attacks. The focus of this work will be on techniques to DPA attack on cryptographic algorithms AES implemented in hardware.

## II. CRYPTANALYSIS

The science of cryptology is driven by two opposed, but complementary disciplines, cryptography [4] and cryptanalysis. While cryptography is the science of developing mathematical algorithms and protocols for protecting information, cryptanalysis is the science of breaking cryptographic schemes. The cryptographic system is no longer considered as a black-box with two inputs and one output, but as complex system that has other inputs and other output as well. As these additional input and output channel are considered are more attack platform become possible. These additional input and output channel called side-channel. Side channel attacks

are the attacks that are based on “Side Channel Information”. Side-channel information is information that can be manipulated from cryptographic devices. Side-Channel attacks exploit information that leaks from a cryptographic device.

### A. Side-Channel Analysis (SCA)

SCA is a method for extracting information from electronic devices through analyzing their physical characteristics. In other word, Side-channel analysis is observing the behaviors of the electronic circuit. The analysis is often applied to encryption devices in order to investigate the leakage of secret data, such as keys and PIN codes. Side-channel analysis is an important device in the product development and estimate cycle of Smart Cards. Common cryptographic implementations software, dedicated hardware and Smart Card.

In side channel cryptanalysis, the attacker objective is key recovery.

### B. DPA Attack on AES S-box

AES operates on 128-bit data blocks and supports three key sizes (128, 192, and 256 bits). AES has been designed with the limited resources of a typical 8-bit processor of a Smart Card in mind and most operations are byte oriented. Therefore, 8 state bits can be predicted using a guess on 1 key byte. For the attack, the influence of 1 state byte in the datapath on the power consumption of the AES core is estimated [5].

DPA attacks to block cipher are usually performed separately on each operation block, guessing the secret key used in the first or last state round. In the AES algorithm, since most operation are byte oriented, the key is guessed 8 bits per time and represent a part of the round key, which is given by the key scheduler. For simplicity reason, the attacker attacks as target byte the output of an S-Box, so that its value can be easily predicted from the plain text in input and the key guess.

After successful attacks to every S-Box at the first or last round algorithm the secret key is completely known. Therefore in this project we describe differential power analysis attacks to a single S-box using simulated power trace.

## III. STATISTICAL POWER ANALYSIS

Statistical Power Analysis is a more powerful platform of power analysis attack. It can be further sub-categorized into Differential Power Analysis and Correlation Power Analysis.

Statistical power attacks are based on the relationship between the power consumption and the Hamming weight of data being manipulated at a given point in time. The differences in power consumption are potentially extremely small, and cannot be identified individually, as the information will be lost in the noise incurred during the acquisition process [6].

The first implementation of this type of attack is known as Differential Power Analysis [1], and works by amplifying the differences in the power consumption.

#### A. Model of the Power Consumption

Abstract model of the instantaneous power consumption of a CMOS circuit. The power consumption of a circuit at a particular time  $t$  is the sum of the power dissipated by all gates at this time. We can state this more formally as :

*Simple Power Model:* let  $t$  denote the time, and  $n(t)$  be a normal distributed random variable which represents the noise components. Let  $f(a,t)$  denote the power consumption of gate  $a$  at the time  $t$ . Then a simplified power model for the power consumption is the function

$$P_{los} = \sum_g f(a,t) + n(t) \quad (1)$$

If we consider the function  $f(a,t)$  as random variable from an unknown probability distribution, what can we say about  $P_{los}(t)$ ? If all  $f(a,t)$  are randomly and independently drawn from this probability then the Central Limit Theorem says that  $P_{los}(t)$  is normally distributed. In a DPA attacks the attacker divides the power measurements in two or more difference between these sets in order to verify the oracle [8].

*The Hamming Weight Model:* The simplest model, initially proposed in [6], [14] attaches an affine relationship between the power consumption and the Hamming weight of the data being manipulated at a given point in time. This is largely caused by the bus on the microprocessor that consumes a vast amount of power compared to any other single feature on the chip. This can be expressed as,

$$p_{los} = dH(I) + n \quad (2)$$

Where  $p_{los}$  is the power consumption,  $I$  is the value of the data being manipulated, and  $H$  is a function that calculates the Hamming weight. The variable  $n$  includes the acquisition noise and the variation from one clock cycle to another, as the commands being executed by the CPU change [6].

*The Hamming Distance Model:* Another model was proposed in [6], [15] that modified the Hamming weight model. This model is similar but with an addition of another variable,  $J$ . The model is

$$p_{los} = dH(J \oplus I) + n \quad (3)$$

This model operation is the same as described in the Hamming weight model, and the value  $J$  represents some previous state. The justification for this model is that the Hamming weight model is only valid if the bus is set to 0 between each value sent. If this value is not set to 0 then there

is a transition from state  $J$  state  $I$ , so the change in value of the bus can be modified by XOR between the two values.

Both models are used in Statistical Power Analysis. In Differential Power Analysis it is not necessary to consider the Hamming Distance Model, as this will only change the sign of the peak produced. The later model becomes more necessary when use Correlation Power Analysis, where a more complete model of the chip is required.

#### B. Differential Power Analysis (DPA)

By the word "differential", the DPA takes the differences of power consumptions among different samples with different operations or with different data processed [8], [9], [17]. For Differential Power Analysis, an attacker does not need information about the analyzed hardware nor about the points in time during processed of information. Using DPA, the adversary doesn't have to know implementation details of the attacked device. If Smart Cards are unprotected against DPA, it is possible to reveal the secret key by measuring and analyzing the power consumption .

#### C. Cryptanalysis Procedures

Having a look at the first round of AES, one can see that a bitwise XOR operation followed by substitutions with the S-boxes are performed. As described in [16], AES is defined with keylength of 128. So the first XOR and respectively the substitution will, in case of the 128 bit Key, be performed with the whole secret key. As an attacker only needs to attack an 8-bit subkey at once, which leads to a key space of  $2^8$  that means 256 possible keys, he only has to repeat his attack for the remaining 15 S-boxes to get the full 128 bit key.

An adversary can gain partial key value and then the entire key value can be extracted. The algorithm for performing a DPA attack is given now [10]:

- 1) An attacker mounts a DPA attack by running the cryptosystem for  $M$  random plaintext values with the same key  $K$ . The plaintext value is denote as  $I_i$ , where  $i$  from 1 to  $M$ . For each plaintext input, the power consumption curve  $C_{ij}$  is measured in discrete time. The power trace  $C_{ij}$  is the power consumption while processing the cryptographic algorithm. The  $i$  index corresponds to the input  $I_i$  and  $j$  index corresponds to the processing time.
- 2) Considering applied cryptographic algorithms and effects upon the secret key, a selection function is chosen. Therefore, one target bit is determined, which becomes the value of isolated bit. The particular bit  $a$  depends on some bits of the secret key in evidence.
- 3) Under the hypothesis using the selection function, attacker can calculate the theoretic value of  $a$  by known plaintext  $I_i$  and partial guessed key. By exploiting the theoretic value, the collected power traces  $C_{ij}$  are separated into two sets. The sets including all the traces are presented as the following eq<sup>n</sup>.

$$S_0 = \{C_{ij} | a = 0\} \quad (4)$$

$$S_1 = \{C_{ij} | a = 1\} \quad (5)$$

Accordingly,  $S_0$  and  $S_1$  are the sets of power traces being separated by isolated bit 0 and 1 respectively. The amount of power curves in  $S_0$  is denoted  $|S_0|$ . The sum of  $|S_0|$  power curves is denoted  $\sum c_{ij}$  and the sum of the traces from each set is equal to the number of random plaintext values. Thus

$$|S_0| + |S_1| = M. \quad (6)$$

- 4) The average power trace for each set,  $C_{a0j}$  and  $C_{a1j}$  is computed.

$$C_{a0j} = \frac{1}{|S_0|} \times \sum C_{ij} \text{ for some } C_{ij} \in S_0 \quad (7)$$

$$C_{a1j} = \frac{1}{|S_1|} \times \sum C_{ij} \text{ for some } C_{ij} \in S_1 \quad (8)$$

After subtracting the averages, differential power signal  $D_j$  is obtained in discrete time

$$D_j = C_{a0j} - C_{a1j} \quad (9)$$

The selection function is selected for the intermediate bit  $a$ . While the data including such bit is calculated, the amount of power consumption will become a little difference  $\in$  owing to the bit being zero or one. Provided that instructions computing the bit occur at time  $j'$ , the following expected differential equation results in time  $j = j'$ .

$$E\{C_{ij}|a = 0\} - E\{C_{ij}|a = 1\} = \in \quad (10)$$

When time  $j$  is not equal to  $j'$ , the cryptosystem is calculating other bits and power consumption is independent of this bit. While  $j \neq j'$ .

$$E\{C_{ij}|a = 0\} = E\{C_{ij}|a = 1\} = E\{C_{ij}\} \quad (11)$$

then

$$E\{C_{ij}|a = 0\} - E\{C_{ij}|a = 1\} = 0 \quad (12)$$

- 5) For each guessed key, attacker performs a new separation for the power traces and obtains a new bias signal. If correct value for key bit is selected, the differential power signal will present significant peaks. If guesses key value is not selected correctly, the differential power signal will not show any biases. Increasing the number of plaintexts, the differential power signal converges to the expectation equation.

$$\lim D_j = E\{C_{ij}|a = 0\} - E\{C_{ij}|a = 1\} \quad (13)$$

With enough measurements, the differential power signal is  $\in$  at time  $j'$  and will converge to zero all other times. Consequently, average can recover a very weak signal that is in the presence of noise. In reality, the DPA bias signal will not always converge to zero, but the largest biases will occur at times  $j'$ .

The actual results from monitoring Smart Card power signals is presented by D. Messerges *et al.* and the techniques improve maximize power signals is introduced by J. Kelsey *et al.* The DPA attack has been successfully applied to many existing implementations of cryptographic algorithms [10].

#### D. DPA-Prevention Techniques

Here, we discuss the different methods that have been proposed in order to mask implementations of the AES. In our masked software implementation of AES the inputs and outputs of each operation are masked additively. In the following functions we briefly describe how the masked operation of the four AES operations have been implemented.

**Masked AddRoundKey:** The AddRoundKey operation does not change the mask and therefore it does not require special attention in our masked implementation. Essentially, we use the same AddRoundKey operation as in an unmasked implementation.

**Masked SubBytes:** We mask the SubBytes operation  $SUB$  with values  $B$  and  $B'$  (the masks). Therefore, we have to derive a new masked S-box  $SUB'$  with the property that

$$SUB'(A \oplus B) = SUB(A) + B' \quad (14)$$

#### IV. DPA COUNTERMEASURE

This section discusses countermeasures that have been proposed in the literature since the announcement of DPA. Most solutions proposed involve changes in software, and to date very few have addressed solutions in hardware.

##### A. Masking Method

Random masking operations process masked input data and result in masked output data. For example, the following illustrates secure implementations of two operations, where data  $D$  masked with random value  $R_d$  results in a secure data  $D'$ .

**Boolean mask:** For bitwise Boolean function, the XOR operation is in the use of Boolean masking operation as shown in eq<sup>n</sup> 15.

$$D' = D \oplus R_d \quad (15)$$

**Arithmetic mask:** Addition or subtraction module  $2^{32}$  is employed as arithmetic masking operation as shown in eq<sup>n</sup> 16.

$$D' = (D \pm R_d) \bmod 2^{32} \quad (16)$$

By using masking operations to construct cryptosystems, the transaction between two different masking operations must be performed with secure and efficient algorithm.

##### B. Secure Masking of S-box

This section also discusses masking techniques with a special focus on the AES S-box. Masking Advanced Encryption Standard means masking the intermediate bytes which are executed in an AES mathematical calculation [11].

AES consists of four transformations, SubBytes(), ShiftRows(), MixColumns() and AddRoundKey(). Except for SubBytes(), all transformations are linear, i.e., they have the property that  $f(a+m)=f(a)+f(m)$  [11], [12]. Hence, for such transformations, it is an easy task to compute how the mask  $m$  has changed during the transformation. As a consequence, it is also simple to re-establish the original mask  $m$  after a (sequence of) transformation(s). is non-linear,

re-establishing the mask after a SubBytes() transformation is difficult [11]. This is because

$$SubBytes(a + m) = SubBytes(a) + m' \quad (17)$$

and

$$SubBytes(a + m) \neq SubBytes(a) + SubBytes(m) \quad (18)$$

Here we review the different methods that can be used to mask SubBytes().

*Re-computation of SubBytes Function:* The SubBytes() transformation can be implemented as a table lookup[12]. Suppose we want to define a masking such that

$$SubBytes(a + m) = SubBytes(a) + m \quad (19)$$

In order to achieve this with a table lookup we have to compute a corresponding table MasSubBytes() for the mask  $m$ .

---

**Algorithm 1** MasSubBytes Function

---

```

INPUT: m
OUTPUT: MasSubBytes(a+m) = SubBytes(a) + m,
for $l \leftarrow 0$ to 255 do
 MasSubBytes($l+m$) = Subbytes(l) + m
end for
return MasSubBytes

```

---

Such a table needs to be computed for each mask value  $m_l$ . If  $l$  different masks  $m$  are used, then the complexity of this procedure is  $l*256$ . If we ensure, that the same masks  $m_l$  are re-established before the SubBytes() operation in each round, the same  $l$  tables can be used throughout the complete AES calculation [11].

*Implementation of Countermeasure:* The proposed algorithm is given below

---

**Algorithm 2** MasSubBytesby using Arithmetic Mask

---

```

INPUT: m
OUTPUT: MasSubBytes
for $l \leftarrow 0$ to 255 do
 MasSubBytes (l, m) = ((SubBytes(l) $\pm$ m) $\pm$ $RND()$)
 mod 2^{32}
end for
return MasSubBytes

```

---

The above figure indicate that the algorithm of the conversion data from Boolean masking to arithmetic masking.

This algorithm process input  $m$  where

$$m' = m \oplus rnd_m. \quad (20)$$

and performs output  $Y$ , where

$$Y = m - rnd. \quad (21)$$

Above eq<sup>n</sup> employing a random select bit, the algorithm randomly processes  $m$  or ones complement of  $m$  to prevent power analysis.

## V. SIMULATION MODELS

### A. Experimental Simulation Results

In the MatLab experiment simulation environment, DPA attacks are presented and 5000 power signal are collected. For accomplishing DPA attacks, the experimental result is exhibited to monitor at least 5000 power traces. The vertical axis indicates the differential power trace (voltage). The horizontal axis indicates the amount of data obtained by MatLab simulation. The value of horizontal coordinate divided by MatLab workplace's sample creates equal execution time.

1) *Experimental result of Straightforward AES implementation:* In this attack, the differential power signals have obvious peaks. Since the guessed key is wrong, the calculated bit for partitioning power traces is not the exact intermediate data. After performing the statistic analysis, the average power consumptions for both sets are balanced and then the differential power signal will not have any obvious peaks.

#### Attacking a key Bit in a S-Box:

The most simple Differential Power Analysis attack exploits a decision function that uses a single target bit of the predicted output of an S-Box. The guess however have to consider all key values that are involved during the output computation. However it does not produce effective attacks because the single target bit is too weakly related to power consumed.

#### Attacking Several Key Bits in a S-Box:

The decision function on a single key bit can be naturally extended to work with bit pairs or even more bit values. However, when attacking several bit at the same time, the decision function may not be binary anymore. Because all relevant information for attacking the power consumption model was given in section 3.1 before, here some figures for better understanding are given. Differential curve of the calculated secret key.

2) *Experimental results of masking method:* The design experiment mounting DPA attacks on two encryption algorithms: AES and AES with masking countermeasure. After comparing each experimental result, the AES implementation with masking countermeasure indeed prevents DPA attacks. The experimental results not only demonstrated the security of AES could be achieved by using improved transformed masking method but also examine the practicality of DPA attacks on straight forward AES.

## VI. CONCLUSION

In this project, we showed that massive security increases are possible in small 8-bit processors by using of arithmetic masking method.

In the previous section, we have discussed different Statistical Power Attacks on the unmasked and on the two masked AES implementations on AES S-box output. The targets of these attacks were the S-Box operations in the first round of AES.

We have analyzed this truth based on simulations of MatLab of all S-Box implementations. These simulations have shown that the number of transitions that occur in the S-Boxes

depends on the S-Box input. Even in the masked cases, this dependency has been observed.

If no countermeasures are taken, power analysis may be very effective against many applications. So it is obvious that the DPA of the unprotected AES implementation is successful, since a distinct correlation peak is contain the correct key hypothesis. Finally we conclude that algorithmic countermeasures to secure the AES algorithm against side-channel attacks have not been resistant all first-order differential power attacks.

## REFERENCES

- [1] Johan Borst, "Block Ciphers: design, analysis and side-channel analysis", Katholieke Universiteit Leuven-Faculteit Toegepaste Wetenschappen Kasteelpark Arenberg 1, B-3001 Leuven-Heverlee (Belgium), September 2001, pp-139.
- [2] E. Biham A. Shamir, "*Differential Cryptanalysis of the Data Encryption Standard*", Springer-Verlag, 1993.
- [3] Larry T. McDaniel III, May 29,2003 Blacksburg, Virginia. "An Investigation of the Differential Power Analysis Attacks on FPGA-bases Encryption Systems", 2006.
- [4] A. Menezes, P. van Oorschot, and S. Vanstone, "Handbook of Applied Cryptography", CRC Press, 1996, <http://www.cacr.amth.uwaterloo.ca/hac>.
- [5] Kris Tiri, Electrical Engineering Dept. UC Los Angeles,USA, Ingrid Verbauwhede, Dept. ESAT/SCD-COSIC, K.U.Leuven, Belgium, "Simulation Models for Side-Channel Information Leaks", 2005.
- [6] Michael Tunstall,Royal Holloway, "secure cryptographic algorithm implementation on embedded platforms", University of London, 2006.
- [7] Kai Schramm, "DES Side channel Collision Attacks On Smart Card Implementations", Department of Electrical Engineering and Information Science, Ruhr-University, Communication Security (COSY) Group, August 7th 2002.
- [8] Manfred Aigner and Elisabeth Oswald, "Power Analysis Tutorial". Institute for Applied Information Processing and Communication, University of Technology Graz, "Inffeldgasse 16a, A-8010 Graz, Australia", "[www.iaik.tugraz.at/aboutus/people/oswald/papers/dpa\\_tutorial.pdf](http://www.iaik.tugraz.at/aboutus/people/oswald/papers/dpa_tutorial.pdf)".
- [9] Jun-Zhe Chen, "The Research of Rijndael Against Physical Cryptanalyses", Laboratory of Cryptography and Inforamtion Security (LCIS), Department of Computer Science and Inforamtion Engineering National Central University, Chung-Li, Tajwan 320, Republic of China. July 17, 2002.
- [10] Shain-Han Wu, Laboratory of Cryptography and Information Security (LCIS). Department of Computer Science and Information Engineering, National Central University, Chung-Li, Tajwan 320, Republic of China, "The Research on Masking Countermeasure Against Differential Power Analysis". June 21, 2004.
- [11] Elisabeth Oswald, Stefan Mangard and Nobert Pramstaller, "Secure and Efficient Masking of AES- A Mission Impossible?", Technical Report, IAIK-TR 2003/11/1, <http://www.iaik.tu-graz.ac.at/research/scalab/index.php>, 4th June 2004.
- [12] Federal Information Processing Standards Publication 197, Specification for the "advanced encryption standard (aes)", November 26, 2001.
- [13] James Alexander Muir, "Techniques of Side Channel Cryptanalysis", A thesis presenetd to the University of Waterloo in fulfilment of the thesis requirement for the degree of Master of Mathematics in Combinatorics and Optimization, Waterloo, Ontario, Canada, 2001.
- [14] P. Kocher, J. Jaffe, and B. Jun, Differential power analysis, Advances in Cryptology-CRYPTO'99 (M.J Wiener, ed.), Lecture Notes in Computer Science, vol.1666, Springer-Verlag, 1999, pp. 388-397.
- [15] E. Brier, C. Chevier, and F. Olivier, Correlation power analysis with a leakage model, Cryptographic Hardware and Embedded Systems-CHES 2004(M.Joye and J.-J. Quisquater, eds.), Lecture Notes in Computer Science, vol. 3156, Springer-Verlag, 2004, pp. 16-29.
- [16] National Institute of Standards and Technology, Data Encryption Standard (DES) (FIPS-46-3), A22, 1999.
- [17] Elisabeth Oswald, Stefan Mangard, Christoph Herbst, and Stefan Tillich, "Practical Second-Order DPA Attacks for Masked Smart Card Implemenations of Block Ciphers", Institute for Applied Information Processing and Communication (IAIK), Graz University of Technology, Inffeldgasse 16a, A-8010 Graz, Austria, 2006.



# An Energy-Efficient Data Security System for Wireless Sensor Network

Md. Anisur Rahman and Mitu Kumar Debnath

Department of Computer Science & Engineering, Shah Jalal University of Science & Technology,  
Sylhet, Bangladesh. {anis112, debnath68}@gmail.com.

**Abstract** — Wireless sensor networks are a new type of networked systems, characterized by severely constrained computational and energy resources, and an ad hoc operational environment. Because sensor networks may interact with sensitive data and operate in hostile unattended environments, it is imperative that these security concerns be addressed from the beginning of the system design. However, due to inherent resource and computing constraints, security in sensor networks poses different challenges than traditional network security. There is currently enormous research potential in the field of wireless sensor network security. In this paper we introduce a Message Expansion Problem which can be solved by a Data Link-layer security architecture called ‘Cipher-text Stealing method’ for wireless sensor networks. Conventional security protocols tend to be conservative in their security guarantees, typically adding 16-32 bytes of overhead. With small memories, weak processors, limited energy, and 30 byte packets, sensor networks cannot afford this luxury. TinySec addresses these extreme resource constraints with careful design; we explore the tradeoffs among different cryptographic primitives and use the inherent sensor network limitations to our advantage when choosing parameters to find a sweet spot for security, packet overhead, and resource requirements.

**Index Terms** — Security, Wireless, Sensor, Message Expansion, Cipher text Stealing.

## I. INTRODUCTION

Sensor networks consist of hundreds or thousands of self-organizing, low-power, low cost wireless nodes deployed masses to monitor and affect the environment. Potential applications include burglar alarms, inventory control, medical monitoring and emergency response, monitoring remote or inhospitable habitats, target tracking in battlefields, disaster relief networks, early fire detection in forests, and environmental monitoring. Sensor networks are typically characterized by limited power supplies, low bandwidth, small memory sizes and limited energy. This leads to a very demanding environment to provide security. Public-key cryptography is too expensive to be usable, and even fast symmetric-key ciphers must be used sparingly. Communication bandwidth is extremely dear: each bit transmitted consumes about as much power as executing 800–1000 instructions, and as a consequence, any message expansion caused by security mechanisms comes at significant cost.

We foresee TinySec will cover the basic security needs of all but the most security critical applications. As a part of this, we were motivated by an observation about 802.11 wireless networks: several studies report that 50-80% of all 802.11 wireless networks operate in the clear, without any cryptographic protection whatsoever [16, 21 and 22].

## II. SECURITY ISSUES AND THREAT MODELS

Wireless communications are difficult to protect; they are by nature a broadcast medium. In a broadcast medium, adversaries can easily eavesdrop on, intercept, inject, and alter transmitted data. In addition, adversaries are not restricted to using sensor network hardware. They can interact with the network from a distance by using expensive radio transceivers and powerful workstations. Adversaries can repeatedly send packets to drain the nodes' batteries and waste network bandwidth. Since sensor networks will be deployed in a variety of physically insecure environments, adversary can steal nodes, recover their cryptographic material, and pose as authorized nodes in the network. However, we do not address these threats. Our focus is on guaranteeing message authenticity, freshness, integrity, and confidentiality. We do not address resource consumption attacks, physical tamper resistance, or node capture attacks.

## III. MOTIVATION FOR LINK LAYER SECURITY IN SENSOR NETWORKS

In conventional networks, message authenticity, integrity, and confidentiality are usually achieved by an end-to-end security mechanism such as SSH [5], SSL [3] or IPSec [4] because the dominant traffic pattern is end-to-end communication; intermediate routers only need to view message headers and it is neither necessary nor desirable for them to have access to message bodies. This is not the case in sensor networks. The dominant traffic pattern in sensor networks is many-to-one, with many sensor nodes communicating sensor readings or network events over a multihop topology to a central base station. To prune these redundant messages to reduce traffic and save energy, sensor networks use in-network processing such as aggregation and duplicate elimination [17,18]. Since in-

network processing requires intermediate nodes to access, modify, and suppress the contents of messages, it is unlikely we can use end-to-end security mechanisms between each sensor node and the base station to guarantee the authenticity, integrity, and confidentiality of these messages. A link-layer security architecture can detect unauthorized packets when they are first injected into the network. For the above reasons, we decided on a link-layer security architecture for TinySec. Link-layer security mechanisms guarantee the authenticity, integrity, and confidentiality of messages between neighboring nodes, while permitting in-network processing.

#### IV. SECURITY GOALS

A link layer security protocol should satisfy some basic security properties such as:

- Access Control and Message Integrity
- Message Confidentiality
- Data Authenticity
- Data Freshness

#### V. OUR SECURITY PLATFORM

We expect higher level security protocols will rely on the link-layer security architecture as a primitive. For example, key distribution protocols, some of which utilize public key cryptography, could use Tiny-Sec to create secure pair wise communication between neighboring nodes. To reduce the effort in implementing these protocols, Tiny-Sec must provide the right set of interfaces to facilitate their development.

#### VI. OUR SECURITY DESIGN

##### A. Encryption

Using semantically secure encryption typically requires two design decisions: selecting an encryption scheme and specifying the IV format. In our design of TinySec, we use a specially formatted 8 byte IV, and we use cipher block chaining (CBC). In this section, we introduce the structure of our IV format and argue why CBC[8] is the most appropriate encryption scheme for sensor networks.

##### B. TinySec Initialization Vector (IV) Format

The length of our IV [2], and the way we generate IVs, can have a dramatic effect on both security and on performance. If the IV is too long, we will add unnecessary bits to the packet, which translates to a significant cost in

overall throughput and in energy drain. At the same time, if the IV is too short, we run the risk that the IV will repeat, and then our security warranty is void. How long is long enough? By the pigeonhole principle, a  $n$ -bit IV is guaranteed to repeat after  $2^n + 1$  packets are sent, no matter how we choose the IV. If we use a  $n$ -bit counter, repetitions will not occur before that point. However, for some IV generation strategies, repetitions may occur earlier. If we choose each IV as a random  $n$ -bit value, then by the birthday paradox, we expect (probabilistically) to see the first repetition after roughly  $2^{n/2}$  packets have been sent. Therefore, we use a counter in our IV, and we transmit it in the packet so that the receiver can learn the value of the counter. The structure of the IV is

$$\text{dst} || \text{AM} || l || \text{src} || \text{ctr}$$

where *dst* is the destination address of the receiver, *AM* is the active message (AM) handler type, *l* is the length of the data payload, *src* is the source address of the sender, and *ctr* is a 16 bit counter. The counter starts at 0, and the sender increases it by 1 after each message sent.

#### VII. MAJOR COMPONENTS OF RESEARCH: SECURITY ANALYSIS

##### A. Message Integrity and Authenticity

The security of CBC-MAC is directly related to the length of the MAC. Conventional security protocols use MACs of 8 or 16 bytes, again erring on the side of caution. We show here that our choice of a 4 byte MAC is not detrimental in the context of sensor networks. We can model TinySec's CBC-MAC as a function that produces 4 bytes of output [9]. Given a 4 byte MAC, then, an adversary has a 1 in  $2^{32}$  chance in blindly forging a valid MAC for a particular message. If an adversary repeatedly attempts blind forgeries, she should succeed after about  $2^{31}$  tries. Note that adversaries cannot determine offline if a forgery will be successful or not; an adversary can only test the validity of an attempted forgery by sending it to an authorized receiver. This implies she must send about  $2^{31}$  packets before she can succeed at forging the MAC for a single malicious packet. However, in sensor networks, this may provide an adequate level of security. Adversaries can try to flood the channel with forgeries, but on a 19.2kb/s channel, one can only send 40 forgery attempts per second, so sending  $2^{31}$  packets at this rate would take over 20 months. Battery-operated sensor nodes do not have enough energy to receive that many messages.

##### B. Confidentiality

The security of CBC mode encryption reduces to the length of the IV, but this security assumes no IV reuse. With an 8 byte counter or 16 byte random IV, avoiding repetition is relatively easy. Although TinySec uses an 8 byte IV, we limited ourselves to 4 additional bytes of overhead in the

packet to represent the IV. The other 4 bytes of the IV borrow from the existing header fields: the destination address, the AM type, and the length. TinySec partitions the last 4 bytes of the IV into  $\text{src}||\text{ctr}$ , where  $\text{src}$  is the source address of the sender and  $\text{ctr}$  is a 16 bit counter starting at 0. Having every node generate IVs from a 4 byte counter starting at 0 is bad idea; the first packet sent from one node will reuse the same IV as the first packet sent from all other nodes. Generating IVs randomly is also a poor choice; the birthday paradox implies we can expect a collision after  $2^{16}$  total packets in the network. Our format for the last 4 bytes strives to maximize the number of packets each node can send before there is a global repetition of an IV value. The  $\text{src}||\text{ctr}$  format of the last 4 bytes guarantees each node can send at least  $2^{16}$  packets before IV reuse occurs. For a network of  $n$  nodes all sending packets at approximately the same rate, this results in about  $n \cdot 2^{16}$  total packets before we expect an instance of IV reuse. At one packet per minute per node, IV reuse will not occur for over 45 days. IV reuse is only a problem when we reuse the same IV with the same key. The first 4 bytes of the IV,  $\text{dst}||\text{AM}||\text{l}$ , help prevent information leakage during the unfortunate event of a counter on a node repeating. If a counter value for a particular source address is reused, there is only potential information leakage when the  $\text{dst}||\text{AM}||\text{l}$  values are exactly the same for both messages. CBC mode for encryption enables Tiny-Sec to provide strong confidentiality guarantees for applications in sensor networks.

#### C. Message Expansion problem in TinySec using CBC mode

Naively using CBC mode for encryption with a 8-byte block cipher results in cipher-texts which are multiples of 8 bytes. This may result in message expansion[7], which increases power consumption. We use a technique known as cipher-text stealing to ensure the cipher-text is the same length as the underlying plaintext. Encrypting data payloads of less than 8 bytes will produce a cipher-text of 8 bytes because cipher-text stealing requires at least one block of cipher-text. However, the fixed overhead of sending a message generally discourages short messages.

#### D. Analytical result for message expansion overhead

Our communication bandwidth is extremely short. So communication is more expensive. It has seen that each bit transmitted consumes power as 800-100 CPU instructions. So if we send a 30 byte packet and for encryption process if the overhead is 8 byte per packet, then the percentage of overhead for the whole packet is approximately 26% which is too much expensive in case of energy efficiency.

### VIII. PROPOSED TECHNOLOGY: CIPHER-TEXT STEALING METHOD

#### A. Cipher text stealing

In cryptography, ciphertext stealing (CTS) [6] is a general method of using a block cipher mode of operation that allows for processing of messages that are not evenly divisible into blocks without resulting in any expansion of the ciphertext, at the cost of slightly increased complexity of the MAC.

#### B. General characteristics

Ciphertext stealing is the technique of altering processing of the last two blocks of plaintext, resulting in a reordered transmission of the last two blocks of ciphertext and no ciphertext expansion. This is accomplished by padding the last block (which is possibly incomplete) with the high order bits from the second to last ciphertext block (stealing the ciphertext from the second to last block). The (now full) last block is encrypted, and then exchanged with the second to last ciphertext block, which is then truncated to the length of the final plaintext block, removing the bits that were stolen, resulting in ciphertext of the same length as the original message size. In all cases, the processing of all but the last two blocks is unchanged. Cipher text stealing for CBC mode doesn't necessarily require the plaintext to be longer than one block.

#### C. Cipher text stealing mode description

In order to encrypt or decrypt data, use the standard block cipher mode of operation on all but the last two blocks of data. The following steps describe how to handle the last two blocks of the plaintext, called  $P_{n-1}$  and  $P_n$ , where the length of  $P_{n-1}$  equals the block size of the cipher in bits,  $B$ , the length of the last block,  $P_n$ , is  $M$  bits, and  $K$  is the key that is in use.  $M$  can range from 1 to  $B$ , inclusive, so  $P_n$  could possibly be a complete block. The CBC mode description also makes use of the ciphertext block just previous to the blocks concerned,  $C_{n-2}$ , which may in fact be the IV if the plaintext fits within two blocks.

For this description, the following functions and operators are used:

- $\text{Head}(\text{data}, a)$ : returns the first  $a$  bits of the 'data' string.
- $\text{Tail}(\text{data}, a)$ : returns the last  $a$  bits of the 'data' string.
- $\text{Encrypt}(K, \text{data})$ : use the underlying block cipher in encrypt mode on the 'data' string using the key  $K$ .
- $\text{Decrypt}(K, \text{data})$ : use the underlying block cipher in decrypt mode on the 'data' string using the key  $K$ .
- XOR: Bitwise Exclusive-OR. Equivalent to bitwise addition without use of a carry bit.
- $||$ : Concatenation operator. Combine the strings on either side of the operator.
- $0^a$ : a string of  $a$  0 bits.

#### D. CBC ciphertext stealing

In CBC, there is already interaction between processing of different adjacent blocks, so CTS has less conceptual impact in this mode. Error propagation is affected.

##### a) CBC encryption steps

- $X_{n+1} = P_{n+1} \text{ XOR } C_{n+2}$ . Exclusive-OR  $P_{n+1}$  with the previous ciphertext block,  $C_{n+2}$ , to create  $X_{n+1}$ . This is equivalent to the behavior of standard CBC mode.
- $E_{n+1} = \text{Encrypt}(K, X_{n+1})$ . Encrypt  $X_{n+1}$  to create  $E_{n+1}$ . This is equivalent to the behavior of standard CBC mode.
- $C_n = \text{Head}(E_{n+1}, M)$ . Select the first  $M$  bits of  $E_{n+1}$  to create  $C_n$ . The final ciphertext block,  $C_n$ , is composed of the leading  $M$  bits of the second-to-last ciphertext block. In all cases, the last two blocks are sent in a different order than the corresponding plaintext blocks.
- $P = P_n \parallel 0^{B \cdot M}$ . Pad  $P_n$  with zeros at the end to create  $P$  of length  $B$ . The zero padding in this step is important for step 5.
- $D_n = E_{n+1} \text{ XOR } P$ . Exclusive-OR  $E_{n+1}$  with  $P$  to create  $D_n$ . For the first  $M$  bits of the block, this is equivalent to CBC mode; the first  $M$  bits of the previous block's ciphertext,  $E_{n+1}$ , are XORed with the  $M$  bits of plaintext of the last plaintext block. The zero padding of  $P$  in step 4 was important, because it makes the XOR operation's effect on the last  $B \cdot M$  bits equivalent to copying the last  $B \cdot M$  bits of  $E_{n+1}$  to the end of  $D_n$ . These are the same bits that were stripped off of  $E_{n+1}$  in step 3 when  $C_n$  was created.
- $C_{n+1} = \text{Encrypt}(K, D_n)$ . Encrypt  $D_n$  to create  $C_{n+1}$ . For the first  $M$  bits, this is equivalent to what would happen in CBC mode (other than the ciphertext ordering). For the last  $B \cdot M$  bits, this is the second time that these data have been encrypted under this key (It was already encrypted in the production of  $E_{n+1}$  in step 2).

##### b) CBC decryption steps

- $D_n = \text{Decrypt}(K, C_{n+1})$ . Decrypt  $C_{n+1}$  to create  $D_n$ . This undoes step 6 of the encryption process.
- $C = C_n \parallel 0^{B \cdot M}$ . Pad  $C_n$  with zeros at the end to create a block  $C$  of length  $B$ . We are padding  $C_n$  with zeros to help in step 3.
- $X_n = D_n \text{ XOR } C$ . Exclusive-OR  $D_n$  with  $C$  to create  $X_n$ . Looking at the first  $M$  bits, this step has the result of XORing  $C_n$  (the first  $M$  bits of the encryption process'  $E_{n+1}$ ) with the (now decrypted)  $P_n \text{ XOR } \text{Head}(E_{n+1}, M)$  (see steps 4-5 of the encryption process). In other words, we have CBC

decrypted the first  $M$  bits of  $P_n$ . Looking at the last  $B \cdot M$  bits, this recovers the last  $B \cdot M$  bits of  $E_{n+1}$ .

- $P_n = \text{Head}(X_n, M)$ . Select the first  $M$  bits of  $X_n$  to create  $P_n$ . As described in step 3, the first  $M$  bits of  $X_n$  contain  $P_n$ . We queue this last (possibly partial) block for eventual output.
- $E_{n+1} = C_n \parallel \text{Tail}(X_n, B \cdot M)$ . Append the tail ( $B \cdot M$ ) bits of  $X_n$  to  $C_n$  to create  $E_{n+1}$ . As described in step 3,  $E_{n+1}$  is composed of all of  $C_n$  (which is  $M$  bits long) appended with the last  $B \cdot M$  bits of  $X_n$ . We reassemble  $E_{n+1}$  (which is the same  $E_{n+1}$  seen in the encryption process) for processing in step 6.
- $X_{n+1} = \text{Decrypt}(K, E_{n+1})$ . Decrypt  $E_{n+1}$  to create  $X_{n+1}$ . This reverses encryption step 2.  $X_{n+1}$  is the same as in the encryption process.
- $P_{n+1} = X_{n+1} \text{ XOR } C_{n+2}$ . Exclusive-OR  $X_{n+1}$  with the previous ciphertext block,  $C_{n+2}$ , to create  $P_{n+1}$ . Finally, we reverse the XOR step from step 1 of the encryption process.

#### E. Ciphertext stealing encryption and decryption

##### a) CBC cipher text stealing encryption using a standard CBC interface:

- Pad the last partial plaintext block with 0.
- Encrypt the whole padded plaintext using the standard CBC mode.
- Swap the last two ciphertext blocks.
- Truncate the ciphertext to the length of the original plaintext.

##### b) CBC cipher text stealing decryption using a standard CBC interface:

- $D_n = \text{Decrypt}(K, C_{n+1})$ . Decrypt the second to the last ciphertext block.
- $C_n = C_n \parallel \text{Tail}(D_n, B \cdot M)$ . Pad the ciphertext to the nearest multiple of the block size using the last  $B \cdot M$  bits of block cipher decryption of the second-to-last ciphertext block.
- Swap the last two ciphertext blocks.
- Decrypt the ciphertext (except the now last block) using the standard CBC mode.
- Truncate the plaintext to the length of the original ciphertext.

#### F. Keying Mechanism

The appropriate keying mechanism for a particular network depends on several factors such as the target threat model, ease of use, and the networking and security requirements of applications. In cryptographic design, a good rule of thumb is to use different keys for different applications. When we refer to a TinySec key, we mean a pair of Skipjack [8] keys, one for encrypting data, and one for computing MACs. The simplest keying mechanism is to

use a single network wide TinySec key among the authorized nodes.

#### IX. OUR PROPOSED IMPLEMENTATION STRATEGY

TinySec currently runs on the Mica, Mica2, and Mica2 Dot platforms, each using Atmel processors; the Mica sensor node uses the RFM TR1000 radio, while the Mica2 and Mica2Dot nodes use the Chipcon CC1000 radio [15]. Additionally, TinySec is integrated into the TOSSIM simulator [12], which runs on an Intel x86 platform. Others have ported TinySec to a Texas Instruments microprocessor.

##### A. Components of a Sensor node

Researchers at UC Berkeley were the first to come up with the concept of intelligent wireless sensors and sensor networks. They have developed sensor devices called 'motes' and an operating system called 'Tiny OS' [9] which is specifically designed to run on the motes. The term 'mote' is used to refer the sensor node as whole. However, it's the hardware component of the sensor node. The software component includes the TinyOS operating system and various applications and protocols developed on top of it.

##### B. Motes

A mote is essentially a microcontroller along with a number of sensors attached to it. Some motes have integrated sensors. These motes are very small in size due to the nature and purpose of sensor networks applications. Some examples of motes are Telos, Rene, Mica and Mica2. Mica2 is the most representative of all and the majority of the research done today in the field of wireless sensor networks use Mica2 motes [19].

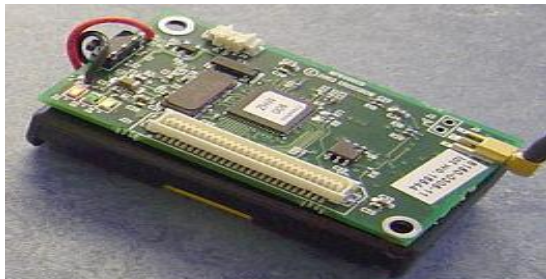


Fig. 1. Mica2

##### C. Mica2 Components

The main block of Mica2 is the microcontroller or processor - Atmel Atmega128 AVR [11]. AVR is an 8-Bit Harvard architecture, with separate instruction and data memory. AVR micro controllers provide several sleep modes. The purpose of these modes is to provide a way of suspending program execution when necessary, thereby reducing power consumption. The microcontroller unit (MCU)[10] is

responsible for control of the sensors and the execution of communication protocols and signal processing algorithms on the gathered sensor data. The microcontroller interfaces with Radio, LEDS, Flash Memory and Sensor board/Programming interface [14].

##### D. Sensing Hardware

The continuous size and cost reduction of electronic devices is gradually making the vision of ubiquitous wireless sensors networks a reality [1]. The modular design of the motes allows a wide range of analog and digital sensors to be attached to the mote. However, the reference sensor board for the mica platform is the "Mica Sensor board". A variety of these sensor boards are available. They allow for a range of different sensing modalities as well as interface.

##### E. TinyOS

TinyOS is a small event-driven, component based operating system, designed specifically for supporting the concurrency intensive operations required by networked sensors with minimum hardware requirements. The TinyOS framework contains numerous pre-built sensor applications and algorithms e.g. multihop ad-hoc routing [13] and supports different sensor node platforms. TinyOS is optimized in terms of memory usage and energy efficiency. It provides defined interfaces between the components which reside in neighboring layers.

#### X. CONCLUSION

We explored some of the tradeoffs between performance, transparency, and cryptographic security, and propose a design that meets the needs in sensor networks. After going through it we have found that it has been experiencing a drawback of message expansion which in turn reduces the power of resource-constrained motes. As a result, whole sensor network will starve for adequate energy and it will gradually become a vulnerable network. We proposed a system that will retain the size of message after encryption. And it is "ciphertext stealing system". This way unnecessary power consumption due to encryption will be reduced.

#### ACKNOWLEDGEMENT

All complements to the almighty Allah who gives us the ability and knowledge to do this hard work. We express our greatest gratitude, appreciation and profound respect to our supervisor Md. Khairullah, Lecturer, Dept. of CSE, Shah Jalal University of Science & Technology, Sylhet, Bangladesh for his continuous, proper & perfect guidance and invaluable suggestion.

## REFERENCES

- [1] Crossbow technology inc. <http://www.xbow.com>.
- [2] [http://en.wikipedia.org/wiki/Initialization\\_vector](http://en.wikipedia.org/wiki/Initialization_vector)
- [3] OpenSSL. <http://www.openssl.org>.
- [4] Security architecture for the Internet Protocol. RFC 2401, November 1998.
- [5] <http://www.ssh.com>.
- [6] <http://everything2.com/e2node/ciphertext%2520stealing>.
- [7] <http://www.win.tue.nl/~sjouke/publications/papers/sensorauth.pdf>.
- [8] <http://www.tropsoft.com/strongenc/skipjack.htm>
- [9] <http://www.cs.berkeley.edu/~pal/research/tossim.html>.
- [10] <http://archive.chipcenter.com/circuitcellar/march01/c0301cg1.htm>, <http://www.auditmypc.com/acronym/MCU.asp>, <http://www.freepatentsonline.com/5371869.html>.
- [11] <http://www.atmel.com/products/AVR/>, [http://cgi.ebay.com/ATMEL-Atmega-AVR-AVRISP-V2-0-STK500-USB-ISP-Programmer\\_W0QQitemZ290268160609QQihZ019QQcategoryZ4661QQtcZphotoQQcmdZViewItem](http://cgi.ebay.com/ATMEL-Atmega-AVR-AVRISP-V2-0-STK500-USB-ISP-Programmer_W0QQitemZ290268160609QQihZ019QQcategoryZ4661QQtcZphotoQQcmdZViewItem).
- [12] <http://www.tinyos.net/tinyos-1.x/doc/tutorial/lesson5.html>, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.2.6175>.
- [13] <http://www.isi.edu/~johnh/PAPERS/Xu01a.pdf>.
- [14] <http://news.thomasnet.com/fullstory/488967>.
- [15] [http://findarticles.com/p/articles/mi\\_hb4804/is\\_ai\\_n174439](http://findarticles.com/p/articles/mi_hb4804/is_ai_n174439)
- [16] Chris Hurley. The worldwide wardrive: The myths, the misconceptions, the truth, the future. In Defcon 11, August 2003.
- [17] Samuel R. Madden, Michael J. Franklin, Joseph M. Hellerstein, and Wei Hong. TAG: A tiny aggregation service for ad-hoc sensor networks. In The Fifth Symposium on Operating Systems Design and Implementation (OSDI 2002), 2002.
- [18] Samuel R. Madden, Robert Szewczyk, Michael J. Franklin, and David Culler. Supporting aggregate queries over ad-hoc wireless sensor networks. In Workshop on Mobile Computing and Systems Applications, 2002.
- [19] David Malan, Matt Welsh, and Michael D. Smith. A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography. In First IEEE International Conference on Sensor and Ad Hoc Communications and Networks, October 2004.
- [20] Adrian Perrig, Robert Szewczyk, Victor Wen, David Culler, and J.D. Tygar. SPINS: Security protocols for sensor networks. In The Seventh Annual International Conference on Mobile Computing and Networking (MobiCom 2001), 2001.
- [21] Peter Shipley. Open WLANs: the early results of wardriving, 2001.
- [22] Peter Shipley, 2003. personal communication.
- [23] Robert Szewczyk, Joseph Polastre, Alan Mainwaring, and David Culler. Lessons from a sensor network expedition. In First European Workshop on Wireless Sensor Networks (EWSN '04), January 2004.

# MIMOME channel secrecy capacity

Mohammad Rakibul Islam <sup>1</sup>, Jinsang Kim <sup>1</sup>, Md. Shamsul Arefin <sup>2</sup>

<sup>1</sup>Dept. of Electronics and Radio Engineering, Kyung Hee University  
1 Seocheon, Kihung, Yongin, Gyeonggi, 449-701, Korea

<sup>2</sup>Dept. of Electrical and Electronic Engineering  
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh  
Email: rakibultowhid@yahoo.com

**Abstract**—Although conventional cryptographic security mechanisms are essential to the overall problem of security, the openness of wireless medium poses both threats and opportunities for securing transmission. Information theoretic security is attracting researchers for its robust nature. In this paper, a Gaussian multiple input multiple output multiple eavesdropper (MIMOME) channel is considered where a transmitter is communicating to a receiver in the presence of an eavesdropper. The transmitter is equipped with multiple antennas, while the receiver and the eavesdropper also contains multiple antennas. We present a technique for determining the secrecy capacity of the MIMO channel under Gaussian noise. To do so, we transform the channel into multiple single input multiple output (SIMO) Gaussian wiretap channel and then use scalar approach using standard techniques of communications theory. Then we consider the effect of fading and found out the limiting values of the secrecy capacity.

**Index Terms**—Secrecy capacity, MIMO, wiretap channel, covariance matrix, fading.

## I. INTRODUCTION

Wireless media is an open medium for communication and is vulnerable to eavesdropping. The openness of the transmission medium makes eavesdropping extremely easy and anyone within communication range can listen to the traffic in the air, and possibly extract information. However, the unique properties of wireless medium might provideways of combating such security threats. For example, due to random fading, the intended receiver will have a channel different from an eavesdroppers when she is at a reasonable distance away. This difference can be utilized to secure the communication between the transmitter and the intended receiver. The eavesdropping attack was first studied by Wyner in [1], where he considers a single-user wire-tap channel. The measure of secrecy is the message equivocation rate at the wire-tapper, which is defined as the entropy of the message at the wire-tapper, given the wire-tappers observation. Wyner models the wire-tappers channel as a degraded version of the channel from the transmitter to the legitimate receiver, which is a reasonable assumption in a wired channel. For this channel, Wyner

identifies the rate-equivocation region and therefore, the secrecy capacity. Wyners result was extended to the Gaussian wire-tap channel in [2], and it was shown that Gaussian signalling is optimal. The secrecy capacity was found to be the difference between the capacities of the main and the eavesdropping channels. Csiszar and Korner [3] studied the general, i.e., not necessarily degraded, single-transmitter, single-receiver, single-eavesdropper, discrete memoryless channel with secrecy constraints, and found an expression for the secrecy capacity, in the form of the maximization of the difference between two mutual informations involving an auxiliary random variable. The auxiliary random variable is interpreted as performing pre-processing on the information. The explicit calculation of the secrecy capacity for a given channel requires the solution of this maximization problem in terms of the joint distribution of the auxiliary random variable and the channel input. The use of multiple transmit and receive antennas has been shown to increase the achievable rates when there are no secrecy constraints [4]. The Gaussian multiple-input multipleoutput (MIMO) wire-tap channel is a special case of the single-transmitter, single-receiver, single-eavesdropper wire-tap channel. Since the Gaussian MIMO channel is not degraded in general, finding its secrecy capacity involves identifying the optimum joint distribution of the auxiliary random variable representing pre-processing and the channel input in the Csiszar-Korner formula. However, solving this optimization problem directly for non-degraded channels is difficult, forcing researchers typically to follow a two-step solution, where in the first step a feasible solution is identified (an achievable scheme), and in the second step a tight upper bound that meets this feasible solution is developed (tight converse). The first paper studying secrecy in MIMO communications is [5], which proposes an achievable scheme, where the transmitter uses its multiple transmit antennas to transmit only in the null space of the eavesdroppers channel, thereby preventing any eavesdropping. Reference [6] studies the Gaussian single-input multiple-output (SIMO) wire-tap channel, and shows that it is equivalent to a scalar

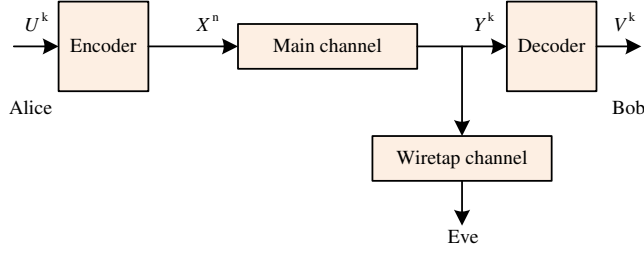


Fig. 1. Wiretap channel

Gaussian channel, and gives the secrecy capacity using the results of [2]. An achievable scheme has been proposed for the Gaussian multipleinput single-output (MISO) wiretap channel in [7], and independently and concurrently in [8]. In both of these papers, the achievable secrecy rate is obtained by restricting the 2 channel input to be Gaussian, with no pre-processing of information. The secrecy rate found in [7,8] is shown to be the secrecy capacity of the Gaussian MISO wire-tap channel in [9,10]. Further, [9, 10] allow the eavesdropper to have multiple antennas (MISOME). In all these papers, no one has tried to find the secrecy capacity of MIMO wiretap channel using the scalar gaussian channel.

In this paper we are trying to find out the secrecy capacity of a MIMO channel using a scalar gaussian approach. We consider the MIMO channel a combination of several SIMO channel and find out the secrecy capacity. We use the following notations throughout this paper: Bold face lower and upper case letters are used to represent vectors and matrices, respectively.  $\mathbf{x}^\dagger$  denotes the conjugate transpose of the complex vector  $\mathbf{x}$ . Whether a variable is deterministic or random will be clear from the context.

The remainder of this paper is organized as follows: In section II, wiretap channel is analyzed where gaussian wiretap channel is also discussed. In section III, System model for the MIMO channel is introduced where the mathematical model is also developed. In section IV, we consider a fading channel and develop a boundary condition for secrecy capacity. Then section V concludes this paper.

## II. WIRETAP CHANNEL

The wiretap channel is a degraded form of a broadcast channel, where the goal is to maximize the transmission rate in the main channel while making negligible the amount of information leaked to the cascade (wiretapper) channel. The secrecy capacity is the maximal rate at which this goal is achieved. We are considering two discrete memoryless channels (DMC) as the ones depicted in Figure 1. The encoder takes the input sequence  $U^k = (U_1, \dots, U_k)$  and transforms it into a  $n$  symbols sequence  $X^n = (X_1, \dots, X_n)$ . The rate of the code is  $R = k/n$ . The security engineer must design an encoder/decoder

pair that ideally maximizes the transmission rate of the legitimate user, subject to the constraint that the rate at which the wiretapper learns the sequence is as small as possible. The wiretapper knows the encoding used, and its ignorance about the source depends only on the noise realization present in the channels. The source is assumed to be stationary and ergodic, and takes its values over a finite alphabet. The probability of block decoding error is denoted by

$$P_e = \Pr\{U^k \neq V^k\} \quad (1)$$

The wiretapper uncertainty about the source is measured by the equivocation

$$H(U^k|Z^n), \quad (2)$$

after observing the output of the channel.

Definition 2.1: The fractional equivocation

$$\delta = \frac{H(U^k|Z^n)}{H(U^k)} \quad (3)$$

The rate of the code

$$R = \frac{H(U^k)}{n} \quad (4)$$

Definition 2.2: (Achievability) The pair  $(R^*, d^*)$  is said to be achievable if for all  $\epsilon > 0$  there exists an encoder/decoder pair such that

$$\begin{aligned} R &\geq R^* - \epsilon \\ \delta &\geq d^* - \epsilon \\ P_e &\leq \epsilon \end{aligned} \quad (5)$$

Theorem 2.1: [1]. The set of achievable pairs  $(R, d)$  can be characterized as follows

$$\mathcal{R} = \{(R, d) : 0 \leq R \leq C_M, 0 \leq d \leq 1, R_d \leq C_S\} \quad (6)$$

where  $C_S$  is the secrecy capacity and its value is

$$C_S = C_M - C_W \quad (7)$$

where  $C_M$  is the main channel capacity and  $C_W$  is the eavesdropper channel capacity.

Although Wyner's work only considered discrete-time channels, Leung and Hellman [11] proved that the results also hold in a particular case of continuous-time channel. A Gaussian wiretap channel is wiretap channel where the noise is additive white and Gaussian, such that the



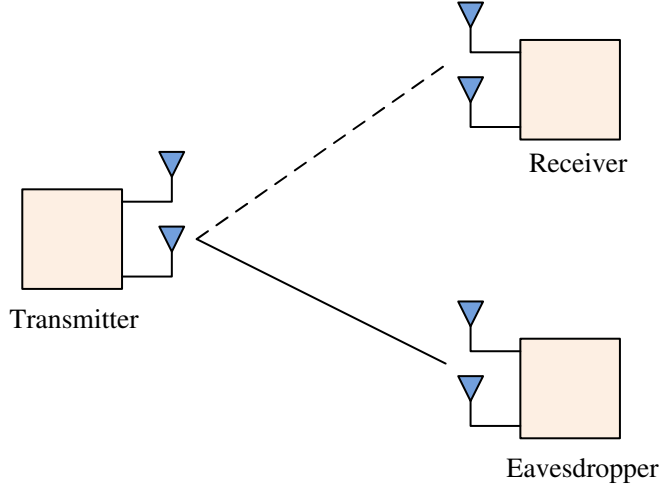


Fig. 2. A communication system with multi-antenna transmitter, receiver and eavesdropper

channel is power limited (P) and the noise processes are independent and have components that are i.i.d.  $\mathcal{N}(0, \sigma_1^2)$  and  $\mathcal{N}(0, \sigma_2^2)$  respectively. The achievable region of the Gaussian wiretap channel is the same as defined in eq. 6 with

$$\begin{aligned} C_M &= \frac{1}{2} \log\left(1 + \frac{P}{\sigma_1^2}\right) \\ C_W &= \frac{1}{2} \log\left(1 + \frac{P}{\sigma_1^2 + \sigma_2^2}\right) \end{aligned} \quad (8)$$

### III. SYSTEM MODEL

Figure 2 shows a communication system, with a transmitter equipped with a transmitter, a receiver and an eavesdropper, each with multiple antennas. The user and eavesdropper channel attenuations can be represented by  $N_t \times N_r$  and  $N_r \times M_r$  vectors  $\mathbf{H}$  and  $\mathbf{G}$ , where  $N_t$  is the number of transmit antennas whereas  $N_r$  and  $M_r$  are the number of receive antenna at the legitimate receiver and eavesdropper. We are considering the degraded multiple input multiple output (MIMO) channel depicted in Figure 3. We will determine the secrecy capacity of the MIMO Gaussian wiretap channel. For every  $N_t$  transmitting antenna, Alice sends symbols that have limited average power  $P > 0$ , i.e.,

$$\frac{1}{K} \sum_{k=0}^{K-1} E\{x^2[k]\} \leq P \quad (9)$$

Bob uses  $N_r$  receive antennas and Eve uses  $M_r$  receive antennas to recover Alice's message. Alice sends no message to Eve, so there are no common messages. The channel parameters are represented by  $\mathbf{H}$  with dimension  $N_r \times N_t$  and  $\mathbf{G}$  with dimension  $M_r \times N_r$ .

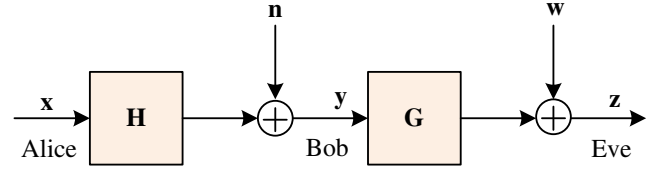


Fig. 3. Degraded broadcast channel

The received signals at the receiver and the eavesdropper at  $k$ -th time slot are

$$\mathbf{y}[k] = \mathbf{H}\mathbf{x}[k] + \mathbf{n}[k] \quad (10)$$

$$\mathbf{z}[k] = \mathbf{G}\mathbf{y}[k] + \mathbf{w}[k] \quad (11)$$

where  $\mathbf{n}$  and  $\mathbf{w}$  are independent random vectors, each one of them being complex and jointly Gaussian distributed with mean 0 and non-singular covariance matrices  $\Sigma_1$  and  $\Sigma_2$  respectively.  $\mathbf{H}$  and  $\mathbf{G}$  are known and fixed. When the main and eavesdropper channel experiences fading,  $\mathbf{H}$  and  $\mathbf{G}$  are assumed to be a vector of i.i.d. zero mean unit-variance complex circularly symmetric Gaussian random variables.

Our goal is to show that the channel described by eq. (10) can be represented by a summation of several scalar channels and that such representation is in fact a parallel Gaussian wiretap channel. Eq. (10) can be the addition of  $N_t$  parallel single input multiple output (SIMO) channel and can be written as

$$\begin{aligned} \mathbf{y}[k] &= \sum_{i=1}^{N_t} \mathbf{y}_i[k] \\ &= \sum_{i=1}^{N_t} (\mathbf{h}_i \mathbf{x}_i[k] + \mathbf{n}_i[k]) \end{aligned} \quad (12)$$

After putting the value of  $\mathbf{y}[k]$  in the eq. (11) we get

$$\mathbf{z}[k] = \mathbf{G} \sum_{i=1}^{N_t} (\mathbf{h}_i \mathbf{y}_i[k] + \mathbf{n}_i[k]) + \mathbf{w}[k] \quad (13)$$

The multiple input multiple output multiple eavesdropper (MIMOME) channel can be seen as a multiple parallel single input multiple output multiple eavesdropper (SIMOME) channel. Every SIMOME channel can be converted to a corresponding scalar channel [6]. The  $i$ th main channel and the eavesdropper channel at SIMOME looks like the following

$$\mathbf{y}_i[k] = \mathbf{h}_i x_i[k] + \mathbf{n}_i[k] \quad (14)$$

Replacing  $\mathbf{y}[k]$  by  $\mathbf{y}_i[k]$  in eq. (11) we get

$$\begin{aligned} \mathbf{z}_i[k] &= \mathbf{G}(\mathbf{h}_i x_i[k] + \mathbf{n}_i[k]) + \mathbf{w}[k] \\ &= \mathbf{G}\mathbf{h}_i x_i[k] + \mathbf{G}\mathbf{n}_i[k] + \mathbf{w}[k] \end{aligned} \quad (15)$$

Multiply the equation (14) by  $\mathbf{h}_i^\dagger \Sigma_{1i}^{-1}$  we get

$$\mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{y}_i[k] = \mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{h}_i x_i[k] + \mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{n}_i[k] \quad (16)$$

both the sides are now  $1 \times 1$  in dimension and in fact scalar. Now taking  $\mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{y}_i[k] = y_i[k]$ ,  $\mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{h}_i = h_i^2$  and  $\mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{n}_i[k] = n_i[k]$  we get the equation (16) like the followings

$$y_i[k] = h_i^2 x_i[k] + n_i[k] \quad (17)$$

Again multiplying eq. (15) by  $(\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1}$  we get

$$\begin{aligned} (\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{z}_i[k] &= (\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{G}\mathbf{h}_i x_i[k] \\ &\quad + (\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} (\mathbf{G}\mathbf{n}_i[k] + \mathbf{w}[k]) \end{aligned} \quad (18)$$

both the sides are now scalar as it was in the previous case. Now taking  $(\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{z}_i[k] = z_i[k]$ ,  $(\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{G}\mathbf{h}_i = g_i^2$  and  $(\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} (\mathbf{G}\mathbf{n}_i[k] + \mathbf{w}[k]) = w_i[k]$  we get

$$z_i[k] = g_i^2 x_i[k] + w_i[k] \quad (19)$$

Eq. (17) and eq. (19) shows the scalar representation of this MIMOME channel. In the next section, we will derive the secrecy capacity using these results.

#### IV. SECRECY CAPACITY

The equivocation rate  $R_e$ , is the conditional entropy of the transmitted message, conditioned on the received signal at the eavesdropper. The equivocation rate is a measure of the amount of information that the eavesdropper can attain about the message, and quantifies the level of secrecy in the system. The secrecy capacity,  $C_S$ , is the largest rate  $R$  achievable with perfect secrecy, i.e.,  $R_e = R$ .

The main channel capacity for SIMOME can be written like the following

$$\begin{aligned} C_{M_i} &= \frac{1}{2} \log(1 + h_i^2 P) \\ &= \frac{1}{2} \log(1 + \mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{h}_i P) \end{aligned} \quad (20)$$

Again the eavesdropper channel capacity for SIMOME is

$$\begin{aligned} C_{W_i} &= \frac{1}{2} \log(1 + g_i^2 P) \\ &= \frac{1}{2} \log(1 + (\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{G}\mathbf{h}_i P) \end{aligned} \quad (21)$$

So the main channel and eavesdropper channel capacity for MIMOME channel is given by

$$\begin{aligned} C_M &= \sum_{i=1}^{N_t} C_{M_i} \\ &= \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{h}_i P) \end{aligned} \quad (22)$$

$$\begin{aligned} C_W &= \sum_{i=1}^{N_t} C_{W_i} \\ &= \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + (\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{G}\mathbf{h}_i P) \end{aligned} \quad (23)$$

Positive secrecy capacity indicates successful secret communication and the secrecy capacity of the MIMO Gaussian wiretap channel can be written as

$$\begin{aligned} C_S &= C_M - C_W \\ &= \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \mathbf{h}_i^\dagger \Sigma_{1i}^{-1} \mathbf{h}_i P) \\ &\quad - \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + (\mathbf{G}\mathbf{h}_i)^\dagger (\mathbf{G}\Sigma_{1i}\mathbf{G}^\dagger + \Sigma_2)^{-1} \mathbf{G}\mathbf{h}_i P) \end{aligned} \quad (24)$$

Considering  $\Sigma_{1i} = \sigma_{1i}^2 \mathbf{I}_{N_r}$  and  $\Sigma_2 = \sigma_2^2 \mathbf{I}_{M_r}$ , eq. 24 is rewritten as

$$\begin{aligned}
C_S = & \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \mathbf{h}_i^\dagger (\sigma_{1i}^2 \mathbf{I}_{N_r})^{-1} \mathbf{h}_i P) \\
& - \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + (\mathbf{G} \mathbf{h}_i)^\dagger (\mathbf{G} (\sigma_{1i}^2 \mathbf{I}_{N_r}) \mathbf{G}^\dagger \\
& + (\sigma_2^2 \mathbf{I}_{M_r}))^{-1} \mathbf{G} \mathbf{h}_i P)
\end{aligned} \quad (25)$$

Again  $\mathbf{G}$  is unitary [6], we get

$$\begin{aligned}
C_S = & \sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\|\mathbf{h}_i\|^2}{\sigma_{1i}^2} P \right) \\
& - \sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\|\mathbf{h}_i\|^2}{\sigma_{1i}^2 + \sigma_2^2} P \right)
\end{aligned} \quad (26)$$

Taking  $\frac{P}{\sigma_{1i}^2} = \text{SNR}_i$  we get

$$\begin{aligned}
C_S = & \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \|\mathbf{h}_i\|^2 \text{SNR}_i) \\
& - \sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\sigma_{1i}^2 \|\mathbf{h}_i\|^2}{\sigma_{1i}^2 + \sigma_2^2} \text{SNR}_i \right)
\end{aligned} \quad (27)$$

From the previous equation it can be shown that  $\sigma_2^2 > 0$  makes  $\frac{\sigma_{1i}^2}{\sigma_{1i}^2 + \sigma_2^2} < 1$  and in turn makes  $\sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \|\mathbf{h}_i\|^2 \text{SNR}_i) > \sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\sigma_{1i}^2 \|\mathbf{h}_i\|^2}{\sigma_{1i}^2 + \sigma_2^2} \text{SNR}_i \right)$ . So the secrecy capacity  $C_S$  becomes positive. An interesting conclusion is that the secret communication is possible when the eavesdropper channel is noisy.

## V. FADING CHANNEL

We assume that the channel of the intended receiver is complex Gaussian fading, while the eavesdroppers channel is constant. Considering the main channel as a fading channel eq. (27) becomes

$$\begin{aligned}
C_S = & E_{\mathbf{h}_i} \left[ \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \|h_i\|^2 \text{SNR}_i) \right] \\
& - E_{\mathbf{h}_i} \left[ \sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\sigma_{1i}^2 \|h_i\|^2}{\sigma_{1i}^2 + \sigma_2^2} \text{SNR}_i \right) \right]
\end{aligned} \quad (28)$$

It can be observed that, since the function  $f(x) = \log(1 + ax)$  is concave, using Jensens inequality [12], we have

$$\begin{aligned}
E_{\mathbf{h}_i} \left[ \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \|h_i\|^2 \text{SNR}_i) \right] & \leq \\
\sum_{i=1}^{N_t} \frac{1}{2} \log(1 + E_{\mathbf{h}_i} \|h_i\|^2 \text{SNR}_i)
\end{aligned} \quad (29)$$

And similarly

$$\begin{aligned}
E_{\mathbf{h}_i} \left[ \sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\sigma_{1i}^2 \|h_i\|^2}{\sigma_{1i}^2 + \sigma_2^2} \text{SNR}_i \right) \right] & \leq \\
\sum_{i=1}^{N_t} \frac{1}{2} \log \left( 1 + \frac{\sigma_{1i}^2}{\sigma_{1i}^2 + \sigma_2^2} E_{\mathbf{h}_i} \|h_i\|^2 \text{SNR}_i \right)
\end{aligned} \quad (30)$$

if  $\sigma_{1i}^2 \ll \sigma_2^2$  then the secrecy capacity  $C_S$  in eq. (28) leads to

$$C_S = E_{\mathbf{h}_i} \left[ \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + \|h_i\|^2 \text{SNR}_i) \right] \quad (31)$$

The boundary value of secrecy capacity for this particular case will be

$$C_S \leq \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + E_{\mathbf{h}_i} \|h_i\|^2 \text{SNR}_i) \quad (32)$$

But if  $\sigma_2^2 = 0$  ie. the eavesdropper channel is noiseless, the secrecy capacity  $C_S$  will be equal to 0. So, the limiting value of the secrecy capacity depending on the noisy condition of the eavesdropper channel can be shown as

$$0 \leq C_S \leq \sum_{i=1}^{N_t} \frac{1}{2} \log(1 + E_{\mathbf{h}_i} \|h_i\|^2 \text{SNR}_i) \quad (33)$$

## VI. CONCLUSION

A MIMOME channel is considered where a transmitter is communicating to a receiver in the presence of an eavesdropper. The transmitter is equipped with multiple antennas, while the receiver and the eavesdropper also contains multiple antennas. We present a technique for determining the secrecy capacity of the MIMO channel under Gaussian noise. To do so, we transform the channel into multiple SIMO Gaussian wiretap channel and then use scalar approach using standard techniques of communications theory. We also have introduce the effect of fading in this MIMO channel and have derived the limiting value for secrecy capacity. One particular direction corresponds to the case where the broadcast channel is no longer degraded, such as the one considered by I. Csiszar and J. Korner [3]. Our next step will be to work with non degraded broadcast channel.

## ACKNOWLEDGMENT

This work was supported by the u-LCRC at Kyung Hee University, Korea, under the ITRC program of MIC, Korea (IITA-2008-(C1090-0801-0002)).

## REFERENCES

- [1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, 54(8):210, October 1975.
- [2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. on Information Theory*, 24(4):451-456, July 1978.
- [3] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. on Information Theory*, 24(3):339-348, May 1978.
- [4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommunications*, 10:585-595, November 1999.
- [5] R. Negi and S. Goel, "Secret communication using artificial noise," *IEEE Vehicular Technology Conference*, Toulouse, France, May 2006.
- [6] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," *IEEE International Symposium on Information Theory*, Adelaide, Australia, September 2005.
- [7] S. Shafiee and S. Ulukus, "Achievable rates in Gaussian MISO channels with secrecy constraints," *In IEEE International Symposium on Information Theory*, Nice, France, June 2007.
- [8] Z. Li, W. Trappe, and R. D. Yates, "Secret communication via multi-antenna transmission," *41st Conference on Information Sciences and Systems*, Baltimore, MD, March 2007.
- [9] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," *IEEE International Symposium on Information Theory*, Nice, France, June 2007.
- [10] A. Khisti and G. Wornell, "Secure transmission with multiple antennas: The MISOME wiretap channel," *Submitted to IEEE Trans. on Information Theory*
- [11] S. Leung-Yan-Cheong and M. E. Hellman, "The gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451-456, 1978.
- [12] T. Cover and J. Thomas, *Elements of Information Theory* Wiley, 1991.
- [13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1987.
- [14] X. Li and E. P. Ratazzi, "MIMO transmissions with information-theoretic secrecy for secret-key agreement in wireless networks," *IEEE Military Communications Conference (MILCOM2005)*, 2005. Atlantic City, NJ.
- [15] X. Li, M. Chen, and E. P. Ratazzi, "Space-time transmissions for wireless secret-key agreement with information-theoretic secrecy," *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC05)*, 2005. The Italian Academy at Columbia University, New York.

# A Practical Analysis of the Robustness and Stability of the Network Stack in Smartphones

Sheikh Mahbub Habib, Cyril Jacob and Tomas Olovsson

Department of Computer Science & Engineering, Chalmers University of Technology, Gothenburg,  
SE-41296, Sweden

**Abstract** - Smartphones are widely used nowadays and their popularity will certainly not slow down in the near future due to improved functionality and new technology improvements. Becoming more and more similar to PCs and laptops, they will also begin to face the same security problems especially in terms of network security. In this paper, we provide an overview of security issues for smartphones and give a brief introduction of the network stack architecture of Windows Mobile 5.0 platform in order to motivate and plan for efficient penetration tests against Windows Mobile 5.0 powered smartphones. Furthermore, a number of attacks have been done against different layers of the network stack and a list of supporting penetration tools are provided for interested practitioners. Detailed results are provided from the penetration tests performed which should be very useful for security vendors, researchers and OS vendors to give more attention to security architectures and development of security software (e.g. firewalls, antivirus) for the mobile operating systems. In a whole, robustness and stability will be assessed for the mobile OSs used by the smartphones throughout the paper.

**Index terms** - Network stack, Network scanning, Penetration testing, Smartphones, Vulnerability scanning.

## I. INTRODUCTION

Smartphones combine the functionality of mobile phones and Personal Digital Assistants (PDAs). These types of devices are becoming increasingly popular in our daily life and decrease the need to carry several devices to get similar functionality. Almost all the latest smartphones implement network technologies such as IEEE 802.11 (WLAN), Bluetooth, Infra-red and GSM/3G. When compared to regular phones, these devices have extra functionality and offer more services which will make them more vulnerable. Smartphones also enable service providers to quickly introduce more and new services to mobile users to satisfy their needs, often without really considering security. Security in these devices is often neglected due to too much attention given to the development and quick release of new functionality and services. As a result, smartphones are likely to be vulnerable both through their services and due to their relatively new and untested network stacks.

Since most operators today use NAT for connections through GSM, GPRS, 3G, etc., it is not possible for

attackers to directly scan and find the devices and the services they may offer from the Internet. This may change in the future, but at present, this offers some protection for the users. WLAN access, though, is different and the IP address being used on the WLAN may be accessible for users all over the Internet. In addition, because of vulnerabilities in the link layer and above, for example when using WEP encryption [1], attackers may be able to spawn link-level attacks via a wireless access point as well.

We should not only consider attacks against the Smartphone but also consider that it may be used as a bridge to connect to systems on the networks behind the device. It may, for example, be possible to reach corporate networks where the Smartphone is used as man-in-the-middle bridging two networks together, a scenario which was shown in enterprise networks using Blackberry devices [2].

Smart phones are used in many situations by professionals in situations where good and reliable communication is a key factor. Typical examples include voice and data communications in crisis management, for example as the only means for communication between rescue workers. It is therefore essential to know what level of stability, robustness and security these devices can provide. A typical problem we don't want to see is that a single malicious network packet from the Internet causes the devices to freeze, crash or even enable attackers to install their own software on the devices. The work provided here should be seen as a first step in the direction to develop procedures for testing and evaluation of the services offered by Smartphones.

The main target of the analysis in this paper is the Windows Mobile 5.0 operating system's network stack. Testing includes fingerprinting, port scanning, and running exploit tools in order to test the different layers of the stack.

## II. TCP/IP Architecture of Windows Mobile 5.0

Since this paper is targeting the network stack of Windows Mobile 5.0 which in turn is based on the Windows CE 5.1 platform, an understanding of the TCP/IP stack is useful when planning penetration tests and in order to analyze and understand the implications of possible security problems. All operating systems prior to Windows Vista and Windows 2008 server including the Windows

Mobile 5.0 platform use the same TCP/IP stack that was designed in the early 1990's by Microsoft [3][4]. Although this stack has been patched a number of times since then, vulnerabilities are still found in the different desktop operating systems using it. For this reason it is evident that Windows Mobile 5.0 may suffer from the same type of vulnerabilities as the other desktop operating systems like Windows XP/2000/ME/NT/98/95 have had before being patched. Moreover, Windows CE 4.1 and later versions include a dual stack architecture [5] that makes them vulnerable to attacks not only against IPv4 but also the IPv6 implementation.

In the next section, we will discuss test methodologies in order to develop a suitable methodology to test Smartphones, not only Windows Mobile based but also other platforms will be tested with the same methodology in the future.

### III. TEST METHODOLOGY

When evaluating and testing the devices, we have based our work on a scheme developed by NIST [8] containing nine techniques useful when evaluating security for a network. These are network scanning, vulnerability scanning, penetration testing, password cracking, log review, integrity checkers, virus detection, war dialing and war driving. Some of the tests are manual where an individual has to initiate and conduct the test and some of the tests are automated requiring no human involvement. Several of the above mentioned tests are often used together to get a more complete assessment of the achieved level of security. However, in this study, the test environment will mainly focus on the network stack of Smartphones, so password cracking, log review, integrity checkers, virus detection, war driving and war dialing will be irrelevant regarding practical analysis of network stack.

*Network scanning, vulnerability scanning and penetration testing* will be enough for the security testing of the network stack performed here but they will not give the full picture when the device is used in a real life scenario. These three testing techniques are discussed more in detail in the following paragraphs.

#### A. Network Scanning

In network scanning, a port scanner is used to identify all targeted hosts on a network and the services they offer to other hosts on the network. There are some specific services which are very common in smartphones such as netbios-ns (netbios name service) and the wap-push protocol. The port-scanning tool makes it possible to create a complete list of active ports and thus list all services running on the device. A high-level human interpretation is then needed to analyze the results which will be used when planning the vulnerability scanning. The Nmap tool [9] was used to perform port scanning of the targeted Smartphones.

#### B. Vulnerability Scanning

Vulnerability scanners use a similar concept as port scanners but try to go deeper and get a more realistic view about the target device. These types of scanners identify open ports and hosts and also provide information regarding possible vulnerabilities through the different ports, often application level protocol problems or problems in the application itself. Moreover, vulnerability scanners help to identify out-of-date software versions, lack of applicable patches or system upgrades and applications running on devices are compared with known vulnerabilities. Vulnerability scanners often have large databases of vulnerabilities used to identify flaws in the operating systems and applications. The results from these type of tests are valuable when selecting the right tools for actual penetration tests. The results from our port scanning tests showed that the Nessus vulnerability scanner [10] would be a good choice in order to check what vulnerabilities are actually present in our Smartphones.

#### C. Penetration testing

Penetration testing is the third and final step in the security testing where specially designed tools, often targeted against a special network stack, its implementation and other services offered on the platform, is used. In order to further analyze the possible vulnerabilities indicated by the vulnerability scanning process, commonly known attack tools were used to do the actual attacks against the smartphones. The services detected during the Nmap and Nessus testing guided us to select other penetration tools and exploit tools. In addition, Wireshark [11] was used as a network packet capture tool to analyze the results.

### IV. PENETRATION TEST RESULTS AND DISCUSSION

All the tests are initiated from two clients, one running Windows XP and one with Ubuntu Linux. The target was an XDA Exec device with Windows Mobile 5.0 OS. Wireshark was used as packet capture software to analyze the results. In the next few sections, results will be reported according to the test methodology described earlier.

#### A. Network scanning

In the network scanning performed, the Nmap "host discovery technique" [9] was used to find the MAC address and IP address of the Smartphone and to find all open/filtered/closed ports on the device. This is a technique which may reveal services even if protected by a device's firewall or built-in security barrier, if present. Moreover, a fingerprinting technique was used to more exactly identify the operating system running on the target, something which enabled us to more closely see how close its behavior is to other Windows operating systems (such as TTL, window size, initial sequence numbers, etc.). These tests gave several hints about the internal network stack

TABLE I  
NETWORK SCANNING RESULTS

| Name of the scan         | Port    | Command                  | Results                                                                                                                                                                               |
|--------------------------|---------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Host discovery-ping scan | ICMP    | # nmap -sP 192.168.1.102 | Target device is found alive and MAC address shown.                                                                                                                                   |
| Host discovery-SYN ping  | TCP     | # nmap -PS 192.168.1.102 | None of the tcp ports were open. Target host was shown alive in the network.                                                                                                          |
| Ports discovery-SYN scan | TCP     | # nmap -sS 192.168.1.102 | Same result as SYN ping.                                                                                                                                                              |
| Ports discovery-UDP scan | UDP     | # nmap -sU 192.168.1.102 | <b>PORT - STATE - SERVICE</b><br>137/udp open filtered netbios-ns<br>138/udp open filtered netbios-dgm<br>1034/udp open filtered activesync-notify<br>2948/udp open filtered wap-push |
| OS fingerprinting        | TCP,UDP | # nmap -O 192.168.1.6    | Device type: webcam switch general purpose. Running: AXIS embedded, Cisco embedded, IBM MVS, Microsoft Windows 2003/.NET NT/2K/XP 95/98/ME.                                           |

architecture of the device and the services it offers. The detailed test results are given in Table I.

In Table I, the port discovery technique shows that the system has four services running which are *netbios-ns*, *netbios-dgm*, *activesync-notify* and *wap-push*, all using UDP ports. The tests show that the Netbios-ns service discloses the netbios-name of the machine, the netbios-datagram service(netbios-dgm) offer support for connection-less file sharing activities, activesync-notify is a service used to synchronize the device with workstations or laptops and wap-push is used to handle the multimedia message service (MMS). All these services are enabled by default.

By examining the results of the OS fingerprinting test from Table I, according to database of Nmap, an exact match of the OS cannot be found due to absence of the fingerprint of Windows Mobile OS. But after examining the output more in detail, Nmap suggests that it can be a Microsoft Windows Operating System. Moreover, some packets were captured to investigate the TCP and IP headers. This investigation showed that it is reasonable to

assume that the device has a network stack architecture which is similar to other Microsoft Windows operating systems for desktop systems. Therefore, similar attacks to those being successful in the past against Windows systems should be done further on when performing the penetration testing.

In the next section, Nessus [10] vulnerability scanner will be used to list the possible known vulnerabilities in the targeted smartphone via different ports and services running.

#### B. Vulnerability scanning

To perform vulnerability scanning, Nessus Server and Client were used to scan the remote device in order to find different ports and services running in that particular device. We have summarized the results regarding smartphone's vulnerability in Table II.

TABLE II  
VULNERABILITY SCANNING RESULTS

| Port name           | Synopsis                                                                                           | Description                                                                                                                                                                                                                                                                                                                                                                                                       | Risk factor & Reference ID                                                                                                           |
|---------------------|----------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|
| netbios-ns(137/udp) | It is possible to obtain the network name of the remote host                                       | The remote host listens on udp port 137 and replies to NetBIOS nbtscan requests.                                                                                                                                                                                                                                                                                                                                  | Risk: None.<br>CVE : CVE-1999-0621<br>Other references :<br>OSVDB:13577<br>Nessus ID : 10150                                         |
| general/udp         | Traceroute                                                                                         | Makes a traceroute to the remote host.                                                                                                                                                                                                                                                                                                                                                                            | Risk: Low<br>Nessus ID : 10287                                                                                                       |
| general/tcp         | i)It is used to ping the remote host.<br>ii)Also, used for resolving the IP address to device name | This script displays, for each tested host, information about the scan itself:<br>-The version of the plugin set<br>-The type of plugin feed (Direct, Registered or GPL)<br>-The version of the Nessus Engine<br>-The port scanner(s) used<br>-The port range scanned<br>-The date of the scan<br>-The duration of the scan<br>-The number of hosts scanned in parallel<br>-The number of checks done in parallel | Risk: None<br>Nessus ID : 19506                                                                                                      |
| general/icmp        | It is possible to determine the exact time set on the remote host.                                 | The remote host answers to an ICMP timestamp request. This allows an attacker to know the date which is set on your machine.<br>This may help him to defeat all your time based authentication protocols.                                                                                                                                                                                                         | Risk: None<br>CVE : CVE-1999-0524<br>Nessus ID : 10114                                                                               |
| general/icmp        | The remote host leaks memory in network packets.                                                   | The remote host is vulnerable to an 'Etherleak' - the remote ethernet driver seems to leak bits of the content of the memory of the remote operating system.                                                                                                                                                                                                                                                      | Risk: Medium / CVSS Base Score : 5.0<br>(CVSS2#AV:N/AC:L/Au:N/C:P/I:N/A:N)<br>CVE : CVE-2003-0001<br>BID : 6535<br>Nessus ID : 11197 |

### C. Penetration testing

The most revealing problem detected here is the Etherleak vulnerability [12]. Etherleak used to be a common problem present in many operating systems early 2003, and the target system still has this vulnerability. In short, this means that the Ethernet driver leaks information found in the memory of the operating system, particularly contents from a previously received packet. Attackers on the same subnet can collect this information from the target, for example by sending ICMP echo messages. Desktop operating systems have received patches to this problem, but not Windows Mobile 5 devices.

In the penetration testing phase, we have performed many attacks against the target. The following paragraphs show the detailed results from these tests.

#### 1. ARP spoofing

We observed that the target is vulnerable to a denial-of-service attack when receiving a gratuitous ARP for its own IP address. Also, we observed that traffic can be hijacked between the target and router/access point using forged ARP reply packets.



Table III

## PENETRATION TESTING RESULTS

| Attack name & network stack level                     | Command to initiate the attack            | Result                                                                                                                                        |
|-------------------------------------------------------|-------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| ARP spoofing(IP level attack)                         | # arpspoof -t 192.168.1.100 192.168.1.100 | DOS attack observed. Cannot assigned IP address to the device until the attack tool stopped.                                                  |
| Session hijacking using ARP Spoofing(IP level attack) | # arpspoof -t 192.168.1.100 192.168.1.1   | Hijack the traffic between router (AP) and target. Here, DOS attack was observed after some time. No warning message shown in the smartphone. |
| ARP spoofing using Broadcast(IP level attack)         | # arpspoof 192.168.1.100                  | No changes observed in the target.                                                                                                            |
| IP Options attack (IGMPv3 exploit)                    | # ./exploit 192.168.1.100 192.168.1.4     | Target device completely hanged after receiving the crafted IGMP packet. Soft reset needed to put the target into normal state.               |

We used the arpspoof tool from the dsniff 2.3 suite [13] for the IP tests. Testing was performed on the Ubuntu Linux host as the attacker, one LinkSys router/gateway as victim (whose ARP entry was being spoofed), and one smartphone to which the gratuitous ARPs were sent to. The hosts were 192.168.1.104 (attacker), 192.168.1.100 (target), 192.168.1.101 (victim). Results are shown in Table III.

## 2. IP Options attack

The system also turned out to be vulnerable against a well-known exploit, the “IGMP V3 DoS vulnerability” [14]. We used the Linux machine as the attacker with the smartphone as the target in a LAN environment where both machines were connected to internet through a LinkSys router. Moreover, a real exploit tool [15] was used after being modified to gcc compatible code to suite our Linux system. The result is shown in Table III.

The device hangs after receiving a specially crafted IGMP packet. This can be sent to the target system and cause it to hang until it is reset. It is easy for an attacker to continuously send such packets to the target and render it completely useless. Some reports also suggest that the vulnerability could be used to execute arbitrary code on affected systems. There is no patch available in Microsoft’s security bulletin website related to this attack for Windows Mobile 5.0 [19].

## 3. Other attacks

Some other well-known attacks such as Land attack, Ping of Death, Teardrop, IP/TCP Fuzzing, Blat, Bonk, Boink, Naptha, Newtear/Open tear, Nuke and many other attacks

were performed in the testing of the network stack’s stability and robustness. These tests were successfully handled by the device.

Since the target device has a dual stack architecture, Neighbor discovery spoofing can be performed. It is an attack where a malicious user conducts a spoofing attack through the Neighbor discovery protocol present in IPv6. The vulnerability discloses sensitive information and results in a DOS (Denial of Service). The vulnerability is caused due to a bug in the implementation of Neighbor discovery protocol while processing neighbor solicitation request. It can be exploited by someone adding a faked entry to the neighbor’s cache via an ND solicitation request, which may contain spoofed IPv6 address.

## VI. LIMITATIONS

This paper is just a beginning of an important research area which needs a lot of attention from researchers and developers. First of all, the tested Smartphone had no antivirus or firewall installed during the tests. A personal firewall on the device may have stopped some of the vulnerabilities found here, but this has not been tested by us.

The tests mentioned in this paper have only focused on the WLAN connectivity of these kinds of devices. Many smartphones today communicate not only using WLAN but also via GSM, GPRS and other more “phone-related” protocols as well as via Bluetooth. All these protocols should also be tested in order to have a full picture of the device.

## VII. CONCLUSIONS

In this paper, we have analyzed the stability and robustness of a Windows Mobile 5.0 Smartphone. We have also conducted tests to find necessary tools to aid in this work. The paper also contains a description of the TCP/IP stack architecture of the device since it has affected the planning phase of the security analysis work and guided us in the tool selection part.

We have included a discussion about test methodology and how we have acted in order to find out suitable tests for Smartphones. Careful selection of tools and attacks is important both for our own analysis as well as for future work with analysis of other devices. Detailed test results from our penetration tests are given here revealing some problems resulting from lack of robustness in the network stack of the tested device. Some rather well-known vulnerabilities which used to be present in desktop operating systems, like Etherleak, ARP spoofing, session hijacking using ARP spoofing and the IP options attack (IGMP V3 vulnerability) were found to be present in the tested system. It seems that the network stack in this type of device contains several vulnerabilities and is not fully capable to operate in complex and hostile network environments.

We will continue this work with more tests which will include penetration tests against Smartphones from other vendors like SonyEricsson and Nokia with completely different operating systems and different network architectures. Even though we have not yet performed any tests against these devices, we believe that they are likely to contain many more vulnerabilities than the tested system since Windows Mobile systems are based on rather well-tested implementations used many years in desktop operating systems.

We believe that our findings will be beneficial to the vendors to release the patches, security researchers and practitioners and security vendors who are developing firewalls, antivirus, and intrusion detection systems for the smartphones. The work should also be useful for end-customers who need to know what level of security and stability these devices offer and it makes it possible for them to know in what situations these devices should or should not be used.

## REFERENCES

- [1] Andrew Bittau, Mark Handley, Joshua Lackey, "The Final Nail in WEP's Coffin", <http://tapir.cs.ucl.ac.uk/bittau-wep.pdf>, Accessed on 7th July 2008.
- [2] Jesse D'Aguanno, "Blackjacking-Owning the Enterprise via the Blackberry", <http://www.praetorian.net/presentations/blackjack.html>, presented at Defcon 14-Las Vegas, NV 2006.
- [3] "Next Generation TCP/IP stack in Windows Vista and Windows Server 2008", [http://technet.microsoft.com/sv-se/library/bb878108\(en-us\).aspx](http://technet.microsoft.com/sv-se/library/bb878108(en-us).aspx).
- [4] "TCP/IP Architecture Model", <http://msdn.microsoft.com/en-us/library/ms880936.aspx>
- [5] "Dual Stack Architecture", <http://msdn.microsoft.com/en-us/library/ms881720.aspx>
- [6] CVE ID: CVE-2008-2160, "Multiple unspecified vulnerabilities in the JPEG (GDI+) and GIF image processing in Microsoft Windows CE 5.0 allow remote attackers to execute arbitrary code via crafted (1) JPEG and (2) GIF images", <http://nvd.nist.gov/nvd.cfm?cvename=CVE-2008-2160>, Source: US- CERT/NIST, Release date: 12<sup>th</sup> May 2008.
- [7] <http://www.microsoft.com/technet/solutionaccelerators/mobile/maintain/SecModel/aff7cf7f-0e11-4ef4-8626-f33bd969b35a.msp?mfr=true>, Accessed on 21<sup>st</sup> April 2008.
- [8] John Wack, Miles Tracy, Mrurgiah Souppaya, "Guideline on Network Security Testing", *NIST Special Publication 800-42, October 2003*.
- [9] Nmap security scanner v4.20, <http://nmap.org/>, 2008.
- [10] Nessus Vulnerability Scanner v3.2.1, <http://www.nessus.org/nessus/>, released on 30<sup>th</sup> May 2008.
- [11] Wireshark v.1.0.2, <http://www.wireshark.org/>.
- [12] Ofir Arkin, Josh Anderson, "Ethernet frame padding information leakage", Copyright @stake Inc. 2003.
- [13] Dug Song, "Dsniff 2.3", <http://monkey.org/~dugsong/dsniff/>.
- [14] CVE ID: CVE-2006-0021 "IGMP v3 DoS Vulnerability", <http://nvd.nist.gov/nvd.cfm?cvename=CVE-2006-0021>, National Vulnerability Database, National Institute of Standards and Technology, Published on February 14, 2006.
- [15] Alexey Sintsov, "Microsoft Windows XP/2003 (IGMP v3) Denial of Service Exploit (MS06-007)", [http://secwatch.org/exploits/2006/03/MS06-007\\_exp.c.info](http://secwatch.org/exploits/2006/03/MS06-007_exp.c.info), Published on March 22, 2006.
- [16] Collin Richard Mulliner, "Security of Smart Phones", *Master's Thesis submitted in University of California, Santa Barbara*, June, 2006.
- [17] Tim Newsham Jim Hoaglund, "Windows Vista Network Attack Surface Analysis: A Broad Overview", *Symantec Advanced Threat Research*.
- [18] "TCP/IP Architecture", <http://www.microsoft.com/technet/archive/winntas/maintain/tcparch.msp?mfr=true>
- [19] "Vulnerability in TCP/IP Could Allow Denial of Service (913446)", <http://www.microsoft.com/technetsecuritybulletin/ms06-007.msp>, 14<sup>th</sup> March 2006.

# Sensitivity Analysis in Robust and Kernel Canonical Correlation Analysis

Md. Ashad Alam<sup>†</sup>, Mohammed Nasser<sup>‡</sup> and Kenji Fukumizu<sup>#</sup>

<sup>†</sup>Hajee Mohammad Danesh Science and Technology University, Dinajpur-5200, Bangladesh

<sup>‡</sup>Rajshahi University, Rajshahi-6205, Bangladesh

<sup>#</sup>Institute of Statistical Mathematics, Tokyo 106-8569, Japan

Emails: kakon\_bud@yahoo.com; mnasser.ru@gmail.com; fukumizu@ism.ac.jp

**Abstract**—A number of measures of canonical correlation coefficient are now used in pattern recognition in the different literature. Some robust forms of classical canonical correlation coefficient are introduced recently to address the robustness issue of the canonical coefficient in the presence of outliers and departure from normality. Also a few number of kernels are used in canonical analysis to capture nonlinear relationship in data space, which is linear in some higher dimensional feature space. But not much work has been done to investigate their relative performances through simulation and also from the view point of sensitivity. In this paper an attempt has been made to compare performances of kernel canonical correlation coefficients (Gaussian, Laplacian and Polynomial) with that of classical and robust canonical correlation coefficient measures using simulation and influence function. We investigate the bias, standard error, MSE, qualitative robustness index, sensitivity curve of each estimator under a variety of situations and also employ boxplots and scatter plots of canonical variates to judge their performances. We observe that the class of kernel estimators perform better than the class of classical and robust estimators in general and the kernel estimator with Laplacian function has shown the best performance for large sample size.

**Index Terms**—Pattern Recognition, Robust CCA, KCCA, Monte Carlo Simulation, Influence Function and Sensitivity Curves.

## I. INTRODUCTION

The main goal of pattern recognition is supervised or unsupervised classification. It is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make a sound and reasonable decisions about the categories of the patterns. The statistical approach has been most intensively studied to meet up this goal properly [1]. In this approach each pattern is represented in terms a number (such as  $s$ ) of features or measurements and is viewed as a point in a  $s$ -dimensional space. The aim is to choose those features that allow pattern vectors belonging to different categories to occupy compact and disjoint regions in a  $s$ -dimensional feature space. The effectiveness of the representation space (feature set) is determined by how well patterns from different classes can be separated. To do this, the measure of the relation between distinct features is essential. Classical

CCA is a multivariate procedure for assessing the linear relationship between two sets of variables (or features). Canonical correlation was developed by Hotelling [2]. His basic elegant paper on the topic, “Relations between two sets of variables” appeared in 1936. In this paper, Hotelling described an example where one set of variables consists of mental tests and the other of physical measurements on a group of people. Then it acquired popularity and many statisticians, biometricians, economists, social scientists, and many other researchers applied it in diverse field of knowledge. During the last one decade the use of CCA is utmost popular method in the area of computer science such as generic object recognition, image analysis, image retrieval, machine translation, image segmentation, computer vision, pattern recognition, and facial expression recognition. Now-a-days KCCA are used in different fields such as computer science, information technology and so on. There are some limitations of KCCA as complexity of transformation, dimensionality (dimension is larger than sample size) and time complexity to calculate. Most of the KCCA researchers, namely [3], [4], [5] etc. tried to overcome limitations of KCCA. In a recent work [6], the influence function of kernel PCA and a robust KPCA has been theoretically derived. One observation of their analysis is that KPCA with a bounded kernel such as Gaussian is robust in that sense that the influence function does not diverge, while for KPCA with unbounded kernels such as polynomial the IF goes to infinity. This can be understood by the boundedness of the transformed data in the feature space by a bounded kernel. While this is not a result for CCA but for PCA, it is reasonable to expect that KCCA with a bounded kernel is also robust. This consideration motivates us to do some empirical studies on the robustness of KCCA. It is very important to know how KCCA is effected by outliers and to develop measures of accuracy. Therefore, we do intend to study a number of conventional robust estimates and KCCA with different functions.

In this article we consider five canonical correlation coefficients and investigate their performances by simulation, influence function and sensitivity curve. We

briefly discuss the classical, robust and kernel canonical correlation coefficients that we consider in our study in section II and III. Section IV offers the simulation as well as influence function results of the canonical correlation coefficients and finds the overall best performer. The final section gives conclusion.

## II. CLASSICAL AND ROBUST CANONICAL CORRELATION ANALYSIS

Classical CCA can be seen as the problem of finding basis vectors for two sets of variables such that the correlation between the projections of the variables onto these basis vectors are mutually maximized. The classical covariance and correlation matrices as well as eigenvector and values are highly sensitive to outlying observations as was shown in the context of CCCA. An obvious approach to robustify CCA is to estimate the population, sample covariance or correlation matrix. A number of approaches for robust canonical correlation analysis except the one just mentioned were compared and discussed by Branco[7] and Taskinen [8] obtained influence function and asymptotic distributional properties of CCCA based on robust estimates of the covariance matrix. Many researchers [7], [8] and [9] developed a few number of robust methods in CCA and suggested that from the view point of robustness and computation the performance of minimum covariance determinate (MCD) estimator is the best [7]. In our paper, we consider MCD estimator as an estimator of class of robust method.

## III. KERNEL CANONICAL CORRELATION ANALYSIS (KCCA)

At the beginning of this century there have been advances in pattern recognition of data using kernel methods [10], [11]. It is a method that generalizes the classical CCA to nonlinear setting in a nonparametric way. The kernel-based canonical correlation analysis no longer requires the Gaussian distribution assumption on observation and, therefore, enhances the applicability of CCA greatly. Basics idea of KCCA is as follows: first the original pair of multivariate variables  $x$  and  $y$  are mapped into  $\phi_x(x)$  and  $\phi_y(y)$ , elements of some high (possibly infinite) dimensional Hilbert spaces  $\mathcal{H}_x$  and  $\mathcal{H}_y$  respectively through a suitable kernel function without calculating  $\phi(\cdot)$ 's in order to get benefit of the simple fact-what is non-linear in a vector space may be linear in a transformed high dimensional space. Resorting to kernel trick through forming kernel matrix linear transformation similar to classical case is implemented in the transformed high dimensional space. Using about similar calculation that, in fact, can find more complex transformation than CCA in original space (summarized in Fig. 1).

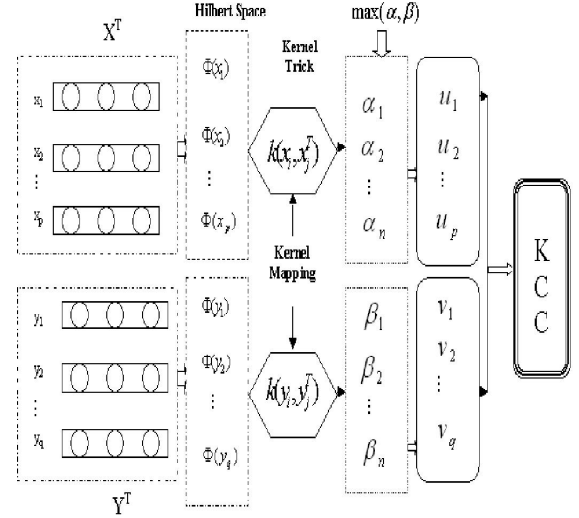


Fig. 1. System of Kernel CCA.

### A. Derivation of the Kernel Canonical Correlation

Let us consider linear dependence of two multivariate data set  $(x_i, y_i)$  be the  $p$  dimensional and  $q$  dimensional  $x$  and  $y$  sets of variables having  $i = 1, 2, \dots, n$  observations respectively. The classical canonical correlation is find the directions  $a$  and  $b$  so that the correlation between the projections of  $X$  onto  $a$  and that of  $Y$  onto  $b$  is maximized such that

$$\begin{aligned} \rho &= a \in \mathbb{R}^p, b \in \mathbb{R}^q \frac{Cov[a^T X, b^T Y]}{\sqrt{Var[a^T X] Var[b^T Y]}} \\ &= a \in \mathbb{R}^p, b \in \mathbb{R}^q \frac{a^T \hat{V}_{xy} b^T}{\sqrt{a^T \hat{V}_{xx} a} \sqrt{b^T \hat{V}_{yy} b}} \end{aligned} \quad (1)$$

where  $\hat{V}_{xx}$ ,  $\hat{V}_{yy}$  and  $\hat{V}_{xy}$  are the sample variance and covariance matrices respectively. Classical canonical correlation is highly depend on linearity assumption, because of this assumption it gives us naive measurement in case of nonlinear data set. Even robust method can fail to find worthy relationship in such type of data. We have to seek such a method that gives us accurate measurement in case of nonlinear data. Kernel canonical correlation is a such type of powerful tool that offers us an alternative solution for nonlinear data by first projecting the data into a higher dimensional feature space (Hilbert space). Let  $\{(x_i, y_i) ; i = 1, 2, \dots, n\}$  be the training sample with  $x_i$  and  $y_i$  taking values in  $\mathcal{X}$  and  $\mathcal{Y}$  respectively. Let us prepare kernels  $k_x$  on  $\mathcal{X}$  and  $k_y$  on  $\mathcal{Y}$  as  $x_1 \dots x_n \rightarrow \phi_x(x_1), \dots, \phi_x(x_n) \in \mathcal{H}_{k_x}$  and  $y_1 \dots y_n \rightarrow \phi_y(y_1), \dots, \phi_y(y_n) \in \mathcal{H}_{k_y}$ . Now we

apply classical CCA on Hilbert space  $(\mathcal{H}_x, \mathcal{H}_y)$

$$f \in \mathcal{H}_x, g \in \mathcal{H}_y \frac{\sum_{i=1}^n \langle f, \tilde{\phi}_x(x_i) \rangle_{\mathcal{H}_x} \langle g, \tilde{\phi}_y(y_i) \rangle_{\mathcal{H}_y}}{\sqrt{\sum_{i=1}^n \langle f, \tilde{\phi}_x(x_i) \rangle_{\mathcal{H}_x}^2} \sqrt{\sum_{i=1}^n \langle g, \tilde{\phi}_y(y_i) \rangle_{\mathcal{H}_y}^2}} \quad (2)$$

where  $\tilde{\phi}_x(x_i) = \phi_x(x_i) - \frac{1}{n} \sum_{i=1}^n \phi_x(x_i)$  and  $\tilde{\phi}_y(y_i) = \phi_y(y_i) - \frac{1}{n} \sum_{i=1}^n \phi_y(y_i)$  Without any lose of generality, we can assume that  $f = \sum_{i=1}^n \alpha_i \tilde{\phi}_x(x_i)$  and  $g = \sum_{i=1}^n \beta_i \tilde{\phi}_y(y_i)$

$$\rho = \alpha \in \mathbb{R}^n, \beta \in \mathbb{R}^n \frac{\alpha^T \tilde{K}_x \tilde{K}_y \beta}{\sqrt{\alpha^T \tilde{K}_x^2 \alpha} \sqrt{\beta^T \tilde{K}_y^2 \beta}} \quad (3)$$

where  $\tilde{K}_x^2$  and  $\tilde{K}_y^2$  are the centered Gram matrices. It is equivalent to performing CCA on two vectors of dimension  $n$  with covariance matrix equal to  $\begin{bmatrix} K_x^2 & K_x K_y \\ K_y K_x & K_y^2 \end{bmatrix}$ . Thus we observe that we can perform a kernelized version of CCA by solving the following generalized eigenvalue problem.

$$\begin{bmatrix} 0 & K_x K_y \\ K_y K_x & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \rho \begin{bmatrix} K_x^2 & 0 \\ 0 & K_y^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (4)$$

based on the Gram matrices.

#### B. Different Kernel Functions for KCCA

The kernel function is used in training and predicting. The parameter of this function can be set to any function of class kernel that computes the inner product in feature space between two vector arguments. In this work we consider, Gaussian RBF kernel,  $k(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{\sigma^2})$ , Laplacian kernel function,  $k(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|}{\sigma})$  and Polynomial function,  $k(x_i, x_j) = (\langle x_i, x_j \rangle + c)^d$ .

### IV. RESULTS

In this section we consider the simulation, influence function, and estimate of influence function results respectively. We generate multivariate normal (MVN) data by considering two covariance matrices  $CVM_1$  and  $CVM_2$  that are given below. We also generate data from uniform distribution having  $[-\pi, \pi]$ . After that we transform generated data by sin and cos function in different way. We find bias, standard error, mean square error, qualitative robustness index(QRI)[12] of the canonical correlation estimators at different models as well as transformed data. The results are represented in different tables and boxplots.

$$\begin{bmatrix} 1.0 & 0.8702 & -0.3657 & -0.3896 & -0.4931 & -0.2263 \\ 0.8703 & 1.0 & -0.3529 & -0.5522 & -0.6456 & -0.1915 \\ -0.3657 & -0.3529 & 1.0 & 0.1506 & 0.2250 & 0.0349 \\ -0.3896 & -0.5522 & 0.1506 & 1.0 & 0.6957 & 0.4957 \\ -0.4938 & -0.6456 & 0.2250 & 0.695 & 1.0 & 0.6692 \\ -0.2263 & -0.1915 & 0.0349 & 0.4957 & 0.669 & 1.0 \end{bmatrix}$$

and  $CVM_2 = \begin{bmatrix} 1.000 & 0.505 & 0.569 & 0.602 \\ 0.505 & 1.000 & 0.422 & 0.467 \\ 0.569 & 0.422 & 1.000 & 0.926 \\ 0.602 & 0.467 & 0.926 & 1.000 \end{bmatrix}$

#### A. Simulation Results

Here we report a simulation study of different correlation coefficients. In our study we consider three different sample sizes (50, 500, and 1000) with different types of Multivariate models. To perform simulation study we consider three experiments as follows.

*Experiment-1 (Data generated from multivariate normal(MVN) and MVN with 6%contamination):* In this experiment we generate data from MVN as ideal model and MVN with 6% fixed outlier as contaminated model. We consider three sample size  $n = 50$ ,  $n = 500$  and  $n = 1000$  and simulated 1000 times. In case of  $n = 1000$  we take fewer replications in order to avoid much calculation. After that we calculate bias, standard error, MSE as well as QRI which are given in the Table I and Table II respectively. From Table I, we see that, although in case of bias (Minimum values as bold) the performance of classical and robust are better but over all, we may conclude that kernel estimators are better than classical as well as robust estimators specially at contaminated models and large sample size.

*Qualitative Robustness Index (QRI):* To measure the effect of contamination on different estimators at different contamination models we use qualitative robustness index;

$$QRI = \frac{1}{\sum |\xi(s)_{100\alpha} - \xi_{100\alpha}^{(s)c}|} \quad (5)$$

where  $\xi_{100\alpha}^{(s)} = 100^{th}$  percentile of the simulated sampling distribution of the estimators at standard model, and  $\xi_{100\alpha}^{(s)c} = 100^{th}$  percentile of the simulated sampling distribution of different estimators at contaminated model. We consider  $\alpha = 0.005, 0.01, 0.025, 0.05, 0.1, 0.5, 0.9, 0.95, 0.975, 0.99, 0.995$  for our standard model and contaminated model. By considering qualitative robustness index (Maximum values as bold)in Table II, we see that both robust and kernel method perform better than classical method.

*Experiment-2 (Data generated from multivariate normal and transformation with  $CC = 0.79$  and  $CC = 0.63$ ):*

In this experiment we generate data by using  $CVM_1$  and  $CVM_2$  and take transformation in the following way. In case of  $CVM_1$ , the data of x-set with three variables  $x_1, x_2$  and  $x_3$  transform by  $\sin(2x_1), \cos(x_2)$  and  $\sin(3x_3)$  as well as y-set as  $\cos(2y_1), \sin(y_2), \cos(3y_3)$ . For  $CVM_2$  we take transformation as  $\sin(x_1), \sin(3x_2)$  and  $\cos(2y_1), \cos(3y_2)$  for both x-set and y-set data respectively. By using canonical coefficient we sketch the boxplots of five estimators, which are given in Fig. 2a, and Fig. 2b. having first population  $CC = 0.79$  and  $CC = 0.63$  respectively. From the Fig.2, we observe that in both cases classical and robust method give very naive estimates over all sample sizes. But the performace of kernel method is the best. For

TABLE I  
Bias, standard error and mean square error vlues of simulated data

| Estimators | Model | Bias          |               |               | SE       |               |               | MSE           |                |               |
|------------|-------|---------------|---------------|---------------|----------|---------------|---------------|---------------|----------------|---------------|
|            |       | $n = 50$      | $n = 500$     | $n = 1000$    | $n = 50$ | $n = 500$     | $n = 1000$    | $n = 50$      | $n = 5000$     | $n = 1000$    |
| $CC$       | MVN   | <b>0.0183</b> | <b>0.0171</b> | <b>0.0011</b> | 0.0486   | 0.0503        | 0.0123        | 0.0489        | 0.0506         | 0.0123        |
|            | MVNC  | 0.6658        | 0.6713        | 0.6628        | 0.0311   | 0.0098        | 0.0069        | 0.4744        | 0.4604         | 0.4462        |
| $RC$       | MVN   | 0.0554        | 0.0460        | 0.0016        | 0.0657   | 0.0751        | 0.0136        | 0.0688        | 0.0772         | 0.0136        |
|            | MVNC  | <b>0.0473</b> | <b>0.0022</b> | <b>0.0012</b> | 0.0717   | 0.0193        | 0.0132        | 0.0739        | <b>0.0193</b>  | 0.0132        |
| $KG$       | MVN   | 0.1887        | 0.1887        | 0.10137       | 0.0134   | 0.0139        | 0.0144        | 0.0495        | 0.0495         | 0.0146        |
|            | MVNC  | 0.1909        | 0.0343        | 0.0160        | 0.0117   | 0.0246        | <b>0.0120</b> | 0.0481        | 0.0258         | <b>0.0123</b> |
| $KL$       | MVN   | 0.1044        | 0.2044        | 0.0591        | <b>0</b> | <b>0</b>      | 0.0212        | <b>0.4180</b> | <b>0.04180</b> | 0.0247        |
|            | MVNC  | 0.2044        | 0.1764        | 0.0659        | <b>0</b> | 0.0215        | 0.0206        | <b>0.0418</b> | 0.0526         | 0.0249        |
| $KP$       | MVN   | 0.0924        | 0.0938        | 0.0047        | 0.0372   | 0.0385        | <b>0.0121</b> | 0.0457        | 0.0473         | <b>0.0121</b> |
|            | MVNC  | 0.1809        | 0.0286        | 0.0454        | 0.0153   | <b>0.0019</b> | 0.0136        | 0.0480        | 0.0198         | 0.0157        |

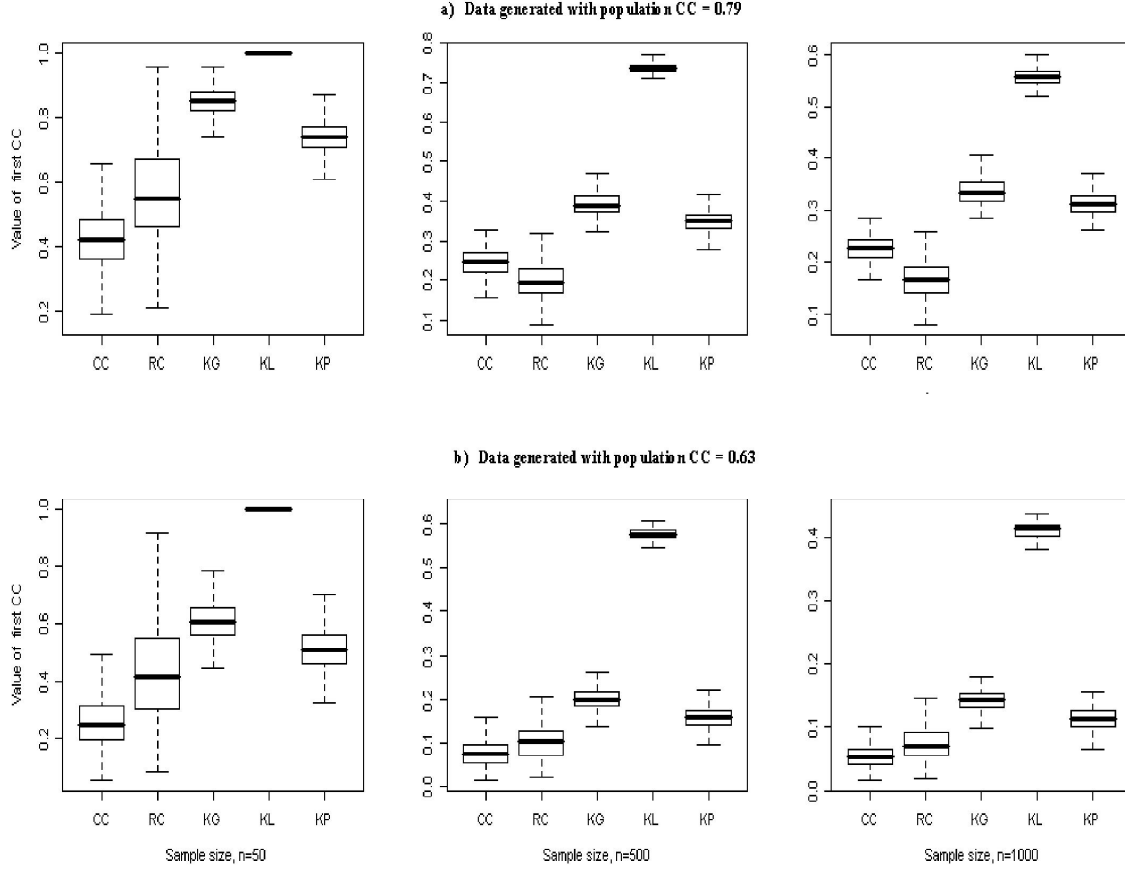


Fig. 2. Boxplots of canonical correlation coefficient of five estimators with population  $CC = 0.79$  and  $CC = 0.63$ .

sample  $n = 50$  the kernel estimator with Gaussian (KG) function gives smallest bias, on the other hand, for large sample Laplacian function (KL) gives smallest bias along with minimum variation in the both cases.

#### Experiment-3 (Generated from uniform distribution ):

In this section we generate data according to [4], [5] but size is 100. At first  $\theta$  is generated from the uniform distribution on  $[-\pi, \pi]$  after that a pair of two dimensional variables  $x$  and  $y$  are generated by  $x = [\theta, \sin 3\theta]$  and  $y = e^{\theta/4}[\cos 2\theta, \sin 2\theta]$ . Form Fig. 3, we see that classical

method does not give linear relationship, robust method gives higher coefficient but relation for first canonical variates are nonlinear.

TABLE II  
Qualitative Robustness Index

| Estimators | $n = 50$      | $n = 500$     | $n = 1000$    |
|------------|---------------|---------------|---------------|
| $CC$       | 0.0015        | 0.0014        | 0.0015        |
| $RC$       | <b>3.4483</b> | 0.0245        | <b>0.0025</b> |
| $KG$       | 0.0833        | 0.0066        | <b>0.0025</b> |
| $KL$       | 0.1           | <b>0.0412</b> | 0.0024        |
| $KP$       | 0.0102        | 0.0083        | 0.0024        |

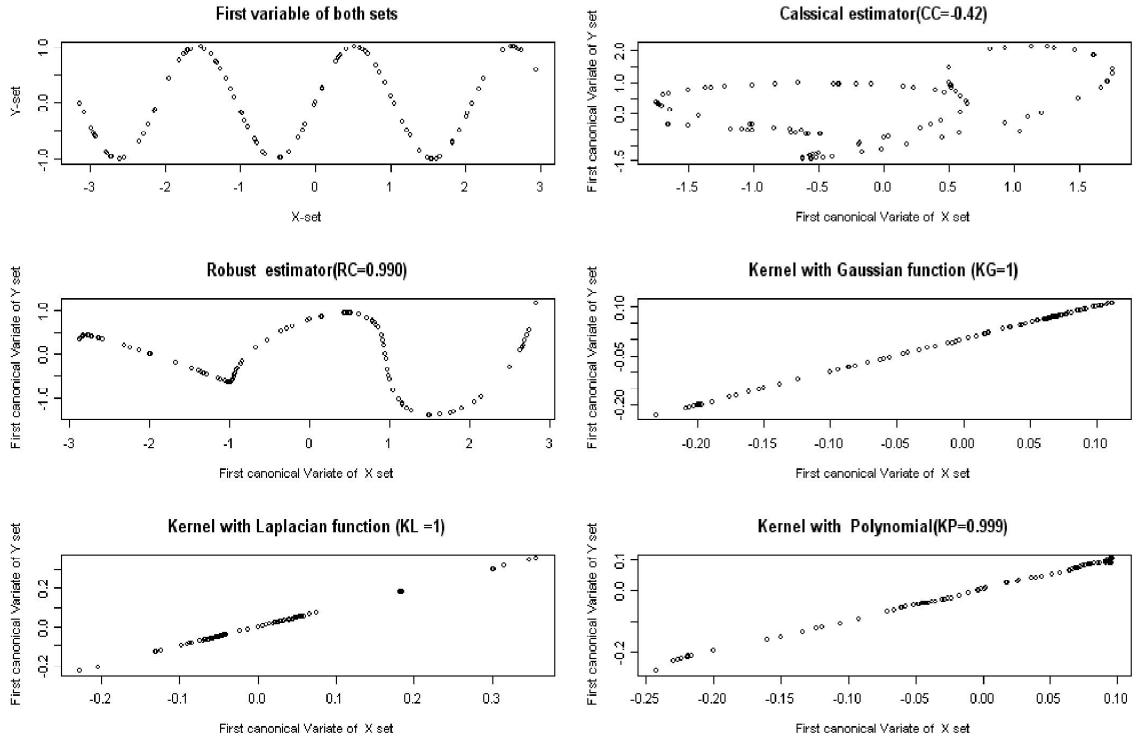


Fig. 3. Scatter plot of first canonical variate of data generated from uniform distribution with size  $n = 100$ .

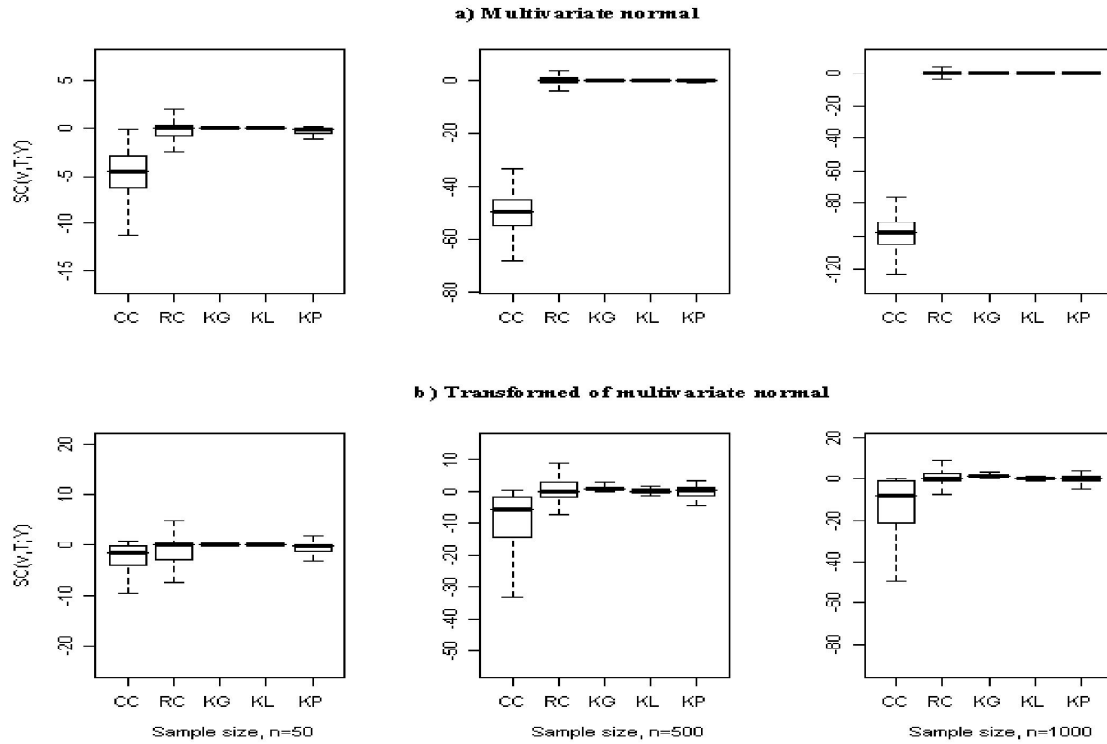


Fig. 4. Box plot of 200 numbers generated from  $SCn(v; T_n : V)$  where follows a) Multivariate normal and b) Transformation of Multivariate normal having  $n = 50, 500$  and  $1000$ .

On the other hand, kernel method shows the best performance providing linear relationship of canonical variates. Especially kernels with Gaussian (KG) and Laplacian function are found as the best ones in offering linear relationship. We have also observed that similar scatter plot for second canonical variate.

### B. Sensitivity Analysis

In this section following [13], we perform a simulation study to confirm some aspects of our findings with the help of sensitivity curves. Firstly we generate 200 samples of size  $n = 50$ ,  $n = 500$  and  $n = 1000$  from a MVN using  $CVM_1$  and compute  $SC_n(v, T_n; V)$  for each of these samples. We also consider an outlier values  $v = 500, 1000, 1500$  and  $0.5, 0.1, 0.15$  for x-set and y-set respectively. That is, the sample sizes are  $n = 51, n = 501$  and  $n = 1001$ . Boxplots of these 200 numbers are shown in Fig. 4a, for MVN using  $CVM_1$  and we repeat the work by taking a transformation as  $\sin(2x_1)$ ,  $\cos(x_2)$ ,  $\sin(3x_3)$  and  $\cos(2y_1)$ ,  $\sin(y_2)$ ,  $\cos(3y_3)$  for x-set and y-set respectively; boxplots of the values target are given in Fig. 4b, of classical estimator (CC), robust estimator (RC) kernel with Gaussian function (KG), kernel with Laplacian function (KL) and kernel with polynomial (KP).

## V. CONCLUSION

In this paper we compare the performances of five (classical, robust and kernel with three functions) estimators of canonical correlation coefficient that are commonly used in the statistical literature as well as in pattern recognition. Their performances are investigated through simulation and influence function. We report a simulation where sampling distributions are taken as multivariate normal with two covariance matrices,  $CVM_1$  and  $CVM_2$  and its contaminated forms, and we also consider uniform distribution with three different sample sizes (50, 500 and 1000) and in different way of transformation. The estimators are compared with regard to bias, standard error, MSE and QRI. Over all, the class of kernel estimators, especially the Gaussian function for small sample and Laplacian function for large sample perform better than the class of classical as well as robust estimators. A study based on sensitivity curves shows that outliers have the least effect on kernel estimator. We recommend using as a measure of kernel canonical correlation coefficient especially when the sample is large and contamination is likely.

## ACKNOWLEDGMENT

The authors wish to acknowledge the various helpful comments, assistance and support of the reviewers and Publication Committee.

## REFERENCES

- [1] A. K. Jain, R. P. W. and J. Mao. Statistical Pattern Recognition: A Review, *IEEE Transactions on Pattern analysis and Machine Intelligence*, vol. 3, no. 1, pp. 4-37, January, 2000.
- [2] H. Hotelling. Relations between two sets of variables, *Biometrika*, vol. 28, pp. 321-377, 1936.
- [3] D. R. Hardoon, S. Szedmak, J. Shawe-Taylor. Canonical correlation analysis: an overview with application to learning methods, *Technical Report, Royal Holloway University of London*, CSD-TR-03-02, 2003.
- [4] Akaho, S. (2001). A kernel method for canonical correlation analysis. In Proceeding of the 2000 Workshop on Information Based Induction Science (IBS2000), vol.17-18, pp.123-128, 2000.
- [5] S. Akaho. A kernel method for canonical correlation analysis, *International meeting of psychometric Society*, 2007.
- [6] S.Y. Huang, Y.R. Yeh, S. Eguchi. Robust Kernel Principal Component Analysis, Research Memorandum 1057, Institute of Statistical Mathematics, Japan 2008.
- [7] A. J. Branco, C. Croux, P. Filzmoser, M.R. and Oliviera. Robust Canonical Correlations: A Comparative Study, *Computational Statistics*, vol. 20, pp. 203-229, 2005.
- [8] S. Taskinen, C. Croux, A. Kankainen, E. Ollila, and H. Oja. Canonical Analysis based on Scatter Matrices, *Journal of Multivariate Analysis*, vol. 97, no.2, pp. 359-384, 2006.
- [9] D. Skocaj, A. Leonardis, and S. Fidler. Robust Estimation of Canonical Correlation Coefficients, *Proceedings of OAGM04*, pp.15-22, 2004.
- [10] F. R. Bach and M. I. Jordan. Kernel independent component analysis, *J. Mach. Learning Res.*, vol. 3, pp.1-48, 2002.
- [11] K. Fukumizu, F. R. Bach, and A. Getton. Consistency of kernel canonical correlation, Research Memorandum 942, Institute of Statistical Mathematics, Japan 2005.
- [12] M. Huda. *Bootstrapping and Robust Estimation in Factor Analysis: Application in Stock Price Data*, M.Sc. thesis, Department of statistics, University of Rajshahi, Bangladesh, 2005.
- [13] M. A. Alam, M. Nasser, and A.H.M. R. Imon. Sensitivity and Influence Analysis of Estimators of Correlation Coefficients, *Journal of Applied Statistics and Probability*, vol. 3, pp. 119-136, May, 2008.



# An Efficient System for Recognition of Human Face in Different Expressions by Some Measured Features of the Face Using Laplacian Operator

Abu Shamim Mohammad Arif, G. M. Atiqur Rahaman, Gautam Kumar Biswas, Syed Nurul Islam

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh

E-mail: shamimarif@yahoo.com, atiq@ku.ac.bd, gautam\_cse\_02@yahoo.com,  
nuru\_ku\_02@yahoo.com

**Abstract** — In this paper we present an efficient system for face recognition with high recognition rate. In our proposed method, at first we detect the face from an image, then two main significant edge lines – chin-line and nose-line are determined and next we apply third order polynomial regression on these two lines to get a third order polynomial equation with four coefficients for each line. Here we use Laplacian operator to determine the curve of chin line and nose line. Then we measure distances from the middle point of the nose curve to the chin line horizontally and vertically, the width of the two eyebrows, the width of forehead and the distance from pupil to eyebrow. We also determine two regions – eye-region and nose-region. For each of these regions, we determine the average value of each three basic colors: red, green and blue. We then store the coefficients of the detected edge lines, the average values of the three basic colors and other distances and perform the task of recognition process. The existing face is recognized for which the weighted error is minimum and higher than a predefined threshold value. Experimental results show that our proposed method successfully recognizes face at a very high rate.

**Index Terms** — Gauss Jordan Method, Independent Component Analysis (ICA), Polynomial Regression, Laplacian Operator, Principal Component Analysis (PCA).

## I. INTRODUCTION

Pattern recognition and matching consists of processing the input, classifying, and matching it with a known pattern. Face recognition is a very complex form of pattern recognition. Classifying a pattern with high dimensions requires a restrictively large number of training samples. This is known as the 'Curse of Dimensionality'. A number of ways have been proposed to solve this problem [1]-[7]. Finding an effective means to reduce the dimensionality is the first step in face recognition.

We have proposed a system for face recognition, which is very efficient. Here we have used second order Laplacian operator for determination of curve and then polynomial regression as a curve fitting mechanism. Beside we have used the width of forehead and eye-brow and also determined the distance between the pupil and

eyebrow. Last of all we have determined the average RGB values of forehead and nose region. In our method, we have used the curves of nose line and chin line because the curves of nose and chin are expressions invariant. Our system has a good recognition rate with high performance.

## II. LITERATURE SURVEY

Ashish Kapoor and Rosalind W. Picard in their paper [8] described PCA and their capability of face recognition in different context. In the paper they describe a fully automatic system that tracks eyes and eye-brows in real time. The pupils are tracked using the red eye effect by an infrared sensitive camera equipped with infrared LEDs. Templates are used to parameterize the facial features. The template parameters are recovered by PCA analysis on these extracted images using PCA based method. For fully automatic facial expression their system performs better.

Srinivas Gutta et al. in their paper [9] used partial faces for identification. In their approach they consider an ensemble of Radial Basis Function (RBF) Networks. They showed that the accuracy rate of their approach is (a) 96% on the partial-face identification task and (b) 97% on the full-face identification task. They used RBF of partial faces and proved that using RBF is the most promising method for face recognition.

Qingshan Liu et al. in their paper [1] used Fisher linear Discriminant Analysis (FLDA) for face recognition. FLDA is based on a linear projection from the image space to a low dimensional space by maximizing the between-class scatter and minimizing the within-class scatter. In this paper, they represent to use kernel based FLDA for face recognition. The kernel trick is used firstly to project the input data into an implicit space called feature space by nonlinear kernel mapping, and then FLDA is adapted to this feature space, thus a nonlinear discriminant can be yielded in the input data. They compared its performance with kernel PCA and FLDA and found that the recognition rate is higher than kernel PCA. For face recognition it gives better performance.

Dominique Valentin et al. [10] in their paper have integrated psychological and physiological information for recognizing faces. Their recognition system is view based. Recognition performance increases with the number of views. According to them the recognition or identification of a face from a new orientation can be achieved efficiently by interpolating between two extreme orientations.

Yongbin Zhang et al. [7] in their paper have used subspace projection method using the concept that a test face may have some expressions not present in the training images of the same face. In this case they have separated different portions or subspaces from a face-image and then weighted these spaces such that the subspaces that show little change in different expressions will get greater weights than those which show huge changes. They have used PCA, ICA and LDA for projecting subspaces and proved that LDA is the most promising method.

Bruce A. Draper et al. [11] in their paper have described the comparison between PCA and two architectures of ICA methods. They have shown us that PCA decorrelates input data using second order statistics and generate compressed data with minimum mean-squared re-projection error where ICA minimizes both second order and higher-order dependencies in the input. For some cases PCA gives us better results than ICA and for some other cases ICA performs better. They have shown ICA as the generalization of PCA.

### III. EDGE DETECTION AND FITTING OF CURVE

In our proposed method we have used a second-order Laplacian operator to detect edge. The Laplacian of a function  $f(x,y)$ , denoted by  $\nabla^2 f(x,y)$ , is defined by the equation no.1 as follows:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \quad (1)$$

Here we have used discrete difference approximations to estimate the derivatives and represent the Laplacian operator by the following 3 x 3 convolution mask which is shown in Fig. 1:

|   |    |   |
|---|----|---|
| 0 | 1  | 0 |
| 1 | -4 | 1 |
| 0 | 1  | 0 |

Fig. 1. Convolution Mask

We should note that first derivative operators exaggerate the effects of noise. Second derivatives will exaggerate noise twice as much if no directional information about the edge is given.

#### A. Curve Fitting and Co-efficient Determination

Curve fitting and co-efficient determination is our root method for recognizing faces. By using this method we introduce the process of taking shape knowledge of a face. In our method we have performed several tasks that result the final recognition. For determination of the training value we take N different faces of the same person. From each face we cut the chin-lines and nose-line. For these lines polynomial regression is applied for obtaining third order polynomial equations. Suppose n coordinates for an edge are:  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ .

Then the equation is generated for each of the lines which take the form as:  $Y = A + Bx + Cx^2 + Dx^3$

From this equation the coefficients A, B, C and D are stored. For N faces related to the same person we have determined the edge equations and for each of the edges we have N values relating to the coefficient A, N values relating to the coefficient B and also for C and D. We determine the averages for each of the coefficients that are stored in the database.

Suppose we are at a point x some where in the image. We have three basic colors for this point. Now suppose this point corresponds to the middle element '-4' of the convolution matrix (Fig. 1). We consider 9 pixel elements of the image, the current pixel x being the middle of the 3x3 matrix. Then we add red colors of these 9 pixels with the corresponding coefficients multiplied with the red colors. This resultant color divided by 9 will be the color of the middle point that is the pixel x. Thus we determine the green and blue color of x. Now this RGB color got from the calculation is assigned to the pixel x. This operator does not give the edge lines but gives gray colored edge on a black back ground. A main face, its detected face by matrix convolution method and the detected chin-line are given here in Fig. 2.

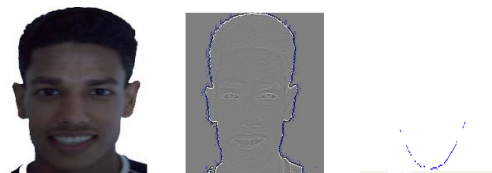


Fig. 2. Main face, Detected face and Chin Line respectively.

TABLE I  
SAMPLE RESULT FOR COEFFICIENTS OF CHIN LINE

| A         | B         | C         | D        |
|-----------|-----------|-----------|----------|
| -1.507236 | 1.3304128 | -0.018014 | 1.085634 |

The four coefficients of the curve which we get from the chin line of the face in Fig. 2 using Gauss Jordan Method is shown in Table I.

We have also determined the nose curve by Laplacian operator. The Fig. 3 shows a main face and the detected nose curve. After detecting the curve the four coefficients are determined by Gauss Jordan method which has been shown in Table II.



Fig. 3. Determination of nose curve.

TABLE II  
VALUES OF FOUR COEFFICIENTS A, B, C AND D OF NOSE LINE

| A         | B         | C         | D        |
|-----------|-----------|-----------|----------|
| -8.935705 | 0.8654327 | -1.080604 | 0.968904 |

#### B. Co-ordinate Equalization System

The same curve placed at two different places in a graph will have two different equations though these two can give the same equation. If we select a direction for the curves to start from and go through, then by subtracting the x and y coordinate values of the first point from all the x and y coordinate values of all the points, we can make the two curves such that they both are generated from the (0,0) point of the coordinate system. Thus the same curves that were giving two different equations will now give identical equations. For edge detection in our approach this technique is used. So the two edges for the same person that might give two different equations for two different templates will give the almost same equations.

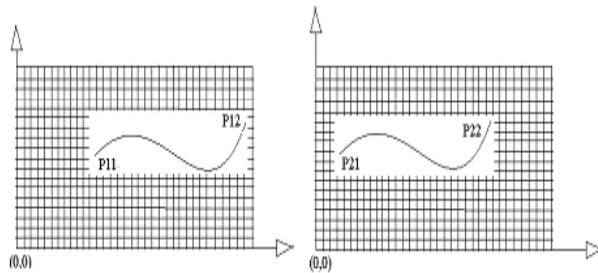


Fig. 4. Coordinate system equalization.

The two curves in the two graphs in Fig. 4 denoted by P11-P12 and P21-P22 will show different equations but by the above coordinate-system these two will give the same equations.

#### IV. ARCHITECTURE OF FACE RECOGNITION SYSTEM

The overall processing steps of the training phase of our proposed system as shown in Fig. 5.

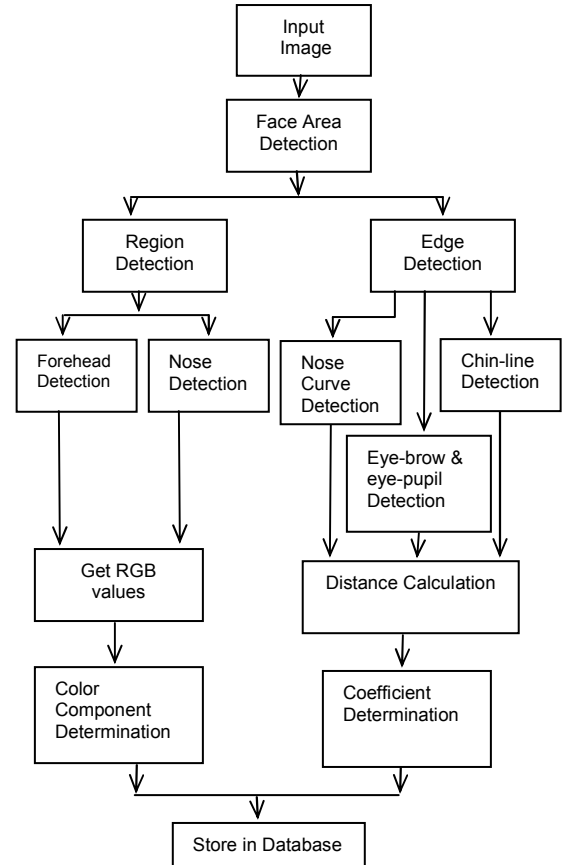


Fig. 5. Architecture of the Training Module of the System.

#### A. Determination of horizontal and perpendicular distances between nose line and chin line

We have deter-mined the horizontal distance between middle point of the nose line and chin line as illustrated in Fig. 6. We have a taken a point on the chin line which has the same x-coordinate value. Then the distance is determined by simply deducting the y-coordinate value between these two points.

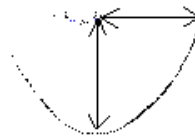


Fig. 6. Horizontal and perpendicular distances between nose and chin line.

Table III shows the experimental values of the two distances for the image shown in Fig. 7.

TABLE III  
EXPERIMENTAL VALUES OF HORIZONTAL AND PERPENDICULAR DISTANCES

|                                                                                                 |    |
|-------------------------------------------------------------------------------------------------|----|
| Horizontal distance from the middle point of the nose curve to chin line                        | 18 |
| Perpendicular distance from the middle point of the nose curve to the lowest point of chin line | 27 |

Here all the distances are in number of pixels.

The perpendicular distance between middle point of the nose line and the lowest point of the chin line is determined by simply deducting the x-coordinate value between these two points. In this way we have got two distances in pixels and then a rational measurement is calculated by dividing the amount of distances by the face length which is calculated as follows:

1. (Horizontal distance / total face length)\*100
2. (Perpendicular distance / total face length)\*100

#### B. Determination the width of forehead and eyebrow

We have calculated these measurements after face detection and for this we have completed the following tasks.

Start from the highest point of the face and determine the RGB color values of the blackest point in the region. The points that are very near to this most black point considering the color values are converted to fully black points by setting their RGB values to (0,0,0). Respectively the vertical beginning and ending points of the forehead are determined as we check black points of the face beginning from the top to bottom. We get two horizontal limits indicating the y co-ordinates of the vertical beginning and ending of the forehead. The distance of these two limits give us the width of the forehead. We get eyebrow region after the forehead region as we proceed from the top of the face. Accordingly we determine the two horizontal limits of this region and from the distance of these two limits we determine the eyebrow-width. The features that we have taken are as below:

1. (Forehead width / total face length)\*100
2. (Eyebrow width / total face length)\*100

The Fig. 7 shows a main face, the detected face and the determined forehead and eyebrow width.



Fig. 7. Determination of forehead and eyebrow region.

#### C. Determining the distance between the pupil and eyebrow

The vertical beginning of the eye region is the ending horizontal line of eyebrow as we check black points of the face beginning from the top to bottom. We also get a horizontal limits which passes through the middle points of pupil and the difference of y coordinates of these two vertical line gives the distance between pupil and eyebrow which is shown in Fig 8.



Fig. 8. Determining the distance between the pupil and eyebrow.

Table III represents the calculated four distant features of the face shown in Fig. 8.

TABLE III  
VARIOUS DISTANCE FEATURES OF A FACE

| Features                                         | Measured Value (in pixels) |
|--------------------------------------------------|----------------------------|
| Horizontal distance from nose curve to chin line | 17                         |
| Vertical distance from nose curve to chin line   | 25                         |
| Width of forehead                                | 26                         |
| Width of eyebrow                                 | 5                          |
| Distance between the pupil and eyebrow           | 8                          |

#### D. Detection of region and determination of color component

We also take into account the nose region like the forehead region. The nose region covers two horizontal lines one at the distance of (total perpendicular length \* 0.54) from the head top position and the other remains at the distance of (total perpendicular length \* 0.71) from the head top. We have determined three average values of red, green and blue colors for these regions according to (2). After determining these three average values, the values are stored in the database.

$$RCA = \frac{\sum_{i=0}^{TP} C(i)}{TP} \quad (2)$$

Where, RCA = Region Color Average, TP = Total number of Pixels, C(i) = i-th pixel color value.

TABLE IV  
RGB VALUES FOR TWO REGIONS

|          | Red | Green | Blue |
|----------|-----|-------|------|
| Forehead | 128 | 128   | 255  |
| Nose     | 153 | 153   | 255  |

Table IV shows an example of calculated RGB values according to (2) of the face shown in Fig. 8.

## V. RECOGNITION

For recognition purpose we have used weighting mechanism. For a face, we determine the four coefficients which we get from each of the chin-line and nose curve, the width of eyebrow and forehead, the horizontal and perpendicular distance between the nose line and chin line, the distance between the pupil and eyebrow and the average RGB values of forehead and nose region. For each of the stored faces we determine the absolute differences for the corresponding values of the face to be recognized. These differences are multiplied with predefined weight values and then are summed together. We determine the exact absolute differences for the chin-line and nose-line coefficients and these differences are added to the weighted-error. We have also defined a threshold value for this purpose which is used in such a way that if the lowest weighted error is less than or equal to the threshold value then the face corresponding to the lowest weighted error is de-fined as the recognized face. The face with the lowest total weighted error is taken as the matched face.

## VI. EXPERIMENTAL RESULTS

### A. Time Complexity

The time complexity of our method is  $O(n^2)$ . Here the execution time of our method is slightly high than real time face recognition using polynomial regression by Mondal [12], which is shown in Fig. 9, but recognition rate is much higher than real time face recognition using polynomial regression method [12] or PCA based method which is represented in Fig. 10.

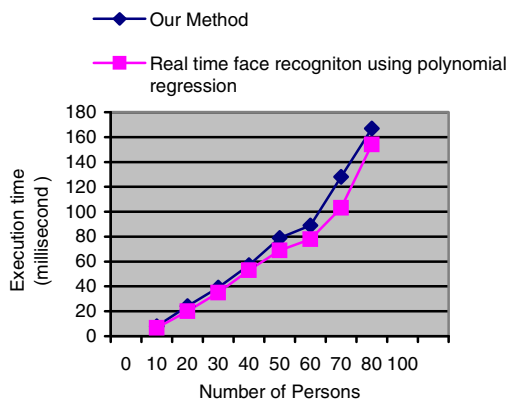


Fig. 9. Comparison of Time performance of the proposed method.

### B. Recognition Efficiency

The detection efficiency of our method is much higher than the real time face recognition using polynomial regression and PCA based method. For different expressions the previous method gives very low performance where as our method gives a relatively high performance. The Table V shows the percentage of detection efficiency of our method, real time face recognition method and PCA based method.

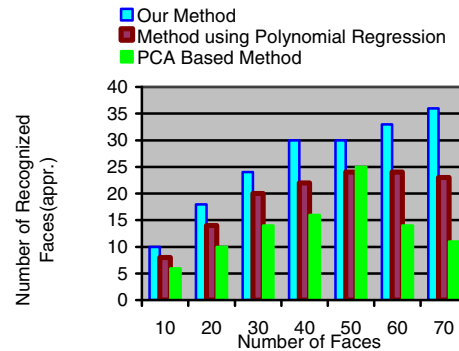


Fig. 10. Detection efficiency of the proposed method.

In Fig. 10, the graphical representation of the efficiency of our method, real time face recognition method and PCA method for about seventy faces is shown.

## VII. DISCUSSION ON PERFORMANCE

Beside a very high detection rate, our method also overcomes some problems that exist in real time face recognition methods which are described in the papers [12] and [6]. If the faces between two or three images are displaced in image template, existing real time face recognition methods give very low performance. Also for different expressions the performance of the existing method is not good. If the faces are placed at angled position at image template then previous methods again show very low recognition rate and at some stage they are totally unable to recognize. But in this case our method survives.

There is a draw back of our method which is that we take the average red, green, and blue color values for each region. As we have taken average values there is a risk of false recognition. The average color values of two distinct regions may be the same or very near. For this issue we have got some wrong results from our method in recognizing faces. But it is the fact that we have removed some drawbacks which were possessed by the previous methods. We have also tried with gradient based method but the method some times produces too much

TABLE V  
DETECTION EFFICIENCY FOR DIFFERENT EXPRESSIONS

| Method                                                            | Number of faces |     |     |       |     |     |     |
|-------------------------------------------------------------------|-----------------|-----|-----|-------|-----|-----|-----|
|                                                                   | 10              | 20  | 30  | 40    | 50  | 60  | 70  |
| Proposed method                                                   | 95%             | 89% | 80% | 75.5% | 59% | 55% | 51% |
| Polynomial regression<br>method for Real Time Face<br>Recognition | 75%             | 70% | 65% | 56%   | 48% | 40% | 33% |
| PCA based method                                                  | 55%             | 50% | 47% | 40%   | 30% | 23% | 15% |

complexity. So our proposed method is certainly an improvement over the traditional techniques.

### VIII. CONCLUSION

Recognition rate of our method is relatively high than other existing methods though its time and run time storage complexity is a little bit higher than polynomial regression method. Our method works properly for different expression of a human face. As our process integrates shape determination technique with the knowledge of color component distribution of sub-regions, this survives in many critical situations where the PCA or the ICA based methods fail to recognize. For improved and more accurate recognition rate fourth, fifth or even any higher order polynomial regression can be used. In this process we have used second order edge detection operator that is Laplacian operator which detect the edges effectively. We have determined two distances from the nose curve to the chin line horizontally and vertically. We have also determined the width of eyebrows and the distance between eye pupil and eyebrow. Since the features that we have used are over all expression invariant hence our method is effective and successful for recognizing faces in different expressions.

### REFERENCES

- [1] Q. Liu, R. Huang and S. Ma, "Face Recognition Using Kernel Based Fisher Discriminant Analysis", *National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Science*, P.O. Box 2728, China.
- [2] T. Heseltine, N. Pears, J. Austin, "Three-Dimensional Face Recognition: An Eigensurface Approach", *Advanced Computer Architecture Group, Department of Computer Science*, The University of York.
- [3] Y. Zhang and A.M. Mart'ínez, "Recognition of Expression Variant Faces Using Weighted Subspaces", *Department of Electrical and Computer Engineering*, The Ohio State University.
- [4] Deepak S. Turaga and T. Chen, "Face Recognition Using Mixtures Of Principal Components", *Video and Display Processing Electrical and Computer Engineering*,

*PhilipsResearch USA Carnegie Mellon University Briarcliff Manor, NY 10510 Pitts-burgh, PA 15213 deepak.turaga@philips.com.*

- [5] L. Torres L. Lorente Josep Vila, "Automatic Face Recognition Of Video Sequences Using Self-Eigenfaces", *Department of Signal Theory and Communications Polytechnic University of Catalonia Barcelona, Spain*.
- [6] M. Mondal, G.M.A. Rahaman, D. Tarafder, "Real Time Face Recognition Using Minimum Measurements When At Least Two Thirds of The Face Is Present in The Image", *Proc. of the 9th International Conference on Computer and In-formation Technology (ICCIT 2006)*, IUB, Dhaka, pp: 247-252.
- [7] Y. Zhang, A. M. Mart'ínez, "Recognition of Expression Variant Faces Using Weighted Subspaces", *ICPR, Proceedings of the Pattern Recognition*, 17th International Conference on (ICPR'04) Volume 3 - pp: 149 – 152.
- [8] A. Kapoor and R.W. Picard, "Real time, Fully Automatic Upper Facial Feature Tracking", *MIT Media Laboratory*, 20 Ames Street, Cambridge, MA 02139.
- [9] S. Gutta, V. Philomin and M. Trajkovic, "An Investigation into the use of Partial Faces for Face Recognition", *Philips Research-USA*, 345 Scarborough Rd. Briarcliff Manor, NY 10510.
- [10] D. Valentin, H. Abdi., B. Edelman "What Represents A Face: A Computational Approach for the Integration of Physiological and Psychological Data", *Perception*, 1997, volume 26(10), pp: 1271-1288.
- [11] W.S. Yambor, B.A. Draper, J.R. Beveridge, "Analyzing PCA-based face recognition algorithm: Eigenvector selection and distance measures", *Christensen, H., Phillips, J. (Eds.), Empirical Evaluation Methods in Computer Vision*. World Scientific Press, Singapore, 2002.
- [12] M. Mondal, M.M. Rahman, M.A. Hossain, K. H. Talukder, "Real-Time Face Recognition Using Polynomial Regression and Sub-region Color Component Distribution", *Proc. of the 8th International Conference on Computer and Information Technology( ICCIT 2005)*, pp. 1177-1182.

# An Effective Term Weighting Method Using Random Walk Model for Text Classification

Md. Rafiqul Islam and Md. Rakibul Islam

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh

**Abstract** — Text classification may be viewed as assigning texts in a predefined set of categories. However there are many digital documents that are not organized according to their contents. So it is difficult task to find relevant documents for a user. Automatic text classification problem can solve this problem. In this paper we introduce a new random walk term weighting method for improved text classification. In our approach to weight a term, we exploit the relationship of local (term position, term frequency) and global (inverse document frequency, information gain) information of terms (vertices). Moreover, we weight terms by considering co-occurrence and semantic relation of terms as a measure of dependency. To evaluate our term weighting approach we integrate it in Rocchio text classification algorithm and experimental results show that our method performs better than other random walk models.

**Index Terms** — Text classification, random walk model, semantic relation, information gain.

## I. INTRODUCTION

Most of the today's organizations consist of unstructured textual information in electronic form. Users are facing the challenge of organizing, analyzing and searching the ever growing amounts of documents. A primary application of text categorization is to assign subject categories to documents to support information retrieval, or to aid human indexers in assigning such categories [5]. Categorization can also be used to build a personalized net news filter, which learns about the news-reading preferences of a user. Text categorization can ease the organization of increasing textual information, in particular Web pages and other electronic form documents [5].

The random walk ranking algorithm on a graph proposed by Brin and Page in [7], has been used in citation analysis, social networks and analysis of the link structure of the web [1]. A graph-based ranking algorithm is a way of deciding the importance of a vertex within a graph, by taking into account global information, recursively computed from the entire graph, rather than relying only on local vertex-specific information. The basic idea implemented by a random walk algorithm is that, when one vertex links to another one, it is basically casting a vote for other vertex [2]. The higher the number of votes that are cast for a vertex, the higher the importance of the vertex. Term weighting is one of the most important considerations for various language

processing applications. The purpose of a term weighting method is to classify the terms by assigning them weights corresponding to how well they improve both recall and precision. In [2] several variations of random walk ranking algorithm are discussed and experimental results show that the new random-walk based approaches outperform the traditional term frequency method. Hassan et al. argue that the bag-of-words model may not be the best technique to capture term importance [2]. Here the new measure of term weighting integrates both the locality of a term and its relation to the surrounding context using a co-occurrence relation in which terms that co-occur in a certain context are likely to share between them some of their importance. Co-occurrence, dependency and semantic networks are three ways to build graph of given text, are discussed in [9]. In co-occurrence network two words are linked if they appear together within the same sentence or predefined window Again in semantic network terms or meanings are the nodes and various semantic relationships are the links. In our approach to weight a term, we exploit the relationship of local (term position, term frequency) and global information (inverse document frequency, information gain) of vertices. Moreover, we weight terms by considering co-occurrence and semantic relation of terms as a measure of dependency. To evaluate our term weighting approach we integrate it in Rocchio [3] text classification algorithm and experimental results show that our method performs better than other random walk models discussed in [2].

## II. RANDOM WALK MODEL - TEXTRANK

The basic idea implemented by a random-walk algorithm is that of "voting" or "recommendation." When one vertex links to another one, it is basically casting a vote for that other vertex. The higher the number of votes that are cast for a vertex, the higher the importance of the vertex [2]. Moreover, the importance of the vertex casting a vote determines how important the vote itself is, and this information is also taken into account by the ranking algorithm. There are several random-walk algorithms that have been proposed in the past and between these algorithms, TextRank is previously found successful in a number of applications, including Web link analysis,



social networks, citation analysis, and more recently in several text processing applications [2].

Given a graph, let  $V_a$  be the set of vertices that point to vertex  $v_a$  (predecessors), and  $V_b$  be the set of vertices that vertex  $v_a$  points to (successors). The TextRank score associated with the vertex is defined using the equation 1.

$$s(v_a) = (1-d) + d \times \sum_{b \in \text{In}(v_a)} \frac{s(v_b)}{|\text{Out}(v_b)|} \quad (1)$$

Where  $s(v_a)$  = score of vertex  $v_a$ ,  $s(v_b)$  = score of vertex  $v_b$ ,  $\text{In}(v_a)$  = indegree of vertex  $v_a$ ,  $|\text{Out}(v_b)|$  = outdegree of vertex  $v_b$  and  $d$  is a parameter that is set between 0 and 1 [2].

The score of each vertex is recalculated upon each iteration based on the new weights that the neighboring vertices have accumulated. Neighbors of a vertex are determined by window size  $N$ . If  $N$  is 3 then consecutive forward and backward three vertices are neighbor. The algorithm terminates when the convergence point is reached for all the vertices, meaning that the error rate for each vertex falls below a pre-defined threshold. This vertex scoring scheme is based on a random-walk model, where a walker takes random steps on the graph, with the walk being modeled as a Markov process. Under certain conditions, the model is guaranteed to converge to a stationary distribution of probabilities associated with the vertices in the graph. Intuitively, the stationary probability associated with a vertex represents the probability of finding the walker at that vertex during the random-walk, and thus it represents the importance of the vertex within the graph [2].

### III. RELATED WORK

Salton is a standard weighting scheme based on “term vector model” to represent documents in terms of vectors [4]. The Salton's approach is used to assign weights to terms by considering local significance of a term exists for individual documents, denoted by  $tf$ , where it indicates term frequency. Global significance of a term appearing across different documents is denoted by  $idf$ , also known as the inverse document frequency. Hence, the weight (term frequency inverse document frequency) is a weight often used to evaluate IR systems [4].

$$W_i = tf_i * \log(D / df_i) \quad (2)$$

Where  $df_i$  = document frequency or number of documents containing term  $i$ ,  $D$  = number of documents in the corpus.

Two new random walk methods discussed by Hassan et al. in [2] used for text classification. Hassan et al. introduce a system that models the weighting problem as a

“random-walk” rather than “random choice” [2]. They assume an imaginary reader who steps through the text on a term by term basis. In this setting, the importance of the term is determined by the probability of the random-walker to encounter the target term in the text during the walk. In this paper, the various best random-walk based term weighting models are summarized

This model represents an undirected graph approach that uses the weighted edge version of PageRank with a variable damping factor. The weight of an edge is calculated by the equation 3.

$$E_{v_1, v_2} = tf \cdot idfv_1 \times tf \cdot idfv_2 \quad (3)$$

Where  $E_{v_1, v_2}$  is the edge connecting  $v_1$  to  $v_2$  and  $tf \cdot idfv_1$ ,  $tf \cdot idfv_2$  represent the term frequency multiplied by the inverse document frequency respect to vertices  $v_1$  and  $v_2$ . The damping factor is expressed as a function of the ‘incoming edges’ weight, calculated as follows:

$$d_{E_{v_1, v_2}} = E_{v_1, v_2} / E_{\max} \quad (4)$$

where  $d_{E_{v_1, v_2}}$  is the damping factor and  $E_{\max}$  represents the highest weight for an edge in the graph. Thus the score of a vertex can be calculated by the formula in [2] is as follows:

$$s'(v_1) = \frac{(1-d)}{|N|} + \sum_{v_2 \in \text{In}(v_1)} C \times \frac{d_{E_{v_1, v_2}} \times s(v_2)}{|\text{Out}(v_2)|} \quad (5)$$

where  $s'(v_1)$  = score of vertex  $v_1$ ,  $N$  = total number of nodes,  $C$  = scaling constant [2]. Here the new measure of term weighting integrates both the locality of a term and its relation to the surrounding context.

$W_{e.oc}$ : This model is similar to the above approach however the damping factor for an edge is estimated in terms of the bigram co-occurrence frequency of the two nodes connected by the edge. For example, if the bigram “free software” occurred four times in a document then the weight of the edge connecting ‘free’ and ‘software’ is four.

### IV. TERM WEIGHTING METHOD USING GRAPH BASED RANDOM WALK MODEL

This section describes the method we use to weight terms to classify text. In our proposed method to draw graph of document's text we consider two relations of terms; co-occurrence and semantic relations. After constructing graph we apply our term weighting method. The details of our method are given below.

#### A. Term Weighting Method



1) To identify the text units (terms), the document is tokenized using text operations - stopword removals, stemming and each token is annotated with part of speech tags. We use the list of stop words enlisted in the Smart System [8] to remove stop word from graph. We add only noun and adjective tagged terms to build text graph. To avoid excessive growth of the graph size by adding all possible combinations of sequences consisting of more than one lexical unit (ngrams), we consider only single words as candidates for addition to the graph.

2) For each term in the processed documents, we calculate and random-walk weights ( ). To calculate the of each vertex, we use the method described by Yue-Heng Sun et al. [6]. The terms in the title should have higher importance than those in the abstract; similarly the terms in the abstract should have higher importance than those in the text, and so on. Therefore in [6], computing method of uses some factors to reflect the importance of term position, shown in the equation 6.

$$tf_{ij} = \alpha tf_{ij1} + \beta tf_{ij2} + \gamma tf_{ij3} \quad (6)$$

Where,  $tf_{ijk}$  is the frequency of a term in the  $k^{th}$  area and  $\alpha, \beta, \gamma$  are the factors that can be adjustable according to pre-experiments. Here  $\alpha > \beta > \gamma > 1$ .

Moreover, to give more importance to a term that has discriminating power to identify a document, Yue-Heng Sun *et al.* introduce information gain ( $IG_j$ ) for a term  $j$  [6]. We calculate  $tf \times idf$  according to the equation 7 given in [6].

$$tf_{ij} \times idf_j = \frac{tf_{ij} \times \log\left(\frac{N}{n_j} + 0.01\right) \times IG_j}{\sqrt{\sum_{i \in d} \left[ tf_{ij} \times \log\left(\frac{N}{n_j} + 0.01\right) \times IG_j \right]^2}} \quad (7)$$

Where  $tf_{ij}$  = frequency of term  $j$  within document  $i$ ,  $idf_j$  = inverse document frequency of term  $j$ ,  $n_j$  = number of documents holding term  $j$ ,  $N$  = total number of documents in the corpus. The  $tf_{ij} \times idf_j$  value of each selected term reflects their importance within the document and corresponds to whole documents corpus.

Weight of an edge is calculated using equation 3. Weight of each edge is required to determine the damping factor  $d_{E_{v1,v2}}$  in equation 4. The value of damping factor  $d_{E_{v1,v2}}$  varies in equation 5 to reflect importance of term. The more the value of damping factor, the term is more important to retrieve the related information. After constructing the graph we iterate the random walk

algorithm until convergence. The threshold value of convergence is 0.0001

3) Vertices are sorted on their final score. For ranking decision, we use the scores associated with each vertex. Once a final score is obtained for each vertex in the graph, vertices are sorted in reversed order of their score.

### B. Method of Building Text Graph

Generally a term can only be represented by one node in the graph. Terms that co-occurs within a given window size are connected by an edge to represent the co-occurrence relation of those terms. For a given term, all the terms that fall in the vicinity of this term are considered dependent terms [2]. This is represented by a set of edges that connect the term to all the other terms in the window. We remove the stop words and stem the remaining words to construct the graph.

To connect the terms using semantic relation, we follow the method described in [9]. In [9] two words in the text are linked only if they have semantic relations and the weights of the links which represent the relatedness of two words can be computed by the following formula:

$$Sim(W_1, W_2) = \max_{i=1 \dots n, j=1 \dots m} Sim(S_{1i}, S_{2j}) \quad (8)$$

$W_1$  has  $n$  senses  $S_{1i}$  ( $i = 1 \dots n$ ) and  $W_2$  has  $m$  senses  $S_{2j}$  ( $j = 1 \dots m$ ). Then semantic similarity between  $W_1$  and  $W_2$  can be approximately computed by maximum similarity between their senses, which can be measured by Wordnet or HowNet [9].

## III. EXPERIMENTAL RESULT AND DISCUSSION

To evaluate our term weighting approach we integrate it in Rocchio [3] text classification algorithm. The Rocchio text classification method uses standard term weighted vectors to represent documents, and builds a prototype vector for each category by summing up the vectors of the training documents in each category [2]. Test documents are then assigned to the category with the closest prototype vector, based on cosine similarity.

In [2] three standard datasets: WebKB, LingSpam and 20Newsgroups are used in text classification evaluations. The WebKB dataset contains seven class labels: Project, Student, Department, Faculty, Staff, Course, and Other. For the evaluation purpose Hassan *et al.* divide the WebKB dataset into two sets - WebKB4 which consists of Project, Student, Faculty, and Course since they represent the largest categories and WebKB6 which additionally consists of Staff and the Department classes. We also use these three datasets organizing WebKB dataset as Hassan *et al.* to compare our result with the result of [2]. In [2] results are given in micro average and macro average and

TABLE I  
MICRO-AVERAGE RESULTS FOR DIFFERENT  
RANDOM-WALK MODELS

| Data Sets    | $rw_e \cdot idf$ | $rw_{e.oc}$ | Our Method |
|--------------|------------------|-------------|------------|
| WebKB4       | 87.7             | 87.5        | 88.3       |
| WebKB6       | 84.2             | 84.3        | 84.6       |
| LingSpam     | 98.3             | 98.5        | 98.5       |
| 20Newsgroups | 93.8             | 94.3        | 94.7       |

best results are obtained using the  $rw_e \cdot idf$  and  $rw_{e.oc}$  random walk models with window size 2. So we do our experiment with window size 2 and compare the obtained results with these two random walk models.

Table I and Table II summarize the experimental results of various random walk models including our method where window size is 2. There are several parameters we use for the evaluation purpose such as *damping factor*  $d = 0.85$ , *scaling factor*  $c = 0.95$ . We set the value of  $\alpha = \beta = \gamma = 1$ , as the data sets we sue are not augmented with position information (title, abstract, body etc). From the result sets we can claim that our method performs reasonably well than other random walk methods. .

## V. CONCLUSION

Text classification, broadly defined as determining and assigning topical labels to content, is a fundamental operation in digital collection. In this paper we propose a new random walk method for term weighting which can be used for text classification. In our approach to weight a term, we exploit the relationship of local (term position, term frequency) and global information (inverse document frequency, information gain) of vertices. Moreover, we weight terms by considering co-occurrence and semantic relation of terms as a measure of dependency. Using WebKB, LingSpam and 20Newsgroups data sets we test our method to compare with previous random walk models discuused in [2]. The experimental result shows that our method achieves better result than other methods.

In this paper we test our method using Rocchio text classification algorithm with window size 2. In future we want to test our approach with SVM [2] text classification algorithm and with window size more than 2.

TABLE 2  
MACRO-AVERAGE RESULTS FOR DIFFERENT  
RANDOM-WALK MODELS

| Data Sets    | $rw_e \cdot idf$ | $rw_{e.oc}$ | Our Method |
|--------------|------------------|-------------|------------|
| WebKB4       | 86.7             | 86.7        | 87.9       |
| WebKB6       | 75.3             | 75.8        | 75.9       |
| LingSpam     | 97.1             | 97.3        | 97.7       |
| 20Newsgroups | 93.7             | 94.3        | 94.6       |

## REFERENCES

- [1] Roi Blanco, Christina Lioma "Random Walk Term Weighting for Information Retrieval" *SIGIR '07* July 23-27, 2007, Amsterdam, The Netherlands. ACM 978-1-595993-597-7/07/0007.
- [2] Samer Hassan, Rada Mihalcea and Carmen Banea "Random Walk Term Weighting for Improved Text Classification" *In Proceedings of TextGraps: 2nd Workshop on Graph Based Methods for Natural Language Processing. ACL*, 53-60, 2006.
- [3] T. Joachims, "A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization". *In Proceedings of the 14th International Conference on Machine Learning*,
- [4] C.M.X. Benjamin, W.L. Woon, K.S.D. Wong, "Customized Term Weighting Scheme for Document Classification". *Proceedings of the International Conference on Computer and Communication Engineering 2008, May 13-15, 2008 Kuala Lumpur, Malaysia*,
- [5] Hayes P., Weinstein S, "A system for content-based indexing of a database of news stories", *In second Annual conference on Innovative Applications of Artificial Intelligence*, 1990.
- [6] Yue-Heng Sun, Pi-Lian He, Zhi-Gang Chen, "An Improved Term Weighting Scheme For Vector Space Model" *Proceedings of the Third International Conference on Machine Learning & Cybernetics, Shanghai, 26-29 August 2004*.
- [7] S. Brin and L. Page "The anatomy of a large-scale hypertextual Web search engine" *Computer Networks and ISDN system*, 1998
- [8] Salton, G. and C. Buckley. "Term-weighting approaches in automatic text retrieval." *Information Processing and Management: an International Journal* 24(5): 513-523.
- [9] Jianyi LIU, Jinghua WANG, "Keyword Extraction Using Language Network." *In Natural Language Processing and Knowledge Engineering*, 2007.

# Vision System for Human-Robot Interface

Md. Ezharul Islam, Nasima Begum and Md. Al-Amin Bhuiyan

Jahangirnagar University, Savar, Dhaka, 1342, Bangladesh

**Abstract** — A vision system has been developed for human-robot symbiosis. The two novel contributions of this paper are 1) Detection of human faces and 2) Localizing the eye. The method is based on visual attributes of human skin colors and geometrical analysis of face skeleton. This paper introduces a spatial domain filtering method named ‘Fuzzily skewed filter’ which incorporates Fuzzy rules for deciding the gray level of pixels in the image in their neighborhoods and takes advantages of both the median and averaging filters. The effectiveness of the method has been justified over human-robot interaction system.

**Keywords** — Human-Machine Interface, Face Detection, Skin Color Segmentation, Eye Localization, AIBO.

## I. INTRODUCTION

When robots are working cooperatively with human beings, it is necessary share and exchange their intentions and interests. Since the communication between humans and robots is established by means of visual perceptions, eye tracking and gaze directions, are important clues for this interaction. Development of an intelligent vision system is, therefore, a vital issue for robotics research.

A fair amount of techniques have been introduced recently literature on face detection, eye localization, eye tracking and gaze detection, for example, geometric modeling [1], auto-correlation [2], neural networks [3], principal component analysis [4] and so on. Model based approaches assume that the initial location of the face is known. Color based approaches reduce the search space in face detection algorithm. The neural network-based approaches require a large number of face and non-face training examples, and are designed primarily to locate frontal faces in grayscale images. Sung and Poggio [5] have developed an example-based approach for locating frontal views of human face in complex scenes. Since this method has been developed for vertical frontal view faces, faces with other orientations cannot be detected. Rowley et al. [3] have developed a neural network-based frontal face detection system where a retinal connected neural network has been employed to justify the small windows of size 20×20 of an image whether it contains a face or not. Lam and Yan [6] have used snake model for detecting face boundary. Although the snake provides good results in boundary detection, but the main problem is to find the initial position. Moghaddam and Pentland

[7] employed principal component analysis for describing the face pattern with lower-dimensional feature space.

This paper explores a face detection and eye localizing system which integrates the detection of human faces in different lighting conditions and localization of eyes on it, as shown in Fig. 1.

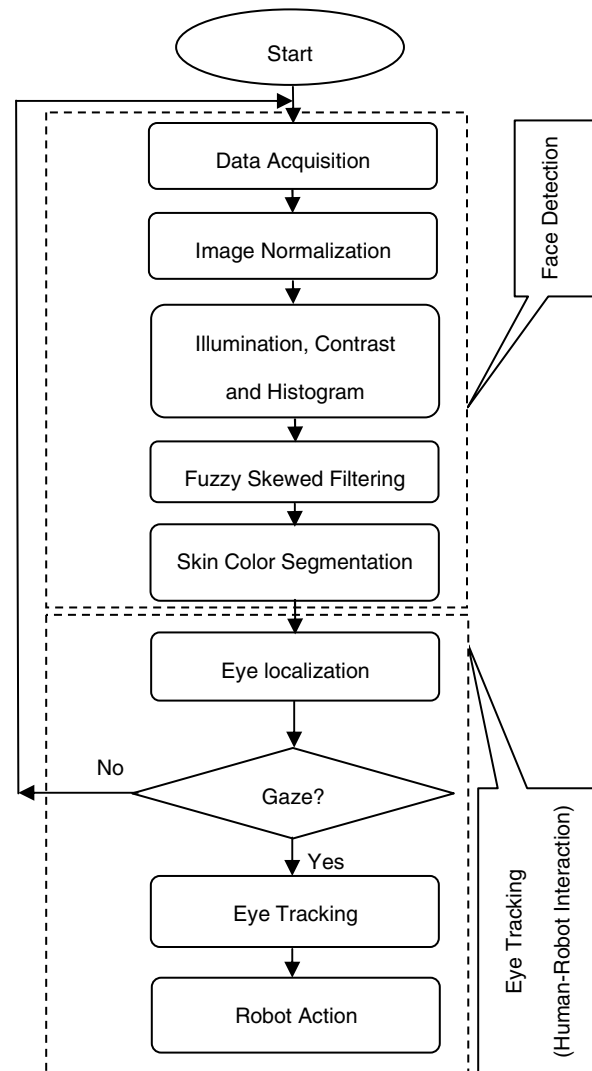


Fig. 1. Architectural design of the human-robot interaction system.

Face detection is established by a two step method: (i) skin color segmentation and (ii) morphological operations. The system is based on the determination of the approximate location of the face area and thresholding the image by color segmentation. The exact location of the face is then determined by employing morphological operations. Then eye is then localized from the knowledge of the face geometry. Experimental results indicate that the system is capable of detecting face and locating the eye position from complex backgrounds with a high degree of variability in lighting conditions, pose and illumination.

The remainder of this paper is organized as follows. Section 2 describes the vision system for human-robot interface. Face detection and eye localization methodologies are also illustrated. Section 3 presents experimental results and performance. Section 4 draws the overall conclusions of this paper.

## II. FACE DETECTION

Face detection is concerned with determining which part of an image contains face. Face detection is achieved by employing by skin color segmentation. The main steps employed for the face detection process is discussed in detail in the following sections.

### A.1 Preprocessing

The original image is obviously a color image. It is first converted into gray scale image and being normalized to same scale, orientation and position. Each normalized image is then subjected to some image processing operations to account for different lighting conditions and contrast.

### A.2 Histogram Equalization

The face images may be of poor contrast because of the limitations of the lighting conditions. So histogram equalization is used to compensate for the lighting conditions and to improve the contrast of the image [8]. Let the histogram  $h(r_i) = p_i / n$  of a digital face image consists of the color bins in the range  $[0, C-1]$ , where  $r_i$  is the  $i$ -th color bin,  $p_i$  is the number of pixels in the image with that color bin and  $n$  is the total number of pixels in the image. For any  $r$  in the interval  $[0,1]$ , the cumulative sum of the bins provides with some scaling constant. Histogram equalization is performed by transforming the function  $s = T(r)$  which produces the mapping with the allowed range of pixel values, i.e., a level  $s$  for every pixel value  $r$  in the original image [9] and  $0 \leq T(r) \leq 1$  for  $0 \leq r \leq 1$

### A.3 Fuzzily Skewed Filter

Various sources of noise may exist in the input image. The fine details of the image represent high frequencies

which mix up with those of noise. So low-pass filters are used to obliterate some details in the image. This paper introduces a new filter named “fuzzily skewed” filter to suppress the noise. This method incorporates the advantages of both the median filter and averaging filter. For  $k=0$ , it acts like median filter and for  $k/2 \leq n/2$ , it acts like neighborhood averaging filter. This method can successfully reduce noise that results to sharp transitions in the gray values, and it does not make the resultant image as blur as the neighborhood averaging filter.

In the fuzzily skewed filter, fuzzy rules are applied for deciding the gray level of a pixel in the image from the neighborhood of that pixel. This is a variation of the Median filter and Neighborhood Averaging filter. The decision process includes the following steps:

- 1) At first the gray values of the neighborhood pixels ( $n \times n$  neighborhood) are stored and then sorted in ascending or descending order.
- 2) Fuzzy membership value is assigned for each neighbor pixels. This step has the following characteristics:
  - (i) A  $\Pi$ -shaped membership function is used.
  - (ii) The highest and lowest gray values get the membership value 0.
  - (iii) Membership value 1 is assigned to the mean value of the gray levels of the neighborhood.
- 3) Consider only  $2 \times k + 1$  pixels ( $k/2 \leq n/2$ ) in the sorted pixels, and they are, the median gray value and  $k$  previous and forward gray values in sorted list.
- 4) Select the gray value that has the highest membership value and place as output.

### A.4 Skin Color Segmentation

Skin color segmentation is based on visual information of the human skin colors from the image sequences in YIQ color space. For this, the dominant and perceptually relevant skin colors extracted in RGB space from images, are converted into the YIQ space. Images are being searched in YIQ space depending on the amount of color content of these dominant colors, that is, whether the skin color value is substantially present in an image or not.

In the YIQ color model a color is described by three attributes: luminance, hue and saturation. The YIQ produces a linear transform of RGB which generates Y representing luminance channel and I, Q representing two chrominance channels to carry color information. The transformation matrix for the conversion from RGB to YIQ is given by [10]:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.320 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

where R,G, B are the red, green, and blue component values which exist in the range [0,255].

Since the human skin colors are clustered in color space and differ from person to person and of races, so in order to detect the hand parts in an image, the skin pixels are thresholded empirically [10]. In this experiment, the threshold value is chosen by the following equation:

$$(60 < Y < 200) \text{ and } (20 < I < 50) \quad (2)$$

The detection of face region boundaries by such a YIQ segmentation process is illustrated in Fig. 2.

The exact location of the face is then determined from the image with largest connected region of skin-colored pixels. The connected components are being determined by applying a region-growing algorithm at a coarse resolution of the segmented image. In this experiment, 8-pixel neighborhood connectivity is employed. In order to remove the false regions from the isolated blocks, smaller connected regions are assigned by the values of the background pixels. After thresholding the image may be encountered by some holes in the face skin region. In order to remove the false regions, the face region is subjected to morphological dilation operation with a 3×3 structuring element several times followed by the same number of erosion operations using the same structuring element. The dilation operation is used to fill the holes and the erosion operations are performed on the dilation results to restore the shape of the face [10]. Then the largest connected component was calculated to find out the actual face area and eliminate other face-like regions.

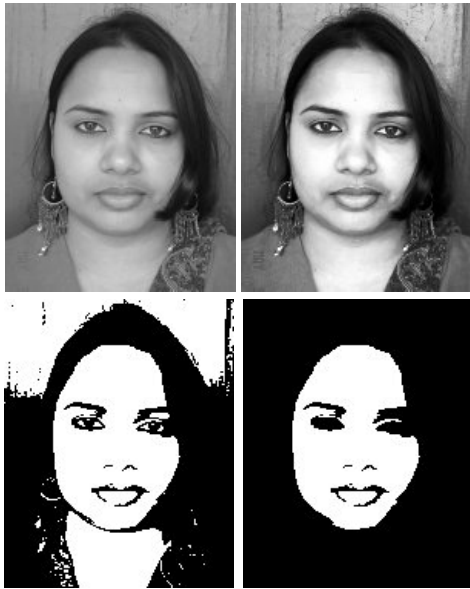


Fig. 2. Skin color segmentation. (Clockwise from top-left) (a) Original image (b) Enhanced image (c) Skin color segmented (d) Face area.

## B. Eye Localization

Once the face is found in an image, the eyes are then searched around the restricted area inside the face. The face image block is then divided into four parts and the left and right eyes are then searched in the upper left and upper right sub-blocks of the facial image, respectively. Here assumed a frontal view of the face and the original face images are subjected to some image processing operations as Histogram equalization and gray level slicing.

After these image processing operations and detecting the face, eyes are then extracted from this face area, assigning a unique tag to each isolated candidate block by labeling the binarized image. These tag points are the representatives of each image feature block.

Let  $P_j = (x_i, y_i), (j \in [1,2])$  be the two tag points of the image blocks of the left eye and right eye, respectively. Since in the binary image, all of the pixels have got the intensity value of 0 and 255 only, these tag points are assigned with their respective representative numbers. The first tag point, i.e., the tag point for the left eye, is determined by searching for the first white pixel (intensity value of 255) from bottom to up and right to left in the top left quadrant of the image window and assigning it with the value of 128. The algorithm for this searching method is as follows:

[Step 1] Compute the center point of the face:

[Step 2] Starting from the center point, search the first black pixel and assign it with the value of '128' (first tag value) as follows:

```
For $i \rightarrow x_{\max} / 2$ to 1 step -1
 For $j \rightarrow y_{\max} / 2$ to 1 step -1
 If $\text{Pixel}[i][j] = 0$ then $P_i = 128$
 return
```

[Step 3] Search the entire window for the tag value '128' and assign all connected pixels with '128'.

Similarly, the above operation is done for the 2nd tag point. These tag points and representative numbers are starting positions of two eyes.

## C. Calculating Gravity Center of Eye

Since the exact location of the pupil is very confusing, so instead of searching the pupil point, we are determining the position of the gravity center of the eyes from the black portions of the irises of the segmented image. If the gray level at each point  $(x, y)$  of the given area E of an eye is considered as the "mass" of  $(x, y)$ , we can define the center of gravity of that component, as well as the moment of inertia about specified points or lines. The pq-th moment of area E about the origin (0,0) is given by:

$$m_{pq} = \sum \sum_{(x,y) \in E} x^p y^q \quad (3)$$

where  $(x, y)$  are the coordinates of the pixel included in area E. The 0-th moment  $m_{00}$  represents the area E and the center of gravity of E is the point  $(\bar{x}, \bar{y})$  whose coordinates are given by [10]:

$$\bar{x} = \frac{m_{10}}{m_{00}} = \frac{\sum_{(x,y) \in E} xy^0}{\sum_{(x,y) \in E} x^0 y^0} \quad (4)$$

and

$$\bar{y} = \frac{m_{01}}{m_{00}} = \frac{\sum_{(x,y) \in E} x^0 y}{\sum_{(x,y) \in E} x^0 y^0}. \quad (5)$$

On detection and localization of eyes, the gaze positions are estimated. Robots are then instructed depending on the movement of the eyes.

### III. EXPERIMENTAL RESULTS AND PERFORMANCE

Face images were analyzed to demonstrate the feasibility of the proposed method. The effectiveness and robustness of this approach is justified using different images with various kinds of lighting conditions.

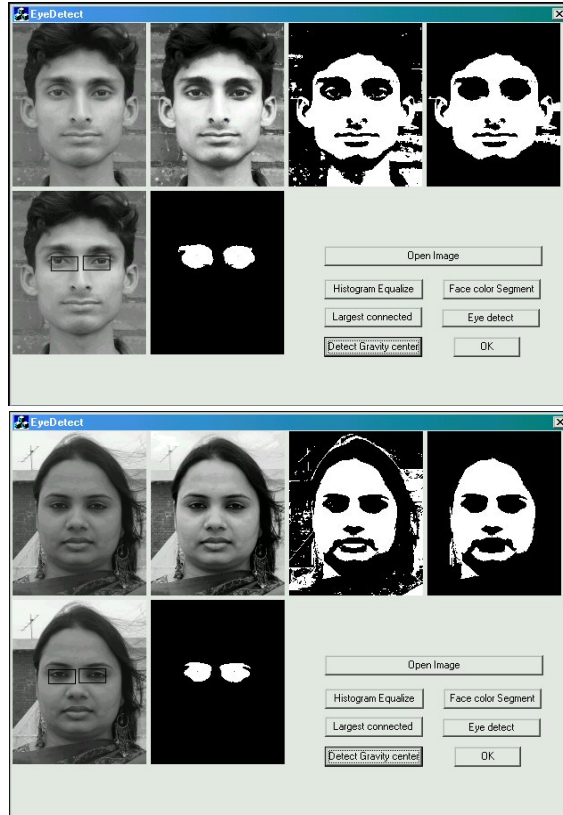


Fig. 3. Eye detection at different environments. Upper image at dark environment and lower image at cloudy environment.

Experiments were carried out on a Pentium IV 2.2MHz PC with 256 MB RAM, and the SONY VISCA camera. The algorithm has been implemented using Visual C++. When a complex image is subjected in the input, the face detection result highlights the eyes of the image, as shown in Fig. 3. Experiments were carried out to justify the system with different input images at different environments and lighting conditions. The main application of this method is, here, face detection and eye locating is done successfully with images of different persons with various kinds of lighting conditions and environment both in shiny and gloomy weather. Most of the images were taken using a digital camera, but some are from scanner, and some from video tapes recorded from different television channels. The algorithm is capable of detecting single face in an image. For multiple faces, the system finds the dominant face only, i.e. the face containing more number of skin colored pixels.

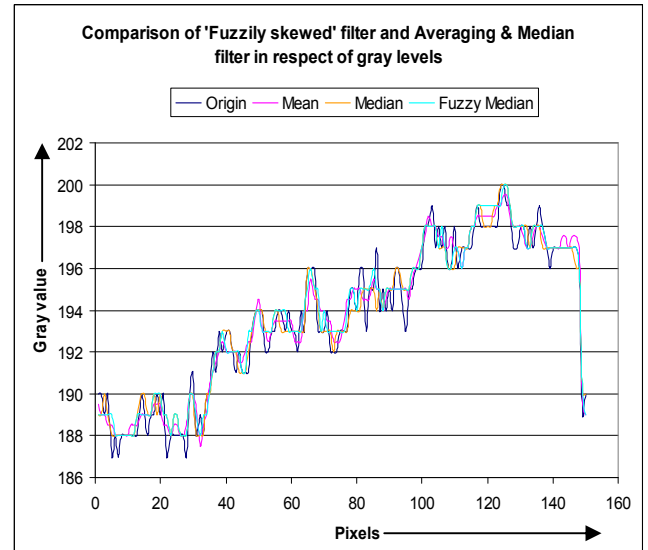


Fig. 4. Comparison between 'Fuzzily skewed' filter and Averaging and Median Filter

TABLE I  
EYE TRACKING COMMANDS TO CONTROL AIBO

| Eye Positions | Actions        |
|---------------|----------------|
| Left Top      | Turn Light Off |
| Center Top    | Move Forward   |
| Right Top     | Turn Right     |
| Left Middle   | Turn Left      |
| Center Middle | Start walking  |
| Right Middle  | Dance          |
| Left Down     | Turn Light On  |
| Center Down   | Move Backward  |
| Right Down    | Sing a song    |
| Eye close     | Stop walking   |

First, the response of the proposed fuzzily skewed filter has been analyzed. A graphical comparison between the 'Fuzzily skewed' filter and averaging and Median filter is shown in Fig. 4. The graphical analysis imply that the fuzzily skewed filter snatches the advantages of both the median filter and averaging filter. This method can successfully reduce noise that results to sharp transitions in the gray values, again, it does not make the resultant image as blur as the neighborhood averaging filter.

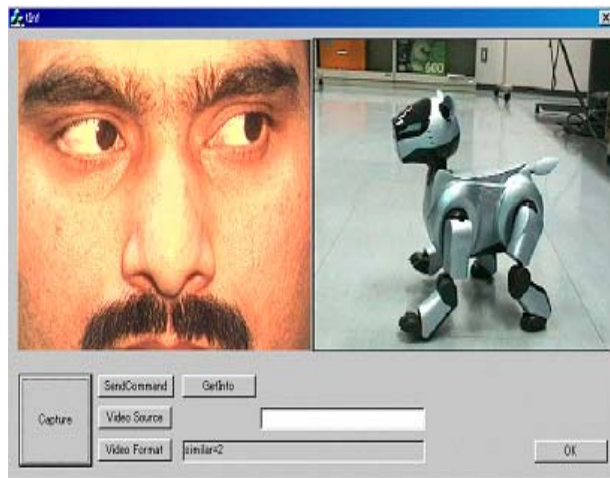


Fig. 5. Eye tracking command to motivate AIBO to start walking.

The system has been implemented by eye tracking program running in server PC and it considered AIBO as a client. The communication link has been established through TCP-IP protocol. The human-robot interaction results are: (i) turn light off, (ii) move forward, (iii) turn right, and so on, according to the eye positions at left top, center top, and right top, respectively. Ten eye tracking commands were executed to justify the robustness of the system and these commands and their corresponding actions are furnished in Table I. Finally, a real time human-robot interaction system has been implemented on eye tracking commands. One of the actions of the robot as a result of eye movement is shown in Fig.5.

## V. CONCLUSIONS

This paper presents a fast face detection and eye localization method depending on color segmentation. Detection of faces and facial features using machine vision techniques has many useful applications. Though human beings accomplish these tasks countless times a day, they are still very challenging for machine vision. In this paper, Face detection is established by skin color segmentation and morphological operations. Here, eye localization is done by calculating largest connected area from the skin color segmented image and then

calculating the tag positions of the two eyes. Finally the gravity centers of eyes are also calculated. Our next approach is to extend the algorithm for multi-face detection and overlapping faces in images and images with obstacles in front of eye (e.g. spectacles) with different orientation and pose and image with more complex backgrounds.

A fuzzy rule based spatial domain filter is also proposed in this project. Experimental result justifies that it incorporates the advantages of median and averaging filter with minimizing their limitations. Our main target is to instruct operations to robots and make them understand the human intensions and interests over facial expressions and gaze detection so that they are capable of grasping with more intelligence while working cooperatively with human beings.

## REFERENCES

- [1] H. Shinn and H. Hui-Ling, "Facial modeling from an uncalibrated face image using a coarse-to-fine genetic algorithm," *Pattern Recognition*, Vol. 34, No. 8, pp. 1015-1031, 2001.
- [2] F. Goudail, E. Lange, T. Iwamoto, K. Kazuo and N. Otsu, "Face recognition system using loacal autocorrelations and multiscale integration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 10, pp. 1024-1028, 1996.
- [3] H. Rowley, B. Shumeet and T. Kanade, "Neural network-based face detection," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, pp. 23-37, 1998.
- [4] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
- [5] K. Sung, and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, pp. 39-50, 1998.
- [6] K. Lam and Y. Hong, "Locating and extracting the eye in human face images," *Pattern Recognition*, Vol. 29, No. 5, pp. 771-779, 1996.
- [7] B. Moghaddam and A. Pentland, "Face recognition using View-Based Modular Eigenspaces," *Proc. of Automatic Systems for the Identification and Inspection of Humans, SPIE* Vol. 2277, 1994.
- [8] Y. Tae-Woong and O. Il-Seok, "A fast algorithm for tracking human faces based on chromatic histograms," *Pattern Recognition Letters*, Vol. 20, No. 10, pp. 967-978, 1999.
- [9] R.C. Gonzalez and R.E. Woods, *Digital image processing*, 2nd Edition, London, Prentice Hall, Inc., 2002.
- [10] M.A. Bhuiyan, V. Ampornaramveth, S. Muto and H. Ueno, "On Tracking of Eye for Human-Robot Interface," *International Journal of Robotics and Automation*, Vol. 19, No. 1, pp. 42-54, 2004.



# Rice Disease Identification using Pattern Recognition Techniques

Santanu Phadikar and Jaya Sil

Dept. of CSE, West Bengal University of Technology, Kolkata 700064, India, [sphadikar@yahoo.com](mailto:sphadikar@yahoo.com)

Dept. of CST, Bengal Engineering and Science University, Howrah-711103, India,  
[jayasil@hotmail.com](mailto:jayasil@hotmail.com)

**Abstract** ---- The techniques of machine vision are extensively applied to agricultural science, and it has great perspective especially in the plant protection field, which ultimately leads to crops management. The paper describes a software prototype system for rice disease detection based on the infected images of various rice plants. Images of the infected rice plants are captured by digital camera and processed using image growing, image segmentation techniques to detect infected parts of the plants. Then the infected part of the leaf has been used for the classification purpose using neural network. The methods evolved in this system are both image processing and soft computing technique applied on number of diseased rice plants.

**Index Terms** -- Fractional zooming, leaf blast (*Magnaporthe grisea*), brown spot (*Cochiobolus Miyabeanus*), rice diseases detection, SOM,

## I. INTRODUCTION

Research in agriculture is aimed towards increase of productivity and food quality at reduced expenditure and with increased profit, which has received importance in recent time. One of the vital components of crop management is accurate diagnosis and timely solution of the field problem. Diagnosis is a most difficult task to perform manually as it is a function of a number of parameters such as environment, nutrient, organism etc. With the recent advancement in image processing and pattern recognition techniques, it is possible to develop an autonomous system for disease classification of crops.

Use of various pesticides for controlling the diseases actually damages the rice field. Most common rice diseases are [1],[2],[12] *leaf blast* [5],[7] (*Magnaporthe grisea*), brown spot (*Cochiobolus Miyabeanus*), leaf blight, sheath rot, steam rot, kernel smart, etc. and most common pests includes yellow stem borer, rice leaf folder, rice bug. The visual symptoms of the damages caused by the pests vary with the severity of the diseases. We have restricted our work within the rice diseases only and considered the most two common diseases in the North East India, namely Leaf Blast (*Magnaporthe grisea*) and Brown Spot (*Cochiobolus Miyabeanus*).

Unprecedented growth in research in the area of image processing and computer vision has been observed during the last two decades, yet no remarkable application of the

techniques found in the field of agriculture. However, the International Rice Research Institute (IRRI) has developed a software named Rice Doctor [2] that suggest name of the diseases by taking the manually observed information about the infected plant. In a recent project a prototype system of automatic identification of cotton insect pest and intelligent decision based on machine vision has been developed [3] while remote sensing technique is being used experimentally for farm pest management [6]. Another effort found in literature that tries to determine the mineral deficiency based on the color texture analysis of rice leaf [4]. The proposed work is to identification of the pest and/or disease depending on the damage symptoms obtained from the infected parts of the plants using image processing and pattern analysis methods. Thus, the system will able to timely diagnosis the field problem and the suggestion helps the farmer to take the appropriate measures to increase the quality of crops and reduce the development cost, which save the environment too.

The paper has been divided into five sections. Section II describes the feature extraction methods that include region segmentation, boundary detection and spot detection of images. Section III narrates the proposed zooming algorithm used to interpolate points in the detected spots, which are subsequently applied for classification of diseases along with the results illustrated in section IV. Conclusions are summarized in section V.

## II. FEATURE EXTRACTION

The samples of the infected rice leaves have been collected from different parts of East Midnapur, a major rice producing district of South Bengal, India.

The images have been grabbed using Nikon COOLPIX P4 digital camera in macro mode. The samples of the rice leaves affected by Blast and Brown spots are shown in the Fig.1.

### A. Segmentation

Acquired images are enhanced by increasing brightness and contrast and then transformed to Hue Intensity Saturation (HIS) model for segmentation. Entropy based bi-level thresholding method [10] has been invoked for



segmenting the images to facilitate identifying the infected parts of the leaf. The segmented images are presented in Fig.2.

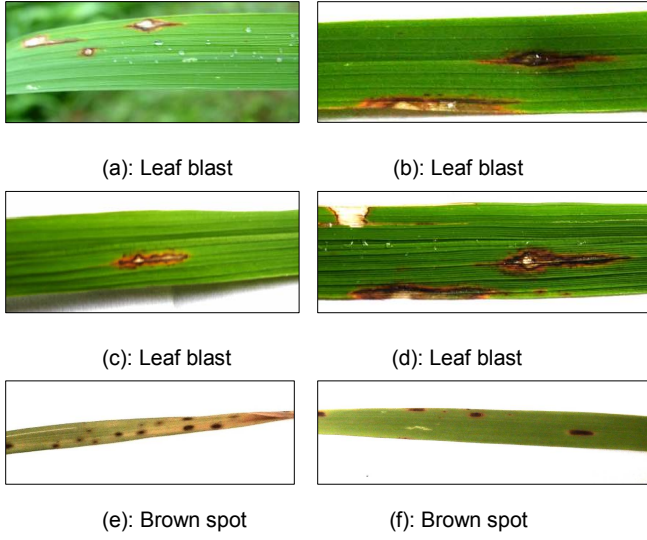


Fig. 1. Original images of infected rice leaves

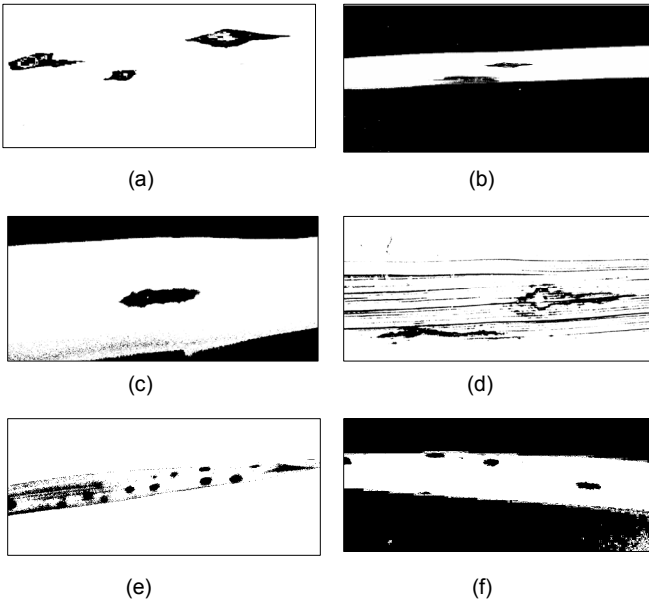


Fig. 2. Segmented Images Corresponding to Fig. 1

### B. Boundary Detection

Boundary detection algorithm using 8-connectivity method [9] has been applied on the segmented images of figure 2(a) and (b), and outputs are shown in Fig.3.

### C. Spot Detection

Boundaries of the segmented images are obtained with varying image sizes 300×300 to 1500×1500. The ratio of the

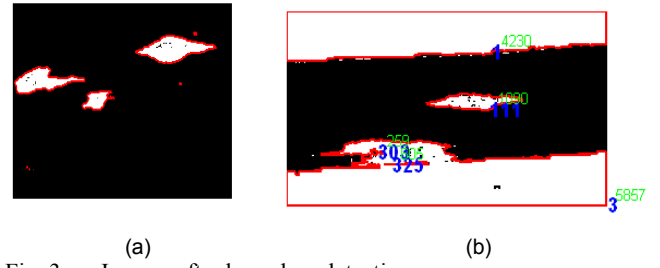


Fig. 3. Images after boundary detection

average value of the green components and intensity of the pixel lies in between 1.15 to 1.40 respectively and the measurement is used to identify the infected spots. The reason of selecting the green component is based on the observation that when the infection occurs in the leaf, the greenness of the leaf is affected most. Instead of considering only the green value we have taken the ratio of green component with the intensity to make the spot detection invariant of the brightness of the images and the age of the leave. The values have been determined heuristically based on the experiments of the hundred of infected leave, shown in Fig. 4.

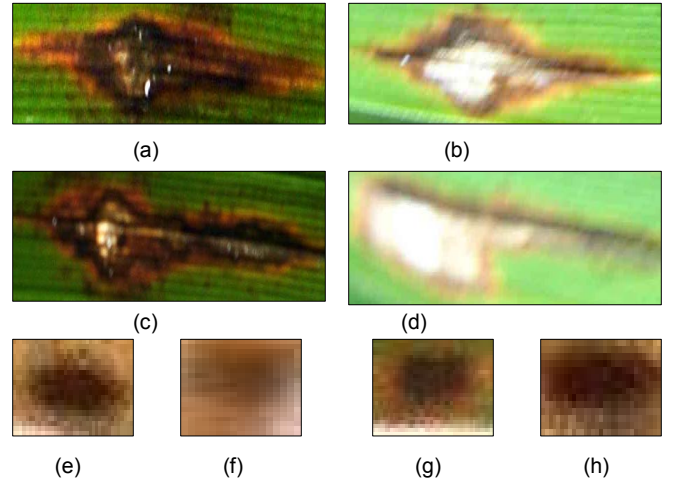


Fig.4. Label (a) to (d) represent the spots of blast and label (e) to (h) represent the brown spots.

## III. ZOOMING ALGORITHMS

Self organizing map (SOM) neural network has been employed for classification [8], [11] of the images where gray value of the pixels of spot images are fed as input. Segmented spots have been normalized to the 100×80, which is determined by the average size of the spots, detected in the previous section to apply it in SOM. There are two possible methods to make the input vector of same size. First one is the padding of zeros in the input vector and the second one is the interpolation of missing points to make the input vector of same size. Since there is wide difference

amongst the spot sizes, the missing points are often more than the original points, which may cause a significant effect in the classification process, if the former method is opted. Thus, interpolation method has been applied for fractional zooming to normalize the spots size.

The interpolation algorithm works as follow:

Step 1. Determine the scaling factor  $s_x=100/m$  and  $s_y=80/n$  where  $m \times n$  is size of the original image.

Step 2.

- (i) All points of the original image located at  $(i,j)$  position are placed in the  $((i*s_x - s_x/2), (j*s_y - s_y/2))$  position as shown by yellow points, with gray values  $a, b, c$  and  $d$  ( figure 5).
- (ii) Scan first row wise, then column wise and calculate the gray value of intermediate points, like  $x$  (shown in figure 5) by using equation  $(a*d_1+b*d_2)/(d_1+d_2)$  where  $d_1$  and  $d_2$  is the distance from  $a$  to  $x$  and  $x$  to  $b$  respectively.
- (iii) Scan row wise and any points with gray value  $y$ , which not computed yet (figure 5) is calculated as the average value of its four neighbor. Note that, only the points whose gray values are already computed in the previous steps, used to compute the gray value of  $y$ .

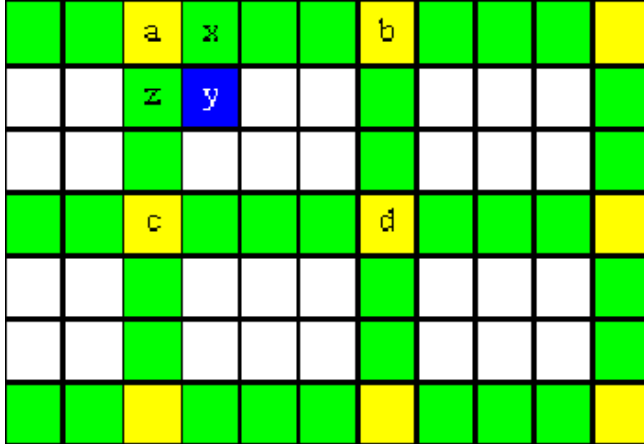


Fig. 5. Illustration of zooming algorithm

The results obtained after zooming, shown in Fig. 6.

#### IV. CLASSIFICATION USING SOM

After obtaining spots of uniform size, unsupervised learning technique SOM is used. Here we have used  $100 \times 80$  input nodes and 2 output nodes, 300 training patterns and 50 epochs to train the network. Here the gray values of the

pixels are applied as the input vector. After performing training with the above mention datasets and 50 epochs, we

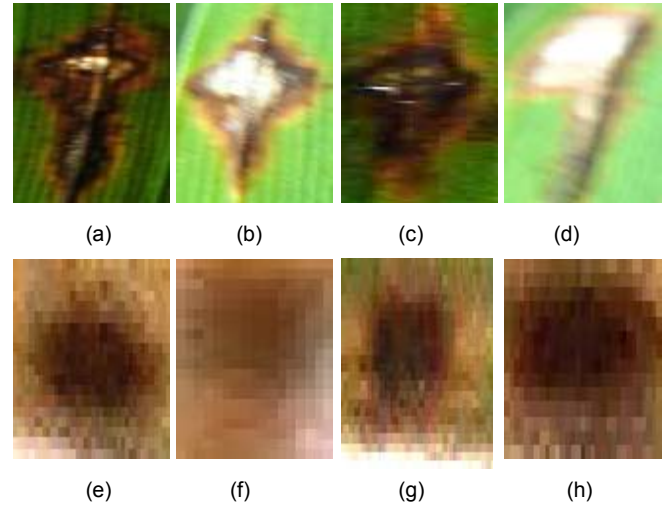


Fig. 6 Zoomed images corresponding to Fig.4

have tested the system with 300 input datasets for each of the cases separately, given below :

- Case 1: Use the RGB of the spots for classification.
- Case 2: Use the Fourier transform of the spot.
- Case 3: Use the Arbitrary rotation of the 50% spot.
- Case 4: Use the Fourier transform of the 50% rotating spots.

The result has been shown below obtained after applying.

TABLE I  
CLASSIFICATION RESULTS

| Category | Percentage of successful classification |
|----------|-----------------------------------------|
| Case1    | 92%                                     |
| Case2    | 84%                                     |
| Case3    | 82%                                     |
| Case4    | 70%                                     |

#### V. CONCLUSION

In the paper, the diseased rice images are classified using SOM neural network where train images are obtained by extracting features of the infected parts of the leave while four different types of images are applied for testing purposes. The zooming algorithm extracts features of the images using simple computationally efficient technique, which results satisfactory classification for test images. It has been observed that the transformation of image in the frequency domain does not yield a better classification compare to the original image.

## REFERENCES

- [1] Srivastava K.P., A Text Book of Applied Entomology, New Delhi: Kalyani Publishers, 1993
- [2] Rice Doctor, International Rice Research Institute, Philipines, [www.iri.org](http://www.iri.org), 2003.
- [3] Li Zhigang, Fu Zetian, Shi Yan, Xia Tiehua, "Prototype system of automatic identification of cotton insect pest and intelligent decision based on machine vision," American Society of Agricultural and Biological Engineers, 2003
- [4] P. Sanyal, U. Bhattacharya and S. K. Bandyopadhyay. "Color Texture Analysis of Rice Leaves Diagnosing Deficiency in the Balance of Mineral Levels Towards Improvement of Crop Productivity," 2007 IEEE 10<sup>th</sup> International Conference on Information Technology, pp. 85-90, 2007
- [5] R.K. Webster, Rice blast disease identification guide, Davis: University of California, 2000.
- [6] Z. Qin, M. Zhang, T. Christensen, W. Li, and H. Tang, "Remote sensing analysis of rice disease tresses for farm pest management using wide-band airborne data", pp.2215-2217, 2003.
- [7] T. Kobayashi, E. kanda, K. Kitada, K. Ishigurd and Y. "Detection of rice panicle blast with multisprectural radiometer and the potential use of multisprectural scanner.' Torigoe-2000
- [8] Laaksonen, J. Koskela, M. Oja, E. , "Self-organizing maps for content-based image retrieval," International Joint Conference on Neural Networks, vol.4, pp.2470-2473 ,1999
- [9] R. C. Gonzalez, R. E. Woods, Digital Image Processing, Delhi: Pearson Education, 2007
- [10] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram", Computer Vision Graphics Image Process. Vol. 29 , pp.273-285, 1985
- [11] H. Ritter and T. Kohonen. "Self-organizing semantic maps", *Biological Cybernetics*, vol.61, pp.241-254, 1989.
- [12] S. H. OU, Rice Diseases,England: Commonwealth Mycological Institute,1972.

# Geospatial Web Services in Environmental Planning

Iftikhar U. Sikder

Department of Computer & Information Science, Cleveland State University,  
2121 Euclid Ave. BU 336, Cleveland OH 44115, USA

**Abstract-** As more and more mapping resources are becoming available, the Web is becoming a confluence of collaborative interactions, negotiations and planning platforms. The integration of Web services with Geographic Information Systems (GIS) has triggered a new wave of interest for using the Web to compose dynamic mapping services for environmental planning. Web services could offer a framework for service integration and composition for collaborative environmental modeling. By wrapping geo-processing services of interoperable standards of XML and SOAP, we demonstrate a prototype implementation of environmental decision support systems GEO- ELCA (Exploratory Land Use Change Analysis) where Web service-enabled middleware adds core functionality to a Web mapping service. The system demonstrates how geo-processing services can be integrated with environmental simulation models using OGC (Open GIS Consortium) compliant connectors that support WMS (Web Mapping Service) and WPS (Web Processing Services).

**Index Terms** — Geographic Information Systems, Web Services, Environmental Planning, Interoperability

## I. INTRODUCTION

In the face of the growing trend of organizations moving towards a flatter structure, environmental resource managers are facing an increasing need to involve multiple stakeholders in spatial decision-making to bridge the gap resulting from differential access to information and resources. The traditional production of geospatial data and models has been relatively centralized [1]. In recent years, there is a growing interest in making access to geospatial information and services available to decision makers and planners to promote Collaborative Spatial Decision Making (CSDM)[2]. CSDM and public GIS involve a "bottom-up" planning model reflecting the stakeholders' perspective to explore scientifically projected planning scenarios. The vision requires a broad understanding of organizational settings against the backdrop of the distributed information architecture for a closer involvement of multiple stakeholders from different geographic locations and social orientations. It presupposes systems that provide stakeholders with the

ability to continuously make plans [3], have it pre-empted by collaborative actions, and evolve based on any new contexts. This service-oriented approach has triggered a new wave of enthusiasm in the composition of complex services in a meaningful way involving not only traditional alphanumeric data but complex geographic data and services [4].

One of the functional requirements of collaborative modeling is the development of reusable generic geospatial procedures, which can be made to communicate each other in distributed systems in a heterogeneous environment. For instance, the user may wish to view relationships among several geographically separate datasets and perform varying degrees of analysis and invoke different geo-processing services (e.g., perform spatial proximity analysis combined with geocoding procedures and specialized cartographic visualization in a given area). From the planning point of view, comparing individual preferences with the collective scenario is a group process that essentially requires geospatial service composition in a collaborative framework and a distributed environment. Hence, the service-oriented architecture (SOA) becomes inevitable for collaborative GIS systems that would then define the use of services to support the requirements. By having such services distributed all over the Internet and accessible in a uniform standard manner, it is possible to envision the integration of several spatial services (chaining of services) to provide higher levels of functionality to existing services[5, 6]. Equally important is the ability to locate and obtain someone else's data with minimal human intervention in the process of system interoperability. In this respect, the OpenGIS Consortium's endorsed standards and specifications for geographic data encoding GML/Geography Markup Language), geographic data publishing (WFS-Web Feature Service), and map publishing (WMS/Web Mapping Service) and Web Processing Service (WPS) are being received in the industry. These standards are becoming key elements of collaborative planning in urban information systems.

While Web services require minimal standards-based protocols such as HTTP or SMTP for use in other applications, they essentially offer communication among systems without requiring centralized planning or control mechanisms to integrate a heterogeneous mix of platforms [7]. This unique advantage of Web services can be well utilized in developing decentralized distributed planning systems as envisioned in collaborative spatial decision making in the NSF-funded Project Varenus [2]. Current researchers lack a theoretical framework for understanding how features of Web services can be related to implementation standards and protocols of distributed planning systems; they also lack a theoretical framework for understanding how such features can be exposed in geospatial service composition for collaborative planning.

## II. MULTI-USER CONTEXT IN SPATIAL PLANNING

In view of decision support frameworks, collaborative modeling presupposes multiple parties with different perspectives working together in a complex emergent environment. These parties (henceforth termed as “agent”, “decision maker”, or “stakeholder” interchangeably) must have integrated data access from heterogeneous sources to integrate with transparent high performance computing resources to compose decision models dynamically. However, in real life situations, it is often difficult to achieve the stakeholders’ views or effective patterns of social interactions in the planning process. For example, decisions on how current land use should be changed depend on legal, environmental, and regulatory constraints, as well as on biases and preferences of different groups or institutions. In particular, understanding urban land use dynamics involves consideration of the complex behavior of individual decision makers and the interaction of local and regional institutions in multiple scales. Moreover, such decisions are inherently spatial in nature because the change in a particular parcel may have direct or indirect consequences to the neighboring parcels.

For example, the EPA’s Brownfields development program involves the expansion, and redevelopment of urban areas that may be complicated by the potential presence of a hazardous substance, pollutant, or contaminant [8]. While cleaning up and reinvesting in these properties takes development pressures off of undeveloped, open land and improves the environment, evaluating a candidate property to determine if it meets the criteria for redevelopment is inherently a collective decision process. For instance, an individual landowner may act from his or her individual interest; however, in

the long run, the overall land use scenario may be undesirable to everyone. Moreover, changes in land use may create concern for environmental impact in the surrounding region. While land-use changes are often identified as a major driving force of ecological changes, with the conversion of land use from one category to another, there is an overall change of hydrological characteristics resulting from the changes of impervious areas. Consequently, the increase in volume and peak flow means a possible increase in the concentration of pollutants that could potentially cause the environment to deteriorate. Hence, a centralized planning process is essentially inadequate to reflect group dynamics. Such group-individual dilemmas make them ideally suited for collaborative planning in a distributed environment, particularly regarding the integration of Web-services in accessing geospatial data and models for environmental monitoring.

## III. WEB SERVICES AND MAPPING STANDARDS

As a precursor to Web services, CORBA and Microsoft’s DCOM or Java-based RMI were often used for distributed access and query of spatial data, sometimes integrated with SDSS. However, being “tightly coupled” with native data structure, the broker-based services are unable to make sure that whenever an organization makes changes in its native data structure or services, the corresponding change is automatically reflected in all other organizations sharing the same resource. Moreover, middleware-based access through a broker relies on a standard definition of “interfaces”. Geo-processing services can become very cumbersome in the absence of such interface. From the decision support point of view, having data access at the client’s end without robust geo-processing capabilities amounts to having little help. In a broker-based solution, the client has to pull a massive amount of data at his/her end and manage it locally. Such approaches assume the client’s explicit ability to manipulate the server connection and remote objects. Thus, frequent spatial processes, such as spatial join between data from two different servers, need to be coordinated at the client’s end. Such object-level manipulations of spatial process often fail to provide high-level views to application developers. Paradoxically, in a spatial decision support system, the user’s or decision maker’s view on spatial features or geometry needs to be realized at a higher level of abstraction while at the same time maintaining the transparency of system processes.

In Web services, a series of protocols such as eXtensible Markup Language (XML); Simple Object Access Protocol (SOAP); Web Service Description Language (WSDL); and Universal Description, Discovery, and Integration (UDDI) provides standards

for communication and collaborative processing capacity among Web-service compliant architecture. A key advantage is that various GIS layers can be dynamically queried and integrated while still maintaining independence in a distributed environment. From an organizational point of view, this feature may be very appealing; for local governments, such as counties and other organizations, can still independently collect and manage data locally and still integrate information and services using Web services. A client, for example, like a transportation company, could directly access a local government's base map without maintaining its own. At the same time, the client can update the local government's data from its own record. As far as data interoperability is concerned, extended collaboration and partnerships using Web Services could provide opportunities to open interfaces and communication mechanisms for distributed computing.

Open GIS Consortium (OGC) has been in the forefront of developing a set of interoperability standards to guide the development of web mapping programs. All service specifications of OGC are collectively known as OGC Web Services (OWS) and adhere to the same communication protocol of HTTP. OWSs are subject to somewhat similar operational interfaces. The interfaces have common operations such as GetCapabilities, which provide service level information; Get operation allows users to access data in a given context and Describe operation provides a description of data or attributes. OGC's Web Map Server (WMS) implementation information specifies that a mapping service should be able to at least produce a map (as a picture, series of graphical elements, feature data) and answers basic queries and finally inform other programs about its capabilities. The three operations defined for a Web Map Service are as follows:

GetCapabilities – obtains machine-readable service metadata

GetMap - returns a map upon receiving request

GetFeatureInfo – optional operation that provides additional feature information (e.g., more information about features in the pictures of maps that were returned by previous map requests).

The Web Processing Service (WPS) interface specifies WPS operations that can be requested by a client and performed by a WPS server. The GetCapabilities and DescribeProcess of WPS are implemented through WSDL which is an XML format for describing Web Services.

### III. WEB SERVICE FRAMEWORK FOR DISTRIBUTED COLLABORATIVE MODELING

In a distributed planning environment, the server side component can expose its functionality through web services, which could provide the benefits of interoperability and open standards. Figure 1 illustrates a stranded-based framework for integrating collaborative models using Web services. In such a system, UDDI will allow any client to discover spatial services (e.g., using the UDDI nodes as FGDC compliant geospatial metadata servers for registered Web services). Typical geoprocessing services may include data management tools like projection and transformation, or topology manipulation, indexing and spatial join, etc.

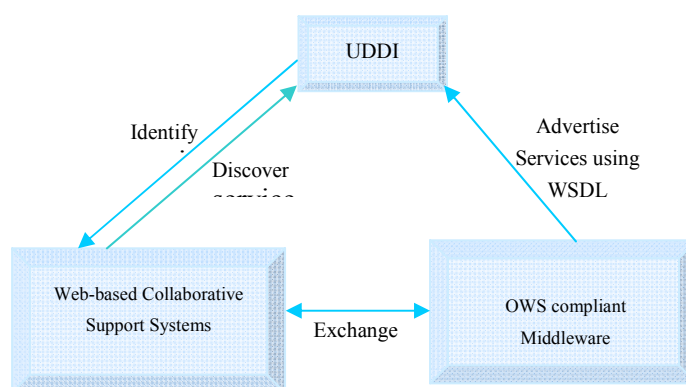


Figure 1. Standard-based Web-service integration framework for collaborative support

Analytical services may include spatial proximity analysis, spatial overlay, feature extraction, geo-statistical analysis services, and so forth. Each geoprocessing service can advertise its XML protocol or SOAP API using WSDL for discovery. To integrate these services with a local application, the Web application developer needs to wrap an application with a SOAP API to invoke or serve a request. Integrating services can be realized by developing OGC compliant middleware that supports access to the Web Mapping Service (WMS) and Web Processing Service (WPS) or Web Feature Services (WFS). Since Web services build stateless or loosely coupled environments, nodes can be dynamically connected whenever needed, thus multi-user access to data and services are basically “just in time”.

### IV. SERVICE ORIENTED ARCHITECTURE: GEO-ELCA

What follows is an illustration of a prototype implementation of a Web service-based spatial decision support system for collaborative planning of urban land use change. GEO-ELCA (Exploratory Land Use Change Analysis) offers a collaborative platform for multiple



stakeholders with exploratory tools to evaluate the environmental consequences of land use changes. GEO-ELCA supports the evaluation of the pollutant flow as a consequence of land use changes. The system allows an exploratory model to evaluate the hydrological scenarios in response to land use changes as a result of stakeholders' different options. The Web interface allows a user to select graphically a land use type and change it to a different category and then visualize the effect of different pollutants. The system allows the dynamic selection of a feature type (i.e., polygon – land use class) interactively so that a user can change attribute items and identify a feature property. A user can initiate a change in land use type by graphically selecting a polygon. The server side application processes the request and makes necessary updates in the database to reflect the corresponding changes of the pollutant coefficients. Every request to change in land use type results in the recalculation of the mass export of pollutants and corresponding statistics. The processed result is sent back to the Web server and then to the client side. The consequences of user decisions initiate the simulation model to estimate the yearly pollution load. The system integrates a simulation model commonly used in urban hydrology--the so-called "Simple Method"[9] for estimating exports of various pollutants' runoff from different land uses. The resulting pollution map can be visualized with multiple theme overlay. The system logs individual user's input preferences into mediating algorithms to resolve conflict among user preferences of land use choice.

(which is typically a light viewer such as a browser); the Web mapping server sends WMS Request (GetMap, GetCapabilities, GetFeatures) and WPS Execute Operation (changeLandUsageType) messages which are XML encoded and enclosed in SOAP 1.1 messages. The construction of SOAP messages is accomplished using client side scripting such as JavaScript, VBScript and can also makes use of advanced features such as AJAX. These messages are received by the web server which exposes a web service interface and has already published its service end point interface in a WSDL file. The SOAP message is decoded and passed on to the middleware solution. The middleware is responsible for decoding the request messages which were in OGC Compliant message formats into a format understood by the web mapping server (see Figure 2). In our case, since the web mapping server is ArcIMS, we need to communicate with it using ArcXML. Hence the middleware solution provides a mapping from the OGC WMS Request message to appropriate ArcXML requests. The ArcXML equivalent of OGC Web Mapping Service Operations is shown in Table 1.

Table 1. OGC's WMS operations and their mapping with ArcXML request

| OGC WMS Operation | ArcXML Request Operation |
|-------------------|--------------------------|
| GetCapabilities   | GET_SERVICE_INFO         |
| GetMap            | GET_IMAGE                |
| GetFeatures       | GET_FEATURES             |

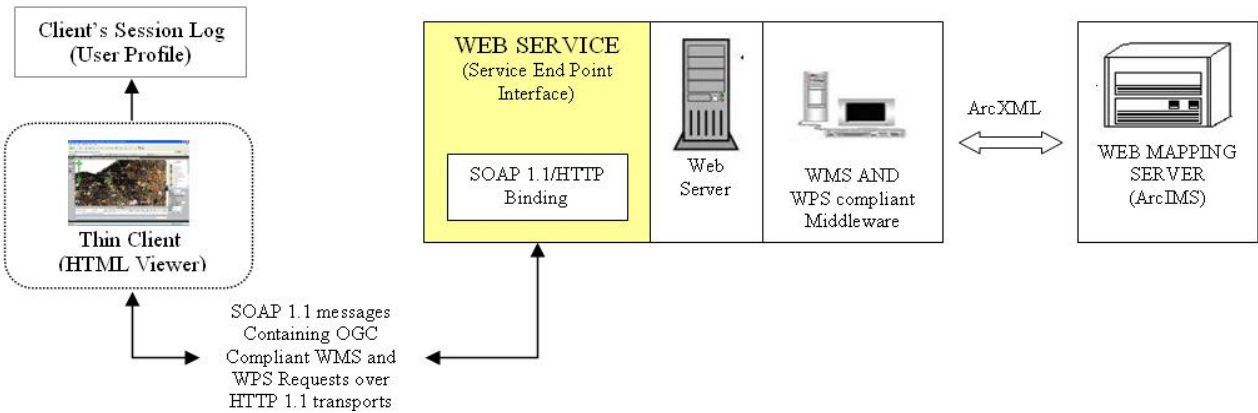


Figure 2. Service Integration Framework in GEO-ELCA

GEO-ELCA uses ArcXML (native XML encoding of spatial object) to communicate between the custom middleware and the Web mapping server of the client

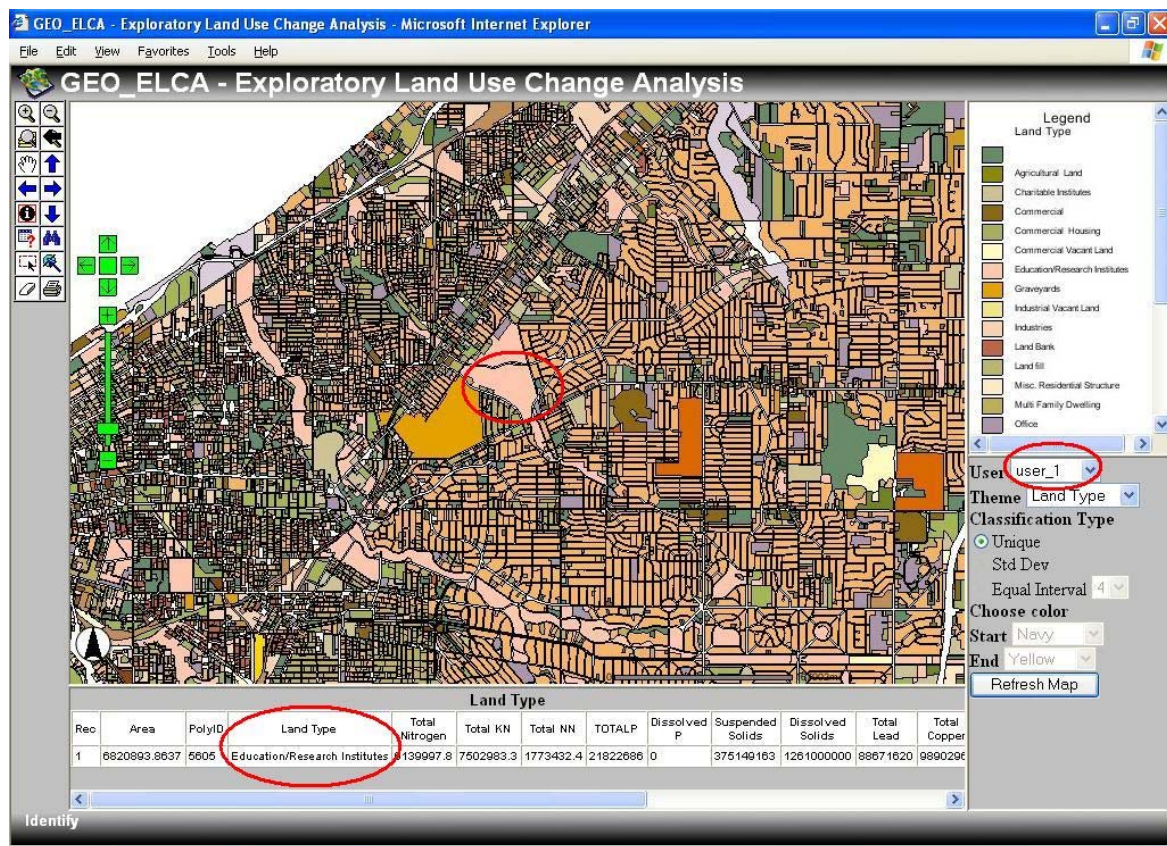


Figure 3. User specific planning scenario of land use change analysis

The response received from the web mapping server is also in ArcXML and needs to be converted into a format compliant with OGC WMS Response message format. The middleware solution performs this mapping. Hence the request issued to the web map server is a GET\_IMAGE request (as always in ArcXML). The web map server renders the map based on the specific user's state of data and this Map Image is returned all the way back to the user. Hence the client visualizes user specific maps. Once again the middleware solution is responsible for ensuring that the response from the web map server (URL of Map Image encoded within an ArcXML response) is returned back to the client in an OGC GetMap compliant response. The mapping performed for WPS messages is similar to that performed for WMS messages with the tasks performed being the same.

## V. THE COLLABORATIVE CONTEXT OF GEO-ELCA

The "virtual" private user workspace provided by GEO-ELCA offers a mechanism to log an individual user session. An individual user session is achieved by means of a relational data model that provides each user a 'virtual private' workspace. Each individual user is

associated with a native profile database, while, in fact, the actual visualization is rendered from a single dataset. Thus, user specific visualization and data editing is possible in real time. For example, when a user's request to change a land use type is received by middleware, it updates the specific user's data. Then it forwards the request for a map to the web map server after converting the request to ArcXML-- a format understandable by the web map server. The web map server queries the data store to retrieve the data it needs to generate the map. The records returned consist of the current user's data and the map. Then a map image is generated using the updated (current user's) data and the map image is returned to the middleware solution, while the master data is restored to original values by updating records from backup data store. Finally, the middleware returns the map image (or URL) of the newly generated user specific map.

Thus, user specific planning profiles can be achieved and communicated in real time. As illustrated in Figure 3., the Web-based planning interface offers user-specific individualized explorative options where a user graphically selects a polygon and chooses a land use category from the pop up list to change its category and simulate the "what if" kind of pollution scenario. The simulation model estimates the pollution potential (e.g.,



total nitrogen, BOD load, etc.) resulting from a user's decision to change land use and graphically display the estimated pollutant. In a multi-user context, each user's planning scenario and corresponding simulation result are achieved in a relational table. As a result, cartographic visualization and representation can be locally customized by user. For example, the same area can be visualized by different user with different options. A user has the option to visualize the confidence limit and standard deviation classification of selected pollutant categories and choose the color scheme or color ramp for classification and hence dynamic map legend rendering. A color ramp is calculated on the fly and a corresponding visualization of the map and the accompanying legend is generated. At present, the system does not offer any group negotiation tool that allows the convergence of a consensus decision. However, such utilities can be easily implemented in the given architecture by integrating multi-objective optimization services (e.g., a discrete decision choice model like the Analytic Hierarchic Process or evolutionary algorithm for group decision making).

## VI. CONCLUSION

Since Web services avoid problems of tightly coupled distributed object techniques and expose an application programming interface over the Web, they hold promise for distributed resource sharing and collaborative environmental planning. By wrapping and dynamically integrating remote geo-processing services from multiple sources, one can develop an emergent collaborative system using interoperable standards of XML and SOAP. We introduced some architecture for Web service-based collaborative modeling applications for environmental planning. The prototype system is based on the conversion of the OGC compatible GIS services to Web Services for flexible collaboration. The significant advantage of the architecture is that it provides a means for self-describing services in which end users through any WMS/WPS client can invoke a set of services (e.g., catalog services of spatial map). Consequently, user specific plan profiling and customized cartographic visualization are rendered dynamically. The consequence of a user's decision is simulated graphically to estimate the yearly pollution load as a result of individual decision. Each user's planning profile can finally be compared for further modification. The collaborative model suggested here is geared towards archiving different opinions in a manner suited to visualization and negotiation.

## REFERENCES

- [1] J. Pickles, "Arguments, Debates, and Dialogues: The GIS-Social Theory Debate and Concerns for Alternatives," in *Geographical Information Systems: Principles, Techniques, Management, and Applications*, P. Longley, M. Goodchild, D. Maguire, and D. Rhind, Eds. New York: John Wiley, 1999.
- [2] M. Goodchild, M. Egenhofer, K. Kemp, D. Mark, and E. Sheppard, "Introduction to the Varenus project," *International Journal of GIS*, vol. 13, pp. 731-745, 1999.
- [3] L. Hopkins, "Structure of a planning support system for urban development," *Environment and Planning B: Planning & Design*, vol. 26, pp. 333-343, 1999.
- [4] I. Sikder and A. Gangopadhyay, "Design and Implementation of a Web-Based Collaborative Spatial Decision Support System: Organizational and Managerial Implications," *Information Resources Management Journal*, vol. 15, pp. 33-47, 2002.
- [5] Z. R. Peng and M. H. Tsou, *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks*: John Wiley, 2003.
- [6] S. Balram and S. Dragic'evic, "Modeling Collaborative GIS Processes Using Soft Systems Theory, UML and Object Oriented Design," *Transactions in GIS*, vol. 10, pp. 199-218, 2006.
- [7] F. Curbera, W. Nagy, and S. Weerawarana, *Web Services: Why and How*: IBM T.J. Watson Research Center, 2001.
- [8] EPA, "Brownfields Cleanup and Redevelopment." vol. 2006, 2006.
- [9] T. Schueler, "Microbes and Urban Watersheds," *Watershed Protection Techniques*, vol. 3, pp. 551-596, 1999.

# Implementation of E-Governance: Only Way to Build a Corruption-Free Bangladesh

S. A. Ahsan Rajon and Sk. Ali Zaman

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh.  
TMRAFM, Khulna-9100, Bangladesh.

**Abstract** — This paper focuses on various aspects of implementing e-governance in developing countries like Bangladesh. In this paper, we present a comparative analysis of present government architecture and the prospects and problems of implementing e-governance in Bangladesh emphasizing on the usage and effectiveness of e-governance to eradicate corruption from various sectors of governance. We especially present the adaptability of e-governance in the prime sectors of government and provide a methodical study on the strategies of involving mass people in the governance process improving information and service delivery with their participation in overall decision-making. The potential to ensure highest level of transparency in all the sectors of government with the implementation of e-governance is also presented in this paper. Moreover, we provide specific recommendations for implementing the e-governance in the most effective and efficient manner.

**Index Terms** — E-Governance, Decision-making, Adaptability, ICT Infrastructure, E-Commerce, E-learning, M-Governance, E-Transparency, E-Administration.

## I. INTRODUCTION

Present century is the age of Information and Communication Technology. The prime concern of this decade is to establish and ensure a better way of management, communication and development with the use of information and information oriented services. Crossing the boundary of personal computations and communications, the uses of digital media for greater levels of management has been initialized early in the century. E-Governance may simply be considered as an extension of this trade. With the enhanced facilities of technology for providing smart and easy living, E-Governance has become a prime demand in the world for ensuring transparency, effectiveness, accountability and efficiency throughout the various sectors of government including decision-making, public policy building, access to information etc.

The paper is organized into eight sections. The following section provides the definition of e-Governance. A study on current status of Bangladesh in the point of view of implementation of e-governance is presented in section three. Section IV presents the necessity of e-Governance and m-Governance in developing countries.

The effectiveness of e-Governance in fighting and eradicating corruption is depicted in section V. Section VI is devoted to provide a framework for implementing e-governance in various leading sectors of the government of Bangladesh. A brief recommendation with concluding remarks is provided in section eight.

## II. E-GOVERNANCE

A number of definitions of e-Governance exist emphasizing on specific activities and functionalities. E-government and e-democracy are the mostly adapted criteria for defining e-Governance. The term e-democracy refers to the processes and structures that encompass all forms of electronic interactions between the government and the citizens in terms of information coverage and decision-making. E-Government is a form of e-business in governance and refers to the processes and structures needed to deliver electronic services to the public (citizens and businesses), collaborate with business partners and to conduct electronic transactions within an organizational entity. Taking the above two into concern, e-Governance is defined as the application of electronic means to improve the interactions between government and citizens and government and business as well as to employ electronic means in internal government operations to simplify and improve democratic government and business aspects of government with the primary goal of boosting administrative efficiency [1]. Overall sectors of e-Governance is basically pointed as G2G (Government to Government), G2B (Government to Business) and G2C (Government to citizens) in terms of service with prime importance in G2C for establishing e-democracy [2][3].

In a very short span, we may define e-Governance as the process of adapting electronic means in possible sectors and stages of government to ensure legitimate mass-access of administrative and service oriented information with the potential to establish transparency and accountability of government activities and maximum service by redesigning and redistributing the administrative and operating system of government. Adaptability, Awareness-building and Accessibility are the most important concern in establishing e-governance.

Here adaptability and accessibility involves the socio-economic status of the citizens especially the economical strength and educational status of the people. Literatures in [1], [3], [4], [5] provide explanations regarding e-governance and related issues with specific sector-based implementation approach.

E-Governance doesn't only mean to have a website in the internet rather to provide online services and information that increase mass-participation in decision making, accountability of the authority, transparency of affairs involving public interests and domestic concerns. Though a good milestone towards establishing e-government has been made in the USA, Europe and other developed westernized countries such as Australia, Singapore, Canada, the developing countries too, are putting steps for implementing e-Governance in various sectors with a number of limitations, challenges and debates.

The prime objective of e-Governance is essentially to ensure good governance and strengthening democratic process providing equal access to information, freedom of constructive expression, greater equity, service efficiency, stable rapid economic growth and social inclusion. The successful integration of e-Governance in the developed countries has already illustrated the demonstrable and tangible evolutionary impact on increasing government functionality, efficiency, improving citizen participation and quality of life as a result of multi-stakeholders partnership. [4]

### III. CURRENT ICT INFRASTRUCTURE OF BANGLADESH

Bangladesh, a small developing country with enormous prospects realizes the importance of implementing e-governance in Bangladesh. However, because of its financial backwardness and the ill-fate of being affected by natural disasters frequently, it is still now in the initializing step of implementing e-Governance.

TABLE I. ICT INFRASTRUCTURE OF BANGLADESH

|                              |                   |
|------------------------------|-------------------|
| Population (Million)         | 150.4 (July 2007) |
| Literacy rate                | 51.6%             |
| Poverty                      | 50%               |
| Tele-density                 | 23.3              |
| Growth of Cellular Phone     | 135%              |
| Household access to computer | 1.36              |
| Household with electricity   | 44.23             |

From the above table it is evident that, the implementation of e-Governance in Bangladesh is not so an easy task, rather it would be a time consuming step for Bangladesh. In Bangladesh ICT has been declared as the thrust sector. In e-Governance introducing step, a number of ministries, which are the highest governing authority,

have already presented their websites [4]. To provide quicker service, Bangladesh government has already made a number of forms available through internet, which is really an appreciable step towards implementing e-Governance [4].

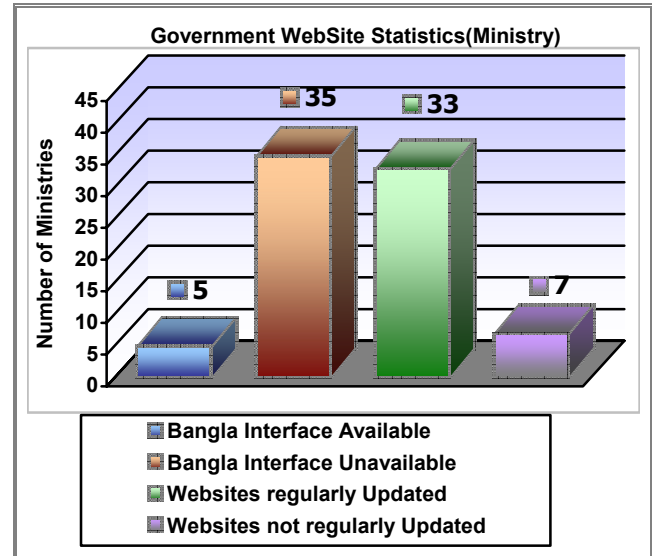


Fig.1. Statistics of the websites of various government organizations in Bangladesh.

### IV. POTENTIAL OF M-GOVERNANCE TO IMPLEMENT E-GOVERNANCE IN BANGLADESH: THE WAY FORWARD

Cellular Phone has revolutionized the telecommunication infrastructure in Bangladesh. It is really amazing that, since 1989; especially within the last decade, mobile phone service has provided about forty million people with telecommunication services resulting a tele-density of 23.3% [4]. The following figure indicates the growth of mobile communication.

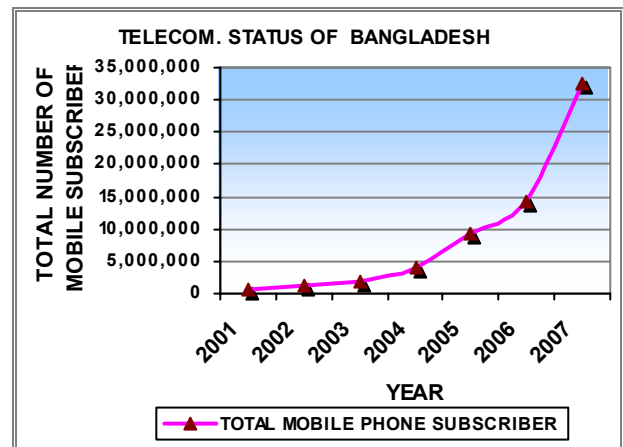


Fig. 2: Telecommunication Status of Bangladesh.

As mobile communication has got a massive use, it may be an effective way to utilize mobile phone oriented services in awareness building for establishing e-governance. In decision making processes and providing essential administrative information we may use cellular phones both in voice and text communication services. Various service-oriented information may also be communicated to all levels of people through mobile communication. The means of seeking services to the government by the citizens may also be ensured through mobile phones. Various add-on services like internet facility using mobile phone may mobilize the process of implementing e-Governance in Bangladesh. It is necessary to mention here that, from the perspective of Bangladesh m-Governance may be considered as the prime media or gateway of implementing e-Governance.

#### V. E-GOVERNANCE: SCOPE AND POTENTIAL TO FIGHT CORRUPTION

The total world, especially the developing countries are in a great challenge to fight corruption that has got rise because of improper administrative transparency, lack of accountability and absence of monitoring processes. The fundamental rights of the citizens are also been suppressed due to manual paper-based administrative system. In this section, a field-specific analysis is provided indicating the scope of implementing e-Governance with its applications.

##### *A. E-Governance in Central Government Process*

The central government, which works as the basis of all prime activities must provide legitimate information on public issues through electronic communications. Major decisions may be taken by conducting online-surveys as well as traditional surveys. This availability of information is a great aspect in establishing accountability and transparency. Publishing accurate fiscal statistics may be the easiest way of restricting corruption in financial sectors, though a number of doubts exist in this concern. Sadly speaking, the constitution and even the basic documents stating public laws are still now not available to the citizens of Bangladesh. Without providing electronic means, it is not at all possible to make these documents available.

##### *B. E-Governance in Public Administration*

Public Administration possesses the power of administering, managing and implementing the decisions-taken by supreme authority. Often this management process is biased and due to corruption, the implementation of development activities derived from public demand are not completely and accurately

performed. Even allegations from citizens regarding the corruption issues do not reach to proper authority. That is, due to the intermediate authority, public concerns are often disregarded. Implementing e-covernance and e-communication may at least echo the public voices to the supreme authority for considerations.

##### *C. E-Governance in Education Sector*

Education, which is the backbone of a nation, is often fatally pointed as a corrupted sector. As, education is the fundamental need, a major portion of budget is spent in development of this sector. But, the history sadly says that, a major portion of this budget often goes into the pocket of some corrupted personnel. Unnecessary and ill-motivated delay in granting various scholarships of the students, meaninglessly harassing the teachers in granting remunerations and other bills are some common objections of the concerned parties. The absence of monitoring in application and implementation is one of the reasons behind these, which may be partially eradicated by providing the information regarding distribution of funds and other operational and statistical data. The debates that often arise in modifying the curriculum may also be solved by integrating the facility of providing public remarks through electronic means before taking final steps of curriculum specification. Publishing the results of public examinations has established a well suited means of automatic verification and hence ensured greater transparency.

##### *D. E-Governance in Judiciary*

Solely administration or solely Judiciary may never ensure a sound running of government. It is must to have proper co-ordination and respect between the parts. It is note-worthy that the independent judiciary is taking initiatives to redesign judiciary with electronic aspects. Still now, there is a lot of scope of ensuring transparency in judiciary by making each step clear to the citizens, which is of no doubt, echoes e-governance.

##### *E. E-Governance in Economy*

The Economy of Bangladesh has faced a lot of difficulties because of corruption. The lack of efforts and weak policy of maintenance with absence of accountability has wasted the major amounts of annual budget. Denying public interest, often political decisions has been implemented, that has simply resulted in weakened economical structure. But, an effective feasibility study maintaining overall robust system analysis and design would demonstrate acceptable and fruitful sectors of economical interest.

#### *F. E-Governance in Agriculture*

Being an agricultural country, Bangladesh should emphasize on enhanced and productive agriculture. Because of the alarming livestock crisis, the whole world is now facing a great challenge regarding the arrangement of food. Price hike is the hottest concern all over the world. The middlemen-ship is also a resultant of poor-governance. In this situation, if we fail to implement enhanced mechanisms of achieving maximal production, the overall existence may be extremely threatened. This enhanced mechanism includes the tasks of making the most fruitful technology available to the users, making the basic operational information available to the agriculturists in the shortest possible time. This dissemination of essential information including market demand-supply analysis is only effectively and efficiently possible through implementation of e-governance. E-agriculture focuses on agriculture informatics, agricultural development and entrepreneurship and provides the agri-services through internet and related technologies. Moreover, if various subsidies on agriculture are provided through internet making all the steps monitor able for public, transparency may be achieved at the highest level saying a goodbye to corruption. This achievement may also bring a horizon in e-marketing or e-commerce attracting the investors of both home and abroad which may enlarge the scope of the growth of economy. Effective study regarding the positive impact of implementing e-agriculture is provided by FAO in [6]. It is now essential for developing countries like ours to present their agricultural products in the world market effectively, which is only possible by adapting e-governance and e-commerce.

#### *G. E-Governance in Bidding*

The dreadful hostile occurrences regarding bidding often costs of valuable lives, which is about a regular headline in the pages of newspapers. The stretch of corruption in analysis and selection process is well known to all. But e-governance may facilitate both the parties by transforming the overall process into an electronic system. If from the invitation of tenders to selecting the best offer, e-governance is employed, the interested parties may get rid from the threats of miscreants in the tender collection and submission process as well as the selection and rejection process may be free and fair.

#### *H. E-Governance in Other Sectors*

Besides the above mentioned sectors, we have a lot of opportunity in establishing transparency and accountability as well a mobilizing the development process by adapting e-governance. For example, providing

a dynamic internet application stating aspects of lands like information regarding proprietary, statistics of the amount of land, map or related statistics describing the perfect and accurate condition of land area may prevent all types of corruptions in this sector. Public safety may also be established in this way. In the election process, which is basically the most important affair for founding a good-governance and sound-democracy, e-governance may facilitate a lot by providing accurate and un-tampered information regarding the candidates. If information regarding criminals, smugglers, loan-infringers, tax-hoaxers is made available through the e-governance supporting processes of law-and order enforcing organizations and other concerned authorities, at least it would be possible to present the real and internal characteristics of the candidates before the citizens to help them in correct and accurate decision making. Thus Citizen Information System *i.e.* e-governance may change the appearance of the whole country giving a new dimension in democracy.

The above-mentioned sectors are presented as the prefect-sectors facilitating overall governance. It is necessary to mention here that, only installation of computers or having a web-site doesn't mean e-governance. It may only be ensured when, the legitimate tasks and affairs is made available through internet and people are accepted to access information regarding those with proper facility of providing constructive remarks is integrated ensuring participation of citizens in decision making and policy building activities. There is no way left of getting relief from the hassle in getting citizen-services from government without implementing e-governance. Specifically, it is of no doubt that, a machine may never provide biased and corrupted effort unless it is instructed by human. Thus, it is a must to have some supreme authority to monitor all the affairs regarding the e-governance initiatives. Information tempering and other issues may be prevented through their monitoring activities. A powerful task force with ICT experts may ensure such activities.

### **VI. PROPOSED APPROACH FOR IMPLEMENTING E-GOVERNANCE IN BANGLADESH**

From the above discussions, it is evident that, there is no feasible and optimal way of fighting the corruption without adapting, implementing and establishing e-governance in Bangladesh. Here we present some specific guidelines for implementing e-governance in Bangladesh.

#### *A. Stage-Setting Steps for Establishing E-Governance*

In order to facilitate e-governance, we must have proper infrastructure. This infrastructure support must be

affordable and effective for the country. In order to establish internet facility, we may employ Broadband over Power Line Communication (BPLC) [7], [8] strategies, as the most available infrastructure is electricity in comparison with tele-density.

In this stage-setting step, we also intend to emphasize on awareness building. E-governance is aimed for good-governance, and this good-governance is aimed for the social, political, economical that is overall uplift of the citizens. This necessity and functionality of e-governance should be communicated to the people so that they may realize and evaluate e-governance. situations.

#### B. Analyzing the sectors for Implementing E-governance

As a developing country, it is not possible for Bangladesh to establish e-governance in all the sectors at a time. It is essential for Bangladesh to implement e-governance hierarchically. Consequently, we are to prioritize the sectors in terms of effectiveness, functionality and liability of providing services. The failure to identify the priority sectors may jeopardize the total scheme of implementation of e-governance. This prioritization process should also consider the interdependence and intercommunication among various service-providing sectors for ensuring a consistent, functional and effective analysis to implement e-governance. This feasibility analysis process should must involve citizens as well as system analysts and government officials of all levels.

#### C. Framework for Implementing E-Governance

One of the main issues for implementing e-governance in developing countries is expense. The huge expense of implementing e-governance often turns the e-governance implementation project into an abandoned one. But, this may be easily handled by involving universities and other organizations in implementation process. The universities and other ICT related organizations may provide major support by providing their experienced IT-experts and concerned human resource. This may also bring a new dimension of research in the organizations. As in Bangladesh, there are more than hundred universities (both public and private) among which about all the universities possess courses and degrees regarding computer science and information technology, there is a great scope to engage large amount of human resource into the design and development of e-governance system for various sectors of government. Selecting renowned universities to perform this job with the supervision of honorable experienced teachers may ensure a successful, effective and efficient design, development and implementation of e-governance system. The universities may also be benefited with the scope of e-governance

analysis and implementation related research support. Basically, this may be the only driving force to install the framework for e-governance in minimum expense.

#### D. Design issues for implementing e-governance

The design and implementation of e-governance should must take the social, economical and educational issues into concern. The adaptability and effectiveness of the e-governance depends mainly on the citizens, as they are the stakeholders of e-governance. It should be taken into consideration that, the first language of Bangladesh is Bengali. So that, it would be most effective to ensure the use of Bengali in implementation of e-governance for conveying information. The use of Unicode in representation of Bengali text is also expected to facilitate uniform and massively accessible platform. The integration of the facility to provide comments and other information through internet may guarantee the participation of citizens in decision-making and various multidimensional service-oriented processes. We propose to use a Phonetic based encoding scheme for obtaining the information from the concerned parties. Figure 5 presents a typical screenshot of proposed phonetic based approach of Bengali text composition.

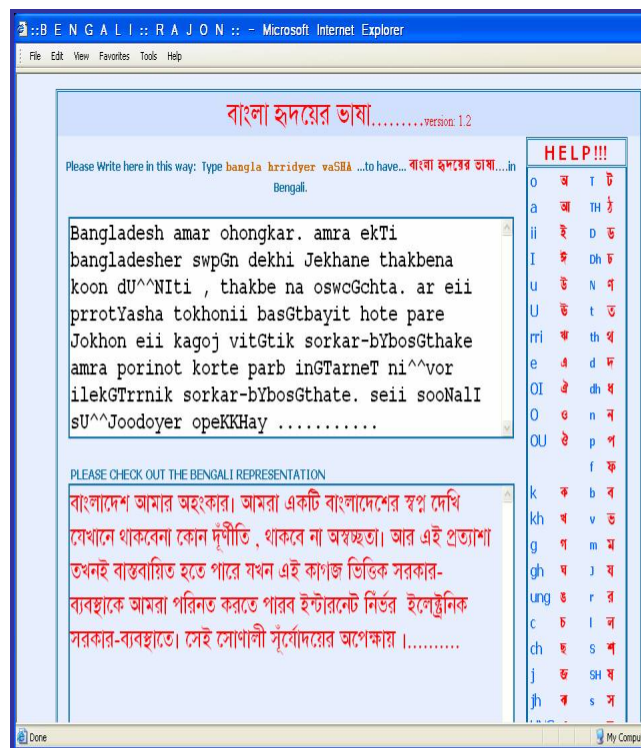


Fig. 3: Keyboard mapping less and Phonetic based approach for online Bengali-Unicode text composition.

E-Governance is a large-scale activity which requires large scale data warehouses to be implemented. For faster, reliable and consistent database applications regarding development of e-governance, distributed database management may be the most effective. It is also necessary to consider WAP-support for all levels of design and development of services relating e-governance.

#### E. Communicating E-governance

E-Governance is dedicated to establish the participation of the citizens in decision-making, policy building, objective selection and quickest service delivery fostering good governance in corruption-free and progress and success-oriented environment. This goal can only be achieved if people are informed about the sectors and services of e-governance. This communication can be made using traditional communication medias. SMS may be an effective approach in communicating the adaptation of e-governance in various sectors. Providing citizen-incentive lucrative facilities and services electronically through e-governance may bring rapid attention of the citizens.

#### VII. E-GOVERNANCE TO M-GOVERNANCE: THE JOURNEY BEGINS

The increase of mobile communication in Bangladesh has made a silent revolution in all the sectors. The spread of the networks has now reached even into the extreme-rural areas. As, people are now more flexible with mobile oriented services, the adoption of m-governance which is to some sense a bridge between e-governance and mobile communication with the use of WAP or SMS [10]. M-Governance is also a great concern because of its extreme growth, availability and flexibility. Considering mobiles as a communication tool for all types of service-request and information possessing from citizens may mobilize the e-governance process.

#### VIII. CONCLUSION

In this paper, we have presented an overview on the sectors for adapting e-governance from the point of view of the democratic governance in Bangladesh. We also describe the specific issues of corruption in various government sectors and have provided the feasibility study for reducing corruption from the corresponding sectors by adapting e-governance. The overall analysis on the usefulness and effectiveness to establish e-Government, e-Administration and e-Democracy in the form of e-Governance has been elaborately provided here.

A detailed approach for implementing e-governance has also been presented in this paper. We also provide specific recommendations for designing, developing and managing the e-governance system in Bangladesh. The implementation of e-Governance is the only way left to uplift the country into a corruption-free developed nation.

#### REFERENCES

- [1] Michiel Backus, "E-Governance and Developing Countries: Introduction and Examples", Research Report, *International Institute for Communication and Development (IICD)*, No. 3, April 2001.
- [2] World Summit on the Information Society, *Plan of Action*, 12 December 2003
- [3] Farooq Sobhan, M. Shafiullah, Zahid Hossain, Mridul Chowdhury, "Study of e-Government in Bangladesh", *Research Report, Bangladesh Enterprise Institute*, Bangladesh. April 2004.
- [4] "e-Governance Horizon Scan Report: An assessment study of e-Governance in Bangladesh", *Access To Information (A2I) Program, Chief Adviser's office, Government of the People's Republic of Bangladesh*, December 2007.
- [5] Iftekhar Ali Chowdhury, Md. Shahidul Islam, and Zahid Hasan Mahmood, "Journey Towards E-Governance: A Bangladesh Perspective", *Proceedings of 7th International Conference on Computer and Information Technology*, 26-28 December 2004, Dhaka, Bangladesh.
- [6] e-Agriculture: A Definition and Profile of its Application, Bridging The Rural Digital Divide, *FAO*, November 2005.
- [7] Dr. Md. Abdul Mottalib, Sheikh Faridul Hasan, Syed Khairuzzaman Tanbeer, "BPLC: Proposed Approach to Implement E-governance in Remote Areas of Developing Countries", *International Academy of CIO (IAC)*, Japan, 2006.
- [8] S. A. Ahsan Rajon, "Broadband Over Power Line Communication : Prospects and Problems in the Context of Bangladesh ", *Proceedings of International Conference on Electronics, Computer and Communication*, Rajshahi, Bangladesh, June 2008.
- [9] Shah Murtaza Rashid Al-Masud, "Does the Digital Divide Still Exist", *Proceedings of 7th International Conference on Computer and Information Technology*, 26-28 December 2004, Dhaka, Bangladesh.
- [10] Melissa Teasdale, *Government in the Palm of Your Hand*, *Proceedings of EURO m-GOV 2005*, Brighton, UK, July 10-12 2005.
- [11] Stephen A. Ronaghan, "Benchmarking E-government: A Global Perspective - Assessing the UN Member States", *Final Report, United Nations Division for Public Economics and Public Administration and American Society for Public Administration*, New York, May, 2002.
- [12] A Y M Ekramul Hoque, Abdur Rashid Sikder, Md. Shamsul Alam, "Statistics for Monitoring Attainment of MDGs in Bangladesh", *Bangladesh Bureau Statistics*, Bangladesh



# File Sharing in Advanced Collaborating Environment: Node Registration & Cache Optimization Techniques

Md. Rezwanur Rahman, Md. Ali Ashik Khan, Tamkin Khan Avi, Mohammad Rezwanul Huq

Department of Computer Science and Information Technology (CIT),  
Islamic University of Technology (IUT),  
Gazipur, Dhaka, Bangladesh.

mrrcit@gmail.com, anondo9@yahoo.com, tamkin\_iut@yahoo.com, md.rezwan.huq@gmail.com

**Abstract** — A major challenge in a collaborative environment is to devise an efficient file sharing and adaptation mechanism. Through this, the users at different nodes can achieve the highest degree of collaboration. Some authors have already proposed file sharing and adaptation framework. As per their proposed framework, users are allowed to share adapted files among them, invoking their file sharing and adaptation service built on the top of advanced collaborating environment (ACE). For file adaptation, a hybrid approach has been mentioned for adapting files which considers user's preferences as well as user's device capabilities. The goal of this adaptation approach is to provide the best possible adaptation scheme and convert the file according to the user's preferences and device capabilities.

In this paper, we propose some new features for file sharing and adaptation framework to have faster and more efficient and meaningful collaboration among users in advanced collaborating environment. We are proposing a Key based registration system as well as an algorithm for Key generation to automate the registration process for slave ACE nodes. We also propose an automated mechanism to distinguish trusted (registered) Slave ACE nodes, which are privileged for p2p file sharing under certain conditions. This approach will completely eliminate the manual intervention in the purposive node selection process. Moreover there would be a network traffic monitoring process which will radically reduce the network overhead to the master ACE node and introduce faster network collaboration in a distributive manner. Consequently, the approach leads towards better collaboration among the users of ACE.

**Index Terms** — P2P File Sharing, Node Registration, Authentication, Collaboration, Cache Optimization.

## I. INTRODUCTION

The notion of advanced collaborating environment is essential to provide interactive communication among a group of users. The advancement in the field of networking and multimedia technology has outmoded the traditional video conferencing concepts. The 3R factor that is- Right People, Right data and Right time, is the major concern of ACE, in order to perform a task, solve a problem, or simply discuss something of common interest [1].

Figure 1 depicts the concept of Advanced Collaborating Environment (ACE), where media, data, applications are shared among participants joining a collaboration session via multi-party networking [2]. Each ACE node has a couple of audio/video/interaction devices,

LCD/projector/tiled display systems, and a number of support machines. Some early prototypes of ACE have been mainly applied to large-scale distributed meetings, seminar or lectures and collaborative work sessions, tutorials, training etc [3], [4].

Access Grid is a group-to-group collaboration environment with an ensemble of resources including multimedia, large-format displays, and interactive conferencing tools. It has very effectively envisioned the implementation of ACE in real life scenario. Venue server and venues concepts come from the Access Grid multi party collaboration system [4], [5]. Venue server is the server component for venues. Node represents the aggregation of hardware and software for participating in a meeting which is same as ACE nodes in ACE environment. Moreover, Venue Server has Data Store for each venue registered to it. Venue Data Store holds all the shared files from different users in the venue.

We would like to define two new terms which will be used for the rest of our paper.

- **Master ACE Node:** It is a kind of ACE node which has the capability to directly communicate to Venue through venue client as well as Venue Server. From the device configuration point of view, it has higher configuration as accessories like HD camera, microphone etc. are attached to this node.

- **Slave ACE Node:** Slave ACE node has less capabilities compared to Master ACE Node in terms of device configuration as well as it cannot communicate to Venue and Venue Server directly. Slave ACE node will be connected to a Master ACE node and all the functionalities needed for users in Slave ACE node are achieved through Master ACE node.

Services built for ACE environment also have been integrated to Access Grid. Here, it may be noted that, we use both the term 'node' and 'device' which are actually contain the same meaning in the later portion of the thesis. The rest of the paper is organized as follows. Section 2 begins the discussion of related work. In Section 3, the Problem Statement is specified clearly. In section 4, our contribution through this paper is evidently described. Section 5 describes our proposed mechanism followed by implementation progress and issues in section 6. Finally, we draw conclusion and talk about some of our future works in section 7.



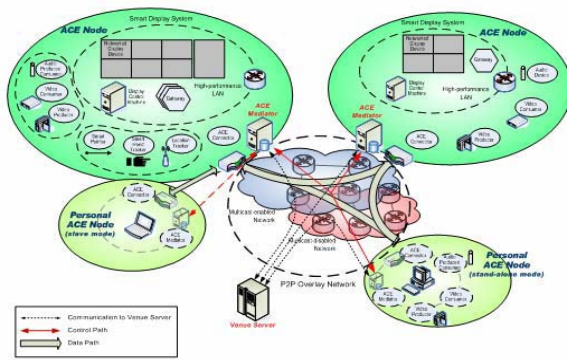


Fig. 1 Advanced Collaborating Environment (ACE)

## II. RELATED WORK

Much research has been initiated in the area of context-aware computing in the past few years. Many projects have been initiated for developing interactive collaboration. These projects enable users to collaborate with each other for sharing files and other media types.

The Gaia [8], [9], [10] is a distributed middleware infrastructure that manages seamless interaction and coordination among software entities and heterogeneous networked devices. A Gaia component is a software module that can be executed on any device within an Active Space

Gaia a number of services, including a context service, an event manager, a presence service, a repository and context file system. On top of these basic services, Gaia's application framework provides mobility, adaptation and dynamic binding of components.

Aura [11], [12] allows a user to migrate the application from one environment to another such that the execution of these tasks maximizes the use of available resources and minimizes user distraction. Two middleware building blocks of Aura are Coda and Odyssey. Coda is an experimental file system that offers seamless access to data [10] by relying heavily on caching. Odyssey includes application aware adaptations that offer energy-awareness and bandwidth-awareness to extend battery life and improve multimedia data access on mobile devices.

The work on user-centric content adaptation [13] proposed a decision engine that is user-centric with QoS awareness, which can automatically negotiate for the appropriate adaptation decision to use in the synthesis of an optimal adapted version. The decision engine will look for the best trade off among various parameters in order to reduce the loss of quality in various domains. The decision has been designed for the content adaptation in mobile computing environment.

## III. PROBLEM STATEMENT

In this thesis, we have tried to enhance the efficiency of file sharing and adaptation framework described in [6], [7] by introducing Key based registration and automated trusted node identification to facilitate systematic file uploading and sharing, cache optimization techniques for

effective and efficient handling of files. These new features will establish the framework as an organized system and allow users at Slave ACE nodes to share files in an efficient and faster way. Thus our problem statement may be summarized as follows:

*To provide secured registration as well as to set observation based privileges to nodes for efficient file uploading and sharing. Also to provide effective cache optimization techniques for better and intelligent collaboration among the users and efficient handling of files.*

## IV. OUR CONTRIBUTION

File Sharing and Adaptation Framework illustrated in [6], [7] include file sharing and data adaptation service. File sharing service is demonstrated by the realization of data adaptation service. These two services are very much necessary to provide effective collaboration among users in advanced collaborating environment. But there are some problems associated with this approach. As for example, in the existing framework, the slave ACE nodes are registered by their MAC addresses which are not unique in every cases. Again another drawback is, the node registration for uploading files was a completely manual process executed by the network admin, which doesn't have strong ground on historical data analysis. Moreover, frequent requests from different nodes could potentially increase the network overhead towards Master ACE node and there is no systematic solution of this problem.

In this paper, we tried to identify these problems and provide some effective solutions to address these issues. To ensure a secured registration we propose a **key based** mechanism for **user registration** to solve the problem with the MAC address. To increase the efficiency of the whole system we introduce **automated trusted node registration** which encompasses different observation phases. For solving the **cache optimization** issues, we have proposed some methods related to cache optimization so that the overall framework can work more efficiently.

Thus, the comparison shows that our work will definitely encompass a meaningful advancement over the aforementioned work in the issues of secured user registration, node registration automation, reducing the network traffic at a central point of the system and increased collaboration among users through P2P file sharing to some extent.

## V. PROPOSED MECHANISM

### A. Definitions:

#### User Registration:

In Access Grid Collaboration Environment the User Registration is normally done by providing Email address along with other necessary information. As per our proposed system, each user will be given a Key for

registration to avoid the complexity with the MAC address based registration.

#### Algorithm for Key Generation:

For generating a unique key we simply append the registration date and time with the given MAC address of the user. Thus the string becomes:

Unique Key = Registration date-time + MAC address.

#### Observation for trust-based nodes:

In our proposed mechanism the observation is done upon some specific criteria.

**Pre-trusted Node:** The system keeps track on the login period and downloaded file amount of a registered node. According to this data analysis some nodes get identified as Pre-trusted Nodes. Now they can upload files to the venue server up to a limited extent.

**Trusted Node:** For trusted node identification the system keeps track on the login period, downloaded file amount along with upload amount and frequency. Being a Trusted Node, a user can upload unlimited amount of files to the venue server also get privileged for p2p file sharing.

#### Requested node:

A Slave ACE node that requests for a file to Master ACE node is termed as Requested Node.

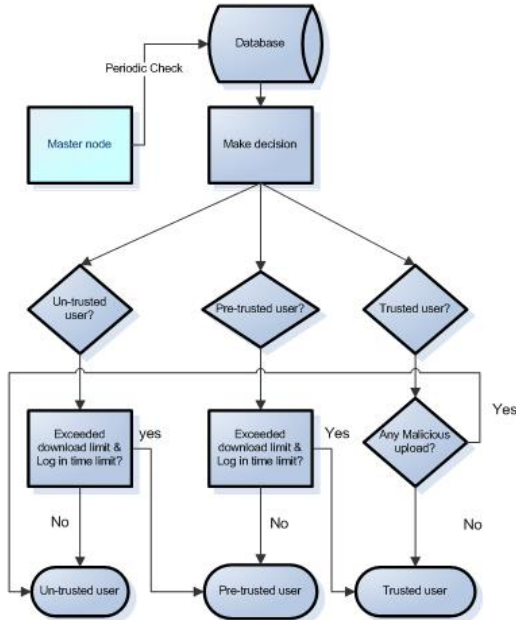


Fig. 2 Action flow of Updating node privilege.

#### B. Action Flow of Prototype Implementation

##### Action Flow for Updating Node Privilege:

In this process the master node performs a periodical check to the database to update the list of the trusted nodes. There are three different conditions. Firstly, for an un-trusted node, if the login time and download limit exceeds, it gets a promotion to be noted as pre-trusted node. Secondly, for a pre-trusted node, if the download limit and login time exceeds the given bound then it is promoted as trusted node. Finally, in case of a trusted node, if the user uploads any malicious file ever, it is demoted to the un-trusted state. Otherwise the state remains same. Figure 2 depicts the mechanism for updating the node privilege.

##### Action Flow for File Uploading Mechanism:

At first, user connects from Slave ACE node to Master ACE node. S/he needs to provide e-mail address, the assigned unique key and Master ACE node address. Then, the system checks whether the user already has a valid account with the system or not. If not then a **unique key**, generated for that particular node and user profile form are provided to that user for getting registered.

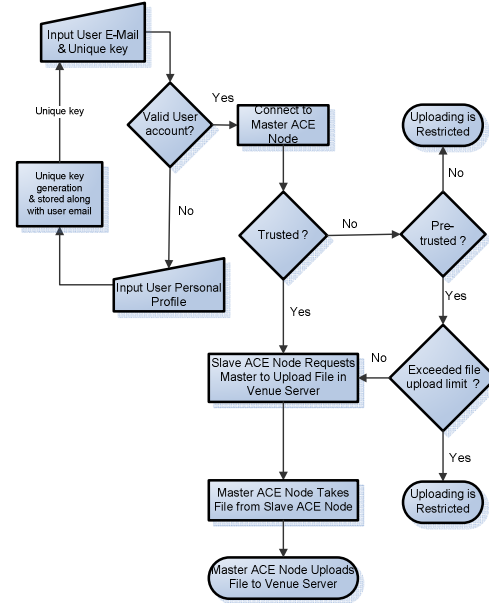


Fig. 3 Action flow of File Uploading mechanism.

If the node has already been registered then the system checks whether the node is trusted or not. If trusted then the user can upload any amount of files to the venue server. If not trusted then the system checks whether it is pre-trusted or not. If not then, the user is identified as a normal node and cannot upload any file.

Now if the user is a pre-trusted node, then the system checks if it is below the assigned limit for uploading files then allowed uploading new files. Otherwise upload is restricted. Figure 3 depicts the File Uploading mechanism of the system.

##### Cache Optimization Mechanism:

According to [7] file cache was maintained considering the hit ration of any requested files. A periodic checking (once per week) was mentioned for accomplishing the deletion of the files. The problem is an important factor, Cache size was not taken into account in that instance. Basically the deletion of the files from cache should be based on the availability of space in the cache.

When the number of hot files is increased then the cumulative size of these files would be an issue. For handling such circumstances a ranking using the hit ratio formula  $H = (\text{File counter value} / \text{Total Number of requests arrived}) \times 100\%$  should be maintained. These

files are sorted according to the rank and maintained in a queue.

The size of the cache is to be given by the system Admin. A list is maintained and the cache is filled up according to the descending rank order of the list. The list is updated each time a new file is requested and as its hit ratio increases the order may reshuffled. The least popular files are deleted from the cache. Thus the cache is always occupied by the most popular adapted files.

Previously in [7] all the adapted versions of a single file were kept in the cache. In our proposed system only the most frequently requested adapted versions will be kept in the cache. Other versions would be removed from the cache. To clarify this thing an example is given in Table 1.

TABLE I  
Example of hit\_counter (see figure 6 ) Table

| hit_file_id | counter | file_id | adapted_scaling | adapted_sharpness | adapted_color |
|-------------|---------|---------|-----------------|-------------------|---------------|
| 1           | 18      | 5       | 30              | 25                | 30            |
| 2           | 2       | 5       | 40              | 25                | 28            |
| 3           | 1       | 10      | 30              | 25                | 55            |

Suppose, a file (file\_id 5) with dimensions (30,25,30) is requested 18 times for download. Then same file (file\_id 5) with dimension (40,25,28) is requested twice. Since the dimension is different there is a new entry in the table, though the file is same. Now as the adapted version (30,25,30) is more frequently requested, so the other entry of the file is removed from the cache. So the status on\_cache for the adapted version (40,25,28) is given false.

## VI. IMPLEMENTATION PROGRESS & ISSUES

### A. Currently Implemented Modules

**Master File Sharing Service:** It had been implemented as an AG shared application. That is why, it implemented in Python. This module retrieves file from Data Store at venue server and sends this file for appropriate adaptation.

**Slave File Sharing Service:** It had been implemented as stand-alone application. It was implemented in Python. It provides the interface to the user for entering venue URL and selecting the desired file.

**Data Adaptation Service:** It is a stand-alone application. It had been also implemented in python. That module has a decision engine which provides the appropriate adaptation scheme for converting the original file. At the beginning, the user enters his/her e-mail address and preferred file type for sharing. Then, he/she presses the appropriate button for connecting Master ACE node (see figure 5). After pressing the connect button the master file sharing service will connect to the venue data store and retrieves file of user mentioned type. Then, the user will select one of the files. The specified file will be downloaded through the master file sharing service, the decision engine takes the decision of selecting appropriate adaptation method and then it will be passed to the data adaptation service. The function DataAdapter() takes the file and convert it based on the decision provided by the decision engine. Then adapted file will be sent to Slave

### B. Methods to Follow for Implementing New Features

Figure 6 shows the overall design of our proposed framework to enhance the already developed framework described earlier. The basic requirement of the file upload feature is the trusted node authentication. As per our proposed mechanism the Pre-trusted and trusted nodes are automatically validated by the system based on some periodic data analysis. As depicted in the schema diagram (see figure 7), the unique\_key is the main factor which is set by the key generation algorithm, used to identify each of the nodes separately. The Pre-trusted nodes can upload the files up to a limit. The limit can be set depending on the application environment. The trusted nodes are facilitated with an unlimited amount of file uploading capability.

There are features of the frame work given in [7] like, hot file and cold file identification based cache updating, hit ratio based advertise multicasting etc are kept intact in our system. Additionally here the system keeps track of the downloaded file in the downloaded\_file\_tracking table.

Another feature of our proposed system is to optimize the cache automatically. This method has been stated clearly in the earlier section. Cache optimization technique will surely reduce the burden to the system to store so many files as well it will increase the throughput of the overall system.



Fig. 4 Screen Shot of Previous work in [6]

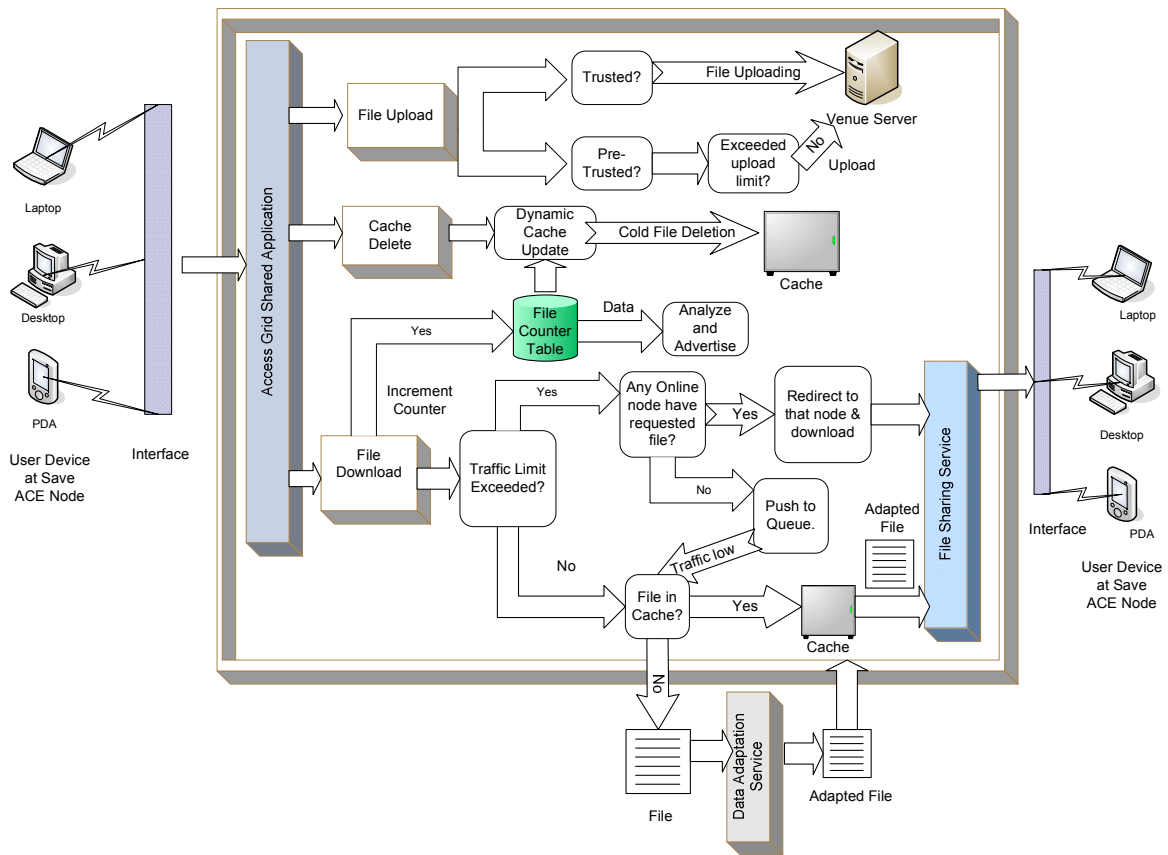


Fig. 5 Block Diagram of the Proposed System

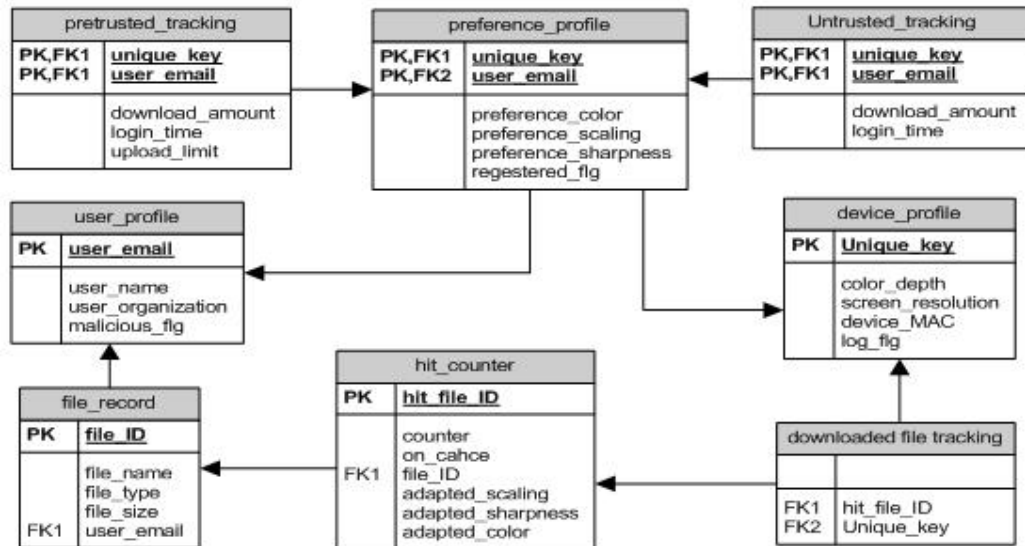


Fig. 6 Schema Design for Backend Database

## VII. CONCLUSION & FUTURE WORK

In this paper, we have presented automated and network efficiency related features for file sharing and data adaptation framework in ACE. Our proposed features have made the framework more organized and enables the users to share adapted files faster and intelligently. Moreover, cache optimization technique has added a new dimension for effective handling and storage of files. These features together realize the improved file sharing service.

Lots of interesting works are to be done in the near future for efficient file sharing and adaptation service. An interesting filed can be Context Aware System that deals with issues like pattern matching and user requirement prediction, which will ensure higher degree of user collaboration.

We believe that our effort will certainly play a leading role for overcoming the deficiencies of this framework and also break new ground for more advancement in this field of research. Hopefully, our prototype implementation will be considered as a leading work in this domain in near future.

## REFERENCES

- [1] B. Corri, S. Marsh and S. Noel, "Towards Quality of Experience in Advanced Collaborative Environments", In Proc. of the 3rd Annual Workshop on Advanced Collaborative Environments, 2003.
- [2] Sangwoo Han, Namgon Kim, and JongWon Kim, "Design of Smart Meeting Space based on AG Service Composition", AG Retreat 2007, Chicago, USA, May 2007.
- [3] R. Stevens, M. E. Papka and T. Disz, "Prototyping the Workspaces of the Future", IEEE Internet Computing, pp. 51-58, 2003.
- [4] L. Childers, T. Disz, R. Olson, M. E. Papka, R. Stevens and T. Udeshi, "Access Grid: Immersive Group-to-Group Collaborative Visualization", In Proc. of Immersive Projection Technology Workshop, 2000.
- [5] Access Grid, <http://www.accessgrid.org/>
- [6] Mohammad Rezwanul Huq, Young-Koo Lee, Byeong-Soo Jeong, Sungyoung Lee, "Towards Building File Sharing and Adaptation Service for Advanced Collaborating Environment", In the International Conference on Information Networking (ICOIN 2008), Busan, Korea, January 23-25, 2008.
- [7] Mohammad Rezwanul Huq, Md. Abdul Mottalib, Md. Ali Ashik Khan, Md. Rezwanur Rahman, Tamkin Khan Avi, "Enhanced File Sharing Service among Registered Nodes in Advanced Collaborating Environment for Efficient User Collaboration", In the International Conference on Grid Computing & Applications (GCA 2008), WORLDCOMP'08, Las Vegas, USA, July 14-17, 2008.
- [8] Anand Ranganathan and Roy H. Campbell, "A Middleware for Context-Aware Agents in Ubiquitous Computing Environments", In ACM/IFIP/USENIX International Middleware Conference, Brazil, June 2003.
- [9] The Gaia Project, University of Illinois at Urbana-Champaign, <http://choices.cs.uiuc.edu/gaia/>, 2003.
- [10] M. Roman, C. Hess, R. Cerqueira, A. Ranganathan, R. Campbell and KNahrstedt, "A Middleware Infrastructure for Active Spaces", IEEE Pervasive Computing, vol. 1, no. 4, 2002.
- [11] M. Satyanarayanan, Project Aura, <http://www-2.cs.cmu.edu/~aura/>, 2000.
- [12] M. Satyanarayanan, "Mobile Information Access", IEEE Personal Communications, <http://www-2.cs.cmu.edu/~odyssey/docdir/ieeepcs95.pdf>, Feb. 1996.
- [13] W.Y. Lum and F.C.M. Lau, "User-Centric Content Negotiation for Effective Adaptation Service in Mobile Computing", IEEE Transactions on Software Engineering, Vol. 29, No. 12, Dec. 2003.



# A New Approach to Wrap Legacy Programs into Web Services

<sup>1</sup>Saeed.Parsa, <sup>2</sup>Leila.Ghods

<sup>1</sup>Associate Professor, Computer Engineering Department, Iran University of Science and Technology (IUST), Tehran, Iran, Parsa@iust.ac.ir

<sup>2</sup>Master Student, Computer Engineering Department, Iran University of Science and Technology (IUST), Tehran, Iran, LeilaGhods@gmail.com

**Abstract** — Reengineering a legacy system to provide Web Services is a great challenge. Wrapping legacy systems is a proper solution to expose legacy program functionalities as Web Services. In this paper we propose a new approach in two phases to integrate legacy programs written in legacy languages such as C or Pascal into Web Services. This will be done by analyzing the legacy program and creating a Service Bus Class. Also, an automatic wrapping tool is developed to generate Web Services and related source code in Microsoft .Net environment. To evaluate this approach a legacy student learning program was wrapped into Web Services.

**Index Terms** — Legacy Programs, Service Bus Class, SOA, Web Service, Wrapping.

## I. INTRODUCTION

The software portfolio of many organizations consists of legacy systems which cannot be simply discarded [1]. These systems are valuable assets for organizations. They have been evolving with new emerged technologies in rapidly changing business environment. On the other hand, Web Services technology and Service-Oriented Architectures (SOA) are rapidly developed and widely supported. However SOA architecture implementation requires IT organizations to evaluate their systems architectures and determine how they will deliver Web Services. One aspect of this great issue is to deal with legacy applications which are running smoothly and performing critical tasks, but were built without taking the advent of new technologies. These applications built at high cost during several years have knowledge locked within them and need to be adapted to SOA.

Indeed, the more application software is used the more reliable it will be. Apparently, organizations prefer not to miss their existing reliable application software when applying new technologies to extend their software capabilities. On the top of that it is not economical to discard application software which has been developed and completed within a period of time by spending high expenses. Therefore, it is observed that application software developers seek for cost-effective approaches to support new business needs by leveraging and extending

existing software systems. Organizations should allow flexible and loosely coupled integration and communication between business entities. Tightly coupled applications should be transformed into flexible reusable building blocks of future systems. To achieve this, web service layers could be wrapped around existing program components, such that the component could be accessed through the web services [1][2][3]. An interesting classification of approaches for integrating legacy systems in modern environments has been presented by Zhang and Yang [18] who distinguished between the classes of black-box re-engineering techniques, which integrate systems via adaptors that wrap legacy code and data and allow the application to be invoked as a service and the class of white-box re-engineering techniques, which require analysis and modification to existing code in order to expose the system as Web Services.

Here, the main challenge is to remove the burden of handing code custom wrappers and connectors for legacy programs. Custom coding needs time, resources, and money. Although, there are some automatic and semi-automatic tools to wrap legacy programs, these tools are mostly dependent on their vendors and there is not enough accessible information about them. Also they usually work on specific programming languages.

The importance of Wrapping and Web Services motivated us to propose a new approach and an automatic wrapping tool to expose legacy program functionality written in legacy languages such as Pascal into Web Services. The proposed approach consists of two stages of analyzing and web service wrapping of legacy code. In the analysis stage reverse engineering techniques is applied to extract the functions within a given code. The extracted methods are then presented to the user to select a number of them. In the wrapping stage, a class comprising the selected methods is created. This class is then scanned by our web service wrapping tool, to convert the class methods into web services.

The paper is organized as follows: The proposed approach is presented in Section 2; the design and

implementation of Wrapper tool is demonstrated in Section 3; an example of wrapping a legacy program into Web Services and evaluation is presented (given) in Section 4; section 5 concludes the paper with suggestion for the future work.

#### A. Related Works

There are some works done on wrapping mechanism of existing programs to work in modern environments. For example, authors in [3] analyzed the different ways in which a web-based solution could be provided for the technological requirement of legacy systems and focused on the web-based development solutions. Some of the challenges inherent in migrating to Web Services are outlined in [4].

In [5] Authors used XWrap software migrating web sites and generating web services, automatically. A mechanism of Web Service gateways in which existing web sites are wrapped by several Web Service wrappers is proposed in [6]. Some of the approaches convert web sites by the means of XML and the legacy valuable functions are extracted and wrapped [7].

The efforts on the AgentEvo project, which develops an evolution approach to reengineer legacy systems into agent-based Web Services are reported in [8]. Also, Microfocus has developed a DLL object. The object can be uploaded in MS IIS to supply ASP programs with the access and call to COBOL subroutines [9].

Canfora Method can be used for the system includes user interface and has appropriate interaction with users. In this approach, a user interaction model is created and a wrapper is generated that uses a Finite State Automata to simulate different system states on the basis of the user interaction [10].

Indeed, the authors in [11] uses a gradually migration strategy to keep object oriented programs and legacy system components all together, simultaneously. In this strategy, a wrapper presents a user interface that enables new programs to communicate with legacy system components by message interchanging.

Sneed in [12] has defined a methodology to wrap legacy applications for reuse in a Service Oriented Architecture. He has approved that business functions can be extracted from existing programs, wrapped and integrated into an e-business application framework.

The authors in [13] designed a Web service conceptual architecture to be used as a template to make legacy applications available as Web Services. They have shown how it is possible to change proprietary interface into those based on open standards as needed for dynamic e-business. They also described how transaction management can be handled, without changing the existing system used by the legacy application.

A web service wrapper is implemented in [14] to extend the functionality of a JAVA RMI application hosted in an organization Intranet such that it can be executed by clients outside the Intranet using standard Protocols such as HTTP and SOAP.

## II. APPROACH REVIEW

For companies to stay competitive, it has become important to interconnect seamlessly their databases, applications and legacy systems to a coherent IT infrastructure [15]. Web Services are regarded as the key technology for implementing widely distributed applications based on loosely coupled systems [16]. It is a very efficient way for developers to reuse existing core business in a legacy system. The summary of the proposed approach is shown in Fig.1.

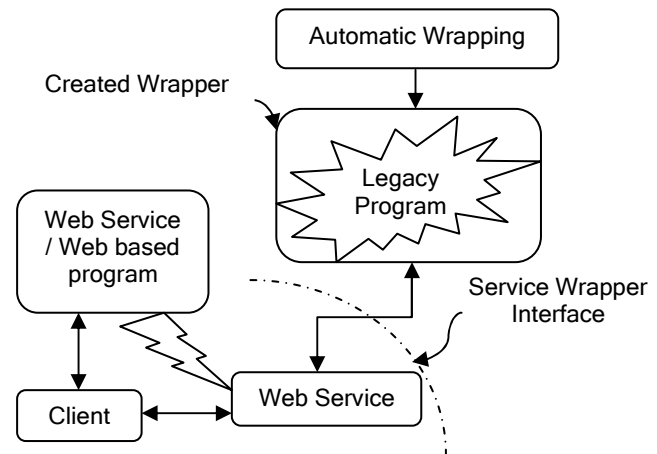


Fig.1. Wrapping Legacy Program Review

The main goal is to surround legacy programs with a wrapper and make it accessible in SOA.

#### A. Proposed Approach for Wrapping

In order to make legacy programs to be accessible with modern web based application, we propose a new approach. Our approach is proposed in two phases: The first phase includes analyzing the legacy program and determining the way of using its functions to the web environment. There are four steps to do this: Evaluation, Conversion, Reengineering and Web Service Generation. The result of this phase is a new code that contains the valuable functions of the legacy program. The analysis of the system in White Box method is shown in Fig.2.

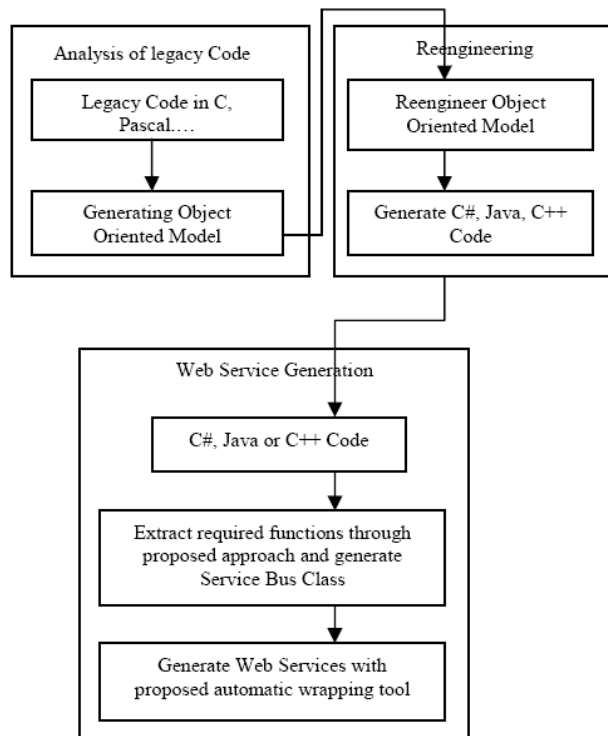


Fig. 2. Migrate legacy code in White Box method

This phase can be done through eVolution 2000 – the product of TSRI Company [17]. If the Black Box method and reverse engineering is used, the above process will be changed and there is no need to reengineer and generate object oriented model.

In the second phase, by using Visual Studio .NET and a proposed developed tool, wrapping legacy program into Web Services can be done. The Service Bus Class prepares the interested functions to be wrapped into Web Services. This process is shown in Fig.3.

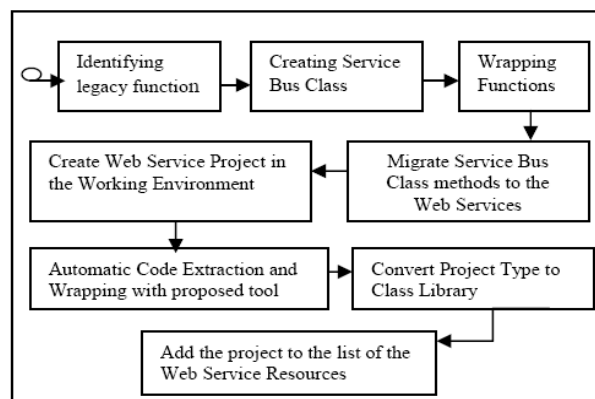


Fig. 3. The proposed process of wrapping legacy function into Web Services

### B. The Service Bus Class

The Service Bus Class acts as an intermediate between legacy system and Web Services. Its main duty is migrating legacy function to external environment. This class must conform to the following characteristics:

- (1) Since only public classes are accessible from outer classes, the accessibility of this class must be public.
- (2) The methods name of this class and legacy functions must be the same. Also, they should have the main parameters of legacy function.
- (3) If legacy function contains global variable and the White Box method is used to migrate legacy code, this variable should be added to the method call parameters.
- (4) All the communication problems with legacy program should be solved in this class and its methods.

### C. Web Service Wrapper Operation

Web Service wrapper is used to enable communication between legacy program and web based one. The task of the wrapper is to wrap the legacy function into Web Services and generate related source code directly. The designed wrapper is divided into two layers: Display Layer and Function Layer. Fig.4 shows the structure of the wrapper.

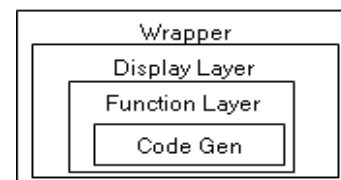


Fig. 4. Wrapper Structure

The Display Layer is to show the web service code created by the function layer. The Function Layer of wrapper implements the wrapping of method. It contains code generator. Code generator generates the Web Services code according to the properties of Web Services and Web Service Methods. Wrapper accepts the Web Service requests as a SOAP message and deserializes it, then calls the corresponding function. It also receives the response and parses and serializes it, then sends it back to the requestor as a SOAP message. Also, this approach enables asynchronous function calls. Web Services can be called asynchronously, by using asynchronous pattern call in .NET.



### III. TOOL IMPLEMENTATION

Visual studio .NET environment is extendable. Add-in projects are one of these extending methods. To create an appropriate tool in developing applications, Visual Studio .NET automation model can be used. This model is also known as extensibility model. It is a programming interface which enables accessibility to the underlying Microsoft Visual Studio procedures and manipulates them. In the other words, Add-in programs are compiled ones changes the programming environment. We used this facility to generate Web Service wrapping code. This tool enables us to select a method in the Service Bus Class and generate Web Service code, automatically. The generated web service code will be inserted to a Web Service project. Executing Web Service project makes Web Service methods to be accessible on the web.

The designed tool consists of two parts: Analyzing area of source code and assistant display area.

Analyzing area: by reading project file, developed tool parses the source code in the selected Microsoft .NET Solution, and specifies the file structure composing of Namespaces, Classes, Methods, Parameters, types of parameters, and return values of methods.

Assistant display area: The tool shows the wrapped source code of a function and displays them in the assistant display area. When a function is wrapped into Web Services, the wrapped Web Services code is displayed here. Fig.5 shows a sample code wrapped into a Web Service.

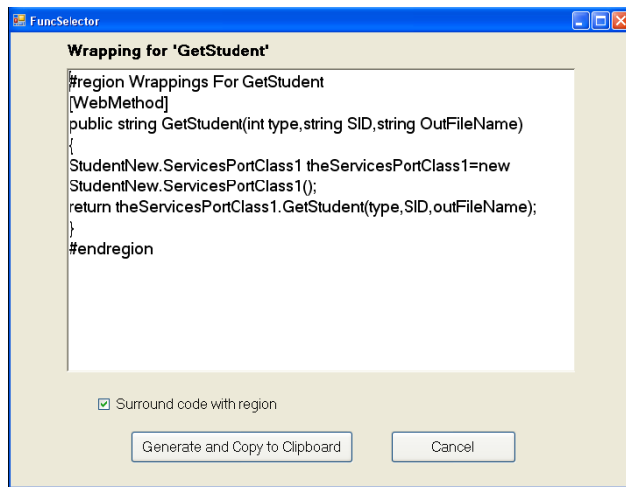


Fig. 5. Web service wrapped code

On the basis of the analysis of the legacy application, developers will decide which methods can be wrapped into Web Services methods according to the needs of

integration. This tool can simply be installed in the programming environment.

### IV. EVALUATION

To evaluate the proposed solution, we took a Student learning system written in Pascal language. Its source code and executable file is also accessible. Based on the analyzing and understanding of the legacy application, we encapsulated the necessary application functions into Web Services and provided the methods for other systems to access these Web Services.

In the student learning system, the methods of inquiring and analyzing the student personal information are encapsulated to the Web Services and are accessible to the modern web based applications. Fig.6 presents this program call of a part of this system in DOS environment.

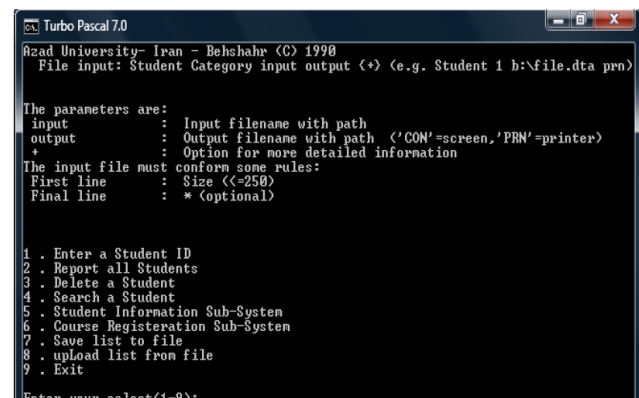


Fig. 6. Dos student learning system call

As the picture shows, the program is command line executable by entering required parameters. Below statement shows this program call.

C:\>StudentNew.exe 4 8404001-3 output.txt

According to the above analysis, there is no need to analyze program source code and the functions are available through that statement. So, we use Black Box method and only program inputs and outputs and their calling method would be specified in Service Bus Class.

After identifying the above result, the program should be entered into Visual Studio environment and wrapped into Web Services. To do this, according to the proposed approach, the functions should be identified. Then a Service Bus Class should be created and the interested functions should be wrapped and conveyed to the web with the proposed tool. At this step, a Service Bus Class named ServicePortClass1 is created. This class has a method named GetStudent. Input parameters of this class are in the form of DOS calling format. The first Parameter

is the type of execution, the second parameter is StudentID and the third one is the response result (Student Information). In this sample the input and output files should be ready for legacy program execution .So, before execution starts the input file is uploaded on the server and new Web Service named UploadFile is added to the Web Services. To generate Web Services the proposed tool will be used. The following code is generated automatically:

```
#region Wrappings For GetStudent
[WebMethod]
public string GetStudent(int type,string SID,string OutFileName)
{
 StudentNew.ServicesPortClass1 theServicesPortClass1=new
 StudentNew.ServicesPortClass1();
 return theServicesPortClass1.GetStudent(type,SID,outFileName);
}
#endregion
```

By adding this code to a Web Service project and executing it, the service will be available on the web. To access the Web Service, different applications in diverse platforms can be used. For example a win32 application in C# language was created to use the produced service. Fig.7 shows the user interface of the client program:

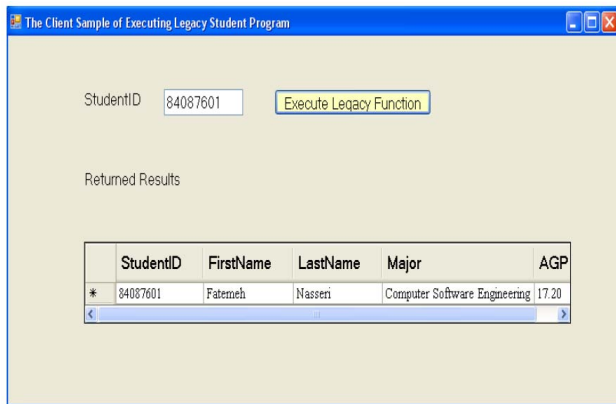


Fig. 7. The sample client program to call produced Web Service

This program calls the generated Web Service and makes the legacy DOS based function to execute.

By using the proposed approach legacy programs can reuse as a service in the web environment. Modifiability, Portability, Flexibility, and Interoperability are the main motivators behind the use of wrapping as Web Services. Developers of Web Services need not know exactly how these services are used in terms of user interfaces. They should focus on implementing business logic or server resources accessing functions with clearly defined program interfaces. Front-end applications may call Web

Services to access these important business functions and resources remotely. By publishing their Web services via UDDI or other registries, companies can offer “software as a service.” They can provide their Web services on a subscription or per-usage charge basis to generate a new revenue stream.

Our approach allows the users to submit a request and then continue his/her work without being blocked. Therefore, the response time for the clients is the sum of the time the request sent to the application server, the time the Application Server allocates a unique handle for the request, and the time for the handle to travel back to the clients. We recorded the response time in the scenario of executing program in its local legacy environment and in web environment by wrapping it into web services through our proposed tool. The analyzed server runs on the windows server platform with 2 GHz CPU and 1 MB RAM. As shown in Fig.8, our approach has some overhead in executing legacy program, because of the network traffic and wrapping operation. This overhead can be ignored because of the benefits of making legacy functionalities accessible to the web environment and provide flexibility for other web services and clients to consume legacy services.

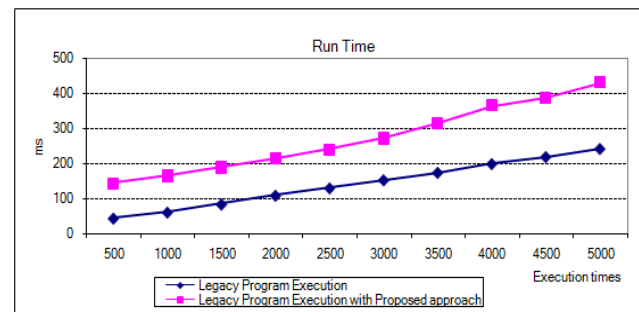


Fig. 8. Running time evaluation of legacy program

## V. CONCLUSION

Vast majority of the enterprises consists of the legacy applications. Re-architecting IT infrastructure to Service Oriented Architecture is impossible without taking these legacy applications along. The evolution of IT industry produced numerous automated tools to design and deploy Web Services. However, the approaches used by those tools were not explicitly pointed out. In order to help software developers to use legacy systems with Microsoft .Net, and generate Web Services or Web Services code according to the properties of Web Services and Web Services method, this paper designs and develops an aided tool to wrap legacy functions as Web Services. According

to the above implementation, the integration cost of legacy systems with modern technologies and programs decreases noticeably, Because the existing tools was used and the black box method which does not contain the source code analysis is also used to expose legacy functions to modern environments. Indeed low migration time increases the performance of the generated Web Services.

There is possible to extend the existing tools. The possibility of making changes in Service Bus Class with modern technologies and advances without making any changes of the old system is the other advantage of this approach.

#### A. Future Works

Although an unified approach for the proposed tool with Microsoft .Net has been presented, there are still issues to be addressed:

- (1) More common programming languages such as Java can be used to extend this process.
- (2) Because of the weak quality of legacy code, the more qualified code generation in automatic wrapping and executing will increase effectiveness.

#### REFERENCES

- [1] K. Bennett (guest editor), *IEEE software Special Issue on Legacy System*, vol.12, no.1, 1995.
- [2] J. Sang, G. Follen, Ch. Kim & I. Lopez, "Development of CORBA-based engineering applications from legacy Fortran programs", *Proceedings of the ELSEVIER information and Software Technology*, pp. 175-184., JAN. 2002.
- [3] R. Agarwal, S. Nayak, and S. Mishra, "Reengineering Legacy System using Web", *The 24th Annual International COMPSAC 2000*, Taipei Taiwan, pp. 182-183, Oct. 2000.
- [4] S. Tilley, J. Gerdes, T. Hamilton, S. Huang, H. Muller and K. Wong, "Adoption Challenges in Migrating to Web Services", *Proceedings of the Fourth International Workshop on Web Site Evolution (WSE'02)*, California, Canada, pp. 21-29, Oct. 2002.
- [5] Y. H. Lu , Y. Hong , J. Varia , D. Lee, "Pollock: Automatic Generation of Virtual Web Services from Web Sites", *SAC'05*, Santa Fe, New Mexico, USA, March 13-17, 2005.
- [6] H. P. Hu y, T. Kawamura and T. Hasegawa, "Web Service Gateway: a step forward to e-business", *Proceedings of the IEEE International Conference on Web Services(ICWS'04)*, California, USA, pp. 648-655, Jul. 2004.
- [7] M. El-Ramly, E. Stroulia and P. Sorenson, "Mining System-User Interaction Traces for Use Case Models", *IEEE* 2002.
- [8] F. Chen, H. Yang, H. G, and B. Xu, "Agentification for Web Services", *Proceedings of the 28th Annual International COMPSAC 2004*, Hong Kong, China, pp. 514-519, Sep. 2004.
- [9] L. Aversano, A. Cimitile, G. Canfora and A. De Lucia, "Migrating Legacy Systems to the Web", *IEEE Comp. Soc. Press*, pp. 148-157, 2001.
- [10] G. Canfora, A. R. Fasolino, G. Frattolillo, P. Tramontana, "A Flexible Wrapper for the Migration Of Interactive Legacy System To Web Services ", *Proceedings of the 10th European Conference on Software Maintenance and Reengineering (CSMR'06)* , *IEEE*, pp. 22, 22-24 March 2006.
- [11] C. Albanese, T. Bodhuin, E. Guardabascio, M.Tortorella "A Toolkit for Applying a Migration Strategy: A Case Study", *In Proceedings of Conference on Software Maintenance and Reengineering*, *IEEE CS Press: CSMR*, Budapest, Hungary, Los Alamitos CA, pp.154-163, March 2002.
- [12] Harry M. Sneed, "Integrating legacy software into a service oriented architecture", *Proceedings of the 10th European Conference on Software Maintenance and Reengineering (CSMR'06)* , *IEEE*, pp. 11, 22-24 March 2006.
- [13] Dietmar Kuebler and Wolfgang Eibach,, "Adapting Legacy Applications as Web Services", *IBM*, JAN. 2002.
- [14] R. Venkatavaradan and S.Srivatsa Sivan, "Design Guidelines: Building Web Service Wrappers for an XMLbased System", April 2005.
- [15] G. Menkhaus and U. Frei, "Legacy System Integration a Grammar-based Transformation", *the 26th International Conference on Information Technology Interface (ITI2004)*, Cairo, Egypt, Jun. 2004, pp. 119-124.
- [16] M. Debusmann and K. Geihs, "Towards Dependable Web Services", *Proceedings of the 10th IEEE Pacific Rim International Symposium on Dependable Computing*, Papeete, Tahiti, French Polynesia, pp. 5-14, Mar. 2004.
- [17] The Software Revolution, Inc. (TSRI), Available at: <http://www.softwarerevolution.com/>
- [18] Z. Zhang and H. Yang, " Incubating Services in Legacy Systems for Architectural Migration", *Proceeding, 11th Asia-Pacific software engineering conference, IEEE*, Busan Korea, pp. 196-203, DEC. 2004.

# Speech Enabled Operating System Control

Md. Abdul Kader, Biswajit Singha, and Md. Nazrul Islam

Department of Computer Science and Engineering  
Khulna University of Engineering & Technology  
Khulna-9203, Bangladesh

E-mail: {kdr2k4, biswa2k4cse, kuet\_nazrul2k}@yahoo.com

**Abstract** — Our goal is to provide the computer with a natural interface, including the ability to understand human speech. For this purpose, we propose a way how to operate an Operating System with voice command. At first, the user initiates a given command by his voice through the microphone then the recognition software of the proposed system will take over to recognize the command. If the recognition is succeeded or matched with one of the given voice command then it will perform the operation on the Operating System according to speaker's speech command. In our proposed system we have used Markov chain model for voice recognition process and Voice-XML for creating the voice grammar in the software part. It has the flexibility to work with the speech of any user means and remove the problem of 'hesitation' which is very effective for spontaneous speech recognition. The proposed system minimizes the error rate up to 30% which is very precise for any speech recognition system.

**Index Terms** — Voice recognition, Voice-XML Operating System, Speech, Voice command.

## I. INTRODUCTION

Recent years it has been seen that the improvements in the quality and performance of speech-based human-machine interaction is steady. The next generation of speech-based interface technology will enable easy to use automation of new and existing communication services, making human-machine interaction more natural. For the disabilities people the absence of the data bases and diversity of the articulator handicaps are major obstacles for the construction of reliable speech recognition systems, which explains poverty of the market in systems of speech recognition for disabilities people [1].

If someone is unable to handle mouse and keyboard or the ports of mouse and keyboard are not functioning, properly there exist some other means to access operating system. "Speech" may act as one of them. There is a growing demand for systems capable of controlling Operating System using only the voice of a person. And this paper represents a way how to control the OS by using voice command. First of all, speech is recognized by the system software and if the speech or voice command is recognized then each and every command executes the

corresponding operation upon the Operating System. The desirable characteristics of such a system include scalability that is the efficiency of the process of speech recognition and executing the commands. When we speak, a microphone takes the analog signal and converts the analog signal from our voice into digital chunks of data that the computer analyzes. From this data the computer must extract enough information to confidently guess the word being spoken. This is not, however a small task! In fact, in the early 1990s, the best recognizers were yielding a 15% error rate on a relatively small 20,000 dictation tasks, so the controlling process is very poor. But here an efficient method is used for speech recognition, so the recognition process works very efficiently. And after getting the command the software executes the corresponding operation, so the Operating System operated very smoothly with a user's speech.

The later part of this paper shows the efficient recognition process, Markov model, voice command execution process and an example of the process for controlling the Operating System.

## II. METHODOLOGY

### A. Voice Recognition Process:

#### Step 1: Extract Phonemes

The process of speech recognition starts with a sampled speech signal. This signal has a good deal of redundancy because the physical constraints on the articulators that produce speech—the glottis, tongue, Lips, and so on—prevent them from moving quickly [4]. Phonemes are best described as linguistic units. They are the sounds that group together to form our words, although quite how a phoneme converts into sound depends on many factors including the surrounding phonemes, speaker accent and age. English uses about 40 phonemes to convey the 500,000 or so words it contains, making them a relatively good data item for speech engines to work with. Here are a few examples of phonemes and words:

|    |                 |    |                   |
|----|-----------------|----|-------------------|
| Aa | F <u>a</u> ther | T  | T <u>a</u> lk     |
| Ae | C <u>a</u> t    | Th | Th <u>i</u> n     |
| Ah | C <u>u</u> t    | Uh | B <u>o</u> ok     |
| Ao | D <u>o</u> g    | Uw | T <u>oo</u>       |
| Aw | F <u>ou</u> l   | Zh | plea <u>s</u> ure |
| Ng | S <u>i</u> ng   |    |                   |

Fig-1: Some phonemes and words

Phonemes are often extracted by running the waveform through a Fourier Transform. This allows the waveform to be analyzed in the frequency domain. It is probably easier to understand this principle by looking at a spectrograph. A spectrograph is a 3D plot of a waveform's frequency and amplitude versus time. In many cases though, the amplitude of the frequency is expressed as a color (either grey scale or a gradient color). The spectrograph of saying "Generation5" is given by the following Fig-2:

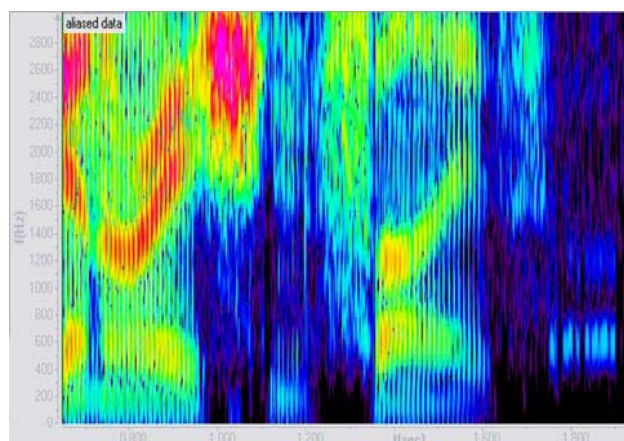


Fig-2: spectrograph of saying "Generation5"

For comparison purpose, another spectrograph of the "ss" bit of assure (this is a phoneme) is also given by the following Fig-3:

Using this, can anybody identify where is "Generation5" and the "sh" of Generation5 in the spectrograph? Note that the timescales are slightly different on the two spectrographs, so they look a little different. It is relatively easy to match up the amplitudes and frequencies of a template phoneme with the corresponding phoneme in a word. For computers, this task is obviously more complicated but definitely achievable [3].

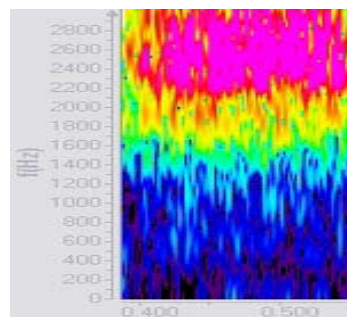


Fig-3: spectrograph of saying "ss"

## Step 2: Markov Models

Now that the computer generates a list of phonemes, then obviously these phonemes have to be converted into words and perhaps even the words into sentences. This can be very complicated indeed, especially for systems designed for speaker-independent, continuous dictation. However, the most common method is to use a Hidden Markov Model (HMM). Basically, think of a Markov Model (in a speech recognition context) as a chain of phonemes that represent a word. The chain can branch, and if it does, is statistically balanced. For example the following Fig-4 shows a Markov chain:

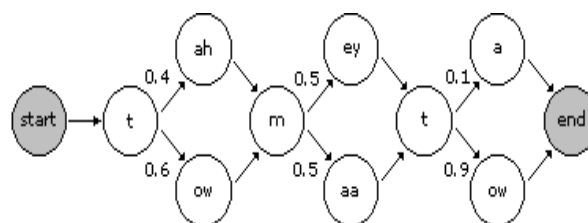


Fig-4: Markov chain

Note that the Markov Model in Fig-4 represents both the American English and the (real) English methods of saying the word "tomato". In this case, the model is slightly biased towards the English pronunciation. This idea can be extended up to the level of sentences, and can greatly improve recognition. For example:

Recognize speech  
Wreck a nice beach

These two phrases are surprisingly similar, yet have wildly different meanings. A program using a Markov Model at the sentence level might be able to ascertain which of these two phrases the speaker was actually using through statistical analysis using the phrase that preceded it. For speech recognition this process adopts all the above processes.



## B. Command Execution Process:

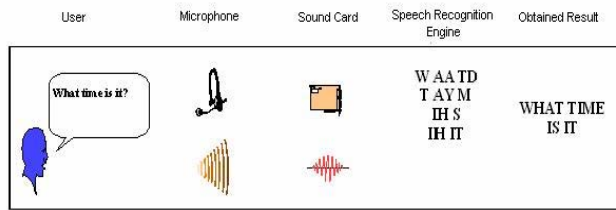


Fig-5: Working Procedure of this software

**User:** At first user is giving command.

**Microphone:** In this step the command is received by microphone and the microphone converts sound wave into electrical pulse.

**Sound Card:** In this step the sound card converts electrical pulse into digital signal.

**Speech Recognition Engine:** In this step Speech Recognition Engine converts digital signals into phonemes.

**Command:** Finally we get text command.

**Operations:** For every command the software performs the corresponding operations.

There are about hundred commands exists for this software, some pre-commands & corresponding operations given in the following table 1:

| Command        | Operation                      |
|----------------|--------------------------------|
| MyComputer     | Explore My Computer            |
| StartMenu      | Explore Start Menu             |
| Up             | Arrow Key up one unit          |
| Down           | Arrow Key down one unit        |
| Left           | Arrow Key right one unit       |
| Right          | Arrow Key left one unit        |
| CursorUp       | Mouse pointer up               |
| CursorDown     | Mouse pointer down             |
| CursonLeft     | Mouse pointer left             |
| CoursorRight   | Mouse pointer right            |
| Back           | Back space                     |
| CloseTheWindow | Close window                   |
| Enter          | Enter                          |
| Ok             | Mouse left button single click |

Table-1: some commands and corresponding action

### Example of some commands action:

At first microphone receives our speech as sound wave. This sound wave is converted into digital signal by the sound card. This digital signal is converted into phonemes by speech engine. This phoneme is our text command. And this text command is executed by our Java program. We can describe this by some examples.

### Example-1:

If we want to explore my computer, we have to give command **“My Computer”**. And after command execution we can see the following Fig-6.

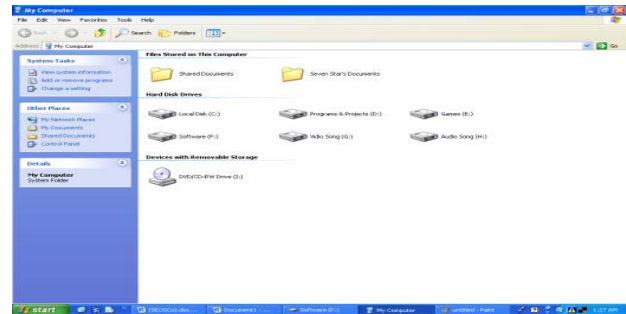


Fig-6: My computer exploration

### Example-2:

If we want to explore start Menu, we have to give command **“Start Menu”**. And if we want to access any kind of programs from the Start Menu, we have to give different kinds of commands such as **“Left”, “Right”, “Up”, “Down”, “Enter”**. If we give **“Up”** command then the selection moves one unit up from its current position. Giving such command we can select our destination program and to explore this program we have to give **“Enter”** command. Thus we can explore other programs according to our choice. Fig-7 shows start menu after the **“Start Menu”** command execution.

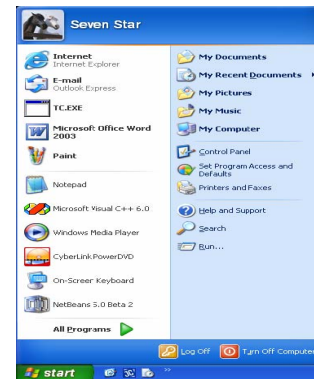


Fig-7: Start Menu exploration

After exploring **“My computer”**, if we want to open different drives we have to access these drives following the process stated in above step i.e. using **“Left”, “Right”, “Up”, “Down”, “Enter”** commands.

### Example-3:

To create a new folder we have to give **“File”** command. After opening **“File Menu”** (Fig-8) we have to give **“Down”** command. By giving **“Down”** command

we have to select the “New”. Then we have to give “Enter” command to create a new folder. Thus we can create a new folder and so on.

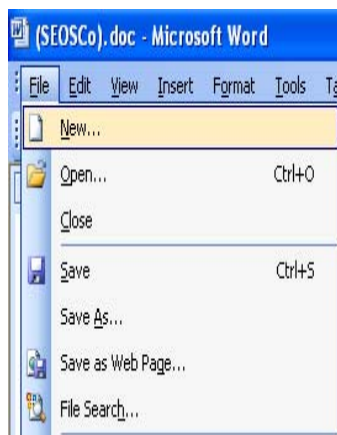


Fig-8: Browse file menu of MS word

To delete a folder we have to give “File” command. After opening “File Menu” we have to give “Down” command. By giving “Down” command we have to select “Delete”. Then we have to give “Enter” command. And thus we can delete a folder. We can also control the different kinds of media player by using voice command in this software.

Example-4:

If we want to search any kind of file or folder, we have to give “Start Menu” command. After exploring “Start Menu” we have to select “search” giving “Up” command. Now we have to give “Enter” command to explore “Search” after select it. We can also see this process following Fig-9.

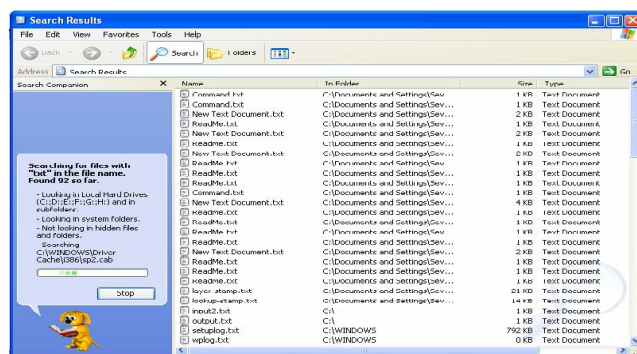


Fig. 8: Searching a file or folder  
Example-5:

We can also write anything opening notepad or Microsoft word document. We can also access any kind of options of Microsoft word document by this software. Fig-9 and Fig-10 shows the text editors.



Fig-9: Writing in notepad

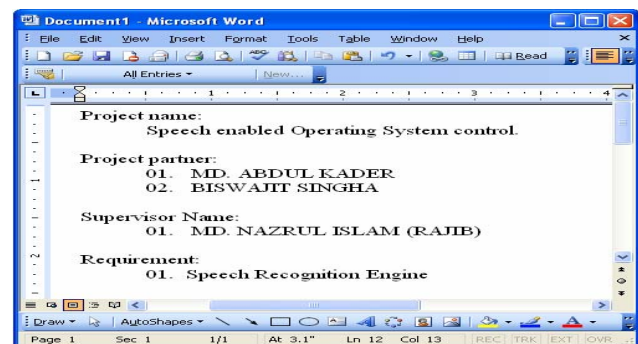


Fig-10: Writing in MS word

Example-6:

We can move the cursor from one position to another position of the window. If we want to move the cursor from:

1. Left to Right we have to give “**Cursor Right**” command.
2. Right to Left we have to give “**Cursor Left**” command.
3. Up to down we have to give “**Cursor Down**” command.
4. Down to up we have to give “**Cursor Up**” command.
5. Left Up corner to Right Down corner we have to give “**Right Down**” command.
6. Left Down corner to Right Up corner we have to give “**Right Up**” command.
7. Right Down corner to Left Up corner we have to give “**Left Up**” command.
8. Right Up corner to Left Down corner we have to give “**Left Down**” command.

After giving any of the above cursor related command we have to give the units (10,20,30,40,50,60,70,80,90, 100, 200,300,400,500) to move the cursor.

We can also maximize and minimize the window giving the “*Maximize*” and “*Minimize*” command respectively. Following such method stated above steps we also access another options of our Operating System.

### III. TECHNOLOGY & FEATURES

We have developed our software in Visual C++ & Java [6]. For voice control purpose we have used Microsoft Speech SDK (Speech Development Kit) [5] SAPI. We have also used the internal Speech Engine for Recognition purpose and XML [2] grammar formation for creating grammar file. There are some features of this software:

1. It is Speech Application Programming Interface (SAPI) implemented software.
2. This software is designed to control Operating System according to the command of a person’s voice.
3. By using this software a number of operations on the Operating System are executed.
4. It can be used as Voice Controlled Operating System software.
5. It is user friendly, so any user can handle it easily.

### IV. UTILITIES & LIMITATIONS

Natural voice recognition system faces the complex problem of spontaneous voice recognition, namely hesitations, out of vocabulary. An efficient dialogue design can greatly improve the performance of the voice interface.

This software helps those people who are physically disable and doesn’t have the liberty of using mouse and keyboard as external input device.

If the ports of mouse and keyboard do not work properly then one can also operate operating system using this software. It may save our cost from purchasing of mouse and keyboard.

Speech Recognition Engine takes some acoustic signals which are not required for our command. For such acoustic signals sometimes our command cannot be recognized properly and is executed in a changed manner. This is the main limitation of this software.

### V. FUTURE PLAN

We believe that we are quite successful in our initial stage. But there are some limitations in our software. We are hoping that we will overcome these limitations and

add extra features in future if possible. Some of the developments that we intend to do in near future are given below that is we are determined:

1. To develop our voice command in Bengali language besides English.
2. To use Bluetooth Technology for the voice command input.
3. To remove the acoustic signals that is not expected in our software.

### VI. CONCLUSION

The interface of the software is user friendly. This software is applicable for Windows XP operating system. As maximum computer users feel comfortable to use this operating system, our software may satisfy their need to access operating system through their voice command which will save time and make life easier. It is capable of meeting the requirements of people unable to use their hands. We are hoping that if we can release our software commercially in software market then it could take a good standing among other similar projects. And if we are provided with sufficient technical and financial help, we can develop this software more effectively and take it to next level. We are looking forward to improve our software’s efficiency and application to make it truly platform independent.

### ACKNOWLEDGEMENT

The authors wish to acknowledge the Professor **Dr. M.M.A. Hashem**, Head of the department of CSE, KUET and lecturer **Tanay Dey**, Dept. of CSE, KUET.

### REFERENCES

- [1] Ben Mosbah, B., “Speech Recognition for Disabilities People” Information and Communication Technologies, 2006. ICTTA apos;06. 2<sup>nd</sup>, Volume 1, Issue , 24-28 April 2006 Page(s): 864 - 869
- [2] Voice XML <http://www.w3.org/TR/2003/CRvoicexml20-20030220/>
- [3] Keller, Eric. “Fundamental of Phonetic Science.” In Fundamental of Speech Synthesis and Speech Recognition, Keller, Eric ed. John Wiley & Sons, Chichester, England, 1994
- [4] Mukund Pabmanabhan, Michel Pichney, “Large-vocabulary speech recognition algorithms”, *IEEE Computer magazine*, 0018-9162/02, pp. 42-50, 2002.
- [5] <http://www.microsoft.com/speech/download/sdk51/>
- [6] <http://java.sun.com/j2se/1.5.0/docs/guide/language/index.html>



# Short Text Compression for Smart Devices

Md. Rafiqul Islam<sup>1</sup>, S. A. Ahsan Rajon<sup>1</sup> and Anonda Podder<sup>2</sup>

<sup>1</sup>Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh.

<sup>2</sup>Leads Corporation, Dhaka, Bangladesh.

**Abstract** — Short Text Compression is a great concern for data engineering and management. The rapid use of small devices especially, mobile phones and wireless sensors have turned short text compression into a demand-of-the-time. In this paper, we propose an approach of compressing short English text for smart devices. The prime objective of this proposed technique is to establish a low-complexity lossless compression scheme suitable for smart devices like cellular phones and PDAs (Personal Digital Assistants) having small memory and relatively low processing speed. The main target is to compress short messages up to an optimal level, which requires optimal space, consumes less time and low overhead. Here we propose a new static-statistical context model to obtain the compression. We also propose an efficient probabilistic distribution based content-ranking scheme for training the statistical model. We analyze the performance of the proposed scheme as well as the other similar existing schemes with respect to compression ratio, computational complexity and compression-decompression time. The analysis shows that, the required number of operations for the proposed scheme is less than that of other existing systems. The experimental results of the implemented model gives better compression for small text files using optimum resources. The obtained compression ratio indicates a satisfactory performance with reduced memory requirements and lower complexity. The compression time is also lower because of computational simplicity. In overall analysis, the simplicity of computational requirement encompasses the compression effective and efficient.

**Index Terms** — Short Text Compression, Syllable, Statistical Model, Text-ranking, Static Coding, Smart Devices.

## I. INTRODUCTION

Twenty-First century is the age of information and communication technology. Science through its marvelous achievements has converted this world into information and communication based global village. The prime aspect of present technology is to ensure a better way of communication throughout the world in a more convenient, easy and cost-effective way. With the aspects of cost, facility and reliability a new trend of introducing small sized devices with some sorts of computing and communicating power has established its place in the arena of research. With the voyage of introducing smart devices, the challenge of adorning them with greater and effective use has come into question. It is now a great concern to embed maximum applications within these smart devices where it is an extreme problem to provide a

low-complex and low-memory consuming version for smart devices of some prime necessary applications like data compression, which generally requires large memory and greater processing speed. Mobile communication that is a great gift of modern technology introducing the era of digital communication also suffers from the same limitation. Though crossing the boundary of voice communication, short messages communication has established its robust place in the arena of digital communication, Short Message Service (SMS) providers (usually Telecommunication Companies) have a constraint that each message should be not more than of 160 characters. This constraint is really a great limitation for frequent communication using SMS. In order to overcome this limitation, compression of the short message is a well policy. That is why; our aim is to make “Short” messages “Shorter”, expressing “Larger” feelings in “Smaller” expense.

Here we introduce a scheme of compressing short English text for smart devices like cellular phones and wireless sensors having small memory and relatively low processing speed communicating with lower bandwidth *i.e.* channel capacity. We have employed a new statistical model with a novel approach of integrating text ranking or component categorization scheme for building the model. Modified syllable based dictionary matching and static coding is used to obtain the compression. Moreover, we have employed a new theoretical concept of choosing the multi-grams, which has facilitated us to obtain mentionable compression ratio using a small number of knowledgebase entries than other methods, consuming less resource.

## II. LITERATURE SURVEY

The following sections give a glimpse of the most recent research developments on short text compression for small devices.

The recent literature regarding short text compression titled “Compression of Short Text on Embedded Systems” by Rein *et al.* proposes a low-complexity version of PPM (Prediction by Partial Matching). They use statistical models with hash table lengths of 16384, 32768, 65536 and 131072 elements requiring two bytes for each element, which results an allocation of 32, 64, 128 and

256 Kbytes of RAM respectively. If this memory requirement may be substantially decreased, we may achieve more efficient compression and hence may make the scheme usable to even very low-quality cellular phones. Another concerned approach by Rein *et al.* is “Low Complexity Compression of Short Message” [1] with “Low Complex and Power Efficient Text Compressor for Cellular and Sensor Networks” [11] are variations of [10].

“Compression of Small Text Files Using Syllables” proposed by Lansky *et al.* [2, 7] concerns on compressing small text files using syllables. To implement their concept they created database of frequent syllables. In this scheme, the primary knowledgebase size is more than 4000 entries initially. For low memory devices, it is obviously difficult to afford this amount of storage as well as to facilitate a well suited mechanism of searching; which leads our proposed scheme to redefine the knowledge base span as short as possible and hence to reduce the scope of loosely choosing the syllables or n-grams. Moreover, it may be a good idea to have specific consideration of syllable involving spaces. In [2, 7], all the training syllable entries are stored without any categorization. This often results for coding redundancy, which can be handled by integrating text ranking or component categorization scheme with syllable selection.

The model proposed by M. R. Islam *et al.* [3] uses the advantages of greedy sequential grammar transform with block sorting to compress data. However, this scheme is highly expensive in terms of memory consumption and thus not suitable for low memory devices.

In [1] the compression starts for text sequences larger than 75 Bytes, and in [10] the starting point is 50 Bytes. If it is possible to make the lower threshold value into less than ten characters, the compression may really be a “very small text” supported one that may place a new milestone in very small text file compression era. That is, our prime aim is to design such an effective and efficient “very short text compression” scheme.

### III. SHORT TEXT COMPRESSION FOR SMART DEVICES

The prime concern of this paper is to implement a lossless compression of short English text for smart devices in a low complexity scheme. The idea behind this is to make it possible to communicate more efficiently by making proper and minimum utilization of bandwidth of wireless and sensor networks for small and embedded devices. More precisely saying, the main concern is to develop a low-complexity compression technique suitable for low-power consuming small-memory smart devices especially for wireless sensors and mobile phones.

The proposed scheme is concerned with two parts. The first one consists of training the statistical model and the second provides a compression-decompression technique. Specifically saying, in analyzing step we count the frequency of representable ASCII characters. Then, we proceed by identifying the syllables (as defined by [2, 7]) of length two, three, four and five. We consider <space> as a distinct vowel and include this while counting the frequency of syllables, that is, searches for <space> either at the beginning or ending of syllable. In the step of boosting the statistical model, we consider the substrings (which were not grabbed in the phase concerning syllables) with length two to five and count frequency of each. In the second step, we employ the provided text-ranking scheme for each entry and calculate the entry index. For entries with same index, we simply sort them. In the phase of choosing entries, we give emphasis on probability distribution based text ranking, which is computed with the help of its neighbor characteristics.

As in most cases, it is unusual to have frequent match of substrings with length more than four characters, we consider maximum of five (extra one for <space>) levels to train the statistical model. In each level, multiple entries with same weightage are simply sorted over. Resultant entries are assigned with non-conflicting binary stream. When we are to compress any text, the input text is successively compared with the statistical model and for any match, the binary stream is returned as output and for mismatch in any level, the level below it is forwarded.

The total compression-decompression is expected to be performed in the following manner.

#### A. Compression Process

In the first step, we plan to employ Modified Multi-grams or Syllable Matching proposed by Lansky *et al.* [2, 7]. A static dictionary based compression scheme uses approximately the same concept as that of character masking. It reads in input data and searches for symbols or groups of symbols that are previously coded or indexed in the dictionary. If a match is found, a pointer or index into the dictionary can be output instead of the code for the symbol. Compression occurs if the pointers or index requires lower space (in terms of memory usage) than the corresponding fragments. [1, 10, 2] Though it is the basic idea behind multi-grams, we use it in a slightly modified fashion. The Knowledgebase for the multi-grams are constructed with the help of the statistics obtained by analyzing the corpuses. We have selected a number of strings and substrings, which are most frequently used in short messages (as well in as usual purposes) and have calculated the probability of using those by considering the average of the probability they appeared with their individual occurrence.

In the second step of compression, we use Static Coding. However, here the modification is made in such a way that, in spite of calculating on-stage codes we use predefined codes in order to reduce the space and time complexity. The codes for dictionary entries are also predefined. The total message is encoded by a comparatively small number of bits and hence, we get a compressed outcome. The prime fact of static coding is that, in case of dynamic coding we are to calculate the frequencies in an elegant manner. Moreover, as the dynamic codes are determined and the lengths are varied through the frequencies of the symbols it is a must to submit the codes along with the compressed data. This is a limitation of dynamic coding if the total arena of compression is low spanned and the total number of symbols is not huge. Compression of short message also suffers in this context while using other similar semi-dynamic coding. Consequently, we are to move towards static coding scheme to obtain compression of short messages. For this concerned study, we have analyzed the corpora- Bib, Book1, Book2, News, Paper1, Paper2, Paper4 and Paper5 from Calgary and Canterbury Corpus.

#### B. Decompression Process

The decompression process is performed through the following steps:

- Step 1: Grab the bit representation of the message,
- Step 2: Identify the character Representation,
- Step 3: Display the decoded message.

As all the symbols are to be coded in such a fashion that, by looking ahead several symbols (Typically the maximum length of the code) we can distinguish each character (with the attribute of Static Coding). In step 1, the bit representation of the modified message is performed. It is simply analyzing the bitmaps. The second step involves recognition of each separate bit-patterns and indication of the characters or symbols indicated by each bit pattern. This recognition is performed on the basis of the information from fixed encoding table used at the time of encoding. The final step involves simply representing i.e. display of the characters recognized through decoding the received encoded message.

#### IV. THEORETICAL ASPECT OF TRAINING THE PROPOSED STATISTICAL MODEL

The proposed scheme achieves better compression ratio with relatively low complexity by means of computational simplicity and effective expert model, which is used to train the statistical context. Firstly, the prime modification is performed in the syllable selection section of [2, 7] proposed by Lansky *et al.* In their papers, only syllables were considered by defining maximal subsequence of

vowels including pseudo-vowel 'y'. Here we propose to consider <blank space> as a *Prime Syllable*. Using <space> as a *prime syllable* may dramatically reduce the total number of characters needed to represent the message through sophisticated encoding. Secondly, the paper [2, 7] does not clearly express the criteria of choosing the training syllables for the model. The term used in [2, 7] to define the criteria of choosing syllables was simply "frequentness". However, we employ a new theoretical perspective on "Text Ranking or Component Ranking" method to choose the syllables for training our proposed model. We, in the first step count the frequency of each vowel from the standard Text Calgary Corpora. In the second step, we count the number of vowels having space either at  $(i - 1)$  or  $(i + 1)$  position with itself at position  $(i)$ . The reason of adding the with-space vowels is simply to increase the weightage of the corresponding vowels. These entries having vowels are also inserted in the knowledgebase as separate entity. In this step, we also count the statistics of consonants in order to build the dictionary or multi-grams entries. In the step of extending primary knowledgebase, proposed in papers [2, 7] the criteria was that, the frequency would be 1:65000, which results a knowledgebase size of more than 4000 entries initially. For low memory devices, it is obviously difficult to afford this amount of storage as well as to facilitate a well suited mechanism of searching, which leads our proposed scheme to redefine the knowledge base span as short as possible and hence to reduce the scope of loosely choosing the multi-grams. In the phase of choosing multi-grams, we give emphasis on Probability Distribution, which is computed with the help of its neighbor characteristics.

We employ a new Text Weighting or Component Ranking scheme to select the multi-grams that facilitates us to efficiently construct the knowledgebase. The proposed text weighting or component ranking is obtained through the following equation:

Suppose we choose a multi-grams  $D$  consisting of characters  $C_1 C_2 C_3 \dots C_n$  of length  $n$ . Thus,

$$\partial(D_{C_1, C_2, C_3, \dots, C_n}) =$$

$$\sum_{i=1}^n (\partial(D_{C_i}) - 1) - \partial(D_{C_1, C_2, C_3, \dots, C_{n-1}}) + \lambda \lambda_{C_1, C_2, C_3, \dots, C_n})$$

for  $i > 1$  and,  $\partial(D_{C_i}) = \lambda(C_i)$  for  $i = 1$ , where  $\lambda(C_i)$  is the frequency of the character  $C_i$  with assumption that each character of the alphabet must exist in the training data. And the value of  $\partial(D_{\phi})$  indicates the multi-grams index of character  $D_{\phi}$ . Here, we refer the multi-gram index obtained from  $\partial(D_{C_1, C_2, C_3, \dots, C_n})$  as the

resultant text-weightage for the text  $C_1 C_2 C_3 \dots C_n$ . For example, we want to have the probability distribution of multi-grams "and".

The steps will be: Let  $\partial(a) = x_1$ ,  $\partial(n) = x_2$ ,  $\partial(d) = x_3$ ,  $\partial(an) = x_4$ .  
Therefore,  $\partial(ana) = (\partial(a) - 1) + (\partial(n) - 1) + (\partial(d) - 1) - \partial(an) + \lambda(C_{abc}) = x_1 + x_2 + x_3 - x_4 + x_5$

Static coding is defined as a set  $S = \{B_1, B_2, B_3, \dots, B_{n-1}, B_n\}$  of binary streams, where each element of the set  $B_1, B_2, B_3, \dots, B_{n-1}, B_n$  are uniquely identifiable providing that,  $B_1, B_2, B_3, \dots, B_{n-1}, B_n$  are not necessarily to be of equal bit length. As all the elements of the sets supports our three basic points of concern- identifiability, uniqueness and variability of any encoding scheme, we are interested to use static coding. This theoretical aspect let us do not consider updating the model throughout the compression, because, firstly, updating a larger knowledgebase using a very short data is not computationally affordable and effective, secondly, a larger portion of the source code may be saved if the update is avoided resulting faster execution [1,10,11] and finally, if the training statistical data is not permitted to be updated as well as to be expanded, we may ensure a constant and consistent memory optimization for the overall compression process. That is, there is no possibility to expand the knowledgebase arbitrarily and thus there is no risk of arising the overloaded memory problem or out-of-memory problem. As the knowledgebase is not updated, the use of static coding is also perfect for the same.

## V. PERFORMANCE ANALYSIS

Though it is a general idea that compression and decompression time should have an inter-relation, the proposed scheme demonstrated a little exception. The points behind that may be summarized through the following discussions.

### A. Performance Analysis of Compression Process

Let the total number of training entries for the statistical model be  $N_g$ , where  $N_g$  is a non-negative integer and the maximum level for statistical modeling is  $L$ . The first level of the statistical model will must contain the single characters, where the total number of distinct character is  $l_1$ . For levels 1, 2, 3, ...,  $n-1$ ,  $n$  the total number of distinct multi-gram entries are  $l_1, l_2, l_3, \dots, l_{n-1}, l_n$  respectively.

When any text is to be compressed, it is hierarchically compared with each level of statistical model starting from the highest order. If there is any match, the corresponding static coding for multi-gram entry is assigned for the text. If the multi-gram entry is not found throughout the level, it is forwarded to the next level. This assignment uses efficient searching procedures. Let the code  $m$  is found at the  $i$ -th level with offset  $k$  resulting a search cost of  $S_m(l_j) + k_m$ , where  $k_m < l_i$  and,  $j=L, L-1, \dots, i-1$  with respect to search space. Here  $j$  limits from  $L$  to  $(i-1)$  instead of  $i$

decreasingly because, as we find the code in somewhere of  $i$ -th level not requiring to search the whole element-space of the  $i$ -th level, rather searching through an offset value  $k$  for  $i$ -th level, the overall search-space is  $L$  to  $(i-1)$ . That is why, for the above consequences, the total searching appears search overhead for  $(i-1)$  number of levels with additional search overhead of  $k$  elements. Here the term "search overhead" stands for the search space complexity as well as other related computational requirements. When the code matches, it is placed in output stream as character representation. This step requires padding the bit-stream and then conversion into character stream. Assume that, the process of overall conversion for each successful entry occurs with the overhead  $B$ . That is, for any multi-gram matching, the required overhead is,

$$C_1 = \sum_{j=L}^{i-1} (S_1(l_j)) + k_1 + B_1$$

Similarly,  $C_2 = \sum_{j=L}^{i-1} (S_2(l_j)) + k_2 + B_2$ , and

$$C_n = \sum_{j=L}^{i-1} (S_n(l_j)) + k_n + B_n$$

In such a way if a total of  $n$  multi-grams are identified and then encoded, the required resultant number of operations in compression process is:

$$T = \sum_{y=1}^n C_y = \sum_{y=1}^n \sum_{j=L}^{i-1} (S_y(l_j)) + \sum_{y=1}^n k_y + \sum_{y=1}^n B_y \quad (1)$$

### B. Performance Analysis of Decompression Process

For the decompression process, the text to be decompressed is converted into binary stream. If the largest code is of length  $c_{max}$  and the smallest code is of length  $c_{min}$  then the decompression process will start the searching with the  $c_{max}$  number of bits and search through the codes up to  $c_{min}$  bits by reducing one bit in each step for unsuccessful match. It is necessary to mention that, the codes with same bit length does not essentially comprise any specific level. So, to reveal the character representation for each entry  $d$  if a switch of  $h$  levels are required, where  $c_{min} \leq h \leq c_{max}$  where the maximum level is  $p$  with the matching offset for corresponding level  $k_d$ , and the assignment of the code with character representation for each successful match requires an overhead of  $B_d$ , then the overall requirement for comparing through the each level settings results (= Overhead of Searching through level + Overhead for Searching through offset + Overhead of Representation). For detecting first character the overhead is,

$$E_1 = \sum_{q=p}^h (S'_1(l'_q)) + k'_1 + B'_1$$

Similarly, for detecting the second character, the level-wise overhead will be:  $E_2 = \sum_{q=p}^h (S'_2(l'_q)) + k'_2 + B'_2$

$$\text{Hence, } E_n = \sum_{q=p}^h (S'_n(l'_q)) + k'_n + B'_n \quad (2)$$

Here,  $p = \text{maximum level}$ , and  $S'$  is a function that denotes the search-overhead for searching in element space provided as parameter of the function and  $h = \text{minimum level}$ . The computation progresses through  $p, p-1, p-2, \dots, h+2, h+1, h$ . Here the subscript  $n$  is used to denote the level-wise overhead for detecting one character representation with respect to level.

In order to detect a single multi-grams  $f$ , the total search-overhead with respect to search space for level-wise calculation is,  $\sum_{q=p}^h (S'_f(l'_q)) + k'_f$

because we are to start with maximum level  $p$  and then proceed decreasingly towards the downward levels  $h$  (as explained above). If  $S'$  is the search-overhead function, then searching from level  $p$  to  $h$  will result  $\sum_{q=p}^h (S'_f(l'_q))$  where  $f$  is the multi-grams, which

is being revealed. For the matching level, as only a partial number of elements are to be searched, the offset  $k$  is used to denote the offset.

After checking through the levels, the procedure follows searching through the bit-wise statistics for any unsuccessful match in level-wise statistics. If there are a total of  $u$  bit-phases, we are to perform searching through the search-space consisting of starting from the maximum bit phase to the minimum bit phase decreasingly. Because of any unsuccessful match in any bit-phase, a bit switch is performed and level wise calculation for that level is forwarded. That is, an overhead of  $\sum_{b=d}^g (E_b)$  is incurred

for each level-wise analysis. Consequently, the overhead of unit step will be,  $C'_1 = \sum_{b=d}^g (E_b)$  where  $d$  and  $g$

are maximum and minimum bit phases respectively and  $d \geq g$ . Substituting the value of  $E_b$ , we get,

$$C'_1 = \sum_{b=d}^g \sum_{q=p}^h (S'_{1,b}(l'_q)) + \sum_{b=d}^g k'_{1,b} + \sum_{b=d}^g B'_{1,b}$$

Similarly, we get,

$$C'_2 = \sum_{b=d}^g \sum_{q=p}^h (S'_{2,b}(l'_q)) + \sum_{b=d}^g k'_{2,b} + \sum_{b=d}^g B'_{2,b}$$

$$C'_n = \sum_{b=d}^g \sum_{q=p}^h (S'_{n,b}(l'_q)) + \sum_{b=d}^g k'_{n,b} + \sum_{b=d}^g B'_{n,b}$$

Here, we use the subscript  $1$  with  $k$  and  $B$  in order to

denote that, the calculations are for detecting unit code only where the calculation is performed starting from  $d$  to  $g$  in decreasing order, that is, in the order of  $d, (d-1), (d-2), \dots, (g+2), (g+1), g$ . If we are to reveal  $n$  number of codes, then the total overhead becomes:  $T' = \sum_{y=1}^n C'_y$

As for each bit wise overhead calculation, we must include level-wise calculations; we may omit the subscript notation for search overhead function for simplicity,

$$T' = \sum_{y=1}^n \sum_{b=d}^g \sum_{q=p}^h (S'_{y,b}(l'_q)) + \sum_{y=1}^n \sum_{b=d}^g k'_{y,b} + \sum_{y=1}^n \sum_{b=d}^g B'_{y,b} \quad (3)$$

## VI. EXPERIMENTAL RESULTS AND DISCUSSIONS

The performance evaluation is performed on the basis of the file “book1”, “book2”, “paper4”, “paper5” and “paper6” from “Calgary Corpus”. As the prime aspect of our proposed Compression Scheme is not to compress huge amount of text rather to compress texts with limited size affordable by the mobile devices i.e. embedded systems, we took blocks of texts less than five hundred characters chosen randomly from those files ignoring binary files and other non-text files and performed the efficiency evaluation.

Recent study involving compression of text data are:

1. “Low Complex and Power Efficient Text Compressor for Cellular and Sensor Networks” (Mode 2) by Rein *et al.* [1, 10] and,
2. “A modification Of Greedy Sequential Grammar Transform based data Compression” by Islam *et al* [3]

We denote the above three methods as DCM-1 and DCM-2 respectively.

The simulation was performed in a 2.0 GHz Personal Computer with 112 MB of RAM with the object oriented programming language Java. The average compression ratio for three random execution results for different size of blocks of text is as follows:

Table 2: Compression Ratio:

| Corpus         | Number of Characters Considered | Compression Ratio (%) for DCM-1 | Compression Ratio (%) for DCM-2 | Compression Ratio (%) for proposed scheme |
|----------------|---------------------------------|---------------------------------|---------------------------------|-------------------------------------------|
| <b>paper 4</b> | 108                             | 44.01                           | 44.03                           | <b>42.94</b>                              |
| <b>paper 5</b> | 061                             | 44.98                           | 45.51                           | <b>44.02</b>                              |
| <b>paper 6</b> | 032                             | 45.22                           | 45.96                           | <b>44.30</b>                              |
| <b>book 1</b>  | 191                             | 46.84                           | 48.11                           | <b>45.97</b>                              |
| <b>book 2</b>  | 104                             | 43.95                           | 46.69                           | 45.27                                     |

The compression ratio is a metric to describe how many compressed units are required to describe one unit of data. The lower the presented value is the better is the compression. A general observation is that higher modes

lead to better compression ratios even if the difference with higher orders becomes smaller.

The prime achievement of our proposed scheme is, it is even capable to compress texts with only five characters. Though the compression ratio is deteriorated for that case, it is an evolutionary step for small text compression.

## VII. CONCLUSION AND RECOMMENDATION

We have presented an effective and efficient approach of compressing short English text message for low-powered embedded devices. Here we have employed modified syllable based dictionary matching and static coding to obtain the compression, which two was not previously adapted in the same model. Moreover, we have used a new theoretical concept of choosing the multi-grams, which has facilitated us to obtain mentionable compression ratio using a small number of knowledgebase entries than other methods consuming less resource. The overall strategy of computational simplicity has also ensured the reduced time complexity for the proposed compression and decompression process. The main aspect of our proposed scheme resides is the text ranking based knowledge-base construction with space integration that initiates a new arena of text compression methodology. A consistent and relevant mathematical analysis of the overall performance also establishes a strong technical basis of the proposed scheme. Moreover, the prime achievement is in the scale of starting threshold of text compression; that we have reduced to less than five characters. With limited knowledge-base size, the achieved compression is of no doubt efficient and effective. As the knowledge base is not accepted to be grown through the continuous applications, we may keep out the low-memory system from the risk of expanding its knowledgebase crossing optimal memory size and thus, the applicability of the proposed system even in any very low memory devices is ensured.

## REFERENCES

- [1] S. Rein, C. Guhmann, and F. Fitzek, "Low Complexity Compression of Short Messages," *Proceedings of the IEEE Data Compression Conference (DCC'06)*, March 2006, pp.123–132.
- [2] J. Lansky, M. Zemlicka, "Compression Of Small Text Files Using Syllables". *Proceedings of Data Compression Conference (DCC'06)*, Los Alamitos, CA, USA, 2006.
- [3] Md. Rafiqul Islam, Sajib Kumar Saha, Mrinal Kanti Baowaly. "A modification of Greedy Sequential Grammar Transform based Universal Lossless data Compression". *Proceedings of 9th International Conference on Computer and Information Technology (ICCIT 06)*, 28-30 December, 2006, Dhaka, Bangladesh.
- [4] Mohammad Golam Robiul Alam, Md. Mosharaf Hossain, Md. Faisal Hussain, Imtiaz Uddin Khan: "Huffman Compression with Diagrams Selected by Association Rule of Data Mining", *Published in the Proceedings of 8<sup>th</sup> International Conference on Computer and Information Technology (ICCIT 05)*, 28-30 December, 2005, Dhaka, Bangladesh.
- [5] David Hertz: "Secure Text Communication for the Tiger XS". *Master of Science Thesis, Department of Electrical Engineering, Linköpings University, Linköping, Sweden.*
- [6] En-Hui Yang and Da-ke He, "Efficient Universal Lossless Data Compression Algorithms Based on a Greedy Sequential Grammar Transform – part Two: With Context Models", *IEEE trans Inform Theory* Vol 49 no: 11, November 2003.
- [7] J. Lansky and M. Zemlicka. "Text Compression: Syllables". In *Annual International Workshop on Databases, TExtS, Specifications and Objects (DATESO)*, Volume 129 of *CEUR Workshop Proceedings*, pp. 32–45. 2005.. *CEUR-WS*.
- [8] Min Chen, "Mobile Entertainment", *T-109.551 Research Seminar on Telecommunications Business II.*, Helsinki University of Technology.
- [9] S. Rein and C. Guhmann, "Arithmetic Coding–A Short Tutorial", Wavelet Application Group, *Technical Report*, April 2005.
- [10] Stephan Rein, Clemens Guhmann, Frank H. P. Fitzek: "Compression of Short Text on Embedded Systems", *Journal of Computers*: Volume 1, No: 06, September 2006.
- [11] Stephan. Rein, F. Fitzek, M. P. G. Perucci, T. Schneider and C. Guhmann, "Low Complex and Power Efficient Text Compressor for Cellular And Sensor Networks", In *15<sup>th</sup> IST Mobile and Wireless Communication Summit*, June 2006.
- [12] Md. Rafiqul Islam, S. A. Ahsan Rajon, Anonda Podder, "Lossless Compression of Short English Text for Low-Powered Deices", in the *proceedings of International Conference on Data Engineering and Management (ICDEM 2008)* Tiruchirappalli, Tamil Nadu, India. February 9, 2008.
- [13] S. A. Ahsan Rajon and Anonda Podder, "Lossless Compression of Short English Text for Low-Powered Devices"- *A non-published undergraduate thesis, CSE Discipline, Khulna University, Khulna, Bangladesh.*

# Smallest and Some New Equiprojective Polyhedra

Masud Hasan<sup>1</sup>, Mohammad Monoar Hossain<sup>1</sup>, Sabrina Nusrat<sup>1</sup> and Alejandro López-Ortiz<sup>2</sup>

<sup>1</sup> Department of Computer Science and Engineering

Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

Email: masudhasan@cse.buet.ac.bd, monowar3306@yahoo.com, nusratsabrina@yahoo.com

<sup>2</sup> Cheriton School of Computer Science, University of Waterloo, Ontario N2L 3G1, Canada

Email: alopez-o@uwaterloo.ca

**Abstract**—A convex polyhedron  $P$  is  $k$ -equiprojective if for all of its orthogonal projections, except those parallel to the faces of  $P$ , the number of vertices in the shadow boundary is  $k$ . Finding an algorithm to construct all equiprojective polyhedra is an open problem first posed in 1968. In this paper we give lower bounds on the value of  $k$  and the size of an equiprojective polyhedron. We prove that there is no 3- or 4-equiprojective polyhedra and a triangular prism is the only 5-equiprojective polyhedron. We also discover some new equiprojective polyhedra.

**Keywords:** Algorithm, equiprojective polyhedra, Johnson solid, lower bound, orthogonal projection

## I. INTRODUCTION

The problem of constructing new polyhedra based on certain mathematical properties has been extensively studied since antiquity in the fields of architecture, art, ornament, nature, cartography, and even in philosophy and literature. (See the book [3] and the web page [6] for some interesting discussions on the history of discovering new polyhedra.) Regularity of faces, edges and vertices, and symmetry are some popular criteria for constructing new polyhedra. For example, in a platonic solid (platonic solids are the most primitive convex polyhedra) each vertex is incident to the same number of identical regular faces [3], [6].

A convex polyhedron  $P$  is  $k$ -equiprojective if its shadow (i.e., the boundary of orthogonal projection) is a  $k$ -gon in every direction, except directions parallel to faces of  $P$ . A cube is 6-equiprojective, a triangular prism is 5-equiprojective (in fact, any  $n$ -gonal prism is  $(n+2)$ -equiprojective [10], [11]), and a tetrahedron is not equiprojective. See Figure 1. In 1968, in a paper [13] entitled, “Twenty problems on convex polyhedra,” Shephard defined equiprojective polyhedra, gave the examples above, and asked for a method to construct all equiprojective polyhedra. Later, Croft, Falconer, and Guy included this problem in their book “Unsolved Problems in Geometry” [2].

So far little progress has been made in this problem. Hasan and Lubiw [10], [11] gave a new characterization and an  $O(n \log n)$  time recognition algorithm for equiprojective polyhedra. Their characterization shows that the class of equiprojective polyhedra is rich. For example, zonohedra, which is an infinite class of convex polyhedra, are equiprojective [10], [11].

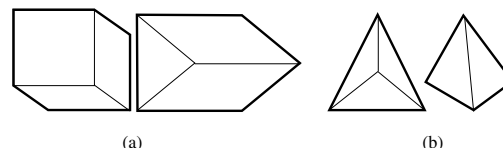


Fig. 1. (a) Some equiprojective polyhedra, and (b) a non-equiprojective polyhedron.

Rahman *et. al.* [12] also proposed some new non-trivial equiprojective polyhedra, which they achieved by cutting and gluing some existing convex polyhedra. In other related work, Hasan, in his PhD thesis [7], extended the idea of equiprojectivity from the number of vertices in the shadow to the number of visible faces, visible edges, and visible vertices in the projection. However, the original problem of constructing all equiprojective polyhedra still remains open.

In this paper we further explore the class of equiprojective polyhedra. As a lower bound on smallest equiprojective polyhedra, we prove that there is no 3- or 4-equiprojective polyhedra and a triangular prism is the only 5-equiprojective polyhedron (Section 3). We also discover some new equiprojective polyhedra by cutting some existing polyhedra (Section 4). Most of our newly discovered equiprojective polyhedra are non-trivial and to our knowledge, none of them has been discovered before as an equiprojective polyhedron.

## II. PRELIMINARIES

For rest of the paper by polyhedra we mean convex polyhedra and by projection (shadow) of a polyhedron we mean its non-degenerate orthogonal projection (shadow) (i.e., the projection direction is not parallel to any face).

The idea of characterization of an equiprojective polyhedron  $P$  is based on partitioning the edge-face pairs of  $P$  into “compensating pairs” [10], [11]. Any edge  $e$  of the shadow of  $P$  corresponds to some edge of  $P$ . As the projection direction changes,  $e$  may leave the shadow boundary. This only happens when a face  $f$  containing  $e$  in  $P$  becomes parallel to the projection direction. In order to preserve the size of the shadow, some other edge  $e'$  must join the shadow boundary. In

order for these events to occur simultaneously,  $e'$  must be an edge of  $f$ , or of a face  $f'$  parallel to  $f$ , and  $e$  and  $e'$  work as compensating each other during the change in the shadow boundary due to  $f$  and  $f'$ . For all such changes in the shadow boundary of  $P$ , each edge in each face, that means each edge-face pair, must be compensated by another edge-face pair.

More precisely, for edge  $e$  in face  $f$ , we call  $(e, f)$  an *edge-side*. Two edge-sides  $(e, f)$  and  $(e', f')$  are *parallel* if  $e$  is parallel to  $e'$  and  $f$  is parallel or equal to  $f'$ . Since in a convex polyhedron, a face can have at most one parallel face, an edge-side has at most three parallel edge-sides. The *direction* of edge-side  $(e, f)$  is the unit vector in the direction of edge  $e$  as encountered in a clockwise traversal of the outside of face  $f$ . Edge-sides  $(e, f)$  and  $(e', f')$  *compensate* each other if they are parallel and their directions are opposite (i.e. one is the negation of the other). In particular, this means that either  $f = f'$  and  $e$  and  $e'$  are parallel edges of  $f$ , or  $f$  and  $f'$  are distinct parallel faces and  $e$  and  $e'$  are parallel edges lying at the “same end” of  $f$  and  $f'$ , where “same end” means a plane through  $e$  and  $e'$  put  $f$  and  $f'$  in the same half space. Note that an edge-side has at most two compensating edge-sides, one within the same face and the other one in the parallel face. When the edges of a face  $f$  are in parallel pairs, they are compensated within  $f$  and in that case  $f$  is called a *self-compensating* face. See Figure 2(a).

**Theorem 1:** [10], [11]  $P$  is equiprojective iff its set of edge-sides can be partitioned into compensating pairs.

Consider a direction  $d$  in the plane of a face  $f$  of  $P$  and its parallel face  $f'$  (if  $f'$  exists) but not in the plane of any other face. While looking from  $d$  the edges of  $f$  are partitioned into two chains, one is visible and the other one is invisible from  $d$ . We call these two chains *lower chain* and *upper chain* of  $f$  and denote them by  $L_d(f)$  and  $U_d(f)$  respectively. Similarly, let  $L_d(f')$  and  $U_d(f')$  be the lower and upper chains of  $f'$  (if it exists) from  $d$ . Let  $\varepsilon$  be a small vector normal to the face  $f$ . Consider the directions  $d + \varepsilon$  and  $d - \varepsilon$ . Choose  $\varepsilon$  small enough such that the above two directions make  $f$  (similarly  $f'$ , if it exists) invisible (visible) and visible (invisible), respectively, and does not change the visibility status of any other face. That means between  $d + \varepsilon$  and  $d - \varepsilon$  the change in the shadow boundary is due to  $f$  and  $f'$  only, and from  $d + \varepsilon$  (similarly from  $d - \varepsilon$ ) the edges of  $f$  and  $f'$  that are in the shadow of  $P$  are  $U_d(f)$  and  $L_d(f')$  (similarly  $L_d(f)$  and  $U_d(f')$ ) respectively. See Figure 2(b) that illustrates the above facts by showing for face  $f$  only. So, the following lemma is obvious.

**Lemma 1:** Assuming both  $L_d(f')$  and  $U_d(f')$  are empty if  $f'$  does not exist, the edges in  $U_d(f) \cup L_d(f')$  have compensating pairs in  $U_d(f') \cup L_d(f)$ .

### III. MINIMALITY

One important subclass of equiprojective polyhedra are *zonohedra*. (There exist more than one definition of zonohedra [1], [4], [5] and we consider the original and “generalized” definition [14] of it). A zonohedron is a convex polyhedron whose edges in each face are in parallel pairs (i.e., each face is self-compensating). A zonohedron has the property that its faces are in parallel pairs with corresponding edges parallel [11], [14].

An *n-gonal prism* is a convex polyhedron with two parallel copies of a convex  $n$ -gon with each pair of corresponding vertices connected by an edge. For an  $n$ -gonal prism, edges in one  $n$ -gonal face are compensated by their parallel pairs in the other  $n$ -gonal face and all other faces are parallelograms (and thus self-compensating). Thus all prisms are equiprojective.

Observe that both a zonohedron and a prism have at least one pair of parallel faces. In fact, in the following lemma we prove that this is true for any equiprojective polyhedron.

**Lemma 2:** Let  $P$  be an equiprojective polyhedron. Then  $P$  has at least one pair of parallel faces.

*Proof:* By way of contradiction assume that  $P$  has no parallel faces. Since  $P$  is equiprojective, by Theorem 1 all of its edge-sides are partitioned into compensating pairs. By definition, an edge-side can be compensated by another one within the same face or in the parallel face. So in  $P$  each edge-side is compensated within the same face, which implies that all faces of  $P$  are self-compensating, further implying that  $P$  is a zonohedron. But the faces of a zonohedron are also in parallel pairs, a contradiction. ■

By Lemma 2, let  $f$  and  $f'$  be two parallel faces of  $P$ .

**Lemma 3:** From direction  $d + \varepsilon$  (similarly from  $d - \varepsilon$ ), in the shadow boundary of  $P$ ,  $U_d(f)$  and  $L_d(f')$  (similarly  $U_d(f')$  and  $L_d(f)$ ) are disjoint.

*Proof:* We prove this for  $U_d(f)$  and  $L_d(f')$ . Proof for the other pair is analogous.

Observe that in a convex polyhedron, two parallel faces are vertex disjoint. So,  $f$  and  $f'$ , as well as  $U_d(f)$  and  $L_d(f')$ , are disjoint in  $P$ . Now we will prove that the disjointness of  $U_d(f)$  and  $L_d(f')$  in  $P$  implies their disjointness in the shadow boundary of  $P$ . Remember that  $P$  is convex and that we only consider its non-degenerate orthogonal projections. It implies that an edge and a vertex in the shadow boundary of  $P$  correspond exactly to an edge and a vertex of  $P$  respectively. (In the contrary, if  $P$  is non-convex and/or the projection of  $P$  is a degenerate one, then a vertex in the shadow boundary of  $P$  could represent an edge, a crossing of two edges, overlapping of two vertices, etc. of  $P$  and similarly, an edge in the shadow boundary could represent a face of  $P$ ). Thus  $U_d(f)$  and  $L_d(f')$  do not overlap or cross each other in the shadow boundary of  $P$ . ■



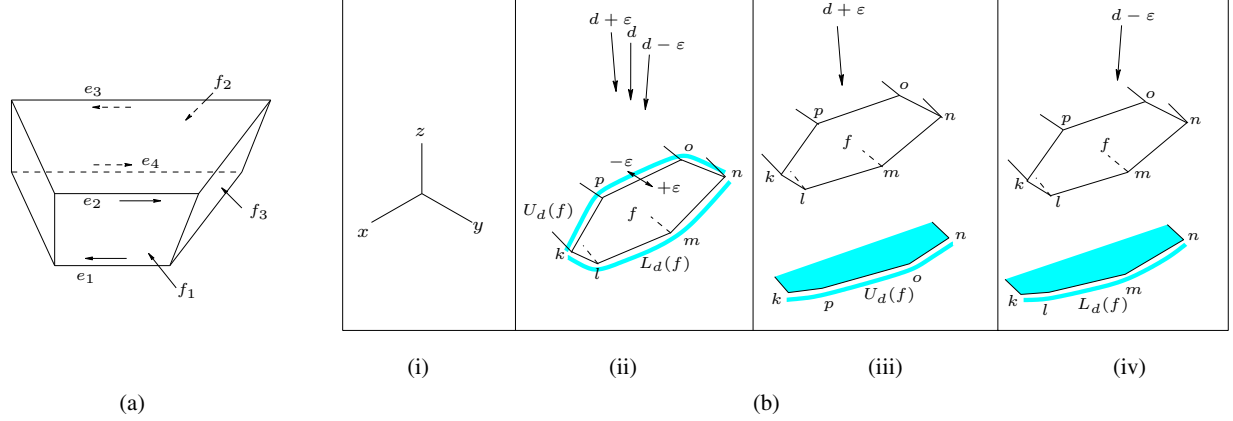


Fig. 2. (a) Examples of some compensating edge-sides.  $(e_1, f_1)$  is compensated by  $(e_2, f_1)$  and by  $(e_4, f_2)$  but not by  $(e_3, f_2)$ . Face  $f_3$  is self-compensating. (b) Illustration of  $U_d(f)$  and  $L_d(f)$ : (i,ii) Face  $f$  lies in the  $x$ - $z$  plane and  $d$  is in the  $z$  direction;  $U_d(f)$  is the chain  $k, p, o, n$ , and  $L_d(f)$  is the chain  $k, l, m, n$ . Vector  $\varepsilon$  is in the  $y$  direction. (iii, iv) From  $d + \varepsilon$  and  $d - \varepsilon$ , edges of  $f$  that are in the shadow of  $P$  are  $U_d(f)$   $L_d(f)$  respectively.

**Corollary 1:** Both  $U_d(f) \cup L_d(f')$  and  $U_d(f') \cup L_d(f)$  have at least three edges.

*Proof:* Both  $f$  and  $f'$  are at least triangles. We know that  $U_d(f) \cup L_d(f) = f$  and  $U_d(f') \cup L_d(f') = f'$ . By Lemma 3,  $U_d(f)$  and  $L_d(f')$  are disjoint and  $U_d(f')$  and  $L_d(f)$  are disjoint. By Lemma 1,  $U_d(f) \cup L_d(f') = U_d(f') \cup L_d(f)$ . So the lemma holds. ■

**Theorem 2:**  $P$  is  $k$ -equiprojective for  $k \geq 5$ .

*Proof:* Consider the projection of  $P$  from direction  $d + \varepsilon$ . By Corollary 1,  $U_d(f) \cup L_d(f')$  has at least three edges. Since by Lemma 3,  $U_d(f)$  and  $L_d(f)$  are disjoint in the shadow boundary of  $P$ , there are at least two edges in the shadow boundary that are not the edges of  $U_d(f)$  or  $L_d(f)$ . So the size of the shadow boundary is at least five. Finally, since  $P$  is equiprojective, it holds for any other direction. ■

**Corollary 2:** If  $P$  is 5-equiprojective, then  $f$  and  $f'$  are triangles.

**Theorem 3:** A triangular prism is the only 5-equiprojective polyhedron.

*Proof:* Let  $P$  be 5-equiprojective. By Corollary 2, two parallel faces  $f$  and  $f'$  of  $P$  are triangles. We first show that  $P$  has no vertex other than the six in  $f$  and  $f'$ . By way of contradiction if  $P$  has an extra vertex  $v$ , then there must be a direction  $d$  in the plane of  $f$  and  $f'$  and not in the plane of any other face such that  $v$  is in the shadow boundary (see Figure 3(a)). Then projections from  $d + \varepsilon$  and  $d - \varepsilon$  have at least six vertices each in the shadow boundary: at least five from  $f$  and  $f'$  (by Lemma 3 and Corollary 1) and  $v$ . See Figure 3(b). This leads to a contradiction that  $P$  is 5-equiprojective.

So the remaining faces of  $P$ , let their set be  $F$ , are formed by the vertices of  $f$  and  $f'$ . We now show that  $F$  has exactly three parallelograms which form a triangular prism with  $f$  and  $f'$ . Let the counter-clockwise vertex-edge sequences of  $f$  (from

outside of  $f$ ) be  $v_1, e_1, v_2, e_2, v_3, e_3$  and the clockwise vertex-edge sequences of  $f'$  (from outside of  $f'$ ) be  $v'_1, e'_1, v'_2, e'_2, v'_3, e'_3$ . So  $e_1, e_2, e_3$  in  $f$  are compensated by  $e'_1, e'_2, e'_3$  in  $f'$  respectively. See Figure 3(c). For any  $i = 1, 2, 3$ , if  $e_i$  is an edge of a face  $g \in F$ , then  $v'_{i+2}$  can not be a vertex of  $g$ , otherwise it would violate the convexity of  $P$ . Now, if  $v'_i$  is a vertex of  $g$ , then  $v'_{i+1}$  must also be a vertex of it, since  $e_i, v'_i, v'_{i+1}$  are coplanar and thus must form a single face of  $P$ . Moreover, no other vertex of  $P$  is coplanar with  $e_i, v'_i, v'_{i+1}$ . Therefore, in summary,  $F$  has only three faces  $f_i$ , for  $i = 1, 2, 3$ , where  $f_i$  has exactly four edges  $e_i, e'_i, (v_i, v'_i), (v_{i+1}, v'_{i+1})$ . Now it is obvious that none of  $f_i$ 's has a parallel pair. Since  $P$  is equiprojective,  $f_i$  must be self compensating. Since  $e_i$  and  $e'_i$  are parallel,  $(v_i, v'_i)$  must be parallel to  $(v_{i+1}, v'_{i+1})$ , which makes  $f_i$  a parallelogram. Therefore,  $P$  is nothing but a triangular prism. ■

#### IV. NEW EQUIPROJECTIVE POLYHEDRA

In this section we obtain four new equiprojective polyhedra from a tetrahedron, a quadrilateral pyramid, a triangular cupola (the third Johnson solid) and a pentagonal rotunda (the sixth Johnson solid), respectively, by cutting them suitably. *Johnson solids* are convex polyhedra with regular faces. We describe our new equiprojective polyhedra as equitruncated tetrahedra, equitruncated pyramid, equitruncated triangular cupola, and equitruncated pentagonal rotunda respectively. For the proof of their equiprojectivity we only give the sketch, rigorous proofs can be found in [8]. We have also implemented these new polyhedra in VRML and their code can be found in [9].

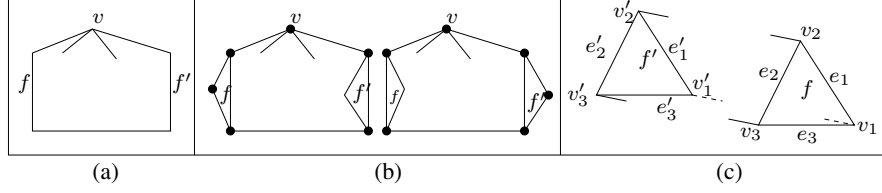


Fig. 3. (a) Projection when  $d$  is in the plane of  $f$  and  $f'$  only. (b) Projections from  $d + \varepsilon$  and  $d - \varepsilon$  have at least six vertices each in the shadow boundary (shown in filled circles). (c) Vertex-edge sequence of  $f$  and  $f'$ .

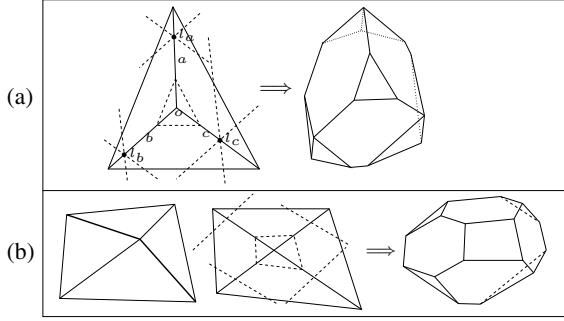


Fig. 4. (a) An equitruncated tetrahedron obtained from a tetrahedron. (b) An equitruncated pyramid obtained from a pyramid.

#### A. Equitruncated tetrahedra

A *cut* to a polyhedron  $P$  we mean a plane that cuts  $P$  into two pieces. By a cut  $C$  through a line  $l$  or a point  $p$  denotes a plane  $C$  which contains  $l$  or  $p$ .

An *equitruncated tetrahedron*  $P$  is obtained by cutting an arbitrary tetrahedron  $P'$ , i.e.,  $P'$  may not be regular. Two types of cuts are applied here. See Figure 4(a). The *top cut* is parallel to the base and cuts the apex into a triangle. All the remaining six *side cuts* are pairwise similar and are defined as follows. Let the apex of  $P'$  be  $o$  and let the three edges adjacent to  $o$  be  $a, b$  and  $c$ . Take three points  $p_a, p_b, p_c$  on  $a, b, c$  and imagine three parallel lines  $l_a, l_b, l_c$  through  $p_a, p_b, p_c$  such that the lines properly intersect the base (e.g., they are perpendicular to the base). Take a pair of parallel side cuts through  $l_a$  and  $l_b$  such that they are parallel to  $c$ . Similarly, take two other parallel pairs of side cuts, where one pair is through  $l_a$  and  $l_c$  and is parallel to  $b$  and the other pair is through  $l_b$  and  $l_c$  and is parallel to  $a$ . That ends the construction of  $P$ . See Figure 4(a).

**Theorem 4:** An equitruncated tetrahedron is 10-equiprojective.

*Proof:* Edges in the top triangle of  $P$  are compensated by the three (old) edges of the base. Each pair of parallel cuts creates a pair of parallel triangles with corresponding edges compensating each other. Six (new) edges of the base are themselves in parallel pairs and thus compensate each other within the base. The remain-

ing three hexagonal faces are self-compensating. By Theorem 1,  $P$  is equiprojective. Finally, the projection of  $P$  in Figure 4(a) shows that it is 10-equiprojective. ■

#### B. Equitruncated pyramid

An *equitruncated pyramid*  $P$  is obtained by cutting an arbitrary quadrilateral pyramid  $P'$ . Again, two types of cuts are applied here. See Figure 4(b). The *top cut* is parallel to the base and creates a quadrilateral at the top. For all the *side cuts*, consider the four edges that are incident to the apex of  $P$ . These four edges can be partitioned into two paths  $\mathcal{P}_1$  and  $\mathcal{P}_2$  where each path  $\mathcal{P}_i$ , for  $i = 1, 2$ , contains two edges from two different faces. Observe that the two edges of  $\mathcal{P}_i$  are coplanar and let this plane be  $\pi_i$ . For each  $i$ , apply two side cuts at the two ends of  $\mathcal{P}_i$  such that they are parallel to  $\pi_{i+1}$ . This ends the construction of  $P$ . See Figure 4(b). We get the following theorem whose proof is similar to the previous one and we omit it here.

**Theorem 5:** An equitruncated pyramid is 10-equiprojective.

#### C. Equitruncated triangular cupola

We obtain an *equitruncated triangular cupola*  $P$  from a triangular cupola  $P'$ . Before we go to the actual construction we review some properties of  $P'$ . See Figure 5(a) for two complementary views of  $P'$ . Beside the regularity of its faces (being a Johnson solid), we observe that the edges of  $P$  that are not edges of the hexagonal base can be partitioned into three similar paths, where each path contains three edges including one from the top triangle. Moreover, edges in each path are coplanar and their plane is parallel with one of the three triangles that are adjacent to the base. See Figure 5(b).

To construct  $P$  we use four different cuts *cut1*, *cut2*, *cut3*, and *cut4*. See Figure 5(b). Let  $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$  be the three paths of  $P'$  (as mentioned above). Let  $t_1, t_2, t_3$  be three triangles adjacent to the base of  $P'$  with  $t_i$ , for  $i = 1, 2, 3$ , parallel to the plane of  $\mathcal{P}_i$ . *cut1* is parallel to  $t_1$ , cuts two end edges of  $\mathcal{P}_2$  and  $\mathcal{P}_3$ , and creates a quadrilateral  $q_1$  with two edges parallel. Next, a parallel pair of *cut2*, which are perpendicular to the base and the plane of  $t_1$ , are applied to the two end edges of  $\mathcal{P}_1$ . These two cuts create two parallel triangles

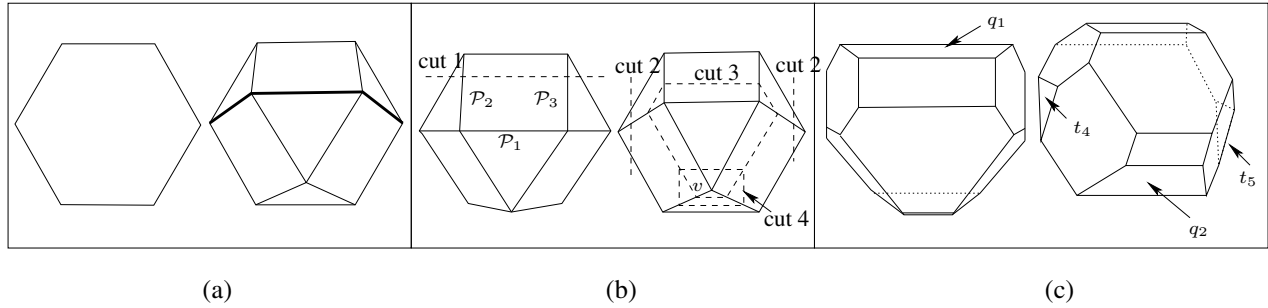


Fig. 5. (a) A triangular cupola  $P'$ . A path is shown in bold. (b) The four cuts applied to  $P'$ . (c) The equitruncated triangular cupola obtained from  $P'$ .

$t_4$  and  $t_5$ . Cut3, which is parallel to the base and the top triangle, is applied at the top and converts the top triangle into a hexagonal face. Next, a cut4 is applied to the common vertex of  $\mathcal{P}_2$  and  $\mathcal{P}_3$ , let this vertex be  $v$ , such that it creates a rectangle. Observe that since the faces around  $v$  are equilateral triangles and squares in alternate fashion and since cut3 is parallel to the top triangle, such a rectangle is guaranteed by the cut4. Cut4 also converts the triangle  $t_1$  into a quadrilateral  $q_2$  with two edges parallel.

**Theorem 6:** An equitruncated triangular cupola is 11-equiprojective.

*Proof:* In  $P$ , all faces, except the triangles  $t_4$  and  $t_5$  and the quadrilaterals  $q_1$  and  $q_2$ , are self-compensating. Edges of  $t_4$  are compensated by the edges of  $t_5$ . In each of  $q_1$  and  $q_2$ , two edges are parallel and thus compensate each other within the same face. The remaining two edges of  $q_1$  are compensated by the remaining two edges of  $q_2$ . Finally, from Figure 5,  $P$  is 11-equiprojective. ■

#### D. Equitruncated pentagonal rotunda

In this section we get two different *equitruncated pentagonal rotunda*  $P_1$  and  $P_2$ , each one from a pentagonal rotunda  $P'$ . Their construction are similar to that of an equitruncated triangular cupola. As before, let us first study some properties of  $P'$ . Figure 6(a) shows two opposite views of  $P'$ . Its base is a regular decagon and its other edges are partitioned into five paths, each of length five and each containing one edge from the top pentagon. Each path lie in a single plane. Let  $e_1, \dots, e_5$  be the ordering of the edges of the top pentagon. We denote the paths as  $\mathcal{P}_1, \dots, \mathcal{P}_5$  assuming that  $\mathcal{P}_i$  contains  $e_i$ . We call two paths  $\mathcal{P}_i$  and  $\mathcal{P}_j$  *non-adjacent* if  $e_i$  and  $e_j$  are nonadjacent. See Figure 6(b).

To construct the first equitruncated pentagonal rotunda  $P_1$ , three types of cuts are used: *cut1*, *cut2*, and *cut3*. See Figure 6(b). A cut4 is applied at each vertex of  $P'$  that is not incident to the base of  $P'$  and it is applied in such a way that it creates a rectangle. Let  $e_1$  and  $e_2$  be two parallel edges of the base decagon. The other two faces adjacent to  $e_1$  and  $e_2$  are a pentagon and a triangle, and let them be  $p$  and  $t$  respectively. A pair

of cut2 and cut3 are applied simultaneously. The cut2 is parallel to  $p$ , cuts  $t$ , and creates a new quadrilateral  $q_1$ . On the other hand, the cut3 is parallel to  $t$ , cuts  $p$ , and creates a new quadrilateral  $q_2$ . This pair of cut2 and cut3 cut the (four) end edges of two non-adjacent paths, and let them be  $\mathcal{P}_1$  and  $\mathcal{P}_3$ . We apply another pair of a cut2 and a cut3 such that they cut the end edges of  $\mathcal{P}_2$  and  $\mathcal{P}_4$ . Finally, a pair of cut1, which are perpendicular to the base and the plane of  $\mathcal{P}_5$ , are applied to the two end edges of  $\mathcal{P}_5$  and create two parallel triangles.

The second equitruncated pentagonal rotunda  $P_2$  is obtained similar to  $P_1$ , except by replacing the second pair of cut2 and cut3 by two parallel pairs of cut1 at the ends of  $\mathcal{P}_2$  and  $\mathcal{P}_3$ , which create two pairs of parallel triangles.

**Theorem 7:** The above two equitruncated pentagonal rotunda are 21- and 23-equiprojective respectively.

*Proof:* Let the two quadrilaterals created by the second pair of cut2 and cut3 be  $q'_1$  and  $q'_2$ , and let the second pair of pentagon and triangles parallel to them be  $p'$  and  $t'$  respectively. For  $P_1$ , all faces except  $p, t, p', t', q_1, q_2, q'_1$  and  $q'_2$  are self-compensating. For each of  $p, p', t$  and  $t'$ , two edges are compensated by two edges of  $q_1, q'_1, q_2$  and  $q'_2$ , respectively. All other edges of  $p, p', t, t', q_1, q'_1, q_2$  and  $q'_2$  are compensated within the same face. Proof for  $P_2$  is similar.

For the equiprojectivity of  $P_1$  to be 21, consider the direction  $d + \varepsilon$ , where  $d$  is perpendicular to the base and does not see the base and  $\varepsilon$  is perpendicular to  $t_1$ . Then for  $\varepsilon$  small enough, the projection boundary of  $P_1$  from  $d + \varepsilon$  consists of six edges from  $q_1, q_2$ , another six edges from  $q'_1, q'_2$ , two edges from  $t_2$  and seven edges from the base.

For  $P_2$ , two pairs of parallel triangles replaces  $q'_1, q'_2$ . As a whole  $P_2$  has three parallel pairs of triangles adjacent to the base. So for a projection of  $P_2$  from a direction  $d + \varepsilon + \varepsilon' + \varepsilon''$ , where  $\varepsilon, \varepsilon', \varepsilon''$  are perpendicular to the three planes of the triangles, respectively, the projection boundary consists of six edges from  $q_1, q_2$ , six edges from the triangles, and eleven edges from the base. ■

At this moment an interesting observation is that if

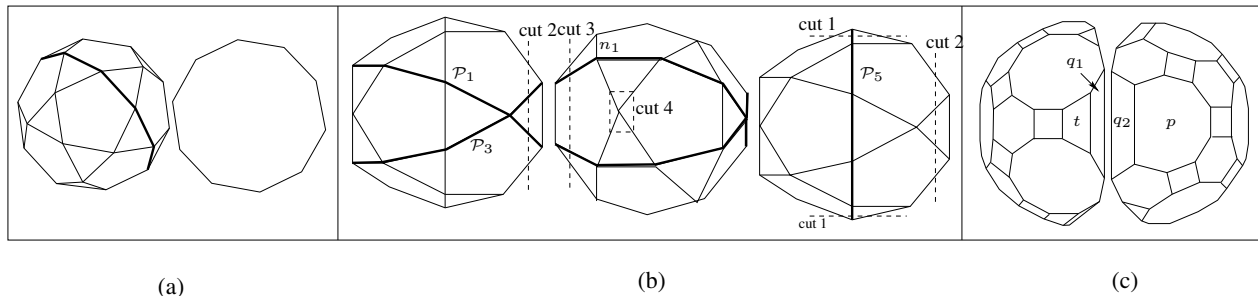


Fig. 6. Obtaining an equitruncated pentagonal rotunda  $P_1$  from a pentagonal rotunda  $P'$ . (a) Two opposite views of  $P'$ . (b) Cuts used in  $P'$  to obtain  $P_1$ . (c) Two complementary views of  $P_1$ .

we apply five pairs of cut1 at the ends of five paths (and thus no cut2 or cut3), then we get a polyhedron which is the half of the *truncated icosahedron*. This polyhedron is a zonohedron and is 25— equiprojective.

## V. CONCLUSION

In this paper we give lower bounds on the value of  $k$  and the size of an  $k$ -equiprojective polyhedron. We prove that there is no 3- or 4-equiprojective polyhedra and a triangular prism is the only 5-equiprojective polyhedron. We also discover some new equiprojective polyhedra.

The problem of constructing all equiprojective polyhedra still remains open. However, we observe that our concept of partitioning the non-base edges into disjoint paths is somewhat similar for all the new equiprojective polyhedra, which might be helpful for deriving a generalized algorithm for constructing equiprojective polyhedra.

## ACKNOWLEDGMENT

We thank Therese Biedl for discussions on the minimality proofs. We thank Anna Lubiw and Craig Kaplan for numerous discussions.

## REFERENCES

- [1] H. S. M. Coxeter. *Regular Polytopes*. Macmillan, New York, 1963.
- [2] C. Croft, K. Falconer, and R. Guy. *Unsolved Problems in Geometry*. Springer-Verlag, New York, 1991.
- [3] P. R. Cromwell. *Polyhedra*. Cambridge University Press, Cambridge, 1997.
- [4] D. Eppstein. The geometry junkyard: Zonohedra. <http://www.ics.uci.edu/~eppstein/junkyard/zono.html>.
- [5] D. Eppstein. Zonohedra and zonotopes. *Mathematica in Education and Research*, 5(4):15–21, 1996. <http://www.ics.uci.edu/~eppstein/pubs/Epp-TR-95-53.pdf>.
- [6] G. W. Hart. Encyclopedia of polyhedra. <http://www.georgehart.com/virtual-polyhedra/vp.html>.
- [7] M. Hasan. *Reconstruction and Visualization of Polyhedra Using Projections*. PhD thesis, School of Computer Science, University of Waterloo, Waterloo, Canada, 2005.
- [8] M. Hasan and M. M. Hossain. Some new equiprojective polyhedra from a tetrahedron, a quadrilateral pyramid and some johnson solids, 2008. Manuscript in preparation.
- [9] M. Hasan, M. M. Hossain, and S. Nusrat. VRML code for some new equiprojective polyhedra, 2008. [http://www.teacher.buet.ac.bd/masudhasan/vrml\\_code.html](http://www.teacher.buet.ac.bd/masudhasan/vrml_code.html).
- [10] M. Hasan and A. Lubiw. Equiprojective polyhedra. In *Proc. 15th Canadian Conference on Computational Geometry*, pages 47–50, Halifax, Canada, August 2003.
- [11] M. Hasan and A. Lubiw. Equiprojective polyhedra. *Computational Geometry: Theory and Application*, 40(2), 2008.
- [12] N. Rahman, S. A. Quader, and M. Hasan. Some new equiprojective polyhedra. In *Proc. International Conference on Computer and Information Technology (ICCIT)*, Dhaka, Bangladesh, December 2006.
- [13] G. Shephard. Twenty problems on convex polyhedra—II. *Math. Gaz.*, 52:359–367, 1968.
- [14] J. Taylor. Zonohedra and generalized zonohedra. *American Mathematical Monthly*, 99(2):108–111, 1992.

# A comprehensive Analysis of Degree Based Condition for Hamiltonian Cycles

Md. Kamrul Hasan<sup>1</sup>, Mohammad Kaykobad<sup>2</sup>, Young-Koo Lee<sup>1</sup>, Sungyoung Lee<sup>1</sup>

<sup>1</sup>Ubiquitous Computing Lab, Department of Computer Engineering, Kyung Hee University, South Korea

<sup>2</sup>Department of Computer Science & Engineering, Bangladesh University of Engineering and Technology, Bangladesh

Email: [hasaniut@yahoo.com](mailto:hasaniut@yahoo.com), [kaykobad@cse.buet.ac.bd](mailto:kaykobad@cse.buet.ac.bd), [yklee@khu.ac.kr](mailto:yklee@khu.ac.kr), [sylee@oslab.khu.ac.kr](mailto:sylee@oslab.khu.ac.kr)

**Abstract** — Rahman and Kaykobad introduced a shortest distance based condition for finding the existence of Hamiltonian paths in graphs as follows: Let  $G$  be a connected graph with  $n$  vertices, and if  $d(u) + d(v) + \delta(u, v) \geq n + 1$ , for each pair of distinct non-adjacent vertices  $u$  and  $v$  in  $G$ , where  $\delta(u, v)$  is the length of a shortest path between  $u$  and  $v$ , then  $G$  has Hamiltonian path. Rao Li proved that under the same condition, the graph is Hamiltonian or belongs to two different classes of graphs. Recently, Mehedy, Hasan and Kaykobad showed case by case that under the condition of Rahman and Kaykobad, the graph is Hamiltonian with exceptions for  $\delta(u, v) = 2$ . Shengjia Li et. al. mentions a graph to be Hamiltonian whenever  $d(u) + d(v) \geq n - 1$ , for all  $\delta(u, v) = 2$ , otherwise  $n$  is odd and the graph falls into a special class. This paper relates the results of Mehedy, Hasan and Kaykobad with the two exceptional classes of graphs introduced by Rao Li and the graph class introduced by Shengjia Li et. al. The paper also provides a thorough analysis of the graph classes and shows the characteristics of a graph when it falls into one of those classes.

**Index Terms** — Hamiltonian cycle, Hamiltonian path, Graphs.

## I. INTRODUCTION

We consider only simple undirected graphs- graphs that do not contain loops or multiple edges. A Hamiltonian cycle is a closed path passing through every vertex of a graph. A graph containing a Hamiltonian cycle is said to be simply Hamiltonian. By Hamiltonicity we mean the virtue of a graph to be Hamiltonian. Naturally every Hamiltonian graph contains a Hamiltonian path and not necessarily vice versa. We define  $\delta(u, v)$  as the shortest distance between  $u$  and  $v$ ,  $K_p$  as the complete graph with  $p$  vertices,  $K_p^c$  complement of  $K_p$  (a set of  $p$  independent vertices),  $K_{p,q}$  as complete bipartite graph with  $p + q$

vertices. ‘ $\vee$ ’ denotes join operator. We term the condition: “ $d(u) + d(v) + \delta(u, v) \geq n + 1$ , where  $|V| = n$  and  $u, v$  are distinct non-adjacent vertices of  $G$ ”, as Rahman and Kaykobad condition.

We also describe three families of graph  $C_n$ ,  $D_n$  and  $L_n$ , defined in [4],[4],[5] respectively.

$C_n$  is defined as  $\{G : V(G) = V(K_1) \cup V(G_1) \cup V(G_2), E(G) = E(G_1) \cup E(G_2) \cup \{ab : a \in V(K_1) \text{ and } b \in V(K_{p_1}) \cup V(K_{p_2})\}\}$ , where  $K_1$  consisting of one single vertex,  $G_1 = K_{p_1} \vee K_{q_1}$  is a complete graph with  $p_1 + q_1$  vertices such that  $p_1 \geq 1$  and  $q_1 \geq 0$ ,  $G_2 = K_{p_2} \vee K_{q_2}$  is a complete graph with  $p_2 + q_2$  vertices such that  $p_2 \geq 1$  and  $q_2 \geq 0$  and  $V(K_1), V(G_1)$  and  $V(G_2)$  are pair wise disjoint

$D_n$  is defined as  $\{G : K_{p,p+1} \subseteq G \subseteq K_p \vee (p+1)K_1\}$ , where  $|V(G)| = 2p+1 \geq 3$ , where  $(p+1)K_1$  denotes  $(p+1)$  isolated vertices.

$L_n$  is defined as  $L_n = L_{2m+1} = \{Z_m \vee (K_m^c + \{u\}) \mid Z_m \text{ is a graph with } m \text{ vertices}\}$ , where  $K_m^c$  is a set of  $m$  independent vertices (also as the complement of the complete graph  $K_m$ ) and  $u$  is another single vertex, furthermore, the edge set of  $L_n$  consists of  $E(Z_m)$  and  $xy \mid x \in V(Z_m) \text{ and } y \in K_m^c \cup \{u\}$

It is to be noted that  $D_n$  and  $L_n$  are actually the same class of graphs with different names. We adapt the notation  $D_n$  in our results to represent both the classes.

## II. EXISTING RESULTS

To prove our results, we need some existing theorems. We list them in this section for the same of completeness of this article.

**Theorem 1.1** (Rahman and Kaykobad [3]). *Let  $G = (V, E)$  be a connected graph with  $n$  vertices such that for all pairs of distinct non-adjacent vertices  $u, v \in V$ , we have  $d(u) + d(v) + \delta(u, v) \geq n + 1$ . Then  $G$  has a Hamiltonian path.*

**Theorem 1.2** (Rao Li [4]). *Let  $G = (V, E)$  be a connected graph with  $n$  vertices. If  $d(u) + d(v) + \delta(u, v) \geq n + 1$  for each pair of distinct non-adjacent vertices  $u, v \in V$ , then  $G$  is Hamiltonian or  $G \in C_n \cup D_n$ .*

**Theorem 1.3** (Mehedy, Hasan, and Kaykobad [1]). *Let  $G = (V, E)$  be a 2-connected graph with  $n$  vertices such that for all pairs of distinct non-adjacent vertices  $u, v \in V$  we have  $d(u) + d(v) + \delta(u, v) \geq n + 1$  then*

- i)  $\delta(u, v) \leq 3$ , for all  $u, v \in V$
- ii) If for any  $u, v \in V$   $\delta(u, v) = 3$ ,  $G$  is Hamiltonian
- iii) If  $\delta(u, v) = 2$  for all non-adjacent vertices,  $G$  may or may not be Hamiltonian. To make the graph Hamiltonian, it needs only one extra edge added to the graph.

Theorem 1.3 is cumulative results of the work [1], to help our analysis.

**Theorem 1.4** (Shengjia Li et. al.[5]). *Let  $G$  be a 2-connected graph with  $n \geq 3$  vertices. If  $d(u) + d(v) \geq n - 1$  for every pair of vertices  $u$  and  $v$  with  $d(u, v) = 2$ , then  $G$  is Hamiltonian, unless  $n$  is odd and  $G \in L_n$ .*

**Theorem 1.5** (Ore [2]). *If  $d(u) + d(v) \geq n$  for every pair of distinct non-adjacent vertices  $u$  and  $v$  of  $G$ , then  $G$  is Hamiltonian.*

We take Theorem 1.3 as the base and fulfill the missing results in the theorem by using other theorems above. We also provide a through analysis of the graph classes and show the characteristics of a graph when it falls into one of those classes.

## III. MAIN RESULTS

This paper relates the results of Mehedy, Hasan and Kaykobad [1] with the two exceptional classes of graphs introduced by Rao Li [4] and the graph class introduced by Shengjia Li et. al. [5] and brings the results of all three works under Rahman and Kaykobad condition. We analyze when a graph falls in the two classes  $C_n$  and  $D_n$  and what the characteristics of a graph instance are. We provide a unified theorem as follows:

**Theorem 2.1.** *Let  $G = (V, E)$  be a **connected** graph with  $n$  vertices such that for all pairs of distinct non-adjacent vertices  $u, v \in V$  we have  $d(u) + d(v) + \delta(u, v) \geq n + 1$  then*

- i)  $\delta(u, v) \leq 4$ , for all  $u, v \in V$
- ii) If there are vertices  $u, v \in V$  with  $\delta(u, v) = 4$ , then  $G \in C_n$
- iii) If there are vertices  $u, v \in V$  with  $\delta(u, v) = 3$  and no other vertices  $u', v' \in V$  with  $\delta(u', v') > 3$ , then  $G$  is Hamiltonian if 2-connected or else  $G \in C_n$
- iv) If  $\delta(u, v) = 2$  for all non-adjacent vertices,  $G$  is Hamiltonian or  $G \in D_n$ , given  $n = \text{odd}$ . If  $G$  is not Hamiltonian, then adding just one extra edge makes the graph Hamiltonian.

**Proof:** Theorem 1.3 along with Lemma 2.1, Lemma 2.2, Lemma 2.3 and Lemma 2.4 proves the theorem 2.1.

**Lemma 2.1.** *If  $d(u) + d(v) + \delta(u, v) \geq n + 1$ , then for all pair of distinct non-adjacent vertices  $u, v \in V$ ,  $\delta(u, v) \leq 4$*

**Lemma 2.2.** *If  $d(u) + d(v) + \delta(u, v) \geq n + 1$ , and there are vertices  $u, v \in V$  with  $\delta(u, v) = 4$ , then  $G \in C_n$*

**Lemma 2.3.** *If  $d(u) + d(v) + \delta(u, v) \geq n + 1$  and for any  $u, v \in V$   $\delta(u, v) = 3$  and no other vertices  $u', v' \in V$  with  $\delta(u', v') > 3$  and  $G$  is not 2-connected, then  $G \in C_n$*

**Lemma 2.4.** *If  $d(u) + d(v) + \delta(u, v) \geq n + 1$  and for every non-adjacent pair  $u, v \in V$   $\delta(u, v) = 2$  and  $n = \text{odd}$ , then  $G \in D_n$ .*

**Proof of Lemma 2.1:** Theorem 1.3[1] proves that under Rahman and Kaykobad condition  $\delta(u, v) \leq 3$  when  $G$  is 2-connected. On the other hand Theorem 1.2[4] proves

that under Rahman and Kaykobad condition  $G \in C_n \cup D_n$  where  $G$  is 2-connected or just connected.

Now lets find out the maximum possible  $\delta(u, v)$  in  $C_n$  and  $D_n$ . In  $C_n$  maximum  $\delta(u, v) = 4$  is possible only when  $u \in K_{q1}$  and  $v \in K_{q2}$ . In  $D_n$  maximum  $\delta(u, v) = 2$ .

So, under Rahman and Kaykobad condition,  $\delta(u, v) \leq 3$ , when  $G$  is 2-connected and  $\delta(u, v) = 4$ , when  $G$  is not two connected. In general,  $\delta(u, v) = 4$ , under Rahman and Kaykobad condition.

**Proof of Lemma 2.2:** In the proof of Lemma 2.1 section we discussed that under Rahman and Kaykobad condition,  $\delta(u, v) = 4$ , only when  $G$  is not 2-connected. If a graph is not two connected, it cannot be Hamiltonian. So, according to Theorem 1.2[4], the graph falls into  $C_n$  or  $D_n$ . But as  $\delta(u, v) = 4$ , the graph cannot fall in  $D_n$  and hence falls into the category  $C_n$ .

**Proof of Lemma 2.3:** We prove the lemma using the same arguments Rao Li [4] used to prove his theorem.

Let  $P: = u s_1 s_2 v$  be the shortest path between  $u$  and  $v$  with length 3. So there exists at least a single pair of distinct non-adjacent vertices in  $G$  (example  $u$  and  $s_2$  or  $s_1$  and  $v$ ) such that the shortest distance between them is 2.

Since  $G$  is not 2-connected (has cut vertex), there exist a vertex say,  $x$  in  $G$  such that  $G[V(G) \setminus \{x\}]$  is disconnected. Let  $G_i$  be the components in  $G[V(G) \setminus \{x\}]$  and the number of vertices in  $G_i$  is  $n_i$ , where  $2 \leq l$  and  $1 \leq i \leq l$ .

So, there must be two vertices in two different components  $u_i \in V(G_i)$ , such that  $u_i x \in E(G)$ , where  $i=1, 2$  and  $\delta(u_1 u_2) = 2$ . Now,

$$n+1 \leq d(u_1) + d(u_2) + 2 \leq n_1 + n_2 + 2 \\ \leq n_1 + n_2 + n_3 + \dots + n_l + 2 = n+1$$

So,  $d(u_1)=n_1$ ,  $d(u_2)=n_2$  and  $n_3=n_4=\dots=n_l=0$ . So, there will be only two components and  $u_i$  will be connected to all the vertices in  $V(G_i) \cup \{x\}$ ,  $i=1, 2$  and  $n=n_1+n_2+1$ . Now let's, define the four set of graphs present in  $C_n$ :

$$U_1 = \{u \in V(G_1) : ux \in E\}, |U_1| \neq 0$$

$$V_1 = \{v \in V(G_1) : vx \notin E\}, |V_1| \geq 0$$

$$U_2 = \{u \in V(G_2) : ux \in E\}, |U_2| \neq 0$$

$$V_2 = \{v \in V(G_2) : vx \notin E\}, |V_2| \geq 0$$

Any vertex  $u'$  other than  $u_1$  in the set  $U_1$  will show  $\delta(u', u_2)=2$  and  $d(u') = n_1$ . So  $u'$  is adjacent to all the vertices in  $V(G_1) \cup \{x\}$ . Symmetrically, it can be

shown that any vertex  $u''$  in  $U_2$  is adjacent to all vertices in  $V(G_2) \cup \{x\}$ .

Let  $v'$  be any vertex in  $V_1$  and so  $\delta(v', u_2) = 3$ .  $n+1 \leq d(v') + d(u_2) + 3 \leq n_1 - 1 + n_2 + 3 = n+1$ , which indicates  $d(v') = n_1 - 1$  and  $v'$  is exactly adjacent to all the vertices in  $V(G_1)$ . Similarly any vertex  $v''$  in  $V_2$  is adjacent to all the vertices in  $V(G_2)$ . Hence  $G \in C_n$ .

The lemma can be proved by the discussion in Lemma 2.1 also. As  $G$  is not 2-connected, it is not Hamiltonian. So, according to Theorem 1.2[4],  $G \in C_n \cup D_n$ . But as maximum  $\delta(u, v) = 2$  in  $D_n$ ,  $G$  is actually of class  $C_n$ .

Now we analyze the characteristics of  $G \in C_n$ . For any non-adjacent vertex pair  $u, v \in E(G)$ ,  $\delta(u, v) = 3$  implies that no vertex  $w \in V$  can be adjacent to  $u, v$  at the same time since then  $\langle u \rightarrow w \rightarrow v \rangle$  would have been a path of length 2.  $\delta(u, v) = 3$ , also implies  $d(u) + d(v) \geq n - 2$ . As,  $u$  and  $v$  cannot have a common neighbor, and the number of vertices that can be connected is  $(n-2)$ . Hence, by contradiction,  $d(u) + d(v) = n - 2$  is proved. So,  $u, v$  connects to all the vertices in the graph, having no common neighbor.

As we assume 1-connected graph (Fig.1.), there should be no cross over edges like  $(1, i+1), (i, n) \in E$ . Because cross over edge will make the graph 2-connected, and also Hamiltonian. In Fig.1, we see that there is two cut vertices  $k$  and  $(k+1)$ .

We know that under Rahman and Kaykobad condition  $\delta(u, v) \leq 3$  for all  $u, v \in V$ . To maintain this, vertices from  $u=1$  to  $(k-1)$ , all should be connected to  $k$ , so than  $\delta(i, v=n) = 3$ ,  $i \in \{1 \text{ to } k-1\}$ . Similarly vertices from  $(k+2)$  to  $v=n$  should be connected to  $(k+1)$  to maintain  $\delta(i, v=1) = 3$ ,  $i \in \{k+2 \text{ to } n\}$ . Now, the non adjacent vertices on the sides  $\{1 \text{ to } k\}$  and  $\{k+1, \text{ to } n\}$ , have a distance 2. So, any two non adjacent vertices  $\{u', v' \in (1 \text{ to } k-1) \text{ or } (k+2 \text{ to } n)\}$  should maintain  $d(u') + d(v') \geq n - 1$ . But maximum possible  $d(u') + d(v') = n - 2$ , only when  $k=2$  or  $k=n-1$ . Hence to maintain Rahman Kaykobad condition, the graph component composed of vertices  $\{1 \text{ to } k\}$  will be complete and similarly the component with vertices  $\{k+1 \text{ to } n\}$  is complete. So, the graph falls in  $C_n$ .

We define this class of graph as  $C_n' \subset C_n$ , where  $\{V(K_{p1}) = \{k\}, V(K_{q1}) = \{1 \text{ to } k-1\}\}$ ,

$V(K_1) = \{k+1\}$  ,  $V(K_{p_2}) = \{k+1 \text{ to } n\}$  ,  
 $V(K_{q_2}) = \{\emptyset\}$   
and  $k \in \{2 \text{ to } n-1\}$

**Proof of Lemma 2.4:** It is to be noted that Rahman and Kaykobad condition fulfills the condition given in theorem

satisfying the above condition is:  $|V(Z_p)|=p$  and  $|(p+1)K_1|=p+1$ ,  $n=2p+1$

In this setup, adding just one edge in  $(p+1) K_1$  makes the graph Hamiltonian, which is a finding of work[1]. It is to be noted that adding that edge reduces the independent set size by two, and becomes  $p-1$  whereas  $|V(Z_p)|$  becomes  $p+2$  ( $|(p+1)K_1| \leq |V(Z_p)|$ ), which implies Hamiltonicity in the graph.

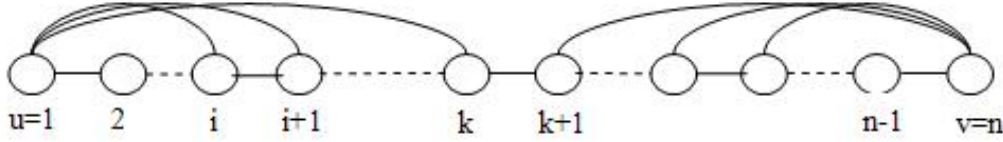


Fig.1. An 1-connected graph G fulfilling Rahman and Kaykobad condition, with  $\delta(u, v) = 3$

1.4 [5]. In Rahman and Kaykobad, for any pair  $u, v \in V$  if  $\delta(u, v) = 2$ ,  $d(u) + d(v) \geq n-1$ . If for all non-adjacent pair  $u, v \in V$   $\delta(u, v) = 2$ , Rahman and Kaykobad condition ensures 2-connectedness. So, the Lemma can directly be derived from Theorem 1.4.

Part (iii) of Theorem 1.3 [1] misses one fact that G is Hamiltonian when  $n$ =even. Lets see why  $n$ =even makes the graph Hamiltonian. We define a graph class  $D_n$   $\{G : K_{p,q} \subseteq G \subseteq K_p \vee qK_1$ , where  $|V(G)| = \text{even}\}$ .  $D_n$  is similar to class  $D_n$ , with the exceptions that  $n$ =even and hence  $q \neq p+1$ . Note that any graph G with  $|V(G)| = \text{even}$ , falls into this class.

To keep  $\delta(u, v) = 2$  for all non-adjacent pair  $u, v \in V$ , the graph  $G \in D_n$  should provide an independent set  $\{V(qK_1)\}$ , to which all other vertices  $\{V(Z_p \subseteq K_p, |V(Z_p)| = p)\}$  should be connected. The independent set should be  $> n/2$  to avoid Hamiltonicity, but doing so breaks the Rahman Kaykobad conditions (one can verify the fact easily by drawing a graph). So  $q \leq p$ , which eventually makes the graph Hamiltonian:  $\{u \in V(Z_p) s_1 \in V(qK_1) s_2 \in Z_p s_3 \in V(qK_1) \dots v \in Z_p, a \in V(qK_1) u, \delta(u, v) = 2\}$ .

Now for  $n$ =odd, the number of independent vertices should be kept larger than the number of vertices in  $Z_p$  and at the same time the Rahman and Kaykobad condition should be satisfied. The only possible configuration

**Corollary 2.1.**  $C_n$  and  $D_n$  have Hamilton-path.

Graph classes  $C_n$  and  $D_n$  fulfills the Rahman and Kaykobad condition but are not Hamiltonian. So, according to Theorem 1.1 [3], they both have Hamilton-path.

**Corollary 2.2.** Under the condition of Rahman and Kaykobad, if every non-adjacent  $u, v \in V$   $\delta(u, v) = 2$ , then the graph is Hamiltonian or the size of the independent set is  $\left\lfloor \frac{n}{2} \right\rfloor$

**Proof:** To avoid Hamiltonicity The size of the independent set should be  $> n/2$ . However, doing so breaks the Rahman Kaykobad conditions, when  $n$ =even. So, the graph is Hamiltonian when  $n$ =even. This is also proven in Theorem 1.4[5] in a different way.

When  $n$ =odd, it is possible to maintain Rahman and Kaykobad condition even making size of

the independent set  $> n/2$  only when the independent set size is  $\frac{n+1}{2} = \left\lceil \frac{n}{2} \right\rceil$ . This can also be obtained by

examining the graph family  $D_n$  found in [4],[5].

Hence the corollary is proved. We also generalize two corollaries derived by rao Li[4].

**Corollary 1.1** (Rao Li[4]). Let G be a 2-connected graph with n vertices. If  $d(u) + d(v) + \delta(u, v) \geq n+2$  for each pair of distinct non-adjacent vertices u and v in G, then G is pancyclic or  $G \in \{K_{n/2, n/2} : \text{where } n \geq 4 \text{ is even}\}$ .

**Corollary 2.3.** Given the condition in Corollary 1.1, G is Hamiltonian.



**Proof:** If for all  $u, v \in V$   $\delta(u, v) = 2$ , the condition given reduces to  $d(u) + d(v) \geq n$  which is actually Ore's condition [2], implying Hamiltonicity (see theorem 1.5). If for any  $u, v \in V$   $\delta(u, v) = 3$ , the condition becomes  $d(u) + d(v) \geq n - 1$ . However, Mehedy, Hasan and Kaykobad [1] (see theorem 1.3) proved that for  $\delta(u, v) = 3$ , with  $d(u) + d(v) \geq n + 1 - 3 = n - 2$ , the graph becomes Hamiltonian. It is also proved in theorem 1.3 [1] that  $\delta(u, v)$  cannot be greater than 3 for  $d(u) + d(v) + \delta(u, v) \geq n + 1$ .  $d(u) + d(v) + \delta(u, v) \geq n + 2$  is a relaxed condition allowing more edges to be present in the graph and hence the finding  $\delta(u, v) \leq 3$ , is also applicable here. So, we say that Corollary 1.1 can be more generalized and the condition given in Corollary 1.1 proves the graph to be Hamiltonian.

**Corollary 1.2** (Rao Li[4]). *Let  $G$  be a 3-connected graph with  $n$  vertices. If  $d(u) + d(v) + \delta(u, v) \geq n + 3$  for each pair of distinct non-adjacent vertices  $u$  and  $v$  in  $G$ , then  $G$  is Hamilton-connected.*

**Corollary 2.4.** *Given the condition in Corollary 1.2,  $G$  is Hamiltonian.*

**Proof:** The same arguments provided for Corollary 2.3 applies here also. It is to be noted that for  $\delta(u, v) = 2$ , the condition is:  $d(u) + d(v) \geq n + 1$ , which is more relaxed than Ore's condition[2]. We also argue that  $\delta(u, v) \leq 3$  and for  $\delta(u, v) = 3$ , the graph is Hamiltonian by using theorem 1.3 [1] and because the graph is 3-connected (2-connected graph is more constrained by number of edges than 3 connected). So, the graph is Hamiltonian.

## IV. CONCLUSION

We relate three seminal works for finding the Hamiltonicity of a graph observing Rahman and Kaykobad condition or a subset of the condition. We determine when the graph is Hamiltonian. If we cannot, we characterize the non-Hamiltonian graph by using two families of graphs and which instances of graphs in those two families, the graph falls in. Our work provides a comprehensive understanding of graphs following Rahman and Kaykobad condition.

## REFERENCES

- [1] Lenin Mehedy, Md. Kamrul Hasan, Mohammad Kaykobad, "An improved degree based condition for Hamiltonian cycles, *Information Processing Letters*", 102 (2007) 108-112.
- [2] Ore O., "Note On Hamiltonian Circuits", *Amer. Math. Monthly* 67(1960), 55.
- [3] Rahman M. Sohel, Kaykobad M., "On Hamiltonian Cycles And Hamiltonian Paths", *Information Processing Letters* 94(2005), 37-51.
- [4] Rao Li, "A New Sufficient Condition for Hamiltonicity of Graphs", *Information Processing Letters, Volume 98, Issue 4, 31 May 2006, Pages 159-161*.
- [5] Shengjia Li, Ruijuan Liand Jinfeng Feng "An efficient condition for a graph to be Hamiltonian", *Discrete Applied Mathematics*, Volume 155, Issue 14, 1 September 2007, Pages 1842-1845.

# Design of a Fractional-order Self-tuning Regulator using Optimization Algorithms

Deepyaman Maiti, Mithun Chakraborty, Ayan Acharya, and Amit Konar

Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata, India

**Abstract** — The self-tuning regulators form an important sub-class of adaptive controllers. This paper introduces a novel scheme for designing a fractional order self-tuning regulator. Original designs for all the sub-modules of the self-tuning regulator are proposed. The particle swarm optimization algorithm is utilized for online identification of the parameters of the dynamic fractional order process while the subsequent tuning of the controller parameters is performed by differential evolution. Results show that the proposed self-tuning regulator is both precise and robust.

**Index Terms** — Controller tuner, differential evolution, fractional order self-tuning regulator, parameter identifier, particle swarm optimization.

## I. INTRODUCTION

Development of a control system involves many tasks such as modeling, design of a control law, implementation and validation. The self-tuning regulator (STR) attempts to automate several of these tasks. This is illustrated in Fig. 1, which shows the block diagram of a process with an STR. It is assumed that the structure of the process is specified.

Parameters of the process are estimated on-line by the block labeled ‘parameter identifier’. The block labeled ‘controller tuner’ contains computations that are required to perform the design of a controller with a specified method and a few design parameters that can be chosen externally. The design problem is called the *underlying design problem* for systems with known parameters. The block labeled ‘controller’ is an implementation of the controller whose parameters are obtained from the ‘controller tuner’ block [1].

In this paper we have presented a scheme for the design of a fractional order STR. This means that the physical process that needs to be controlled as well as the controller that will control it are both of fractional order. The real world objects or processes that we want to estimate are generally of fractional order [2].

A typical example of a non-integer (fractional) order system is the voltage-current relation of a semi-infinite lossy RC line or diffusion of heat into a semi-infinite solid, where heat flow  $q(t)$  is equal to the half-derivative of temperature  $T(t)$ :  $\frac{d^{0.5}T(t)}{dt^{0.5}} = q(t)$ .

So far, however, the usual practice when dealing with a fractional order process has been to use an integer order approximation. Disregarding the fractional order of the

system was caused mainly by the non-existence of simple mathematical tools for the description of such systems. Since major advances have been made in this area recently, it is possible to consider also the real order of the dynamical systems. Such models are more adequate for the description of dynamical systems with distributed parameters than integer-order models with concentrated parameters.

It is necessary to understand the theory of fractional calculus in order to realize the significance of fractional order integration and derivation. Fractional calculus is the branch of calculus that generalizes the derivative or integral of a function to non-integer order, allowing calculations such as deriving a function to  $1/2$  order. For instance  $s^\alpha$  indicates derivation to the order  $\alpha$ . Knowledge in the subject of fractional calculus is essential to design fractional order controllers. Of the several definitions of fractional derivatives, the Grunwald-Letnikov and Riemann-Liouville definitions are the most used [3]. These definitions are required for the realization of discrete control algorithms.

The design of an STR can be divided into sub-tasks such as designing the modules for parameter estimation [4] and controller design [5]. For a chosen structure of a fractional order system, we have designed the ‘parameter identifier’ and ‘controller tuner’ blocks of Fig. 1. As a matter of fact, the controller we have considered is of a fractional order PID type, which differs from the usual integral order PID controller by the property that the orders of integration and derivation are positive real rather than restricted to only positive integers (conventionally unity). Fractional order PID controllers are much superior to their integer order counterparts, especially when controlling fractional order processes [6].

For designing the ‘parameter identifier’ block, we have made use of a stochastic optimization strategy from the family of evolutionary computation, namely particle swarm optimization (PSO) [7] – [8]. The ‘controller tuner’ block utilizes differential evolution (DE) [9] – [10].

Verification of the precision of our design is performed by operating the STR on a system with known parameters, and also by simulation. The robustness of our design is also displayed.

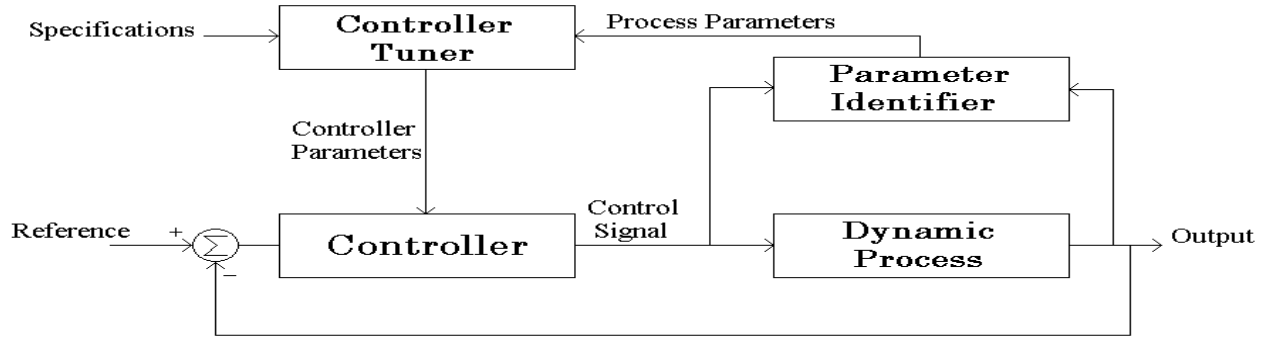


Fig. 1. Block diagram of a self-tuning regulator

## II. STOCHASTIC OPTIMIZATION ALGORITHMS

### A. The Optimization Problem

The optimization problem consists in determining the global optimum (in our case, minimum) of a continuous real-valued function of  $n$  independent variables  $x_1, x_2, x_3, \dots, x_n$ , mathematically represented as  $f(\vec{X})$ , where

$\vec{X} = (x_1, x_2, x_3, \dots, x_n)$  is called the parameter vector. Then the task of any optimization algorithm reduces to searching the  $n$ -dimensional hyperspace to locate a particular point with position-vector  $\vec{X}_D$  such that  $f(\vec{X}_D)$  is the global optimum of  $f(\vec{X})$ .

### B. Particle Swarm Optimization

The PSO algorithm [7] - [8] attempts to mimic the natural process of group communication of individual knowledge, which occurs when a social swarm elements flock, migrate, forage, etc. in order to achieve some optimum property such as configuration or location.

The 'swarm' is initialized with a population of random solutions. Each particle in the swarm is a different possible set of the unknown parameters to be optimized. Representing a point in the solution space, each particle adjusts its flying toward a potential area according to its own flying experience and shares social information among particles. The particles swarm toward the best fitting solution encountered in previous iterations with the intent of encountering better solutions through the course of the process and eventually converging on a single minimum error solution.

Let the swarm consist of  $N$  particles moving around in a  $D$ -dimensional search space. Each particle is initialized with a random position and a random velocity. Each

particle modifies its flying based on its own and companions' experience at every iteration. The  $i^{\text{th}}$  particle is denoted by  $X_i$ , where  $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ . Its best previous solution (pbest) is represented as  $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ . Current velocity (position changing rate) is described by  $V_i$ , where  $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ . Finally, the best solution achieved so far by the whole swarm (gbest) is represented as  $P_g = (p_{g1}, p_{g2}, \dots, p_{gD})$ . The fitness function evaluates the performance of particles to determine whether the best fitting solution is achieved. The particles are manipulated according to the following equations:

$$v_{id}(t+1) = \omega v_{id}(t) + c_1 \cdot \phi_1 \cdot (p_{id}(t) - x_{id}(t)) +$$

$$c_2 \cdot \phi_2 \cdot (p_{gd}(t) - x_{id}(t))$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1).$$

(The equations are presented for the  $d^{\text{th}}$  dimension of the position and velocity of the  $i^{\text{th}}$  particle.)

Here,  $c_1$  and  $c_2$  are two positive constants, called cognitive learning rate and social learning rate respectively,  $\phi_1$  and  $\phi_2$  are two random functions in the range  $[0,1]$ ,  $\omega$  is the time-decreasing inertia factor which balances the global wide-range exploitation and the nearby exploration abilities of the swarm.

### C. Differential Evolution

DE [9] - [10] belongs to the class of evolutionary algorithms guided by the principles of *Darwinian Evolution* and *Natural Genetics* where each time-varying parameter vector (candidate solution) in the population is called a chromosome and each time-step represents a generation. The first step of the algorithm, as usual, is:

*Initialization*: This step is identical to the random initialization of position vectors in PSO.

Each iteration consists of the following three steps:

*Mutation*: For each chromosome  $\vec{X}_i(t)$  belonging to the current generation, three other chromosomes  $\vec{X}_p(t)$ ,  $\vec{X}_q(t)$  and  $\vec{X}_r(t)$  are randomly selected from the same

generation (i, p, q and r are distinct); the scaled difference of  $\vec{X}_q(t)$  and  $\vec{X}_r(t)$  is added to  $\vec{X}_p(t)$  to generate a donor vector  $\vec{V}_i(t+1) = \vec{X}_p(t) + F \times (\vec{X}_q(t) - \vec{X}_r(t))$  where F is a constant scalar in (0,1). We take  $F = 0.8$ .

**Recombination:** A trial offspring vector  $\vec{T}_i(t+1)$  is created for each current-generation parent vector  $\vec{X}_i(t)$  by first choosing a constant CR ( $0 < CR < 1$ ) called the crossover constant and then setting the jth component  $\vec{T}_{i,j}(t+1)$  of  $\vec{T}_i(t+1)$  according to:

$$T_{i,j}(t+1) = \begin{cases} V_{i,j}(t) & \text{if } \text{rand}_j(0,1) < CR, \\ X_{i,j}(t) & \text{otherwise,} \end{cases}$$

where  $\text{rand}_j(0,1)$  is a random number selected from the interval (0, 1),  $j = 1, 2, \dots, n$ . We take  $CR = 0.96$ .

**Selection:** This step is be mathematically expressed as:

$$\vec{X}_i(t+1) = \begin{cases} \vec{T}_i(t+1) & \text{if } f(\vec{T}_i(t+1)) < f(\vec{X}_i(t)), \\ \vec{X}_i(t) & \text{otherwise,} \end{cases}$$

$\forall i = 1, 2, \dots, m$ .

Thus, the next-generation population is generated, keeping the population-size m always unchanged.

### III. DETAILED DESIGN SCHEME OF THE PARAMETER IDENTIFIER BLOCK

We have considered a fractional process whose transfer function is of the form  $\frac{1}{a_1 s^\alpha + a_2 s^\beta + a_3}$ . It should be noted that without loss of generality, we may presume the dc gain to be unity so that the dc gain and its possible fluctuations are included in the coefficients  $a_1$ ,  $a_2$  and  $a_3$ . This fractional order system has five varying parameters, namely three coefficients  $a_1$ ,  $a_2$  and  $a_3$ , and two fractional powers  $\alpha$  and  $\beta$ . (Of course any other structure can be chosen due to the generalness of our design method. This particular structure merely serves as an example.)

#### A. Design Principles of the Identifier

Let  $R(s)$ ,  $C(s)$ ,  $E(s)$  and  $U(s)$  denote respectively the reference input, the output, the error signal (input to the controller) and the control signal (input to the process) respectively.

The controller we have considered (as already mentioned) is a fractional order PID controller, whose transfer function is of the form  $K_p + T_i s^{-\lambda} + T_d s^\delta$ .

$K_p, T_i, T_d$  are the proportional, integral and derivative constants respectively,  $\lambda, \delta$  are the orders of integration and derivation. We can have by simple block diagram algebra,

$$\frac{C(s)}{R(s)} = \frac{K_p + T_i s^{-\lambda} + T_d s^\delta}{K_p + T_i s^{-\lambda} + T_d s^\delta + a_1 s^\alpha + a_2 s^\beta + a_3},$$

$$\Rightarrow E(s) = R(s) - C(s) = R(s) \left[ 1 - \frac{C(s)}{R(s)} \right].$$

For reference input  $R(s) = \frac{1}{s}$  (unit step),

$$U(s) = \frac{(K_p + T_i s^{-\lambda} + T_d s^\delta)(a_1 s^\alpha + a_2 s^\beta + a_3)}{s(K_p + T_i s^{-\lambda} + T_d s^\delta + a_1 s^\alpha + a_2 s^\beta + a_3)},$$

$$\Rightarrow C(s) = \frac{K_p + T_i s^{-\lambda} + T_d s^\delta}{s(K_p + T_i s^{-\lambda} + T_d s^\delta + a_1 s^\alpha + a_2 s^\beta + a_3)}.$$

We sample the output values  $c(t)$  in discrete time domain and store them. The PSO algorithm will search the solution space to come up with a process model which will replicate the observed values of  $c(t)$  for the same signal  $U(s)$ . At the moment of parameter identification, the controller has retained its old values of  $K_p, T_i, T_d, \lambda, \delta$  (fixed the last time the controller was tuned), but the process parameters  $a_1, \alpha, a_2, \beta, a_3$  have changed. So the optimization problem is five-dimensional since we know the controller parameters from its last tuning.

Let, for a process model, the output response is  $p(t)$ . We

will define a parameter  $F = \sum_{t=1}^f [c(t) - p(t)]^2$ , which gives

a measure of the deviation of the output of the trial process model from the output of the actual process. F is the fitness function that the PSO algorithm will try to minimize. At  $F = 0$ , the unknown parameters are optimized. The position vector of the best particle, i.e. the optimized value of  $\{a_1, \alpha, a_2, \beta, a_3\}$  is the identified parameter set. The process model for the optimized solution set should provide output identical to  $c(t)$ .

There are five unknown parameters to be optimized. So the present problem of system identification can be solved by a straightforward application of the PSO algorithm for optimization on a five-dimensional solution space, each particle having a five-dimensional position and velocity vector. This approach gives very accurate results, even when the input data is corrupted to a significant extent. We studied this scheme under both ideal and non-ideal (random error component added to readings) conditions.

#### B. Illustration of the Identifier

Synthetic data for  $c(t)$  is created using  $a_1 = 0.8$ ,  $\alpha = 2.2$ ,  $a_2 = 0.5$ ,  $\beta = 0.9$ ,  $a_3 = 1$ . The controller parameters are:

$K_p=410$ ,  $T_i=650$ ,  $T_d=50$ ,  $\lambda=0.25$ ,  $\delta=1.25$ . Sampling frequency is 20 samples per second. Error in taking readings of  $c(t)$  is simulated by adding a random number in the range  $[-0.05, 0.05]$  to each reading.

The PSO parameters used are: the inertia factor  $\omega$  decreases linearly from 0.9 to 0.4; the cognitive learning rate  $c_1=1.4$ , the social learning rate  $c_2=1.4$ . Number of particles in the population is 50. The PSO algorithm is run for 200 iterations, and this is kept as the stop criterion. The search ranges are:  $a_1, a_2, a_3$ : 0 to 2.0 (for all three parameters),  $\alpha$ : 2.0 to 2.4 and  $\beta$ : 0.7 to 1.1. The velocity ranges are:  $a_1, a_2, a_3$ :  $-0.5$  to  $0.5$  (for all three parameters),  $\alpha, \beta$ :  $-0.1$  to  $0.1$  (for both parameters).

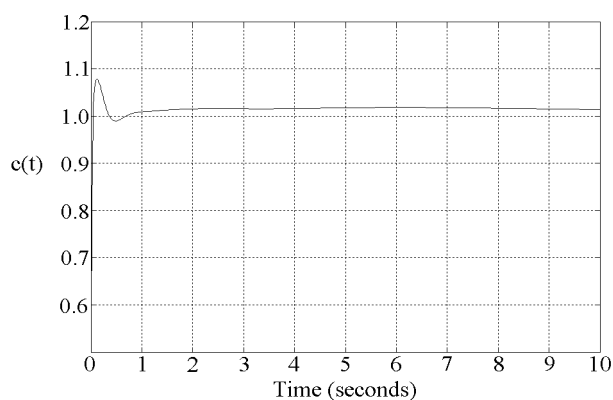


Fig. 2. Output waveform  $c(t)$  used for identification

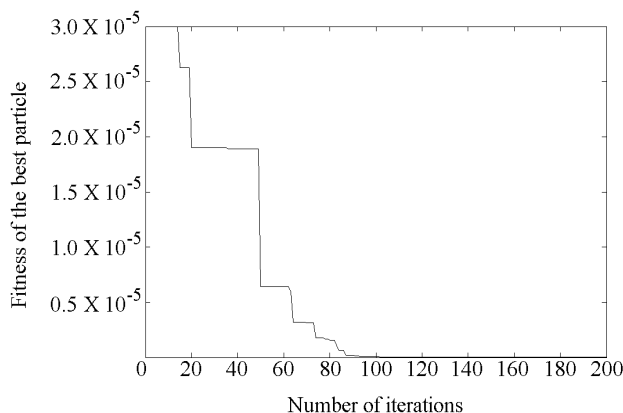


Fig. 3. Variation of best fitness with number of iterations

TABLE I.

IDENTIFICATION AND PERCENTAGE ERRORS

| Parameter | Estimate | Percentage Error |
|-----------|----------|------------------|
| $a_1$     | 0.8000   | 0.0002           |
| $\alpha$  | 2.2000   | 0.0000           |
| $a_2$     | 0.5000   | 0.0035           |
| $\beta$   | 0.9000   | 0.0042           |
| $a_3$     | 1.0000   | 0.0030           |

Table 1 gives results of the identification when the readings are accurate (error component is not added). After 200 iterations, the best fitness is  $2.5946 \times 10^{-14}$ .

Even when the random error component is added, the identification is very accurate.

### C. Analysis

The challenge in fractional order system identification is that the fractional powers are not restricted to assume only discrete integral values, but are distributed in a continuous interval. However, the use of the stochastic optimization strategy makes an accurate identification possible.

So the parameter identifier block can estimate the process parameters with a very high degree of accuracy. Now the controller tuner module should use these values of the identified parameters to tune the controller parameters on-line.

## IV. DETAILED DESIGN SCHEME OF THE CONTROLLER TUNER BLOCK

### A. Design Principles: Formulation of the Objective Function and its Optimization

Our approach is based on the *root locus* method (*dominant poles* method) of designing integral PID controllers [11]. As in the traditional root locus method, the peak overshoot  $M_p$  and rise time  $t_{rise}$  (or, in other words, requirements of stability and damping levels) are specified. From these specifications, we find out the damping ratio  $\zeta$  and the undamped natural frequency  $\omega_n$ , making use of (1), (2):

$$\zeta = -\frac{\ln(M_p)}{\sqrt{\{\ln(M_p)\}^2 + \pi^2}} \quad (1)$$

$$\omega_n = \frac{\pi - \tan^{-1}\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)}{t_{rise}\sqrt{1-\zeta^2}} \quad (2)$$

Using these computed values of  $\zeta$  and  $\omega_n$ , we then determine the desired positions of the dominant poles  $p_{1,2}$  of the closed loop system:

$$p_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2} = -a \pm jb \quad (3)$$

where  $a = \zeta\omega_n$ ,  $b = \omega_n\sqrt{1-\zeta^2}$ ,  $a, b > 0$ .

Let  $G_p(s)$  be the transfer function of the process we want to control,  $G_c(s)$  the transfer function of the controller to be designed and  $H(s)$  the transfer function of the feedback-path. Then, the closed loop transfer function of the controlled system is:  $T(s) = \frac{G(s)}{1 + G(s)H(s)}$ ,

$G(s) = G_c(s).G_p(s)$  being the forward path transfer function:

(4)

Assuming unity feedback, we have  $H(s)=1$ . In this case, the characteristic equation becomes  $1+G(s)=0$ ,

$$\begin{aligned} \Rightarrow 1+G_c(s).G_p(s) &= 0 \Rightarrow 1+G_c(s).\frac{P(s)}{Q(s)} = 0 \\ \Rightarrow Q(s)+G_c(s).P(s) &= 0 \end{aligned} \quad (5)$$

where  $P(s)$  and  $Q(s)$  are respectively the numerator and denominator polynomials of  $G(s)$ ,  $P(s)$  and  $Q(s)$  have no common factor.

As  $p_{1,2}$  must be poles of the closed loop system, each of them must be a root of the characteristic equation and hence must satisfy equation (5). Thus, putting  $s = p_1 = -a \pm jb$  in equation (5), we obtain

$$\begin{aligned} Q(p_1) + G_c(p_1).P(p_1) &= 0 \\ \Rightarrow Q(-a + jb) + (K_p + T_i(-a + jb)^{-\lambda} + T_d(-a + jb)^{\delta})P(-a + jb) &= 0 \end{aligned} \quad (6)$$

Equation (6) is a complex equation in five unknowns, namely  $K_p$ ,  $T_i$ ,  $T_d$ ,  $\lambda$ ,  $\delta$  and our problem of designing a controller that makes the closed loop dominant poles of the system coincide with  $p_{1,2}$  now reduces to determining the set of values of  $\{K_p, T_i, T_d, \lambda, \delta\}$  for which (6) holds good. But as the number of unknowns exceeds the number of equations, there exists an infinite number of solution sets and the equation cannot be unambiguously solved by traditional methods.

This necessitates the application of stochastic global search techniques, which, in turn, requires the formulation of a suitable *objective function* or *cost function*.

Let,  $R$  = real part of the L.H.S. of equation (6),  
 $I$  = imaginary part of the L.H.S. of equation (6),

$$P = \tan^{-1} \frac{I}{R}.$$

We define  $f(K_p, T_i, T_d, \lambda, \delta) = |R| + |I| + |P|$  as our objective function. Clearly,  $f \geq 0$ , in general, and  $f = 0$  if and only if  $R = 0$  and  $I = 0$  and  $P = 0$ , i.e. if and only if equation (6) is satisfied.

So, we now employ DE to scour the five-dimensional search space and home in on the optimal solution set  $\{K_p^*, T_i^*, T_d^*, \lambda^*, \delta^*\}$  for which  $f = f_{\min} = 0$ .

It is germane to mention here that the formulae (1), (2) and (3) hold strictly only for a *second order system* with its *complex conjugate pole-pair* at  $p_{1,2}$ . However, for a higher-order system with a *dominant pole-pair* (real part of this pole-pair is much smaller than those of other poles), these formulae are widely used in control engineering applications with a fair degree of accuracy [11]. Nevertheless, after designing our controller with the

help of these formulae, we perform a simulation to obtain the *unit step response* of the closed-loop control system as a check.

### B. Illustration of the Tuner

The process (control objective) has the transfer function

$$G_p(s) = \frac{1}{0.8s^{2.2} + 0.5s^{0.9} + 1}.$$

We want to design a controller such that the closed loop system has a peak overshoot  $M_p \leq 10\%$  and rise-time  $t_{rise} \leq 0.3$  seconds.

Using the formulae (1) and (2), we obtain  $\zeta = 0.5912$  and  $\omega_n = 9.107 \text{ s}^{-1}$ . Thus the dominant poles for the closed loop controlled system should lie at  $p_1 = (-5.384 + j7.345)$  and  $p_2 = (-5.384 - j7.345)$ .

As per convention, we use unity feedback. The controller transfer function is given by (4). Putting  $s = p_1 = (-5.384 + j7.345)$  in the characteristic equation,

$$1 + \frac{K_p + T_i(-5.384 + j7.345)^{-\lambda} + T_d(-5.384 + j7.345)^{\delta}}{0.8(-5.384 + j7.345)^{2.2} + 0.5(-5.384 + j7.345)^{0.9} + 1} = 0 \quad (7)$$

After separating the real and imaginary parts, we have:

$$R = (K_p + 13.4235) + \frac{T_i}{9.107^{\lambda}} \cos(2.203\lambda) + \quad (8)$$

$$T_d(9.107)^{\delta} \cos(2.203\delta)$$

$$I = -\frac{T_i}{9.107^{\lambda}} \sin(2.203\lambda) + \quad (9)$$

$$T_d(9.107)^{\delta} \sin(2.203\delta) - 98.9237$$

$$P = \tan^{-1} \frac{I}{R}. \quad (10)$$

Thus, (8), (9) and (10) give us our objective function  $f(K_p, T_i, T_d, \lambda, \delta) = |R| + |I| + |P|$ .

We set the limits on the components of the position-vectors of the chromosomes (i.e. the controller parameters) as follows: as a practical consideration, we assume  $0 \leq K_p, T_i, T_d \leq 1000$ ,  $0 \leq \lambda, \delta \leq 2$ .

The stop criterion used is a tolerance value  $f = 0.0001$ .

After the DE algorithm is run, the position vector of the best particle gives the parameters of the optimum PID controller. The transfer function of the identified controller is  $962.80 + 197.55s^{-1.79} + 46.27s^{1.37}$ .

Although we have constructed the objective function  $f$  by making use of the dominant pole  $p_1 = -a + jb$  in the second quadrant, we would arrive at the same  $f$  if we had started with the third-quadrant dominant pole  $p_2 = -a - jb$ . This is because, in the latter case, the imaginary part of the reduced characteristic equation (6) would just be the negative of what we have obtained in (7) so that  $f$ , which involves absolute values only, would

remain unaltered. This is true not only for the particular problem in question but in general as well.

Fig. 4 shows the unit step response of the uncontrolled system as well as the unit step response of the system controlled by the designed STR. The peak overshoot is 5.2%, and the rise time is 0.054 seconds.

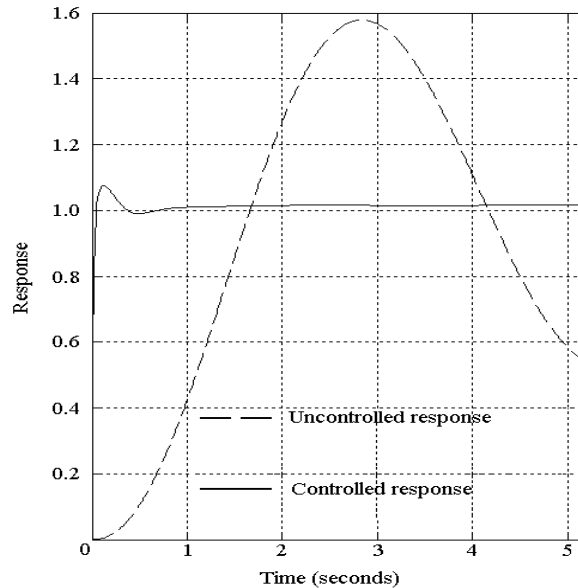


Fig. 4. Unit step responses of the control objective

### C. Analysis of the Tuner Designing Scheme

The controller meets the rise-time and peak overshoot requirements more than satisfactorily. The rise-time performance especially far exceeds the design requirement, which means that the controller tuner is actually over-performing and we can have a significant safety margin. We also considered using PSO in the 'controller tuner' block as well. However we found that DE offers significantly better performance in this case.

An important point is to be noted in this context. For the final system to exhibit the desired performance, it is necessary that the evaluated pole-values  $p_{1,2}$  truly correspond to the *dominant poles* of the closed-loop controlled system. However, the optimization of our cost function ensures only that  $p_{1,2}$  are poles of the system but does not guarantee that these are the dominant poles. In other words, equation (7) embodies a *necessary*, but not *sufficient*, condition for the dominance of  $p_{1,2}$ . Obviously, there may be a number of possible combinations of values of the parameters  $[K_p, T_i, T_d, \lambda, \delta]$  for which the resulting characteristic equation is satisfied by  $p_{1,2}$  but, in each such case, the

characteristic equation will have other roots (closed-loop system poles), too, and these poles will also play an important role in determining the overall system response. This is a possible explanation of the observation that the performance of the final designed system with fractional controller is actually better than desired.

## V. COMMENTS AND CONCLUSIONS

An elegant and effective design scheme for a fractional order STR is proposed. A simple computer program in C or MATLAB using sampled data acquisition and other interfaces with the physical process can actually implement the complete STR. Of course, the same design can easily be employed to implement an integer order STR as well. Apart from the example shown, we also considered several other similar problems and solved them by our proposed method. These have not been included in this paper owing to paucity of space. We also intend to apply other well-known algorithms such as Genetic Algorithm, Bacterial Foraging Optimization and their many variants to the same optimization problems and compare their performance with the presented scheme.

## REFERENCES

- [1] K. J. Astrom, and B. Wittenmark, *Adaptive Control, Second Edition*, Pearson Education, Inc., 1995.
- [2] P. J. Torvik, and R. L. Bagley, "On the appearance of the fractional derivative in the behaviour of real materials," *Trans. of the ASME*, vol. 51, June 1984, pp. 294-298.
- [3] K. B. Oldham, and J. Spanier, *The Fractional Calculus*, Academic Press, New York, 1974.
- [4] J. P. Norton, *An Introduction to Identification*, Academic Press, London, 1986.
- [5] I. Petras, "The fractional order controllers: Methods for their synthesis and application," *Journal of Electrical Engineering*, vol. 50, no. 9-10, pp. 284-288, 1999.
- [6] S. Milos, and C. Martin, "The fractional-order PID controller outperforms the classical one," *7<sup>th</sup> Intl. Scientific-Technical Conf. - Process Control 2006, Czech Republic*.
- [7] J. Kennedy, and R. C. Eberhart, "Particle swarm optimisation," *Proc. of the IEEE Intl. Conf. on Neural Networks*, pp.1942-1948, 1995.
- [8] Y. Shi, and R. C. Eberhart, "A modified particle swarm optimiser," *Proc. of the IEEE Congress on Evolutionary Computation*, pp. 69-73, 1998.
- [9] R. Storn, and K. Price, "Differential evolution - A simple and efficient heuristic for global continuous spaces," *Journal of Global Optimization*, 11(4) (1997) 341-359.
- [10] S. Das, A. Konar, and U. K. Chakraborty, "Two improved differential evolution schemes for faster global search," *ACM-SIGEVO Proc. of Genetic and Evolutionary Computation Conf. (GECCO-2005)*, 2005.
- [11] I. J. Nagrath, and M. Gopal, *Control Systems Engineering, Fifth Edition*, New Age International Publishers, 2007.

# A Concurrent Approach to Clustering Algorithm with Applications to VLSI Domain

Shahid Md. Asif Iqbal, Mohammad Iftekhar Monir, Taufique Sayeed, Abu Hamed Mohammad Misbah-Uddin

Department of Computer Science & Engineering, Premier University, 1/A, O.R.Nizam Road, Probortok Circle, Chittagong-4203, Bangladesh.

**Abstract** — Circuit partitioning plays an important role in physical design automation of very large scale integration (VLSI) chips. In this brief we present a new connectivity based top down as well as bottom up approach to clustering algorithm for VLSI circuit partitioning. The proposed clustering algorithm partitions the circuit by focusing on highly interconnected cell groups. This clustering algorithm leads to a parallel implementation in which multiple processors are used to identify clusters simultaneously. The process starts with forming clusters by grouping the cells that are tightly connected and as well as the cells that are loosely connected. Considering both types of groups has the advantage that clusters formed from this technique will be highly connected and compact too. Therefore the proposed clustering method can reduce the size and also speed-up the large-scale partitioning problem without losing partitioning solution qualities. The performance of the proposed clustering algorithm is evaluated on a standard set of partitioning benchmark-ISPD98 benchmark suite.

**Index Terms** — Base, cluster, partition, push-pop, threshold.

## I. INTRODUCTION

A cluster is a group of strongly connected components in a circuit. The goal of clustering algorithms is to identify the clusters in a circuit. In VLSI layout design, clustering algorithms can be used to construct the natural hierarchy of the circuit. The circuit-partitioning problem can be expressed in graph theoretic terms by modeling each circuit as a hypergraph with circuit components (cells) represented by vertices, and the nets connecting the components represented by hyperedges. A general objective of the circuit partitioning is to minimize the *netcut* or the number of nets that span the partitions. Bottom-up clustering algorithms provide a solution to the problems encountered when partitioning very large circuits [1]. A parallel clustering algorithm can be integrated into the partitioning process by using clustering as a preprocessing step to partitioning. First, clustering is performed on the circuit to obtain a condensed circuit in which each cluster of components has been collapsed to form a single component. Partitioning is then performed on the clustered circuit instead of the original circuit. Since the number of components in a clustered circuit is usually much smaller than that of the original circuit, the

time and space required by the partitioning algorithm is reduced significantly [2]. Clusters with a higher density are considered to be of higher quality.

Here, we propose a new clustering method for VLSI circuit partitioning. The proposed clustering method focuses on finding the natural clusters in a circuit and thus tends to cluster the original circuit without loss of partitioning solution quality. The algorithm first groups a cell and its neighbor cells to form an initial cluster, then the initial cluster is further refined by using a PUSH-POP mechanism to move the cells that have more connections with cells outside the initial cluster. In this way, the cells that are highly connected with one another are identified and merged. To provide a flexible criterion for finding natural clusters and as well as maintaining density of the clusters, the sizes of the clusters are not fixed. Furthermore, instead of ordering cells randomly like most clustering algorithms [2], [3], the proposed clustering algorithm incorporates two nonrandom ordering methods based on local cell connectivity information. Both nonrandom ordering methods can lead to clusters with balanced sizes and high density and consequently good partitioning solutions.

The remainder of this paper is organized as follows. In Section II, a background on the clustering algorithms is given. Section III presents the proposed clustering algorithm in detail. In Section IV, the experimental results of the proposed clustering algorithm on the ISPD98 benchmarks [4] are reported. Finally, conclusions are made in Section V.

## II. BACKGROUND & MOTIVATION

In the past two decades, partitioning problems have been studied by many researchers. The Fiduccia-Mattheyses (FM) [5] heuristic is one of the most widely used iterative improvement approaches. The FM algorithm can provide reasonable solutions for small to medium size circuits, but as VLSI circuits become larger, the solution quality deteriorates [3]. Recent works have focused on multilevel partitioning algorithms that exploit clustering approaches to gradually merge clusters in multiple clustering levels [6], [7]. Currently, hMetis [3], [6] and



MLPart [2] are regarded to be the state-of-the-art partitioners due to their superior performances in terms of speed and the final partitioning solution.

Clustering algorithms play a crucial role in VLSI circuit partitioning. The multilevel partitioning algorithms are based on clustering a large circuit into a sequence of smaller circuits. Clustering is an important factor in determining the performance of partitioning algorithms, where different clustering approaches can lead to different partitioning solutions. A large number of clustering algorithms have been proposed in the past decade [2], [3], [6], [7]. In [2], heavy-edge matching that clusters pairs of cells with highest connectivity is used. The clustering is done based on *edge separability* in [7]. This technique uses a graph model and exploits the global connectivity information for clustering. Edge clustering, hyperedge clustering, modified hyperedge clustering, and FirstChoice clustering schemes are presented in [3] and [6]. The edge clustering scheme groups two nodes that are highly connected to form a cluster. Hyperedge clustering and modified hyperedge clustering group the nodes belonging to a hyperedge to form a cluster. The FirstChoice clustering scheme groups the nodes such that each node in a cluster is highly connected with at least one other node in the same cluster. Among various clustering algorithms, edge clustering is one of the most commonly used clustering approaches and has been successfully used in hMetis and MLPart. However, the edge clustering scheme is likely to ignore natural clusters of a circuit [8]-[10]. Therefore, it may not produce optimal results. An example of a circuit's natural clusters is illustrated in Fig.1 [8]-[9].

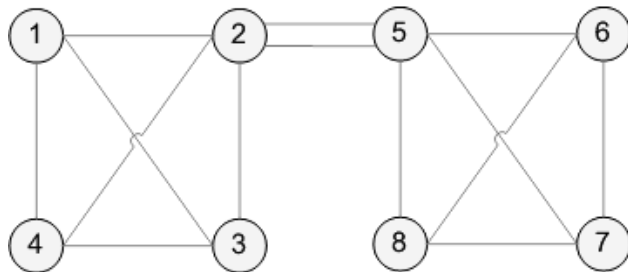


Fig. 1. Example of natural clusters in a circuit

The graph in Fig. 1, contains two natural clusters made of cells ((1),(2),(3),(4)) and ((5),(6),(7),(8)). In the edge clustering scheme, given a cell, the edge clustering will try to find an unclustered cell with the highest connectivity. Therefore, assuming that all cells in Fig. 1, have equal area and no cell already belongs to a cluster, the edge clustering scheme may group cells and together, eliminating the two natural clusters.

### III. PROPOSED CLUSTERING PROCEDURE

Before going to the details let us first define the following terms:

**Base:**-The cell from which clustering can be initiated and which is not a part of any cluster.

**Unclassified cell:**-The cell that is not a part of any cluster.

#### A. Clustering with Parallel Base Selection

1. Sort the cells according to their degree.
2. While any cell is unclassified
3. Select highest and lowest degree cell(s) as base cell(s).
4. Beginning with the base(s) cell start clustering.
5. If area of a cluster is less or equal to 10% of the whole circuit, continue clustering.
6. Else go to step 3.
7. Push & Pop the cells from one cluster to another to balance the cluster size but the number of wire connections should remain same or decline.

In the rest of this section, the preceding steps are discussed in detail. In addition, implementation details and discussion of the proposed algorithm are presented.

In Fig. 2 we present a simple circuit diagram to determine the cell ordering and neighbor relationship table. In Table 1 neighbors of only the base cells are considered.

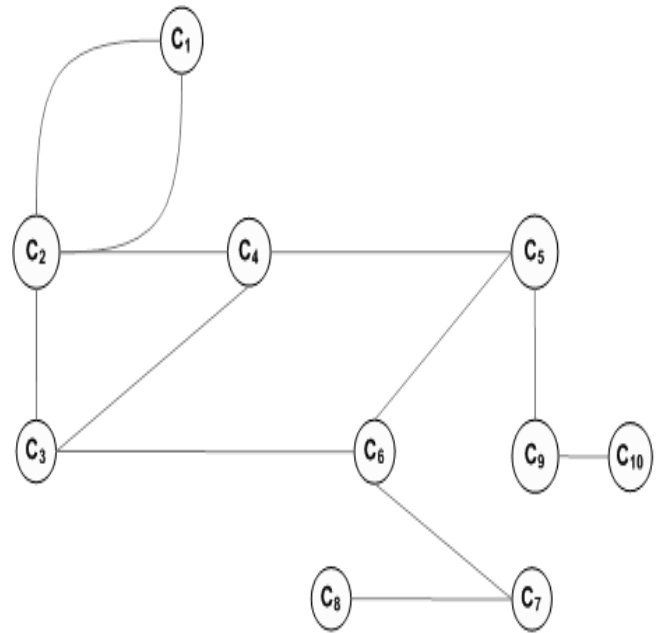


Fig. 2. Graph representation of a circuit

From Fig. 2 we can construct the partial neighbor relationship table for all the cells in the circuit considering neighbors of the base only.

TABLE 1  
NEIGHBOR RELATIONSHIP TABLE

| Cell     | Cell Deg. | 1 <sup>st</sup> level neig. | 2 <sup>nd</sup> level neig. | 3 <sup>rd</sup> level neig. | 4 <sup>th</sup> level neig. | 5 <sup>th</sup> level neig. |
|----------|-----------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| $c_1$    | 2         | -                           | -                           | -                           | -                           | -                           |
| $c_2$    | 4         | $c_1, c_3, c_4$             | $c_5, c_6$                  | $c_7, c_9$                  | $c_8, c_{10}$               | -                           |
| $c_3$    | 3         | -                           | -                           | -                           | -                           | -                           |
| $c_4$    | 3         | -                           | -                           | -                           | -                           | -                           |
| $c_5$    | 3         | -                           | -                           | -                           | -                           | -                           |
| $c_6$    | 3         | -                           | -                           | -                           | -                           | -                           |
| $c_7$    | 2         | -                           | -                           | -                           | -                           | -                           |
| $c_8$    | 1         | $c_7$                       | $c_6$                       | $c_3, c_5$                  | $c_2, c_4, c_9$             | $c_1, c_{10}$               |
| $c_9$    | 2         | -                           | -                           | -                           | -                           | -                           |
| $c_{10}$ | 1         | $c_9$                       | $c_5$                       | $c_4, c_6$                  | $c_2, c_3, c_7$             | $c_1, c_8$                  |

### B. Cell Ordering

Order the cells according to the degree. Here we have  $c_2 > c_3 = c_4 = c_5 = c_6 > c_1 = c_7 = c_9 > c_8 = c_{10}$  from Fig. 2 and Table 1.

### C. Base Cell Selection

The cells with the highest degree as well as lowest degree are base cells. If more than one highest or lowest degree cells exists then all of them are considered as base cells

Here the approach of base cell(s) selection is parallel i.e. the procedure starts from both end of the ordered list of cells. Also the approach considers all the cells (if any) of identical highest and lowest degree.

*Advantages of parallel selection of the base cell (s):* Higher density increases the compactness of each cluster due to highly connected cells accommodated in the same cluster. After PUSH-POP of cells in the initially created clusters, balanced size clusters can be found. Again parallel selection of base cell(s) can faster the clustering procedure.

### D. Initial Clustering

Each base cell initiates clustering with 1<sup>st</sup> level neighbors and then 2<sup>nd</sup> level neighbor (if exist, i.e. a 2<sup>nd</sup> level neighbor can enter the initial cluster if any how it is not a 1<sup>st</sup> level neighbor of the base cell), then 3<sup>rd</sup> level neighbor if they are not the 2<sup>nd</sup> level neighbor and so on up to the threshold. A cell can't be a member of more than one cluster at a time, in that case if more than one base cells try to win the cell than the base cell that have the closer level of connectivity to that will have the cell, if that also doesn't solve the problem i.e. if all the base cells have same level of connectivity to that cell than consider the wirelength between the base and target cell. Again if that also doesn't break the tie than the random assignment of that cell will take place. For the sake of determining the

wirelength we have placed the nodes on a grid and each grid segment assigned a unique weight.

### E. Threshold

Consider each cluster up to the 10% of the total circuit i.e. if there are n number of cells in the total circuit, then each cluster will possess a maximum n/10 cells.

### F. Clustering Remaining Cells

After the 1<sup>st</sup> level of clustering i.e. initial clustering there may remain unclustered cells in the circuit. From the remaining cells construct a list of cells in descending order of node degree and then apply the whole steps that are performed in the initial clustering. Continue the process until all of the cells are being clustered.

### G. PUSH-POP

The initial clusters formed with the cells having highest degree will have larger sizes thus it will not be possible to enter all the nth level neighbors. At the same time clusters formed with lower degree cells will have fewer cells i.e. small size. So to overcome this type of unbalanced problem we apply the PUSH-POP procedure. Where each PUSH will enter a cell from a large size cluster to another large or small size cluster if the no. of connections between those clusters remain same or decrease and at the same time the area of the cluster satisfies the threshold. Each cluster will maintain a stack of cells by its own. The cell having more connections to outside the cluster than inside will be the TOP element of the stack. Remaining cells of the cluster will be stacked in decreasing order of outside connectivity so that the BOTTOM element of the stack will have the cell that have more connections to inside than outside.

### H. Example

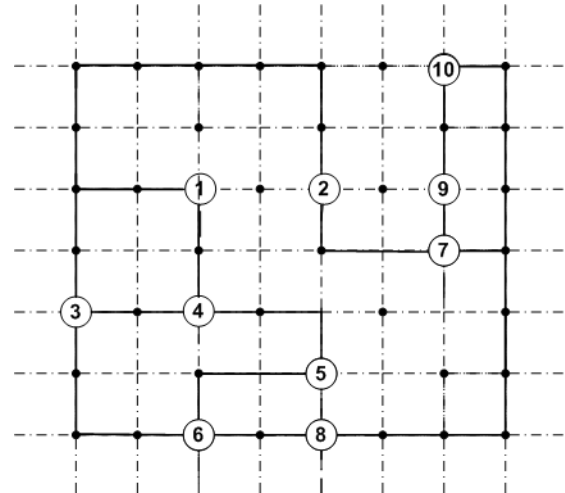


Fig. 3. Grid representation of a circuit

Clustering starts from all the base cells in parallel. Here we have considered the cells as nodes in the grid where each side of a square in the grid is 1 unit. In Fig. 3 Highest degree base cells are 3, 7, and 8 and lowest degree base cell is only 9

Now initial clustering begins with all the base cells at a time. Therefore we will start with four initial clusters. The figure below shows the clusters created from the Fig. 3.

So we have in Fig. 4 cell 2 is the 1<sup>st</sup> level neighbor of base cells 3 and 7, but cell 2 entered the cluster of the base 7 due to wirelength. Similarly 6 entered the cluster of 8 instead of cluster of 3 and 10 entered cluster of 9 instead of 7.

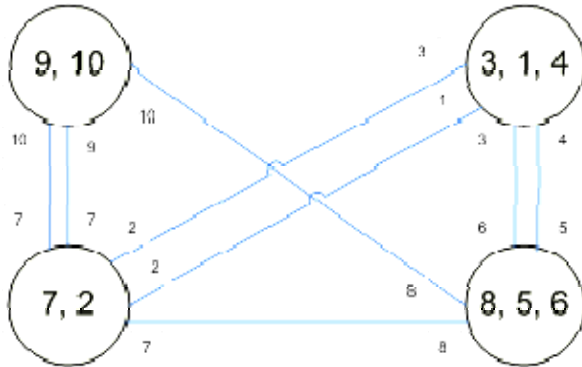


Fig. 4. Initial cluster created from the circuit of Fig. 3

Thus there are four initial clusters more or less balanced size, so PUSH-POP not required here. The connectivity among the clusters also displayed here. Now consider another sample circuit of Fig. 5 to cluster with PUSH-POP function. The circuit has 12 cells. Among them base cells are 7, 3 and 1.

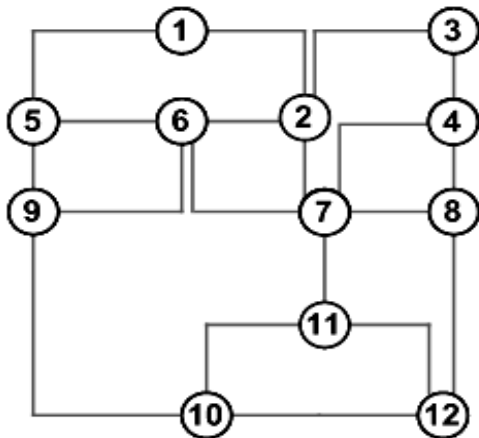


Fig. 5. Example circuit to explain PUSH-POP

Here we have got three initial clusters from the figure beginning with the base cells 7, 3 and 1.

The initial clusters from Fig. 5 are given below

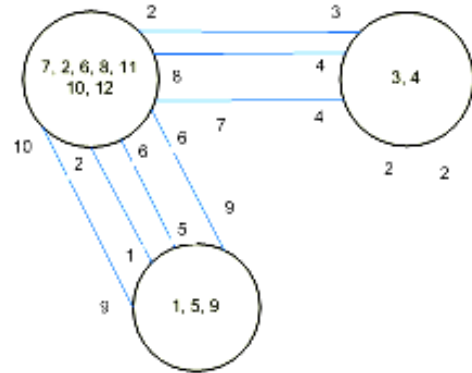


Fig. 6. Initial cluster created from the circuit of Fig. 5

From Fig. 6 we see that the cluster sizes are unbalanced so we can apply the PUSH-POP process here. If we push cell 2 from the cluster of 7 to 3 and cell 6 from cluster of 7 to 1. Then the refined clusters are given in Fig. 7 as follows:

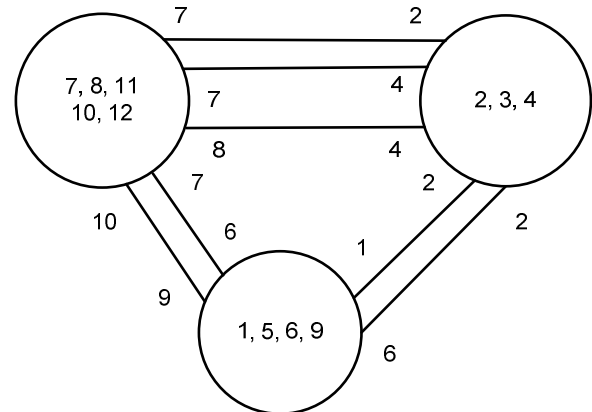


Fig. 7. Refined cluster after PUSH-POP

The number of inter cluster connectivity remain same after PUSH-POP. But the sizes of the cluster are better than the previous one.

#### IV. EXPERIMENTAL RESULTS

The proposed clustering technique has been implemented in C++ and tested as a preprocessing step for multilevel partitioning.

The following table shows a comparison of our proposed algorithm with Net Cluster (NC), best-choice clustering (BC), FirstChoice (FC) clustering. From Table 2 we can say that our algorithm has better runtime. The percentage improve in performance has found by considering the selected IBM benchmarks.

TABLE 2  
COMPARISON OF AVERAGE RUNTIME

| Circuit | Comparison of average runtime |       |       |                    |
|---------|-------------------------------|-------|-------|--------------------|
|         | NC                            | BC    | FC    | Proposed algorithm |
| Ibm01   | 1.024                         | 1.056 | 1.086 | 1.063              |
| ibm05   | 1.028                         | 1.046 | 1.035 | 1.033              |
| Ibm10   | 1.015                         | 1.021 | 1.052 | 1.041              |

The next table shows the improvement of cluster size with no. of runs. After level 3 of clustering we found that the maximum difference between the sizes of cluster is almost near about 10%.

TABLE 3  
ILLUSTRATION OF DIFFERENCES AMONG CLUSTER SIZES  
FROM THREE LEVELS OF PROPOSED PARALLEL  
CLUSTERING PROCEDURE FOR SELECTED IBM  
BENCHMARKS

| Circuit | % reduction of differences between cluster sizes |         |         |
|---------|--------------------------------------------------|---------|---------|
|         | Level 1                                          | Level 2 | Level 3 |
| ibm01   | 50%                                              | 20%     | 10%     |
| ibm05   | 55%                                              | 24%     | 11%     |
| Ibm10   | 47%                                              | 18%     | 9%      |

## V. CONCLUSION AND FUTURE WORK

In this work, a new connectivity-based clustering algorithm is proposed and applied to VLSI circuit partitioning. Experimental results illustrate the effectiveness of the proposed technique. As the cluster size is restricted so there may be a situation that some cells may be unclustered but which is very small in number. Future work involves full implementation of the proposed clustering algorithm within the context of VLSI placement.

## REFERENCES

- [1] N. A. Sherwani, Algorithm for VLSI Physical Design Automation, Third Edition. Massachusetts, USA: Kluwer Academic Publishers, 1999.
- [2] C. J. Alpert, D. Huang and A. B. Kahug, "Multilevel circuit partitioning," IEEE Trans. on CAD, vol. 17, no. 8, pp.655-667,1998.
- [3] G. Karypis, R. Aggarwal, V. kumar and S. Shekhar, "Multilevel hypergraph partitioning: Application in VLSI domain," IEEE Trans. Very Large Scale Integr. (VLSI) Syst. vol. 7, no.1, pp.69-79, Jan. 1999.
- [4] C. J. Alpert, "The ISPD98 circuit benchmark suite," in Procs. ACB/IEEE ISPD, 1998, pp. 80-85.
- [5] C. M. Fiduccia and R. M. Mattheyes, "A linear time heuristic for improving network partitions," in Proc. ACM/IEEE, 1982, pp. 175-181.
- [6] G. Karypis and V. Kumar, "Multilevel k-way hypergraph partitioning," in Proc. ACM/IEEE DAC, 1999, pp. 343-348.
- [7] J. Cong and S. Lim, "Edge separability-based circuit clustering with applications to multilevel circuit partitioning," IEEE trans. on CAD, vol. 23, no. 3, pp. 346-357, 2004.
- [8] Jianhua Li, L. Behjat and Blair Schiffner, "A Connectivity Based Clustering Algorithm With Applications to VLSI Circuit Partitioning," IEEE trans. Circuits and Systems, vol. 53, no. 1, pp. 384-388, May 2006.
- [9] Jianhua Li and L. Behjat, "A Structure Based Clustering Algorithm With Applications to VLSI Physical Design," Proc. of the IDEAS'05, IEEE, 2005.
- [10] J. Huang, J Li, L. Rakai and L. Behjat, "Wirelength Based Clustering Technique for VLSI Physical Design," IEEE ,2007, pp. 655-658.
- [11] G.-J. Nam, C. J. Alpert, P. Villarubia, B. Winter and M.Y. Idiz, "The ISPD2005 placement contest and benchmark suite," in ACM/IEEE ISPD, 2005, pp.261-219.
- [12] J. Cong and M. Smith, "A parallel bottom-up clustering algorithm with applications to circuit partitioning in vlsi design," in ACM/IEEE DAC, 1993, pp.755-760.
- [13] J. Li and l. Behjat, "Net Cluster: A Net-Reduction Based clustering Preprocessing Algorithm," in Proc. ISPD, 2006, pp. 200-205.
- [14] S. Sait and H.Youssef, VLSI Physical Design Automation: Theory and practice. World Scientific, 1998.
- [15] C. Alpert, A. Kahug, G. -J. Nam, S. Reda and P. Villarrubia, "A Semi-persistent Clustering technique for VLSI circuit Placement," in Proc. ISPD, 2005, pp.200-207.

# Proposal for an Efficient Quantum Key Distribution System Using Entanglement

M.A. Bashar<sup>1</sup>, M. A. Chowdhury<sup>1</sup>, R. Islam<sup>1</sup>, M. S. Rahman<sup>1</sup>, S. K. Das<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, Shah Jalal University of Science and Technology, Sylhet-3114, Bangladesh.

<sup>2</sup>Department of Physics, Shah Jalal University of Science and Technology, Sylhet-3114, Bangladesh

E-mail: {basharcse, misbahulalam, kmrafiq}@gmail.com, {rahmanms, skdas-phy}@sust.edu

**Abstract** — We propose a multiuser Quantum Key Distribution system based on entanglement property which is simple, more efficient, cost effective and easy to implement in laboratory environment. Existing protocols uses base communication to provide security, as 50% of the time sender and receiver's base is not agreed there are 50% bit loss during key distribution but our protocol does not need the base communication hence we achieve 3dB performance improvement.

**Index Terms** — QKD, quantum networking, data security.

## I. INTRODUCTION

Quantum Key Distribution (QKD) is more secure and provides some interesting benefits when it is done utilizing entanglement property rather than single photon superposition property. In typical systems, base is communicated after sending a sequence of bits. As Alice and Bob select the base randomly, so in 50% case their base will agree and in other 50% case they will disagree. They discard the bit values in which case they disagree in base selection and keep the bit values in which they agree in base selection. Here we see that they loss 50% of bits that have been generated. But in our system correct bit value is not

dependent on base agreement. Hence, we can save 50% bit loss and our system achieves performance improvement of 3dB (two times) than the present systems. We have done analysis and confirmed that our system does not require base comparison for secure communication.

## II. ARCHITECTURE

Pairs of photons can be produced in various sorts of entangled states, including:

- 1) Polarization entangled [1]
- 2) Time-frequency entangled [2]
- 3) Momentum entangled [3]

In our proposed experiment we have used polarization entangled photons pair. In the recent teleportation field experiment by the Vienna group [2], polarization-entangled photons were transmitted using glass fiber in a channel underneath the Danube River in Vienna. Above the river, there was a direct microwave connection for the classical channel. The two stations were physically separated by approx. 600m. So we hope our architecture will work well.

### A. System Configuration

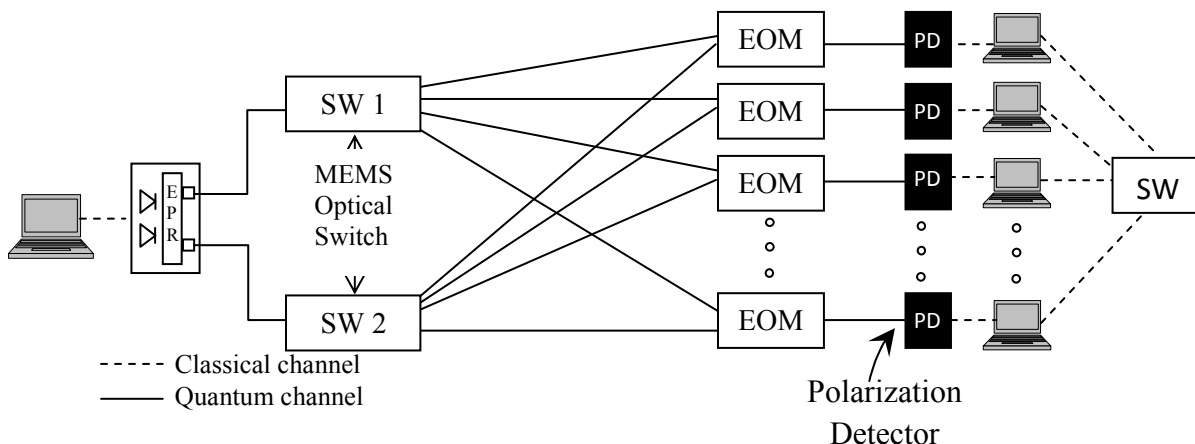


Fig 1: Proposed architecture of multiuser QKD using entanglement.

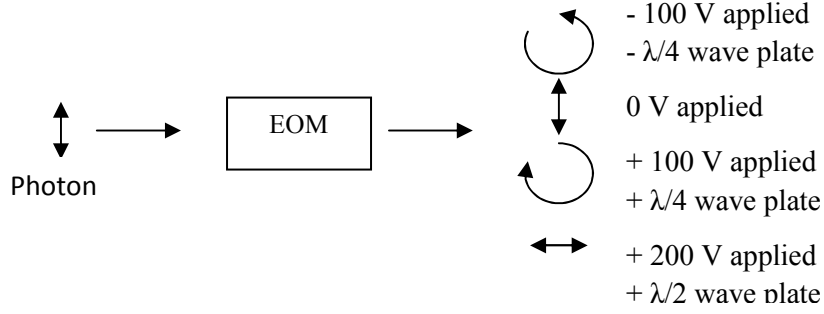


Fig 2: Function of EOM

In our system there is a dedicated server that produces polarization entangled photon pairs. We call the two photons P1 and P2. P1 goes to switch1 and P2 goes to switch2 through two standard telecom optical fiber lines (SMF28). Switch1 and switch2 are two MEMS optical switches which are used for routing and by this way we provide multiuser facility. Each party that engages to the QKD must be connected to both switch1 and switch2 by standard telecom optical fiber line (SMF28). Also they have classical lines to communicate the classical information. We call one party Alice and the other party Bob that are engaged at any particular time in QKD. P1 and P2 propagate and reach Alice's and Bob's EOM (Electro-Optical Modulator) and then polarization detector. High-speed data handling boards, designed and implemented at NIST, are installed in both Alice and Bob's computer. The board collects the detection events and communicates with each other to run the protocol of QKD. Alice and Bob's boards then send the bit value to their own computers for the subsequent error reconciliation and privacy amplification. The computers of Alice and Bob then use the shared secret keys for application-level data encryption as necessary to implement secure communications through an unsecured network.

#### B. Function of Electro-Optical Modulator (EOM) in the setup

The Electro-Optical Modulator - In order to control the polarization of the photons an EOM is used (KD\*P crystal and  $3 \times 3$  mm<sup>2</sup> aperture) which is basically a delay plate controlled by a voltage; it uses the Pockels Effect. For example, at 633 nm, the "half wave plate" voltage is approximately 200 V; if 0 V is applied, the crystal has no effect on the light. We choose polarization between 0°, circular left, 90° and circular right. These circular states are defined by

$$\begin{aligned} |\psi\rangle &= \frac{\sqrt{2}}{2} (|\rightarrow\rangle + i|\uparrow\rangle) \\ |\psi\rangle &= \frac{\sqrt{2}}{2} (|\rightarrow\rangle - i|\uparrow\rangle) \end{aligned} \quad (1)$$

They can easily be obtained by a single EOM applying a "quarter wave plate" or a "minus quarter wave plate" voltage.

### III. PROTOCOL DESCRIPTION

#### A. Possible attacks from Eve

There are many potential attacks for Eve. But quantum mechanics comes into play and guarantees a secure transmission of the key [4, 10]:

- 1) Eve can try to take over the source in the middle and prepare photon pairs with distinct polarization, however Alice and Bob can perform tests of the entanglement. If the Bell's inequality (we have given a short note on Bell's inequality below.) is violated then the source really emits entangled pairs which could have not been tampered with by Eve.
- 2) Eve cannot copy the quantum state of the photon perfectly and then wait for basis reconciliation to measure in the appropriate basis. The no cloning theorem in quantum mechanics prohibits such actions.
- 3) If Eve decides to measure and resend the photon, she will only half of the time pick the right basis. In the other cases Eve will get random information and the photon she sends to Bob is necessarily also uncorrelated to the photon of Alice. Eve therefore introduces an error between otherwise perfectly correlated events and she can be detected. Using the intercept-resend strategy, Eve's tampering will amount to 50% qubit error rate (QBER).

Bell's inequality:

Einstein was challenging the uncertainties of quantum mechanics in the EPR paradox [5] and his famous thought

experiment suggested that for certain quantum states – Schrödinger called them ‘entangle states’ – quantum mechanics seems to require some ‘faster than light’ influence between separated parts of a quantum system. In 1964, nearly 30 years after the EPR paper and 9 years after Einstein’s death, John Bell derived an ‘inequality’ for the correlations originating from ‘common-sense’ pre-determined conditions that was violated by the correlations predicted by quantum mechanics. Bell’s inequality makes it possible to put quantum mechanics to the test. A series of beautiful experiments designed to measure these EPR correlations were performed by Alain Aspect and his group in Paris in the late 1970s and early 1980s [6]. These experiments confirmed that Bell’s inequality was indeed violated and that the predictions of quantum mechanics agreed with the data.

### B. Protocol

- 1) A source of entangled-photon pairs is configured to send one photon to Alice and one photon to Bob.
- 2) Alice and Bob’s detectors are both configured to measure the polarization in a particular base – (0, 90) or (circular left, circular right).
- 3) Randomly deflect the photon (i.e., change polarization) when it passes through EOM.
- 4) Alice and Bob record the bit value and exact time for each detection. Arrival times are used to establish coincident detections.
- 5) Alice and Bob then communicate the measurement time using classical medium.
- 6) Both Alice and Bob keep the bit value that agree in time and discard the other.
- 7) Randomly select some bit positions in the string of quantum bits. Then communicate the bit values in these positions using classical channel to see if they agree hence calculate bit error rate for determining the presence of eavesdropper.
- 8) If no eavesdropper is detected then use bits to make secret key otherwise discard it.

### C. How It Works

Server emits two entangle photons and one of them goes to Alice and other goes to Bob. Both Alice and Bob’s photons goes to polarization detector after passing through the EOM. Alice and Bob change voltage level of their EOM (-100 v, 0 v, 100 v, 200 v) randomly, that is they change polarization of photon randomly. But whatever polarization they impose on one photon, according to quantum mechanics the other photon will be in the same polarization state and before measurement photon’s polarization state is nondeterministic and unknown to them. Whenever measurement is performed (i.e. detection) whatever polarization is detected in one detector, other detector will detect the same polarization (i.e. if one detect vertical other

will also detect vertical). As in our system measurement is performed after the operation of EOM so there is no effect of EOM operation on outcome, that is Alice and Bob will get same output. In this system only one base (0 - 90 or circular left - circular right) will be used for measurement hence they do not need to communicate the base. By excluding base communication from the system we can use 100% data that are communicated and this is two times (3 dB) efficient than the existing system (which needs base communication hence can use only 50% data on average that are communicated). Now the question is how our system will detect the presence of eavesdropper. This can be done very easily. The following example will clear the matter. Let in a situation one of the photons goes to Alice and the other is intercepted by the eavesdropper. After detecting the polarization of the photon eavesdropper will send the photon to Bob using the same polarization that he detected. Whenever eavesdropper detected the photon, entanglement has been collapsed down and Alice and Bob’s photons are no longer correlated with each other. When the photons go through the EOM of Alice and Bob they will be changed in polarization in 75% time. As two photons are no longer correlated so Alice and Bob will get different value on average in 50% time. That is if eavesdropper is not present they will get same value in 100% time and if eavesdropper is present they will get different value in 50% time. We see that presence of eavesdropper introduces certain level of error. So calculating the error level we can detect the presence of eavesdropper.

## IV. PERFORMANCE DISCUSSION

The key rate and quantum bit error rate (QBER) are widely used as the criteria of QKD system performance. For an active QKD network, it is also important to achieve a short system switching time (the time used to establish the connection between Alice and Bob). We discuss these parameters here.

The point-to-point QKD key rate is determined by the mean photon number,  $\mu$ , at server’s output; the loss in the transmission fiber, including fiber attenuation, bending, coupling and connectors; the protocol related loss; and the detection efficiency of the APD. These factors are the principle influence to the key rate in QKD network systems. Besides these factors, the insertion loss in the optical switch should be taken into account.

The mean photon number,  $\mu$ , at Alice is assumed to 0.1 (as it is typical value). The loss in the transmission fiber is 2dB/km, while other losses such as bending, coupling and connection losses in the quantum channel amount to approximately 3 dB. The silicon APD’s detection efficiency is approximately 50%.

### A. Advantages of Using Entanglement:

1) A major benefit of this method is that the server cannot cheat or eavesdrop on the key without being caught with overwhelming probability [7]. In other words, the server does not need to be trusted.

2) There is no encoding of a random number to form the basis of the key, as the randomness comes from the entangled state, e.g.,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2) \quad (2)$$

which is in a superposition of two possible 100% correlated states ( $|1\rangle = |V\rangle$ ,  $|0\rangle = |H\rangle$  in a polarization system).

3) The entanglement allows for “automatic source checking” [8]. In systems in which the various qubit states are produced by several different lasers (or even single-photon generators), information about the state of the qubit can be leaked to other degrees of freedom (thus allowing an eavesdropper to detect the qubit state without inducing any errors). This is prevented if entangled photons are used—any leakage of information to other degrees of freedom of the photon automatically shows up as an increased bit error rate (BER). (Note: leakage of information via some classical means, e.g., detector after pulsing [9], is not eliminated using entanglement.) The security of the key exchange is also not compromised even if the source itself is in the hands of an eavesdropper.

4) It is, in principle, possible to store the photons in some “quantum memory” until the key is required. The key does not exist until the photons are measured. As a corollary, this means that one can, in principle, generate a key even when no quantum channel is available (as long as it was previously available and the quantum bits can be stored). Also, as long as no eavesdropper was present at the time of entanglement distribution, the protocol is secure, even if the measurements are not made—and the secret key not created—until some later time when there is an eavesdropper.

5) The maximum range to date is tens of kilometers in fiber; in free space, only table-top demonstrations have been carried out, though very recently entangled photons were distributed (without key exchange) over ~1km in free space[4]. In principle the range of secure communication is high because the background-induced coincidence rate (leading to errors) can be extremely low [10]. Lumped loss tolerance up to 50 dB is expected, based on a 1ns gate and 1000 counts.sec-1 (this implies a background rate of only 1 millicoincidence.sec-1; however, detector noise will increase this.) The availability of “quantum repeaters” would also increase the usable distance [11].

6) One can remove empty pulses, since the detection of one photon of a pair reveals the presence of a companion. In principle, it is thus possible to have a probability of emitting a non-empty pulse equal to one. It is beneficial only because

presently available single-photon detectors feature high dark count probability. The difficulty to always collect both photons of a pair somewhat reduces this advantage.

7) One frequently hears that photon-pairs have also the advantage of avoiding multi-photon pulses, but this is not correct. For a given mean photon number, the probability that a nonempty pulse contains more than one photon is essentially the same for weak pulses and for photon pairs. However, the fact that passive state preparation can be implemented prevents multiphoton splitting attacks [12, 13], since the different photon pairs are independent. So, even if multiple photon pairs don't increase Eve's information, they lead to an increase in the QBER [12].

8) Using entangled photons pairs prevents unintended information leakage in unused degrees of freedom [14]. Observing a QBER smaller than approximately 15 %, or equivalently that Bell's inequality is violated, indeed guarantees that the photons are entangled and so that the different states are not fully distinguishable through other degrees of freedom.

9) Entangled photons offer interesting possibilities in the context of cryptographic optical networks. The photon pairs source can indeed be operated by a key provider and situated somewhere in between potential QC customers. In this case, the operator of the source has no way to get any information about the key obtained by Alice and Bob.

### B. Drawbacks of Using Entanglement:

1) Pair-photon generation rates limit systems at the moment. Highest rates are  $4 \times 10^5$  to  $1.3 \times 10^6$  coincidences per second [13, 14] measured in the laboratory, significantly lower in fibers.

2) The disadvantage clearly is the increased experimental difficulty.

3) Decoherence is more serious when using photon pairs, because of the larger spectral width. In case of polarization-entangled photons, this gradually destroys their correlation. Hence, although perfectly fine for free-space QC, polarization entanglement is thus not adequate for QC over long optical fibers. A similar effect arises when dealing with energy-time entangled photons. Here, the chromatic dispersion destroys the strong time-correlations between the photons forming a pair. However, it is possible to passively compensate for this effect using either additional fibers with opposite dispersion, or exploiting the inherent energy correlation of photon pairs. Recent studies suggest that faint laser QC is secure over distances of about 50 km that are possible with the detectors available today [15]. Hence in the short term entangled photons won't be used for the realization of industrial prototypes.



C. Our proposed system provides the following extra benefits:

- 1) More efficient (3 dB more efficient than the existing protocols).
- 2) It is cost effective.
- 3) It is simple.
- 4) Easy to implement in laboratory environment.
- 5) Less complex solution to multiuser access.
- 6) More secure distribution system than single photon system as the protocol uses entanglement property.

## V. CONCLUSION

Security in data communication is a major headache now a days. Among the encryption technologies that are available at present, secret key is the most reliable one, which depends on secure key generation and distribution. The quantum key distribution is a perfect solution in this regard because, according to the no cloning theorem of quantum mechanics any attempt to intercept the key by the eavesdropper will be detected immediately. In this paper we have proposed a simple, more efficient and cost effective architecture and corresponding protocol for quantum key distribution. The proposed system achieves 3dB performance improvement than conventional one as it does not depend on communication of measurement base.

## REFERENCES

- [1] Kwiat, P.G., K. Mattle, H. Weinfurter, and A. Zeilinger, "New high-intensity source of polarization-entangled photon pairs", *Physical Review Letters* 75, 4337–4341 (1995); Kwiat, P.G., E. Waks, A.G. White, I. Appelbaum, and P.H. Eberhard, "Ultrabright source of polarization-entangled photons," *Physical Review A* 60, R773–R776 (1999).
- [2] Brendel, J., E. Mohler, and W. Martienssen, *Europhysics Letters* 20, 575 (1993); Kwiat, P.G., A.M. Steinberg, and R.Y. Chiao, "High-visibility interference in a Bell inequality experiment for energy and time," *Physical Review A* 47, R2472–R2475 (1993); Strekalov, D.V., T.B. Pittman, A.V. Sergienko, and Y.H. Shih, "Postselection-free energy-time entanglement," *Physical Review A* 54, R1–R4 (1996).
- [3] Rarity, J.G. and P.R. Tapster, "Experimental violation of Bell's inequality based on phase and momentum," *Physical Review Letters* 64, 2495–2498 (1990).
- [4] Aspelmeyer, M., H.R. Böhm, T. Giatso, T. Jennewein, R. Kaltenbaeck, M. Lindenthal, G. Molina-Terriza, A. Poppe, K. Resch, M. Taraba, R. Ursin, P. Walther, and A. Zeilinger, "Long-distance free-space distribution of quantum entanglement," *Science* 301, 621–623 (2003).
- [5] A. Einstein, B. Podolsky and N. Rosen, "Can quantum mechanical description of physical reality be considered complete?" *Phys. Rev.*, 47, 777–780 (1935).
- [6] [http://en.wikipedia.org/wiki/Alain\\_Aspect](http://en.wikipedia.org/wiki/Alain_Aspect)
- [7] C. H. Bennet, G. Brassard and N. D. Mermin, *Phys. Rev. Lett.* 68, 557 (1992).
- [8] Mayers, D. and A. Yao, "Quantum cryptography with imperfect apparatus," *Proceedings of the 39th Annual Symposium on the Foundations of Computer Science (FOCS98)*, K. Kelly, Ed., (IEEE, Los Alamitos, California, USA, 1998), Catalog #: 98CB36280 pp. 503–509 [quant-ph/9809039].
- [9] Kurtsiefer, C., P. Zarda, S. Mayer, and H. Weinfurter, "The breakdown flash of silicon avalanche photodiodes—back door for eavesdropper attacks?," *Journal of Modern Optics* 48, 2039–2047 (2001).
- [10] Waks, E., A. Zeevi, and Y. Yamamoto, "Security of quantum key distribution with entangled photons against individual attacks," *Physical Review A* 65, 052310 (2002).
- [11] Briegel, H.-J., W. Dür, J.I. Cirac, and P. Zoller, "Quantum repeaters: The role of imperfect local operations in quantum communication," *Physical Review Letters* 81, 5932–5935 (1998).
- [12] G. Ribordy, J. Brendel, J.D. Gautier, N. Gisin, H. Zbinden, Long distance entanglement based quantum key distribution, *Phys. Rev. A* 63, 012309 (2001).
- [13] G. Brassard, N. Lütkenhaus, T. Mor, B.C. Sanders, "Limitations on practical Quantum Cryptography", *Phys. Rev. Lett.* 85, 1330 (2000).
- [14] D. Mayers and A. Yao, "Quantum Cryptography with imperfect Apparatus", *Proc. of the 39th IEEE Conf. on Found. of Computer Science* (1998).
- [15] Stéphane Félix, N. Gisin, A. Stefanov, H. Zbinden, "Faint laser quantum key distribution: Eavesdropping exploiting multiphoton pulses", *J. of Mod. Optics*, 48 (13), 2009–2021 (2001)

# Symbol Timing Estimation using Near ML Techniques and Statistical Performance Evaluation for Binary Communications

Md. Tofazzal Hossain # and Sithamparanathan Kandeepan \*

#National ICT Australia, Canberra Research Laboratory, ACT, Australia

\*Create-Net International Research Centre, Trento, Italy

# \*Research School of Information Sciences and Engineering, Australian National University

#Tofazzal.Hossain@rsise.anu.edu.au, \*kandeepan@ieee.org

**Abstract**—In this paper, we analyze the performance of a non-data aided near maximum likelihood (NDA-NML) estimator for symbol timing recovery in binary digital communications. The analysis of the estimator is performed for an additive noise only channel and the simulation results are extended to a flat fading channel. The probability distribution of the timing estimates is derived, presented and compared with simulation results. The performance of the estimator is presented in terms of the bit error rate (BER) and the error variance of the estimates. The BER is computed when the estimator is operating under additive white Gaussian noise (AWGN) channel and Rayleigh fading channel. The variance of the estimates is computed for the noise only case and compared with the Cramer Rao bound (CRB) and modified Cramer Rao bound (MCRB).

**Index Terms**—BER, CRB, Flat fading channel, Near Maximum Likelihood Estimation, MCRB, Symbol Timing Error, Timing Synchronization.

## I. INTRODUCTION

Timing in a receiver must be synchronized to the symbols of the incoming transmitted signal. In the analog implementation of digital modems a free running clock at the symbol rate is phase adjusted in a feedback manner to find the optimum sampling point of the symbol. In feedforward arrangement, timing wave is generated from incoming signal. Mueller and Muller [1] presented a novel technique for symbol timing recovery in the analog era. With the advancement of the digital and microprocessor systems, came digital signal processors and FPGAs, which allows us to implement complicated open loop systems.

In this paper, we present a non-data aided symbol timing estimation technique and evaluate its performance. Symbol timing can be performed with synchronous sampling where exploiting some sort of error signal, the signal sampling clock is adjusted in a feedback manner and non-synchronous sampling where the sampling is not locked to the incoming signal. Non-synchronous sampling is used in this paper to estimate the symbol timing on received baseband signal. The optimum symbol strobing point is estimated using the samples with minimum intersymbol interference. Obtained samples from nonsynchronously sampled signal may not include the sample which is at the

optimum sampling point and therefore interpolation is used to reduce the error produced by nonsynchronous sampling process. While performing interpolation and upsampling with the available high processor speeds of the latest digital signal processor, no significant limitations are noted. In this paper, interpolations are assumed available for convenience. A good review on interpolation techniques for symbol timing estimation is given by Gardner [2], [3] where some useful interpolation filter functions and the use of them for timing recovery to improve performance of nonsynchronously sampled systems are discussed. Many feedback and/or feedforward symbol timing estimation techniques for synchronous and/or nonsynchronous sampled systems for either continuous time or burst mode transmissions are available in the literature [4]. Techniques for synchronizing receivers can be divided into data aided and non-data aided. The former is the case where synchronization relies on knowledge of the information symbols. The latter is the case where training sequences are impractical or inconvenient, and the decision process is not sufficiently reliable for decision feedback [5].

The performance of the estimator can be evaluated in terms of BER performance and measuring the variance of the estimator and comparing with CRB and MCRB. The BER performance under the assumption of perfect timing estimates is well documented for various modulation formats [6]. In practice timing estimates exhibit small random fluctuations (jitter) about their optimum values that give rise to a BER degradation when compared to perfect synchronization [7], which degrades further when the signal undergoes fading in wireless communications [8]. A joint symbol timing estimation and data detection on flat fading channel based on particle filtering is demonstrated in [9]. We evaluate the performance of the symbol timing estimator operating under a Rayleigh fading channel. We consider the timing estimator as described in [4] and analyze its performance by finding the statistical distribution of the estimates, the distribution can then be used to characterize the average BER theoretically. The analysis are further extended to a flat fading Rayleigh channel using Monte-Carlo simulations. A tighter CRB

and a MCRB are also derived to study the performance of the estimator for the noise only case. Although the considered timing estimator is not optimal we consider it here because of its design simplicity which can be easily used in software defined radios at the baseband. In the following sections we present the timing estimator, the probability distribution of the estimator, the CRB and MCRB, and the performance analysis for a flat fading channel.

## II. SIGNALS AND SYSTEMS MODEL

The system considered here is depicted in figure 1. In our model we consider both the additive noise only channel as well as the fading channel. A conventional baseband pulse amplitude modulation (PAM) signal can be represented as

$$s(t) = \sum_{m=-\infty}^{\infty} c_m g(t - mT) \quad (1)$$

where  $c_m$ 's are the transmitted symbols,  $g(\cdot)$  is the pulse shape used to control the spectral characteristics of the transmitted signal and  $T$  is the signaling interval ( $\frac{1}{T}$  is the symbol rate). The  $c_m$ 's are assumed to be independently and identically distributed, and take on the values  $\pm 1$  with equal probability. The channel introduces a delay of  $\tau$ , and additive white Gaussian noise is considered at the receiver with a double sided power spectral density of  $\frac{N_0}{2}$  W/Hz. Then the signal becomes

$$\begin{aligned} y(t) &= s(t - \tau) + w(t) \\ &= \sum_{m=-\infty}^{\infty} c_m g(t - mT - \tau) + w(t) \end{aligned} \quad (2)$$

where the noise process  $w(t)$  is stationary and Gaussian with a mean of zero. The frequency response of the pulse shaping filter in the transmitter is  $G(f)$ . The received signal is passed through a receiving filter. In general, the optimum filter at the receiver is matched to the received signal pulse  $g(t)$ . Hence the frequency response of this filter is  $G^*(f)$ . Then the received signal is

$$z(t) = \sum_{m=-\infty}^{\infty} c_m g_o(t - mT - \tau) + n(t) \quad (3)$$

where  $g_o$  is the signal pulse response of the receiving filter-that is  $G_o(f) = G(f)G^*(f) = |G(f)|^2$ -and  $n(t)$  is a bandlimited Gaussian noise independent and identically distributed with the same statistics as  $w(n)$ . Now if  $z(t)$  is sampled at times  $kT_s$ ,  $k = 0, \pm 1, \pm 2, \dots$  where  $T_s$  is the sampling interval, we have

$$z(kT_s) = \sum_{m=-\infty}^{\infty} c_m g_o(kT_s - mT - \tau) + n(kT_s) \quad (4)$$

When the symbols are transmitted over flat fading channel

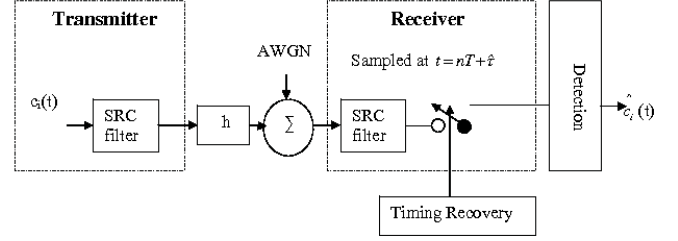


Fig. 1: Block diagram of signal detection using square root raised cosine filter.

$h(t)$ , the faded discrete signal at the baseband can be written as,

$$z^f(kT_s) = \sum_{m=-\infty}^{\infty} h(kT_s) c_m g_o(kT_s - mT - \tau) + n(kT_s) \quad (5)$$

where  $h(kT_s)$  denotes the multiplicative fading process introduced by the flat fading channel sampled at times  $kT_s$ ,  $k = 0, \pm 1, \pm 2, \dots$ . As is shown in figure 1, square root raised cosine (SRC) filter is used as the pulse shaping filter and also as the receiving filter, the overall response is therefore similar to a raised cosine filter. The impulse response of SRC filter is given by:

$$g_{src}(t) = \begin{cases} \frac{\pi(1-\alpha)+4\alpha}{\pi\sqrt{T}} & t = 0 \\ \frac{4\alpha}{\pi\sqrt{T}} \left( \frac{\cos\left(\frac{(1+\alpha)\pi t}{T}\right) + \sin\left(\frac{(1-\alpha)\pi T}{4\alpha T}\right)}{1 - \left(\frac{4\alpha t}{T}\right)^2} \right) & t \neq 0, \pm \frac{T}{4\alpha} \\ \frac{\alpha}{\pi\sqrt{2T}} \left( (\pi - 2)\cos\left(\frac{\pi}{4\alpha}\right) + (\pi + 2)\sin\left(\frac{\pi}{4\alpha}\right) \right) & t = \pm \frac{T}{4\alpha} \end{cases} \quad (6)$$

where  $\alpha$  is called the roll-off factor, which takes values in the range  $0 \leq \alpha \leq 1$  and  $T$  is the symbol period. The theoretical bit error probability  $P_b$  for bipolar signaling is

$$P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (7)$$

where  $Q(x)$  is called the complementary error function or co-error function and is defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{u^2}{2}\right) du \quad (8)$$

## III. TIMING ESTIMATOR

Here a near maximum likelihood (NML) timing estimation technique is described to estimate the symbol timing error as described in [4]. We assume that the sampled received baseband waveform for the additive noise only channel as given by (4) is:

$$z(kT_s) = \sum_{i=0}^{L-1} c_i g_o(kT_s - iT - \tau) + n(kT_s). \quad (9)$$

The noise process  $n(kT_s)$  is a zero mean bandlimited Gaussian process with a variance of  $\sigma^2$  and  $L$  is the number of symbols. In this case we ignore the channel  $h(kT_s)$  to derive the estimator. The probability density function of  $n(kT_s)$  is given by:

$$f_n(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{\{n(kT_s)\}^2}{2\sigma^2} \right] \quad (10)$$

If we let  $\mathbf{R}$  to be a vector of the sample set  $z(kT_s)$  then using (9) and (10)

$$f_{\mathbf{R}}(\mathbf{R}|\tau) = \frac{1}{(2\pi\sigma^2)^{\frac{L}{2}}} \exp \left\{ -\frac{\sum_{i=0}^{L-1} [E - F]^2}{2\sigma^2} \right\} \quad (11)$$

where

$$\begin{aligned} E &= z_i(kT_s) \text{ and} \\ F &= \sum_{i=0}^{L-1} c_i g_0(kT_s - iT - \tau) \end{aligned}$$

We use the Maximum Likelihood (ML) principles to derive the symbol timing estimator assuming no a-priori knowledge of the received symbol sequence  $c_i$ . Let a matrix of samples be given as

$$\mathbf{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_i \\ \vdots \\ Z_L \end{bmatrix} = \begin{bmatrix} z_1(1) & z_1(2) & \cdot & \cdot & z_1(N) \\ z_2(1) & z_2(2) & \cdot & \cdot & z_2(N) \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ z_i(1) & z_i(2) & \cdot & \cdot & z_i(N) \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ z_L(1) & z_L(2) & \cdot & \cdot & z_L(N) \end{bmatrix} \quad (12)$$

where,

$$Z_i = \sum_{i=1}^L z_i(k') \quad (13)$$

and,

$$z_i(k') = \sum_{k'=1}^N c_m g_0(k'T_s - iT - \tau) + n(k'T_s) \quad (14)$$

Here,  $k' = \text{mod}(k, N)$ ,  $L$  is the number of symbols used to estimate the timing offset and  $N$  is the number of samples per symbol. The estimator is then given by:

$$\hat{\tau} = \arg \max \{ \Gamma_1(k', \tau) = \sum_{i=1}^L |z_i(k', \tau)| \} \quad (15)$$

The validity of the technique demands the constraint on number of samples to be integer multiple of symbols. When the estimator operates under fading channel, it becomes,

$$\hat{\tau} = \arg \max \{ \Gamma_2(k', \tau) = \sum_{i=1}^L |z_i^f(k', \tau)| \} \quad (16)$$

where

$$z_i^f(k') = \sum_{k'=1}^N h(k'T_s) c_m g_0(k'T_s - mT - \tau) + n(k'T_s) \quad (17)$$

#### IV. PROBABILITY DISTRIBUTION OF THE TIMING ESTIMATOR

In order to determine the probability of bit error in the receiver, the distribution of the timing estimate errors must be taken into account. When using feedback techniques, the probability density function (pdf) of the timing error may be modelled using a Tichanov pdf [6]. Alternatively if the signal to noise ratio (SNR) is greater than the inverse of twice the variance of the timing offset a Gaussian pdf may be used [10], and since this holds true for the feed forward symbol timing recovery schemes, a Gaussian pdf can be used. However, in our case the timing estimates relate to the maximum energy of the signal. The instant of time at which the signal energy becomes maximum is taken as the estimate. Hence, the distribution of the timing estimates corresponds to the distribution of the maxima of the signal. Considering this, we find an expression for the probability mass function of timing estimates. The analytical expression is then compared with the simulated distribution. Let  $N = 2N' + 1$ , and consider the two vectors,  $\tau = [\tau_{-N'}, \dots, \tau_0, \dots, \tau_{k'}, \dots, \tau_{N'}]$  and  $\Gamma_1(k', \tau) = [\Gamma_1(-N'), \dots, \Gamma_1(0), \dots, \Gamma_1(k'), \dots, \Gamma_1(N')]$  with  $q = \max(\Gamma_1(k', \tau))$ . Then, when the particular time instance  $\tau_{k'}$  is chosen as the timing estimate  $\hat{\tau}$ , that is  $\hat{\tau} = \tau_{k'}$ , the probability can be written as

$$\begin{aligned} P(\hat{\tau} = \tau_{k'}) &= P(q = \Gamma_1(k', \tau)) \\ &= P(q > \Gamma_1(-N'), \dots, q = \Gamma_1(k'), \dots, q > \Gamma_1(N')) \\ &= f_{\Gamma_1(k')}(q) \prod_{j=-N', j \neq k'}^{N'} F_{\Gamma_1(j)}(q) \end{aligned} \quad (18)$$

where,  $f_{\Gamma_1(k')}(q)$  and  $F_{\Gamma_1(j)}(q)$  are the probability density function and cumulative distribution function of  $\Gamma_1(k')$  respectively. The expression for the probability mass function (pmf) for the timing estimates is given by,

$$f_{\hat{\tau}}(\tau) = \begin{cases} P(\hat{\tau} = \tau); & \tau \in [\tau_{-N'}, \dots, \tau_{k'}, \dots, \tau_{N'}] \\ 0; & \tau \notin [\tau_{-N'}, \dots, \tau_{k'}, \dots, \tau_{N'}]. \end{cases} \quad (19)$$

Figure 2 shows the analytical probability mass function and the simulated probability density function for the timing estimates. It is seen that the analytical results closely match with the simulated pdf.

#### V. PERFORMANCE OF THE ESTIMATOR

The statistics of the estimates are the key elements in determining the performances of the symbol timing recovery method. For example, if the estimates are biased, the resulting sampling location selected by the symbol timing estimator will be biased from the true optimum

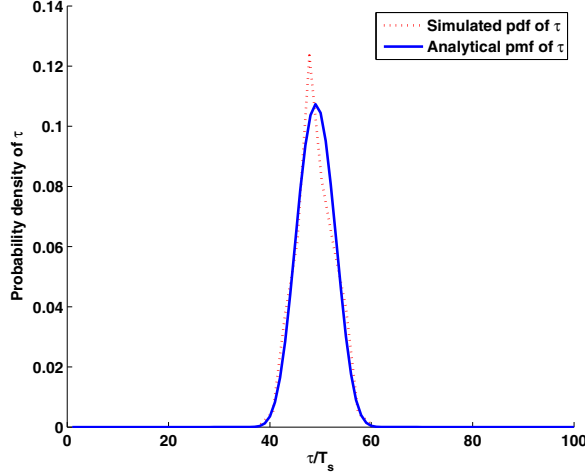


Fig. 2: Analytical and simulated probability distribution function for timing estimates.

and performance will suffer. Similarly, if the estimates have a high variance, the variance of the selected sampling location will also be high. Clearly it is desirable to keep the variance of the estimates as low as possible. The performance bounds for the symbol timing estimator are the CRB and MCRB. We evaluate the performance of the estimator finding bit error probability with matched filter detection and measuring variance of the estimator. Bit error probability is compared to the theoretical bit error probability for AWGN and variance is compared with the lower bounds CRB and MCRB. Figure 3 shows the bit error performance for the NDA-NML estimator when roll-off factor is 0.6 and the number of symbols used for timing estimates is 100. NDA-NML estimator exhibits good bit error performance compared with theoretical BER of AWGN.

#### A. Cramer Rao Bound (CRB)

We consider the additive noise only case to derive the CRB here, assume that the received continuous waveform as given in (3) has complex envelope given by

$$z(t) = s(t) + n(t) \quad (20)$$

which is observed over an interval  $T$ .  $s(t)$  is the information bearing signal and  $n(t)$  represents complex-valued additive white Gaussian noise with two-sided power spectral density. A few of its parameters are unknown. Typically this happens with the carrier phase  $\theta$ , the symbol epoch  $\tau$ , the carrier frequency error  $\nu$  and the transmitted data.

To estimate of a single element of  $\{\theta, \tau, \nu\}$  denoted by  $\lambda$ , which is assumed to be deterministic (non-random), all the other parameters, including the data, are collected in a random vector  $\mathbf{u}$  having a known probability density

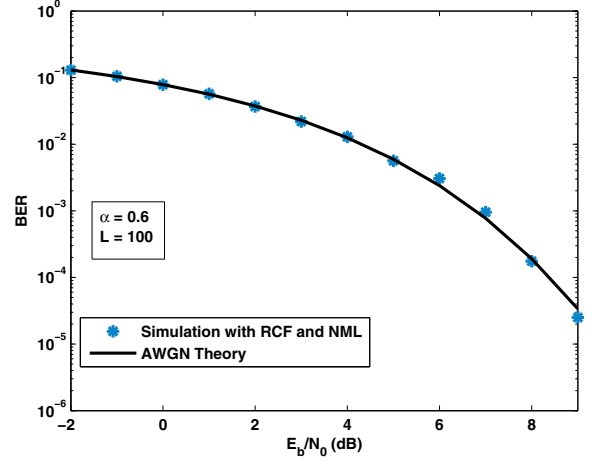


Fig. 3: Bit error performance: NDA-NML estimator.

function  $p(\mathbf{u})$  which does not depend on  $\lambda$ . If  $\hat{\lambda}(\mathbf{z})$  is any unbiased estimator of  $\lambda$ , a lower bound to the variance of the error  $\hat{\lambda}(\mathbf{z}) - \lambda$  is given by the Cramer-Rao formula [11]

$$CRB(\lambda) = \frac{1}{E_{\mathbf{z}} \left\{ \left[ \frac{\partial \ln p(\mathbf{z}|\lambda)}{\partial \lambda} \right]^2 \right\}} \quad (21)$$

where  $E_{\mathbf{z}}$  denotes statistical expectation with respect to the subscripted variable, and  $p(\mathbf{z}|\lambda)$  is the probability density function of  $\mathbf{z}$  for a given  $\lambda$ .

To compute  $CRB(\lambda)$  we need  $p(\mathbf{z}|\lambda)$  which, in principle, can be obtained from the integral

$$p(\mathbf{z}|\lambda) = \int_{-\infty}^{\infty} p(\mathbf{z}|\mathbf{u}, \lambda) p(\mathbf{u}) d\mathbf{u} \quad (22)$$

where  $p(\mathbf{z}|\mathbf{u}, \lambda)$ , the conditional probability density function of  $\mathbf{r}$  given  $\mathbf{u}$  and  $\lambda$  is easily available, at least for additive Gaussian channels. Unfortunately, in most cases of practical interest, the computation of (21) is impossible because either the integration in (22) cannot be carried out analytically or the expectation in (21) poses insuperable obstacles. Using the general form of CRB given by [12], the CRB for timing estimates for AWGN channel can be written as:

$$\overline{CRB}\left(\frac{\tau}{T}\right) = \left[ -\frac{2E_b}{N_0} g_0''(0) L T^2 \right] \quad (23)$$

where  $L$  represents the number of symbols used to make the estimate of the timing offset,  $T$  is the symbol period,  $g_0''(0)$  is the second derivative of the pulse shaping filter with respect to time  $t$  evaluated at  $t = 0$ . For a raised cosine channel when the signal energy normalized to unity,  $g_0''(0)$  is given by [13]

$$g_0''(0) = -\left( \frac{\pi^2}{3} + \alpha^2(\pi^2 - 8) \right) \quad (24)$$



where  $\alpha$  is the roll off factor of the raised cosine filter.

### B. Modified Cramer Rao Bound (MCRB)

The MCRB by [14] to the variance of  $\hat{\lambda}(\mathbf{z}) - \lambda$  is the following

$$MCRB(\lambda) = \frac{1}{E_{\mathbf{z},\mathbf{u}} \left\{ \left[ \frac{\partial \ln p(\mathbf{z}|\mathbf{u},\lambda)}{\partial \lambda} \right]^2 \right\}} \quad (25)$$

This bound is found by [14] observing that

$$\begin{aligned} E_{\mathbf{z},\mathbf{u}} \left\{ [\hat{\lambda}(\mathbf{z}) - \lambda]^2 \right\} &= E_{\mathbf{u}} \left\{ E_{\mathbf{z}|\mathbf{u}} [(\hat{\lambda}(\mathbf{z}) - \lambda)^2] \right\} \\ &\geq E_{\mathbf{u}} \left\{ \frac{1}{E_{\mathbf{z}|\mathbf{u}} \left[ \left( \frac{\partial \ln p(\mathbf{z}|\mathbf{u},\lambda)}{\partial \lambda} \right)^2 \right]} \right\} \\ &\geq \frac{1}{E_{\mathbf{u}} \left\{ E_{\mathbf{z}|\mathbf{u}} \left[ \left( \frac{\partial \ln p(\mathbf{z}|\mathbf{u},\lambda)}{\partial \lambda} \right)^2 \right] \right\}} \\ &= \frac{1}{E_{\mathbf{z},\mathbf{u}} \left\{ \left[ \frac{\partial \ln p(\mathbf{z}|\mathbf{u},\lambda)}{\partial \lambda} \right]^2 \right\}} \end{aligned} \quad (26)$$

where the first inequality derives from application of the CRB to the estimator  $\hat{\lambda}(\mathbf{r})$  for a fixed  $\mathbf{u}$ , while the second is true in view of Jensen's inequality and the convexity of the function  $\frac{1}{x}$  for  $x > 0$ .

In spite of having same structure of CRB and MCRB, MCRB is much easier to use. For Gaussian Channel, the probability density for (25) is a well-known exponential function whose argument is a quadratic form in the difference between  $\mathbf{z}$  and the signal  $\mathbf{s}$ . Thus the logarithm of  $p(\mathbf{z}|\mathbf{u},\lambda)$  equals this quadratic form and the expectation in (25) is readily derived.

The MCRB of symbol epoch ( $\tau$ ) by [5] is given by

$$MCRB(\tau) = \frac{1}{8\pi^2 L \xi} \frac{T^2}{E_s/N_0} \quad (27)$$

where  $L$  is the number of symbol intervals on which receiver bases its timing estimate on the observation on the received signal,  $T$  is the symbol interval,  $E_s$  is the energy of one symbol waveform and  $\xi$  is an adimensional coefficient depending on the shape of  $g_0(t)$ :

$$\xi = \frac{\int_{-\infty}^{\infty} T^2 f^2 |G(f)|^2 df}{\int_{-\infty}^{\infty} |G(f)|^2 df} \quad (28)$$

MCRB for a normalized  $\tau$  is given by [15] where  $\tau$  is between 0 and 1. The MRCB is

$$MCRB(\tau) \geq \frac{1}{2L\Gamma(E_s/N_0)} \quad (29)$$

where

$$\Gamma = \frac{T^2}{E_s} \int_0^T g_d^2(t) dt \quad (30)$$

where  $g_d(t)$  denote the time derivative of  $g_0(t)$ ,  $E_s$  is the symbol energy. For a raised cosine pulse  $\Gamma = \frac{4\pi^2}{3}$ .

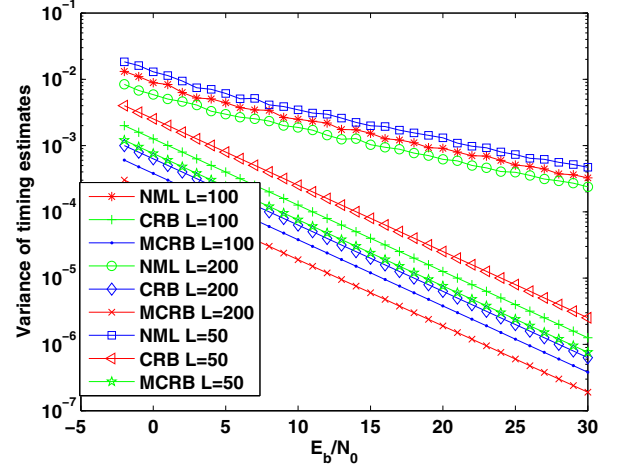


Fig. 4: NDA-NML estimator error variance performance.

Figure 4 shows variance of the NDA-NML estimator with respect to CRB and MCRB. It is demonstrated in [14] that MCRB is generally lower than (at most equal) to the true CRB. The simulation result exhibits that the CRB is always greater than MCRB and the variance of the estimator is greater than CRB and MCRB as well. The variance of the estimator decreases as the number of symbols contributed for the timing estimation increases. It is easy to implement the NDA-NML estimator and the complexity is low as well.

### VI. TIMING ESTIMATOR OPERATING UNDER RAYLEIGH FADING CHANNEL

The statistical characteristics of a fading channel can be modelled using various probability density functions. In mobile radio channels, the Rayleigh distribution is commonly used to describe the statistical time varying nature of the received envelope of a flat fading signal, or the envelope of an individual multipath component. It is well known that the envelope of the sum of two quadrature Gaussian noise signals obeys a Rayleigh distribution [16], [17], [18]. The Rayleigh distribution has a probability density function given by

$$f(r) = \begin{cases} \frac{r}{\sigma_h^2} \exp\left(-\frac{r^2}{2\sigma_h^2}\right), & 0 \leq r \leq \infty \\ 0, & r < 0 \end{cases} \quad (31)$$

where  $\sigma_h$  is the root mean square (rms) value of the received signal before envelope detection. In our system model the sampled received signal experiencing Rayleigh fading can be expressed by (5) where the fading process  $h(kT_s)$  is Rayleigh-distributed. If we use  $h$  in place of  $h(kT_s)$ , it can be expressed as:

$$h = (\sqrt{h_I^2 + h_Q^2}) \quad (32)$$

where  $h_I$  and  $h_Q$  are uncorrelated and independent of each other and follow correlated Gaussian distribution. The correlations among the samples are achieved passing the Gaussian distributed samples through a spectral shaping filter. In our case we use a digital one dimensional filter to correlate the samples. The coefficients of the filter are given by:

$$\mathbf{A} = \left[ 1 \quad -\exp\left(-\frac{\gamma}{Fs}\right) \right], \text{ and } \mathbf{B} = \gamma \quad (33)$$

where  $\gamma$  defines the amount of correlation in  $h(kT_s)$  and  $F_s$  is sampling frequency. The filter is a "Direct form II Transposed" implementation of the standard difference equation:

$$A(1)y(n) = B(1)x(n) - A(2)y(n-1) \quad (34)$$

where  $X$  is the input to the filter and  $Y$  is the filtered output. The channel energy is maintained to be unity so that the received SNR is not changed. Figure 5 shows

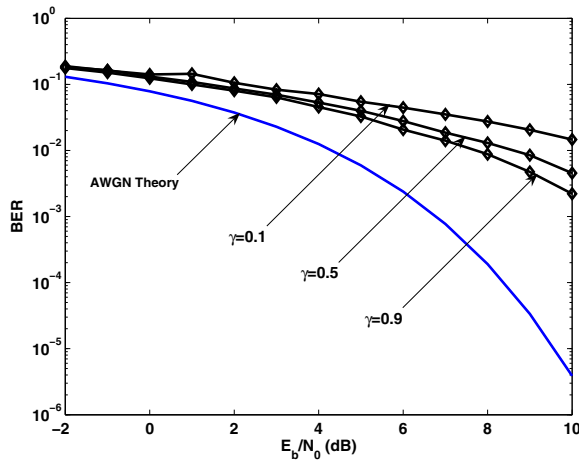


Fig. 5: Bit error performance of NDA-NML estimator under fading conditions.

the bit error performance for timing estimator operating under Rayleigh fading channel. The amount of correlation has a significant effect on BER performance. Figure 5 exhibits that the BER performance for the timing estimator improves as the  $\gamma$  increases under fading conditions.

## VII. CONCLUSION

Statistical performance of a NDA-NML timing estimator was presented. Several interesting aspects of the timing estimator were revealed examining the statistical distribution of the estimator. The expression for the probability mass function of the timing estimates was derived and compared with the simulated density function. The performance of the estimator was assessed considering the BER, and the error variance was compared to the CRB and MCRB for AWGN channel. The performance of the estimator

was also studied when operating under a Rayleigh flat fading channel by calculating the BER for different fading conditions.

## ACKNOWLEDGMENT

The authors would like to thank National Information and Communication Technology (ICT) Australia. National ICT Australia is funded through the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council.

## REFERENCES

- [1] K. H. Mueller and M. Muller, "Timing recovery in digital synchronous data receiver", *IEEE Trans. Communications.*, vol. COM-24, no. 5, pp. 516-531, May 1976.
- [2] F. M. Gardner, "Interpolation in digital modems-part I: fundamentals", *IEEE Trans. Communications*, vol. 41, no. 3, pp. 501-507, Mar. 1993.
- [3] F. M. Gardner, "Interpolation in digital modems-part II: implementation and performance", *IEEE Trans. Communications*, vol. 41, no. 6, pp. 998-1007, June 1993.
- [4] S. Kandeepan and S. Reisenfeld, "DSP based symbol timing estimation techniques", *Proc. IEEE ICICS-PCM'03*, vol.1, pp. 568-572, Dec. 2003.
- [5] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, New York and London: Plenum Press, 1997.
- [6] W. C. Lindsley and M. K. Simon, *Telecommunication Systems Engineering*, New Jersey: Prentice Hall, 1973.
- [7] K. Bucket and M. Moeneclaey, "Timing jitter sensitivity comparison of narrowband M-PSK and bandlimited DS/SS M-PSK signals", *IEEE Trans. Communications*, vol. 1, no. 3, pp. 453-457, May 1993.
- [8] W. K. M. Ahmed and P. J. McLane, "On the error exponent for memoryless flat fading channels with channel-state-information feedback", *IEEE Communications Letters*, vol. 3, no. 2, pp. 49-51, Feb. 1999.
- [9] T. Ghirmai, "Data detection and fixed symbol-timing estimation in flat fading channels", *Proc. IEEE CISS*, pp. 456-461, Mar. 2007.
- [10] J. J. Stiffler, *Theory of Synchronous Communications*, Prentice Hall, 1971.
- [11] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall PTR, 1993.
- [12] T. Jesupret, M. Moeneclaey and G. Ascheid, "Digital demodulator synchronisation-performance analysis", *European Space Agency Rep. ESTEC Contact No. 8437/89/NL/RE*, June 1991.
- [13] C. N. Georgiades and M. Moeneclaey, "Sequence estimation and synchronization from nonsynchronised samples", *IEEE Trans. Inform. Theory*, vol. 37, pp. 1649-1657, Nov. 1991.
- [14] A. D'Andrea, U. Mengali and R. Reggiannini, "The modified Cramer Rao bound and its application to synchronization problems", *IEEE Trans. Communications*, vol. 42, no. 2/3/4, pp. 1391-1399, Feb./Mar./Apr. 1994.
- [15] S. C. White and N. C. Beaulieu, "Improved bounds for timing estimation jitter", *Proc. IEEE GLOBECOM'90*, vol. 1, pp. 609-616, Dec. 1990.
- [16] E. Biglieri, J. Proakis and S. Shamai, "Fading channels: information theoretic and communications aspects", *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.
- [17] C. Loo and N. Secord, "Computer models for fading channels with applications to digital transmission", *IEEE Trans. Vehicular. Tech.*, vol. 40, no. 4, pp. 700-707, Nov. 1991.
- [18] T. S. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall, 1996.

# Layer-2 Protocol Adaptation Method to Improve Fast Handoff for Mobile IPv6 Vertical Handoffs

Vahid Solouk<sup>1,1</sup>, Borhanuddin M. Ali<sup>2</sup>, Sabira Khatun<sup>1</sup>, Daniel Wong<sup>3</sup>, and M. Adzir Mahdi<sup>1</sup>

1. University Putra Malaysia, Serdang, Malaysia,      2. MIMOS Berhad, Malaysian Technology Park, Malaysia,  
3. Malaysian University of Science and Technology, Malaysia

**Abstract** — Inter-technology roaming is known as one of the interesting challenges toward fourth generation of mobile and wireless communication. While FMIPv6 standardizes the fast handoff solutions in IP layer, the issues of media independency are being investigated through IEEE802.21 project. The integration of these two standards is believed to result in solutions for vertical handoffs between different network technologies. This paper presents an improved link layer mechanism to assist FMIPv6 for seamless vertical handoffs. We introduce a new access router discovery method and propose a vertical handoff algorithm accordingly. Further, we report the implementation details performed through simulations. The simulations evidence performance improvements in terms of latency and packet loss. It is also analytically shown that by enabling access router discovery method and improving link layer event services, an MN can be well prepared for handoff and perform faster movements.

**Index Terms** — Vertical handoff, Heterogeneous networks, FMIPv6, MIH, Access router discovery.

## I. INTRODUCTION

The coexistence of various communication systems with different network characteristics (i.e. bandwidth, delay, cost etc) requires seamless mobility solutions to facilitate a Mobile Node (MN) for inter-technology roaming which in turn requires a seamless inter-system handoff. Fast handoff for Mobile IPv6 (FMIPv6) provides solutions for intra-system handoffs in wireless IP networks. IEEE802.21 on the other hand, deploys link layer solutions for heterogeneous media in inter-system handoff. Alongside the standards, the integration of these two protocols has been the concern of several proposals which were based on using link layer triggers for access network independency purpose and preparing information for FMIPv6 protocol which in turn, can result in seamless handoffs for heterogeneous networks. Further to the use of specific link layer messages, we believe that the use of a router discovery method in a proper timing fashion enables MN to learn about necessary IP layer information and the prospective network ahead of time which results in lower handoff latencies together with smaller losses in packet stream. Hence, we introduce an improved access router discovery (i-ARD) method in this paper together with an improved link layer mechanism that assist FMIPv6 and perform seamless vertical handoffs. It is shown through

analysis that by enabling the proposed i-ARD method and the improved link layer event services, an MN can be well prepared for handoff and can perform faster movements. In order to evaluate the performance of the proposed method, we implemented our model and FMIPv6 through simulations and made a comparison in terms of latency and packet loss the two models cause with different types of packet streams. The rest of the paper is organized as follows: Section 2 introduces the structure of FMIPv6 standard and discusses the process IEEE802.21 media independent handoff, followed by a report on the other related researches in integrating these methods. Section 3 describes the proposed model and algorithms of i-ARD and improved layer-2 operations in details. In section 4, we report on the simulation implementations of the proposed method which is followed by the results achieved in the latter part. Section 6 presents an analytical performance inspection of the proposed model and comparisons made to FMIPv6 and finally, the conclusions make the last section.

## II. RELATED WORKS

This part gives a comprehensive explanation of the standards and ongoing works.

### A. FMIPv6

FMIPv6 [1] is a layer-3 solution which deals merely with IP-layer messages and initially intends to overcome the drawbacks of MIPv6 [2] basic handoff addressing the latencies introduced through the operations of movement detection, CoA stateless autoconfiguration including DAD procedure, and BU. The protocol mainly focuses on two problems: start of sending packets by MN immediately after new subnet detection, and sending packets by nAR immediately after MN is detected. This is performed during Handoff Initiation (HI) period while MN is connected to the network through old AR (oAR). During this period, MN sends Router Solicitation for Proxy Advertisement (RtSolPr) message to oAR and receives the advertised prefixes of nAR through Proxy Router Advertisement (PrRtAdv) message. This is a link layer mechanism and delivers the L3 information of the new

<sup>1</sup> Correspondent author email: vsolouk@ieee.org



network from oAR to MN. FMIPv6 protocol assumes the existence of a mapping table of L2 and L3 information in oAR by which neighboring MAC addresses are resolved to the corresponding AR prefixes.

Although FMIPv6 was designed to overcome the major drawbacks of MIPv6 handoff, it still leaves some deficiencies behind together with introducing other shortcomings. Primarily, due to the usage of limited L2 triggers in Handoff Initiation (HI) mode, it lacks a precise anticipation which may cause falsely-picked prospective ARs. While mapping table is supported by AR by exchanging neighboring information between ARs, FMIPv6 does not introduce any protocol or solution of implementing it in oAR. Furthermore, in order to overcome long DAD procedure problem, the whole operation is carried out during HI mode while the old connection is maintained. However, if DAD takes longer than normal, the connection to oAR might fail. This condition is a possibility and FMIPv6 suggests reactive mode of handoff which is equal to basic MIPv6 handoff in terms of the latencies occurred. Another process delay is introduced by FMIPv6 during neighbor discovery procedure during the exchange of NS and NA messages. This process in MN together with the message exchange duration might take several seconds and despite the use of Fast Neighbor Advertisement (FNA) in FMIPv6, the process is likely to take about 800 ms.

### B. IEEE802.21

IEEE802.21 standard [3] is being developed to enable handoff and interoperability in heterogeneous networks as Media Independent Handover (MIH). The proposed handoff procedure utilizes the information achieved from both MN and the network side. The link layer of this framework is focused on providing link intelligence information of the networks in neighborhood to the upper layers for an optimized handoff between heterogeneous networks. The link intelligence information includes network discovery and selection processes. IEEE802.21 introduces a new entity called MIH Function (MIHF) as a

shim layer between layer-2 and layer-3. The main functionality of MIHF is to provide three groups of services to upper layers or MIHF users from the area under the coverage of different networks. This is done by making use of a unified interface which is called Service Access Point (SAP). This unified interface is found in two modes: a Media Dependent mode that interfaces MIHF to heterogeneous link layers, and Media Independent mode which is the interface to upper layers and exposes services to MIHF users. The provided services reflect three groups of information and operations respectively.

The deployment of service primitives described above by IEEE802.21 is claimed to have the considerations of supporting various types of layer-3 mobility management protocols, especially MIP, MIPv6 and SIP. However as MIHF shim layer focuses mainly on solving media independency problem, it operates closer to layer-2 than dealing with mobility management protocols of layer-3. In order to reach an optimized seamless handoff in heterogeneous networks a fair interoperability between layer-2 operations and layer-3 mobility management protocols, particularly handoff management protocols seems to be inevitable.

### C. Integration of FMIPv6 and IEEE802.21

Despite the fact that FMIPv6 and IEEE802.21 are both aiming to improve handoff in different aspects, FMIPv6 merely deals with radio access heterogeneity and layer-2 mechanisms and focuses mostly on IP-layer messages. In contrast to FMIPv6, the aim of IEEE802.21 is to give a solution to vertical handoff while it deals mostly with media independency and improving layer-2 mechanisms. On one hand, FMIPv6 is claimed to be independent of the whole link-layer features and link specifications are considered to be customizable, on the other hand, while MIH services of IEEE802.21 are useful for link layer triggers, improving MIPv6-based vertical handoffs is very limited with the existing MIH primitives due to the fact that 802.21 is merely concerned with detecting layer-2 information.

Integrating these two protocols to optimize vertical handoff has been the concern of few proposals in this regard. Inadequacy of MIH primitives (MIES primitives as claimed) is the motivation of a work proposed in [4] where a new primitive, namely MIH-PrefixInfo including the prospective AR info was adjoined to L2 events and based on modified events, a similar mechanism to FMIPv6 for handoff is proposed. Although this work originally addresses the issues of anticipation and hesitant mobility in FMIPv6, a method was not indicated for AR discovery and information gathering from the neighborhood. In addition to the cost

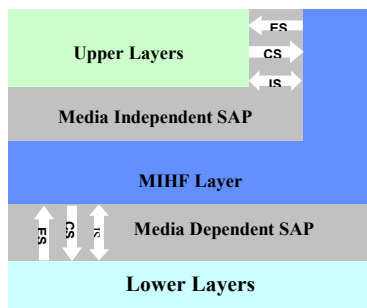


Figure 1: MIHF shim layer interconnection

of improving L2 primitives, the proposed handoff mechanism leaves more deployment issues in AR.

Another solution was suggested in [5] and through a protocol for discovering AR information called Access Router Information Protocol (ARIP). This approach is based on IETF SEAMBOY working group project [6] defined as Candidate Access Router Discovery (CARD). The method uses Radio Access Network Discovery (RAND) protocol to propose a complete layer-2 solution while it still involves layer-3. The method is independent of the mobility management and can be applied to both homogeneous and heterogeneous environments as it does not rely on radio access technology. The information on neighboring ARs (ARIP) is provided at AR and sent to MN. However, the protocol suggests no method on how to collect ARIP info from AR and how the procedure should be initialized. In addition, the protocol still requires AR deployment which is a technology obstacle.

Improving MIIS services was also the concern of few other proposals. Information in MIIS is specified in common formats and as mentioned earlier these information aid handoff decision-making process. Selecting a higher layer mechanism of mobility management to obtain information of neighboring networks from different access technologies is how MIIS information primitives are utilized in [7] and [8]. The higher layer mobility management is a SIP-based mechanism and the method was tested with an MN having two neighboring subnets to select. These approaches suggest the MIH information be obtained by MN through several query/response messages to estimate the network. Once the network is chosen, MIH info reaches MN through request/response messages between MN and nAR. These works are mainly focused on improving SIP-based mobility management particularly handoff and the results were also compared accordingly.

### III. PROPOSED MODEL

As discussed previously, inefficiently performed access router discovery method in FMIPv6 results in considerable delay which contributes to the overall delay of handoff procedure. In addition, more problems such as hesitant mobility (causing ping-pong effect) are bound to happen when the actual target network is not selected precisely. Aiming to reduce the latencies in handoff initiation phase, we propose an improved access router discovery (i-ARD) method to achieve a precise prediction of the prospective serving network. Through this method, the message exchange between ARs is initiated using an established tunnel, whereby neighboring ARs are enquired for

necessary information, and the mapping table at AR is updated according to the MN needs. Consequently, we propose a handoff model in which i-ARD is responsible for acquiring neighboring ARs information as a replacement to MIIS. Therefore, the procedure of information service (MIIS) discovery which consumes much time due to DHCP server delay will be eliminated. For instance, the functionality of delivering the list of all available links which is usually performed through Link-List trigger is replaced by ARID-Info. In case of receiving the link details of candidate AR (cAR) from MN, the serving AR identifies this AR as a prospective network based on its network prefix. This means that the existing mapping table in AR is used for two purposes; the information availability for MN's on-demand requests and to further handoff preparations by AR as the candidate AR for handoff is determined. Based on these functionalities, by using i-ARD, no mapping table or local cache is needed to be maintained in MN since i-ARD protocol is applied mainly between ARs.

In our approach, we use specific service primitives in certain schedules aiming to deliver the necessary information to each node at proper time which in turn leads to fair handoff decision and execution with affordable latency and packet loss. Table 1 outlines the primitives and the related parameters we utilized in the protocol.

#### A. Proposed Algorithm

Figure 2 demonstrates the proposed handoff mechanism using the techniques explained earlier. Prior to any operation, the MN needs to register for MIH services. This can be done while MN has not yet detected its movement and therefore has not initiated any of handoff

process modes. The i-ARD procedure is the next operation which is usually performed at intervals or at any new AR activation. Unlike CARD protocol [6] that specifies the address resolution of ARs based on L2 IDs and initiated by MN, in i-ARD the address resolution procedure is performed in the oAR in which MN is currently served by. During i-ARD phase, the oAR enquires all candidate ARs using iARD-Req multicast

message and receives the information of active and potential ARs through iARD-Rpl reply message. The list of candidate ARs is a combination of two sources; the ARs reported to oAR by MN in certain intervals and the ARs in the vicinity of the reported candidate Access Routers (cAR) which is determined by oAR. Through this advance discovery method, a new entry for each discovered AR is created in oAR and the mapping table will be updated. These entries are available to MN on every request. Based on selected attributes of each network, ARs are classified and sorted in the mapping table by the technology in use.

TABLE 1: MIH PRIMITIVES AND RELATED PARAMETERS

| Primitive  | service | Parameters            |
|------------|---------|-----------------------|
| Link_Avail | ES      | AR prefix , Interface |
| Link_Goin  | ES      | Interface ID, oAR     |
| Link_Down  | ES      | Interface ID, nAR     |
| Link_Swite | CS      | HO mode, Interface ID |
| Link-Up    | ES      | Interface ID, nAR     |

Table 2 indicates the form that AR data are stored and sorted. The location ID code is used to identify the cARs in geographical vicinity which might be different from network neighborhood.

### B. Protocol Description

The actual handoff procedure starts as soon as MN senses an increase in one of the neighboring networks' signal strength. At such a situation, the MN sends ARID-Req message to oAR including L2 ID of the new network. By referring to the mapping table, the oAR specifies the candidate network which MN wants to handoff to. This cAR selection job is performed by oAR based on the information obtained through i-ARD. Since the prospective AR is already specified, the oAR then configures a new care-of address (nCoA) for MN using the new network prefix and MN link ID, initiates an FBU for the new address, and replies the MN with ARID-info including the information of the cAR which corresponds

to the existing L2 ID. These two messages replace router solicitation and advertisement messages (RtSolPr and PrRtAdv) introduced by FMIPv6 protocol indications. By this message, MN is informed of its nCoA together with the prospective AR. Referring to Link Type field included in ARID-info, the MN knows that it should configure its second interface with the nCoA it has received. At the time that MN moves toward the new network, a Link-Going-Down (LGD) trigger which is a result of signal degradation in the old interface will be issued from MIH to inform IP layer of the link degradation. As the signal indication reduces

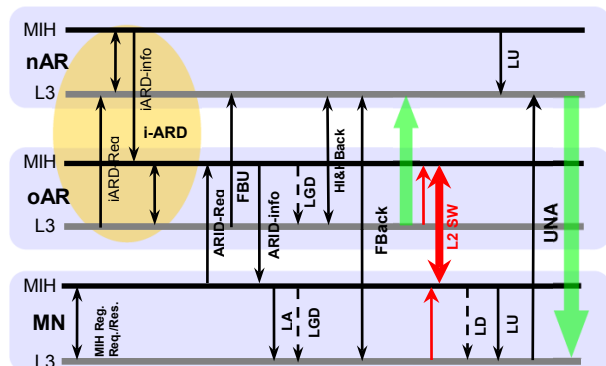


Figure 2: Proposed handoff algorithm

TABLE 2: I-ARD INFORMATION FORMAT

| AR | Link Type | MAC | Prefix | Available BW | Location ID |
|----|-----------|-----|--------|--------------|-------------|
|----|-----------|-----|--------|--------------|-------------|

TABLE 3: SIGNIFICANT SIMULATION PARAMETERS

| Parameter                | Value             |                   |
|--------------------------|-------------------|-------------------|
|                          | Heavy Load        | Medium Load       |
| Inter-Request Time       | Exponential (360) | Exponential (720) |
| File Size (Bytes)        | Constant (50,000) | Constant (5,000)  |
| Start Time               | Uniform (100,110) | Uniform (5, 10)   |
| <b>Model Assumptions</b> |                   |                   |
| Beacon interval          | 0.01 s            |                   |
| AR trans. power          | 0.0005 w          |                   |
| MN rec. threshold        | -95 dBm           |                   |

simultaneously in oAR, an LGD trigger is also issued from MIH to IP layer in oAR and we specify this event as the start of handoff initiation (HI) and tunnel establishment between the two ARs. Since the tunnel establishment and binding update are the most time consuming processes, we introduce an overlapping operation allowing oAR to initiate HI/Hack operation during the waiting time for binding acknowledgement (FBack).

However, the packet forwarding to nAR starts right after the oAR receives and delivers FBack to both oAR and MN. Now that the MN received FBack, it activates the second interface and establishes a low layer connection to the nAR while disconnecting the old connection. This L2 handoff process is performed through a Link switch (LS) command event which is received by MIH layer at both oAR and MN. The events to issue LS command are the start of packet forwarding and receiving FBack at oAR

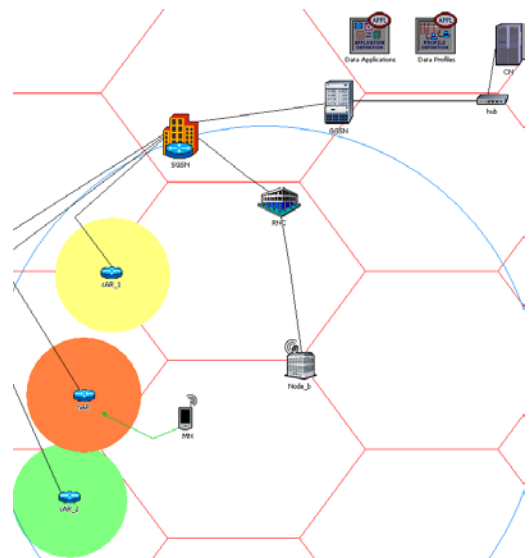


Figure 3: Network topology

and MN respectively.

As indicated in figure 2, there would be enough time for the MN to prepare for L2 handoff and link switch while shorter time is required at oAR for this operation. Thus, we show through simulation that these two events are properly scheduled.

When the Link Up (LU) trigger is received right after new connection establishment, the MN is fully attached to the new network and sends an Unsolicited Network advertisement (UNA) message to the nAR to ensure start of packet forwarding through the new connection. Now the MN can complete binding update with its home network.

#### IV. SIMULATION IMPLEMENTATIONS

In order to evaluate the performance of the proposed algorithm, we developed a model in which new modifications were applied to the current AR and MN modules. Using OPNET Modeler [9], three WLAN access routers together with a Serving GPRS Support Node (SGSN) build the ARs of our heterogeneous environment, while an FTP server plays the role of a Correspondent Node (CN). We also developed a special mobile node (MN) with WLAN and UMTS interfaces. The developed network model using OPNET is shown in figure 3 where the MN moves toward WLAN network through a predefined trajectory while communicating with its current point of attachment. The network condition is set to be free of any congestion so that the delays pertaining to long network access due to queue congestions in AR is eliminated.

The ARs are collocated with Access Points (AP) [10] as a true representation of real life situations while the Node\_b is apart from RNC and covers a large cell. The MN uses two interfaces to WLAN and UMTS networks with the related MAC layer in each interface. The MIH layer was created as a node model and placed before the MIP model. Figure 4 depicts the implemented mobile node model with two interfaces, MIP, and MIH layers. For the purpose of media homogeneity, a node level entity, so called MAC unifier has been created to recognize, specify, and deliver

the messages of each interface prior to entering the MIH module. The same MIH modules together with MIP layer were created on the standard implementation of AR to support MIH messages and triggers and FMIPv6 protocol. Alongside the implemented modules and node models, new process models which implement the new messages and procedures were also created. While ATM links were used between UMTS network entities; because of the nature of device connection in WLAN [11], the connection type between the gateway and AR in WLAN environment was 10Mbps Ethernet.

The radiuses of the areas under the coverage of ARs and the Node\_b were obtained through several simulation tests and were also based on the propagation and receiving parameters set for ARs and MN, respectively. The significant simulation parameters are based on [12] as summarized in Table 3.

#### V. NUMERICAL RESULTS

In order to measure the performance of the proposed protocol, two metrics of handoff latency and packet loss, were obtained from simulation runs based on the mobile node velocity. Roaming toward WLAN network implies the MN was assumed a pedestrian whose walking and running speeds are roughly 3 m/s and 7 m/s, respectively. Therefore, the simulations were run for the MN speeds of various intervals from 1 to 10 m/s for delay calculations and from 1 to 7 m/s for lost packets. Figures 5 and 6 depict the statistics obtained from the simulation runs. The vertical handoff model using i-ARD procedure is clearly comparable with FMIPv6 in terms of handoff delay and packet loss. For heavy FTP load conditions while FMIPv6 shows a latency of around 840 ms at walking speed of 2 m/s, handoff with i-ARD could complete in 372 ms. As a worst case we consider a running user at speed of 7 m/s and it could be seen that the handoff delay exceeds 1.2 s when using FMIPv6 while at the same situation handoff using i-ARD can still be tolerated with approximately 600 ms delay.

The difference is even more obvious with packet loss

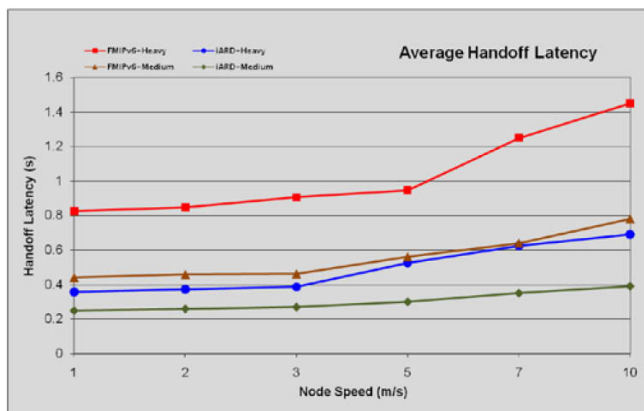


Figure 5: Handoff latency at different node speeds

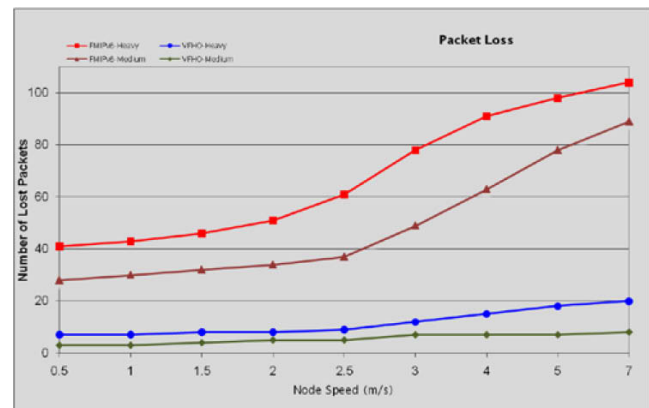


Figure 6: Packet loss

metric. At a walking speed of 2.5 m/s, handoff using i-ARD causes the loss of 8 packets in heavy load whereas, MN using FMIPv6 lost 61 packets which will affect real-time traffic. By increasing mobile moving speed, handoff using i-ARD barely causes 20 packet losses whereas FMIPv6 was measured to introduce over 100 packet loss. In the proposed algorithm, i-ARD procedure occurs prior to any trigger and handoff process therefore, the information is already prepared and ready to deliver to MN. This will omit the process cost in nAR. In reality, the total process of CARD is reported as 500 ms by several experiments [10] from which the major portion belongs to the processing portion of the algorithm. DAD procedure is reported to take around 1000 ms and each message exchange time between fixed nodes is 10 ms. Round-trip transmission between MN and each of the nodes could vary between 50 ms and 500 ms. Assuming equal situations in a comparative scenario for two methods, total handoff process in the proposed method is decreased by five transmission delay times between MN and either of ARs and additional 2 times process delay in oAR, which is a minimum of 300 ms and a worst case of 2500 ms. This difference range was also shown through simulation tests as reported earlier. As mentioned earlier, along with the protocol, one of the significant issues addressed here is process timing, especially while using i-ARD or any other alternative for access router discovery.

## VII. CONCLUSION

In this paper, a method for integrating layer-2 and layer-3 protocols to achieve seamless vertical handoff is proposed. The proposed algorithm improved link layer event services, applied an improved access router discovery mechanism for discovering neighboring access routers, and finally integrated the mechanisms with layer-3 mobility protocol to form a seamless handoff process. It was also shown that proper timing operation sequences can avoid significant processing delays. The proposed algorithm shows lower latency and affordable packet loss compared to FMIPv6. In order for the proposed approach to be applicable to UMTS systems, the event service primitives should be supported by UMTS link layer in real environment.

## ACKNOWLEDGEMENT

The authors wish to acknowledge the assistance and support of OPNET laboratory of University of Malaya and especially Professor Kaharudin Dimyati.

## REFERENCES

- [1] R. Koodli, "Mobile IPv6 Fast Handovers," *IETF*, vol. RFC5268, June 2008 2008.
- [2] D. Johnson, C. Perkins, and J. Arrko, "Mobility Support in IPv6," *IETF*, vol. RFC3775, 2004.
- [3] V. Gupta, "IEEE802.21 Standard and Metropolitan Area Network: Media Independent Handover Services," *IEEE*, vol. Draft, March 2006 2006.
- [4] Y. Y. An, B. H. Yae, K. W. Lee, Y. Z. Cho, and W. Y. Jung, "Reduction of Handover Latency Using MIH Services in MIPv6," in *20th International Conference on Advanced Information Networking and Applications*, 2006, pp. 229 - 234.
- [5] D.-H. Kwon, Y.-S. Kim, K.-J. Bae, and Y.-J. Suh, "Access router information protocol with FMIPv6 for efficient handovers and their implementations," in *IEEE Global Telecommunications Conference, 2005. GLOBECOM '05*, 2005, p. 6 pp.
- [6] M. Liebsch, A. Singh, H. Chaskar, D. Funato, and E. Shim, "Candidate Access Router Discovery (CARD)," *IETF*, vol. RFC4066, July 2005 2005.
- [7] Q. B. Mussabbir and Y. Wenbing, "Optimized FMIPv6 handover using IEEE802.21 MIH services," in *Proceedings of first ACM/IEEE international workshop on Mobility in the evolving internet architecture* San Francisco, California: ACM, 2006.
- [8] A. Dutta, S. Das, D. Famorali, Y. Ohba, K. Taniuchi, T. Kodama, and H. Schulzrinne, "Seamless Handover across Heterogeneous Networks - An IEEE802.21 Centric Approach," 2006.
- [9] "OPNET Modeler."
- [10] L. Dimopoulou, G. Leoleis, and I. S. Venieris, "Analysis and Evaluation of Layer 2 Assisted Fast Mobile IPv6 Handovers in a WLAN Environment " in *10th IEEE Symposium on Computers and Communications (ISCC 2005)*, 2005, pp. 327 - 332.
- [11] V. A. Dubendorf, *Wireless Data Technologies Reference Book*: John Wiley & Sons Canada Ltd, 2003.
- [12] IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," ANSI/IEEE Std. 802.11, 1999.

# BER Performance of Coded MSK/Multi-channel CDMA Fourth Generation Synchronous Downlink Cellular Systems in the Presence of MUI and FO

M. Abou El-Nasr

Arab Academy for Science and Technology (AAST), Alexandria, Egypt  
Department of Computer Engineering

**Abstract**—This paper studies the (BER) performance of synchronous coded Minimum Shift Keying/ Multi-channel Code Division Multiple Access (MSK/MC-CDMA) systems in the presence of Multi-access User Interference (MUI), and carrier Frequency Offset (FO). The theoretical effect of MUI on the BER performance was considered assuming the Standard Gaussian Approximation (GA) in AWGN channel. Furthermore, explicit simulations were carried out in AWGN and frequency selective Rayleigh fading channels for various parameters. The un-coded MSK/MC-CDMA system under investigation outperforms the corresponding system with BPSK modulation. Furthermore, a careful choice of coded MSK/MC-CDMA system parameters can achieve a better performance than the theoretical performance in AWGN channel for a single user and with no frequency offset or extra bandwidth requirements.

**Index Terms**—Algorithms, Downlink, Synchronous, Multi-access User Interference (MUI), Multi-Carrier, CDMA, Coding, and Minimum shift Keying (MSK) Modulation.

## I. INTRODUCTION

Multi-channel code division multiple access (MC-CDMA) scheme combines the high bandwidth efficiency in addition to the robustness against the distortion caused by multipath frequency selective fading channels. Moreover, MC-CDMA is an attractive candidate for future mobile communications. Specifically, it allows for low complexity receivers in the downlink, since the use of multi-carriers combats the effect of inter-symbol-interference (ISI) [1], [2]. However, the performance of this system is sensitive to both Multi-access User Interference (MUI), and frequency offset (FO) [3], [4].

A modulation scheme that is commonly considered for mobile communication systems is the Minimum shift keying (MSK) modulation scheme, as it is a power efficient modulation scheme with constant envelope and has a self synchronizing capability [5], [6].

The use of random spreading gives an upper bound on the bit error rate (BER) performance, and the exact analysis of the BER performance under these conditions seems to be significantly difficult. Hence, efficient approximations are desirable. The simplest well known approximation for the BER performance is the Standard Gaussian Approximation (GA), in which the overall effect of MUI in a

symbol interval is approximated by a Gaussian random variable. Moreover, it can be generalized to linear modulations and error control coding [7], [8], [9], [14], [15], [16], [17], [18].

Error control coding can be used with MC-CDMA systems with no extra bandwidth penalty due to the addition of redundant bits. This makes the application of error control coding very attractive in MC-CDMA systems. The error control coding technique considered in this paper is the cyclic block code, as it is an efficient coding scheme, and is easily implemented [13].

The organization of this paper is as follows. First, Section II includes a description of the system model. Then, Section III studies the BER performance in AWGN channel using GA. The simulation model is given in Section IV followed by numerical results obtained from simulations in AWGN channel and frequency selective fading channels for both un-coded and coded MSK/MC-CDMA synchronous downlink systems for a various number of interferers in Section V. Finally, conclusions are provided in Section VI.

## II. SYSTEM MODEL

The system under investigation, depicted in Fig. 1, has  $R$  users with corresponding  $R$  sequences of symbols. For the un-coded system, the  $R$  sequences are obtained after spreading each with spreading factor ( $S$ ) and MSK modulated. When coding is applied, the sequence is first coded and then MSK modulated after it is being spread. Without loss of generality, the spreading factor is assumed to be equivalent to the number of sub-carriers ( $M$ ). The effect of the  $R - 1$  interferers on the desired user is known as MUI [1]. Random spreading codes are applied under the assumption of perfect synchronization. Furthermore, the spreading codes are assumed to be dynamically changing from one bit to the other, and they are independent and identically distributed (*i.i.d*). All users are assumed to have equal powers  $P_r = P$ , under the assumption of perfect power control. The  $R$  users with MSK modulated symbols are spread and added.



The summation of the  $R$  users'  $m$ -th chip ( $m = 0, 1, \dots, M-1$ ) during the  $i$ -th bit interval is given by:

$$s_m(t) = \sum_{r=1}^R [b_i^r(t-t_0) \cdot c_{i,m}^r(t-t_0) - j b_i^{r-1}(t) \cdot c_{i,m}^{r-1}(t)] \cdot \exp(j\alpha^r) \quad (1)$$

Where,  $b_i^r$  represents the data bit of  $r$ -th user during  $i$ -th data bit, and  $c_{i,m}^r$  is the  $m$ -th chip in the  $i$ -th bit interval for the  $r$ -th user. The parameter represents  $r$ -th user carrier phase and  $t_0$  is the time offset required for MSK.

The baseband received signal impaired with frequency offset is given as [1]:

$$\hat{r} = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} s_m \cdot \exp\left(j2\pi(m+\varepsilon)\frac{h}{M}\right) \quad (2)$$

Where,  $h = 0, 1, \dots, M-1$ , and  $\varepsilon$  is the normalized frequency offset.

### III. THEORETICAL BER PERFORMANCE

This section gives theoretical expressions for the BER performance of the system under study in AWGN channel, and it is compared to BPSK/MC-CDMA system. Moreover, the effect of the frequency offset on the BER performance is studied for both systems.

#### A. BER Performance of Un-coded MSK/MC-CDMA System

The effect of MUI on the BER performance in AWGN channel using GA for BPSK modulation with the assumption of chip and phase alignment for all users is as given in [5]:

$$P_b = Q\left(\sqrt{\frac{(2E_b/N_0)S}{R-1}}\right) \quad (3)$$

Where,  $E_b/N_0$  is the bit energy-to-noise ratio. For the case of MSK modulation, since it is a continuous phase frequency modulation technique, the effect of MUI is equivalent to the case where interfering signals are chip aligned with random phases. Thus, the probability of bit error is influenced by a factor of  $\sqrt{2S}$  instead of  $\sqrt{S}$  for the case of BPSK [5]. Hence, BER is expressed as:

$$P_b = Q\left(\sqrt{\frac{(2E_b/N_0)2S}{R-1}}\right) \quad (4)$$

This means that the BER performance of MSK/MC-CDMA systems outperform BPSK/MC-CDMA systems when the effect of MUI is considered, which is a critical factor that affects the performance of modern cellular systems.

The frequency offset has the same effect on both systems. The probability of bit error with frequency offset of the BPSK modulation case is expressed as [8]:

$$P_b = Q\left(\sqrt{\frac{\sin(\pi\varepsilon)^2(2E_b/N_0)S}{(\pi\varepsilon)^2(R-1)}}\right) \quad (5)$$

The probability of bit error with frequency offset assuming MSK modulation is expressed as:

$$P_b = Q\left(\sqrt{\frac{\sin(\pi\varepsilon)^2(2E_b/N_0)2S}{(\pi\varepsilon)^2(R-1)}}\right) \quad (6)$$

#### B. Performance of MSK/MC-CDMA System with Error Control Coding

Cyclic codes are linear block codes that have the advantage of being easy to encode, and have very efficient decoding schemes. Furthermore, cyclic codes are easy to implement using modern integrated circuit techniques [10], [11], [12], [13]. The codes considered in this paper are cyclic codes with code rates 7/15, 2/6, and 2/10.

The relation of the code-symbol energy per noise spectral density ( $E_c/N_0$ ), and the information bit energy per noise spectral density ( $E_b/N_0$ ) is given as [14]:

$$\frac{E_c}{N_0} = \left(\frac{k}{n}\right) \cdot \frac{E_b}{N_0} \quad (7)$$

Where,  $k$  is the number of message bits, and  $n$  is number of the codeword bits.

### IV. SIMULATION MODEL

A simulation model was carried out based on the parameters described in Section III-A. Moreover, computer simulations for the systems described in Section III-A were carried out in AWGN and frequency selective Rayleigh fading channels. The Frequency selective fading channel simulation was considered assuming that each sub-carrier experiences independent fading. A summary of simulation parameters is given in Table(1).

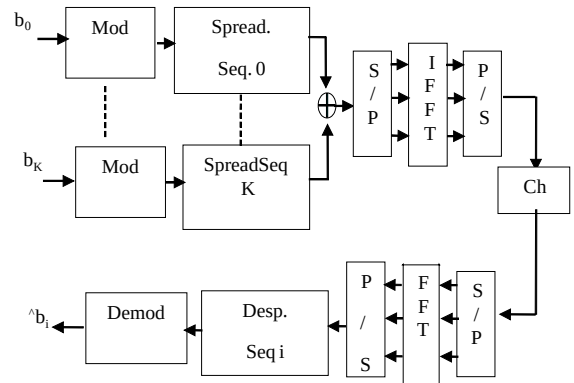


Fig. 1. MC-CDMA System Model.

TABLE I  
SUMMARY OF SYSTEM PARAMETERS

| Parameter                   | Value                                                                          |
|-----------------------------|--------------------------------------------------------------------------------|
| Modulation Schemes          | BPSK, MSK                                                                      |
| Spreading Factors           | $S= 32$ for un-coded system and $S= 16$ for coded system                       |
| Coding                      | Cyclic codes with code parameters $(n, k)$ of values (15,7), (6,2), and (10,2) |
| MUI                         | Synchronous users                                                              |
| Number of interfering users | $R=1, 3, 5$                                                                    |
| FO                          | $\varepsilon = 0, 0.2, \text{ and } 0.3$                                       |

## V. SIMULATION RESULTS

Figure 2 compares the BER performance of the system under study obtained via simulation in AWGN channel to the analytical form given by Eqn. (4), assuming 2, 4 and 6 users. As can be seen, the simulation results agree with the analytical expression. Next, Fig. 3 compares the BER performance of MSK/MC-CDMA and BPSK/MC-CDMA in frequency selective fading channel. As shown, the effect of MUI is less severe when MSK modulation is considered. Furthermore, simulation results assuming various values of the normalized frequency offset were considered in Fig. 4, where the effect of both the frequency offset and MUI on the BER performance is considered. Moreover, simulation results were compared to the theoretical expression defined by Eqn. (6). While increasing the frequency offset degrades the BER performance, its effect is independent of the number of multi-access users in the system. Figure 5 depicts the BER versus the number of users for various values of  $S$ .

Further simulations were carried out to investigate the performance of coded MSK/CDMA systems in AWGN and frequency selective fading channels. The linear block code employed is the cyclic coding with various code rates. Figure 7 compares the simulated BER performance of coded MSK/MC-CDMA to the theoretical performance of the un-coded case for the code (6,2). Moreover, Fig. 8 compares the simulated BER performance of coded MSK/MC-CDMA to the theoretical performance of the un-coded case for the code (10,2). As the code rate decreases, more enhancement is gained in the BER performance. It was noticed that the system with code (10,2) and 6 users outperforms the corresponding un-coded system with a single user.

In addition, simulations were carried out in frequency selective fading channels in order to investigate the effect of fading on the BER performance. Figure 9 compares the simulated BER performance of coded MSK/MC-CDMA to the theoretical performance of the un-coded case for the code (6,2). Moreover, Fig. 10 gives a comparison of the simulated BER performance of coded MSK/MC-CDMA and the theoretical performance of the un-coded case for

the code (10,2). The gain in the BER performance can be noticed with increasing the code rate without further increase in the bandwidth requirement.

Finally, the effect of frequency offset on coded MSK/MC-CDMA systems was investigated via simulation for various values of  $\varepsilon$  and  $R$  in AWGN and frequency selective fading channels. Figure ?? and Fig. 11 show the simulated BER of coded MSK/MC-CDMA using the code (15,7) in AWGN and frequency selective fading channel assuming various values of  $\varepsilon$  and  $R$  respectively. Figure 12 compares the BER performance of coded and un-coded MSK/MC-CDMA systems for the code (6,2) in AWGN channel for various values of  $\varepsilon$  and  $R$ . Furthermore, Fig. 13 and Fig. 14 show the simulated BER performance assuming the code (6,2) in AWGN and frequency fading channels for various values of  $\varepsilon$  and  $R$  respectively. Similarly, Fig. 15 and Fig. 16 depict the simulated BER performance assuming the code (10,2) in AWGN and frequency fading channels for various values of  $\varepsilon$  and  $R$  respectively. The effect of coding is clear in all results, specifically when increasing the code rate. Particular, in Fig. 16, the coded system with code (10,2),  $R = 2$ , and  $\varepsilon = 0.3$  gives a BER of  $1e-4$  at  $\text{SNR} = 8$  dB, whereas the un-coded system with  $R = 1$  and no frequency offset, given in Fig. 6 gives the same BER at  $\text{SNR} \approx 11$  dB in AWGN channel.

## VI. CONCLUSIONS

The effect of MUI on the BER performance of MC-CDMA systems with MSK modulation is less severe than the corresponding systems with BPSK modulation due to phase continuity. This applies in both AWGN and Rayleigh frequency selective fading channels. Nevertheless, the BER performance of these systems is sensitive to carrier frequency offset. However, the application of error control coding enhances the BER performance with no extra bandwidth penalty. A careful choice of the system parameters could result in a better BER performance of coded-MSK/CDMA systems with MUI and frequency offset in frequency selective fading channel than the corresponding un-coded systems with a single user in AWGN channel.



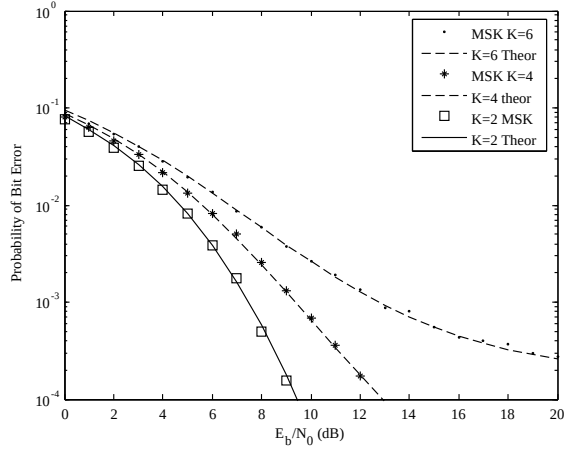


Fig. 2. BER Performance of MC-CDMA with MSK modulation and  $S=32$  in AWGN channel for various values of  $R$ .

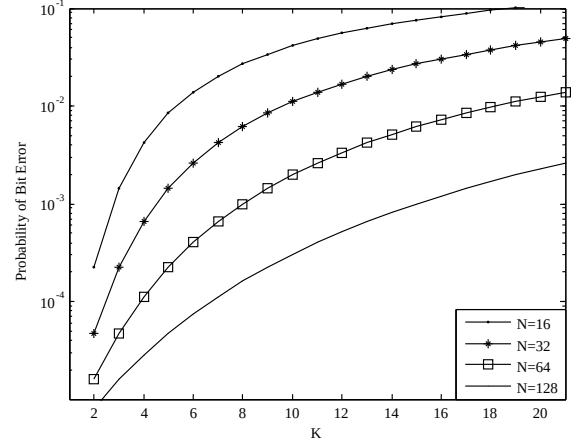


Fig. 5. BER Performance of MC-CDMA, frequency selective channel for  $S=32$  and various values of  $R$ .

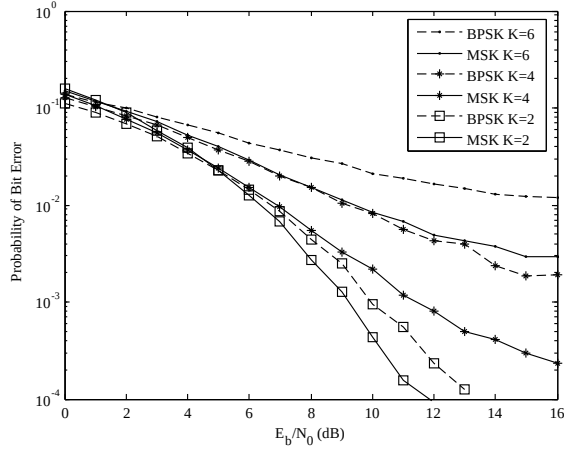


Fig. 3. BER Performance of MC-CDMA, frequency selective channel for  $S=32$  and various values of  $R$ .

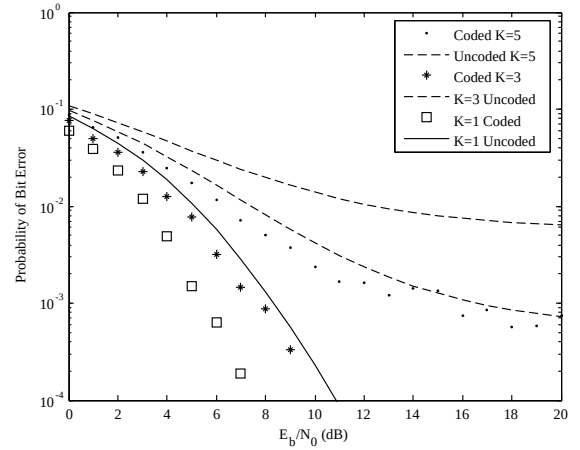


Fig. 6. Simulated BER Performance of coded MSK/MC-CDMA, code (15,7),  $S=16$  in AWGN channel for various values of  $R$ .

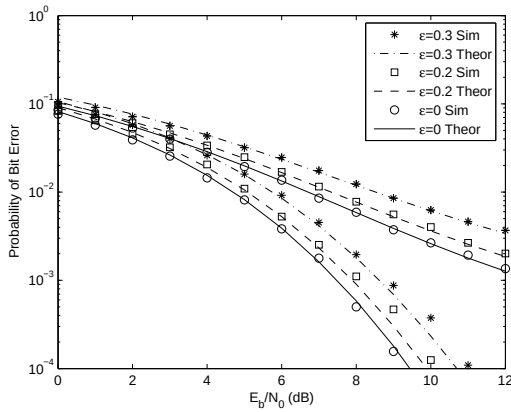


Fig. 4. BER Performance of MC-CDMA with MSK modulation and  $S=32$  in AWGN channel for various values of  $\epsilon$  and  $R$ .

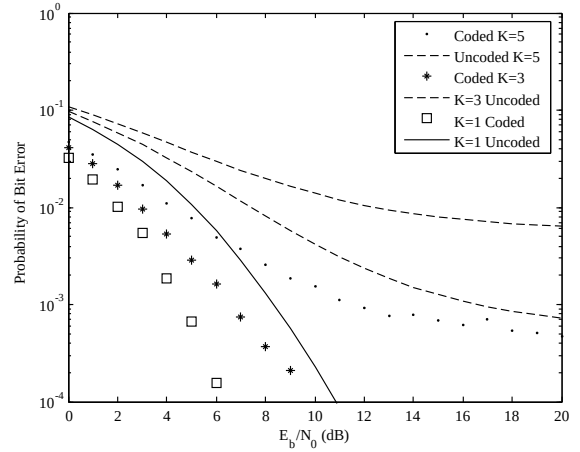


Fig. 7. Simulated BER Performance of coded MC-CDMA with MSK modulation, code (6,2),  $S=16$  in AWGN channel for various values of  $R$ .

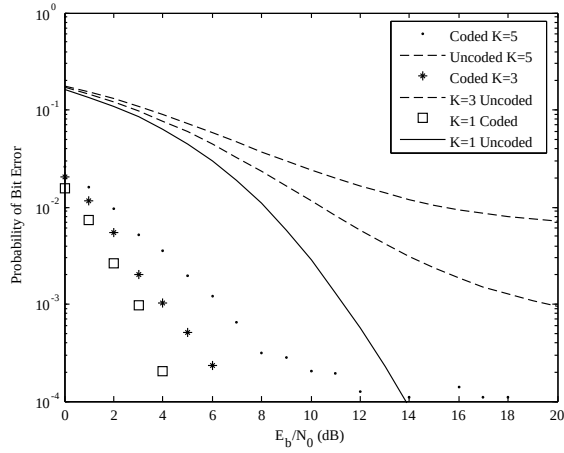


Fig. 8. Simulated BER Performance of coded MSK/ MC-CDMA, code (10,2),  $S=16$  in AWGN channel for various values of  $R$ .

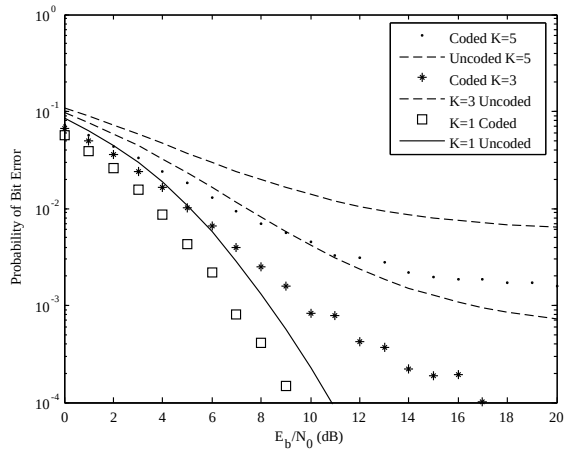


Fig. 9. Simulated BER Performance of coded MSK/ MC-CDMA, code (6,2),  $S=16$  in frequency selective channel for various values of  $R$ .

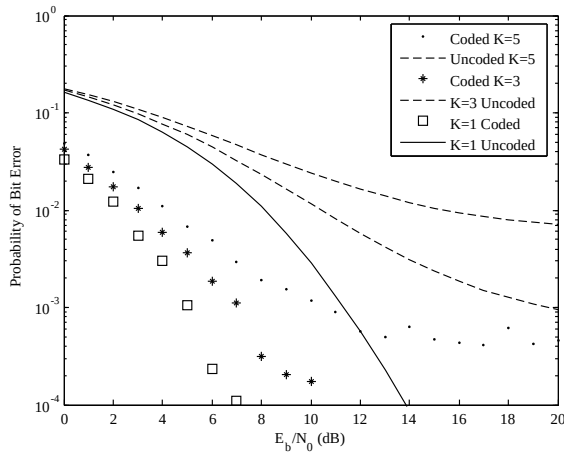


Fig. 10. Simulated BER Performance of coded MSK/ MC-CDMA, code (10,2),  $S=16$  in frequency selective channel for various values of  $R$ .

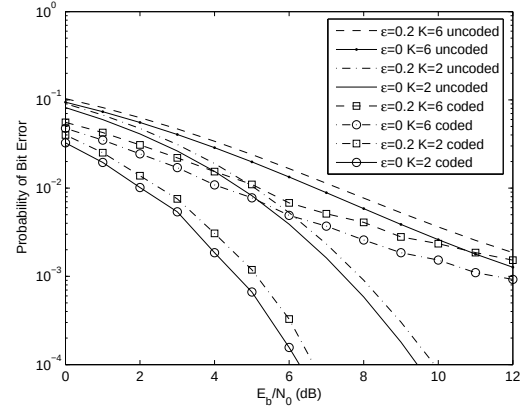


Fig. 11. Simulated BER Performance of coded MSK/ MC-CDMA, code (15,7),  $S=16$  in frequency selective channel for various values of  $\epsilon$  and  $R$ .

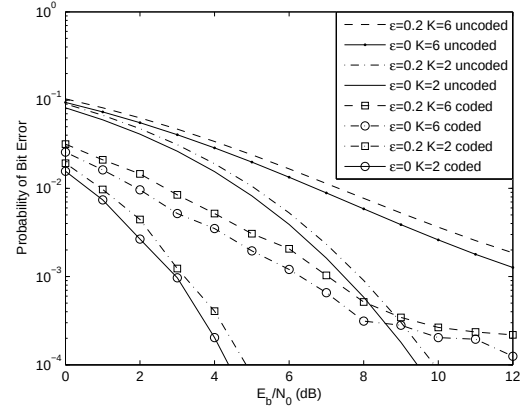


Fig. 12. Comparison of simulated BER Performance of coded MSK/ MC-CDMA, and the corresponding un-coded system for the code (6,2),  $S=16$  in AWGN channel, and various values of  $\epsilon$  and  $R$ .

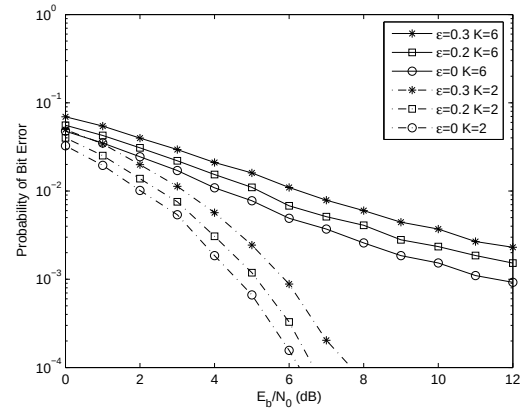


Fig. 13. Simulated BER performance of coded MSK/ MC-CDMA, code (6,2),  $S=16$  in AWGN channel for various values of  $\epsilon$  and  $R$ .

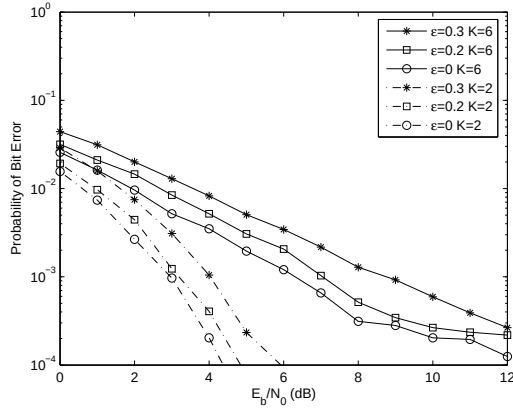


Fig. 14. Simulated BER performance of coded MSK/ MC-CDMA, code (10,2),  $S=16$  in AWGN channel for various values of  $\varepsilon$  and  $R$ .

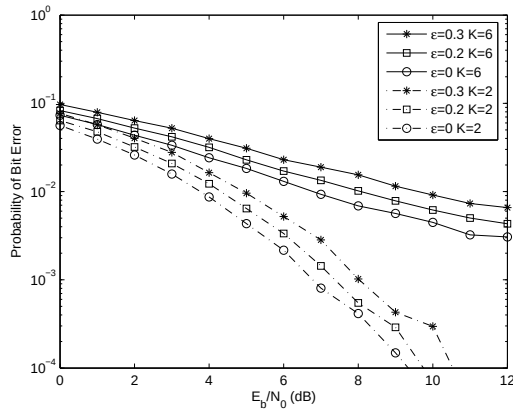


Fig. 15. Simulated BER Performance of coded MSK/ MC-CDMA, code (6,2),  $S=16$  in frequency selective various values of  $\varepsilon$  and  $R$ .

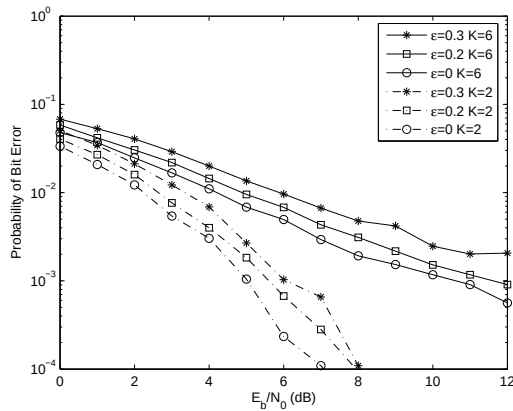


Fig. 16. Simulated BER Performance of coded MSK/ MC-CDMA, code (10,2),  $S=16$  in frequency selective various values of  $\varepsilon$  and  $R$ .

## REFERENCES

- [1] W. Jang, L. Nguyen, and P. Bidarkar. MAI and ICI of Synchronous Downlink MC-CDMA with Frequency Offset. *IEEE Trans. Comm.*, 5(3):693–703, 2006.
- [2] S. Hara. Design and Performance of Multicarrier CDMA System in Frequency-Selective Rayleigh Fading Channels. In *IEEE Trans. VTC.*, 48(5) : 1584– 1595, 1999.
- [3] N. Nasr, H. A. Shaban, and M.A.Mangoud. Coded 4-PAM OFDM for High Rate Data Links. *CISSE 2006.*, 15(5):745–770, Dec. 2006.
- [4] J. S. Lehnert. Serial MSK Spread-Spectrum Multiple-Access Communications. *IEEE Trans. Comm.*, 40 (6): 1119– 1127, 1992.
- [5] M. Mammela and R. Broderick. Constant Envelope Multicarrier Modulation: Performance Evaluation in AWGN and Fading Channels. *MILCOM2005*, 2: 807–813, Oct.2005.
- [6] S. Elnoubi, A. Elshamly. Performance of Multi-carrier CDMA with MSK Modulation and One-Bit Differential Detection in a Multi-Path Fading Channel. *VTC2000*, 3:1041–1048, Sept. 2000.
- [7] R. Morrow and J. S. Lehnert. Bit-to-Bit Error Dependence in Slotted DS/SSMA Packet Systems with Random Signature Sequences. *IEEE Trans. Comm.*, 37(10):1052– 1061, 1989.
- [8] Y. Kim, S. Choi, C. You. BER Computation of an MC-CDMA System with Carrier Frequency Offset. *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, (5): 2555–2558, March 1999.
- [9] S. Kaiser. Trade-off between Channel Coding and Spreading in Multi-Carrier CDMA Systems. *IEEE 4th International Symposium on Spread Spectrum Techniques and Applications Proceedings*, (3):1366–1370, Sept. 1996.
- [10] C. Ibars and B. Ness. Comparing the Performance of Coded Multiuser OFDM and Coded MC-CDMA over Fading Channels. *IEEE Global Telecommunications Conference*, (2):881 – 885, 2001.
- [11] T. Welch and R. Ziemer. Simulation of Coded MC-DS-CDMA Systems. *1998 IEEE Radio and Wireless Conference*, pp: 189–192, Aug 1998.
- [12] T. Lok and J. Lehnert. An asymptotic Analysis of DS/SSMA Communication Systems with General Linear Modulation and Error Control Coding. *IEEE trans. Information Theory*, 44(2): 870–881, March 1998.
- [13] B. Sklar, F. Harris. newblock The ABCs of Linear Block Codes. *IEEE Signal Processing Magazine*, pp: 14–35, July 2004.
- [14] B. Sklar. *Digital Communications: Fundamentals and Applications*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall Inc., 2001.
- [15] J. M. Holtzman. A simple, Accurate Method to Calculate Spread-Spectrum Multiple-Access Error Probabilities. *IEEE trans. Comm.*, 40(3): 461–464, March 1992.
- [16] M. A. Nasr, H. A. Shaban, and W. H. Tranter. Outage Probability and Percentage of Cell Area of OFDMA Cellular Systems with Sectoring. *Communication Networks and Services Research Conference (CNSR 2007)*, pp. 107–116, May 2007.
- [17] M. Abou El-Nasr, H. Shaban, and M. A. Mangoud. Effect of MAI and Frequency Offset on the Performance of Coded MC-CDMA Synchronous Downlink Systems with MSK Modulation. *IEEE 4th International Symposium on Wireless Communication Systems (ISWCS07)*, Oct.2007.
- [18] M. Abou El-Nasr, M.A. Mangoud, and H. Shaban. Performance Comparison of Differential and Non-Differential Synchronous Downlink MC-CDMA Systems with Random Spreading. *The third IEEE International Conference on Communications, Networking, and Mobile Computing (WiCOM2007)*, Sep. 2007.

# A New SS7-SIGTRAN protocol Interchanger Software and Hardware to Implement an Improved Distributed Database Based ATM Network Using Existing IP Network

M. S. Munir, K. Ahmed, A.S.M. Shihavuddin, S. M. Nizamuddin and M. J. Sarker

Electrical and Electronic Engineering Department, Islamic University of Technology (IUT), OIC  
Gazipur-1704, Bangladesh.

{smunir, kabir\_89}@iut-dhaka.edu, {shihav407, anik525, junayed987}@yahoo.com

**Abstract** — This paper presents new Signaling System No. 7 (SS7) to IP based Signaling Transport (SIGTRAN) protocol interchanger software and related hardware to implement an improved distributed database based Automated Teller Machine (ATM) network where the countrywide existing IP network can be used. Currently the ATM network is based on fully centralized database environment. This leads to varying transaction time according to the ATM and central database. Moreover, link failure between central database and ATM leads to transaction failure. To overcome this problem, an improved distributed database system for ATM network is proposed. To deploy such network, a new software and related hardware is also presented which discards the need of costly new leased line or dial up connection. Using the software and hardware, the existing IP network can be used for expanding ATM network. Also the cost and benefits of the network is discussed by allocating fragment on different sites.

**Index Terms** — Automated Teller Machine (ATM), Distributed Database, SS7, SIGTRAN.

## I. INTRODUCTION

Currently all the ATM machines working around the world are based on the concept of centralized database system. Previously a distributed database base ATM network was proposed in [1]. In our paper, we have proposed an improved distributed database based ATM network. Cost and benefit of such network is discussed by allocating fragments on different ATM sites. And a new SS7 to SIGTRAN protocol interchanger software and related hardware is presented to deploy such ATM network using the existing countrywide SS7 network.

An ATM is simply a data terminal with two input and four output devices. Like any other data terminal, the ATM has to connect to, and communicate through, a host processor. The host processor is analogous to an Internet service provider (ISP) in that it is the gateway through which all the various ATM networks become available to the cardholder (the person wanting the cash).[2]

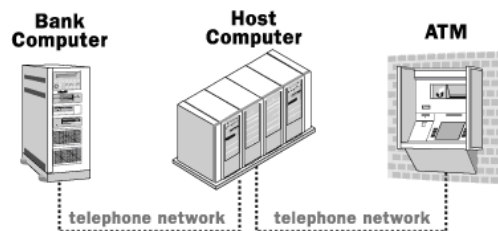


Fig. 1 ATM Connections

Most host processors can support either leased-line or dial-up machines. Leased-line machines connect directly to the host processor through a four-wire, point-to-point, dedicated telephone line. Dial-up ATMs connect to the host processor through a normal phone line using a modem and a toll-free number, or through an Internet service provider using a local access number dialed by modem. Leased-line ATMs are preferred for very high-volume locations because of their thru-put capability and dial-up ATMs are preferred for retail merchant locations where cost is a greater factor than thru-put. The initial cost for a dial-up machine is less than half that for a leased-line machine. The monthly operating costs for dial-up are only a fraction of the costs for leased-line. The host processor may be owned by a bank or financial institution, or it may be owned by an independent service provider.

If we could find a way to connect the leased line or dial up based ATM network to the countrywide existing IP network, then this would result a tremendous reduction of cost. Because leased-line or dial-up ATM network is much most costly than IP network. To use the IP network for ATM, we have designed the new software and hardware. The ATM machine has ports to communicate via lease line or dial up connection. We have designed E1 card that will be used to connect this ATM machine directly to an IP network. The E1 card performs the job of interfacing and the software performs the job of protocol interchanging e.g SS7 to SIGTRAN. An improved distributes database base ATM network is proposed in II, the software architecture is discussed in III while the hardware architecture is discussed in IV.

## II. DISTRIBUTED DATABASE BASED ATM NETWORK

According to current scenario, any request from a user is passed to the client's bank to the institution who issued the card. If a request is issued from a distant place then much time is required to send the request to the cardholder's bank or the institution who issued the card. A much convenient approach would be to store the client's information in multiple branches of the bank. This is done by dividing the global database into fragments. The fragments are allocated to one or more sites. The steps to apply the distributed concept in ATM network are,

- The first issue would be to inquire the client about the frequently accessed places. Then the client information is stored in all the branches those were identified by the client as frequently accessed places. Host processor is updated about which branches contain client information.
- The second issue deals with any request issued by the client to the ATM. The ATM forwards this information to the host processor. The host processor, on receiving the request determines which branch is nearest neighbor that stores information of that client from the place the request is issued. The nearest neighbor based on composite vectors can be determined with the help of the following equation [2]

$$Metric = K_1 * BW + \frac{K_2 * BW}{256 - Load} + K_3 * Delay \quad (1)$$

Where, Metric = factor that store a host processor in its database for its every neighbor.

K1, K2, K3 = Constant used to weight the effect of routing metrics. The default values for K1 and K3 are 1 and for K2 is 0. If we consider the default constant values, the basic metric calculation just reduces to

$$Metric = BW + Delay \quad (2)$$

Regardless of whether we used the default values, the result of the metric calculation is a composite metric that can fairly evaluate potential routes through a highly diverse network. The value of this composite metric is frequently described as the cost of a route. The route with the lowest cost is the best route.

- After determining the nearest neighbor, the host processor routes the transaction request to it and the transaction is performed. Transaction is all other branches at the same account us locked to ensure atomicity. All the other branches that contain the same client information are also updated at the same time by sending the transaction information by the host processor to maintain concurrency control.

So, in order to apply distributed concept, we need to consider two phases which is shown bellow,

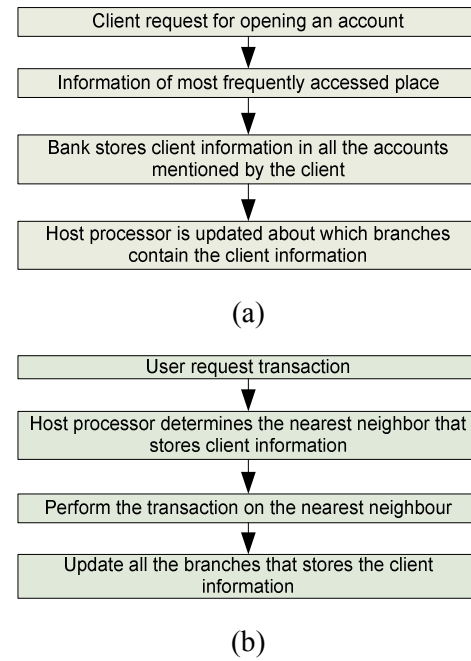


Fig. 2 (a) Initial account opening procedure, (b) Transaction processing

A global relation can be split into several non-overlapping portions called fragments. There are several ways to split the global relation into fragments. The most common of them are Horizontal fragmentation and Vertical fragmentation. Fragments are physically located at one or several sites of the computer network. The allocation scheme defines at which site(s) a fragment is located. All the fragments, which correspond to the same global relation and are located at same site constitute the physical image of the global relation. For our proposed architecture we used horizontal fragmentation. To calculate the cost and benefits of the allocation of fragments of a global relation R, the following definitions are considered,

i is the fragment index, j is the site index, k is the application index,  $r_{ki}$  is the frequency of application k at site j,  $u_{ki}$  is the number of retrieval references of application k to fragment I,  $n_{ki}$  is the number of update references of application k to fragment i. Here,

$$n_{ki} = r_{ki} + u_{ki} \quad (3)$$

1. At first we determined the set of all sites where the benefit of allocating one copy of the fragment is higher than the cost, and allocate a copy of fragment to each element of this set. This method selects "all beneficial sites". We placed  $R_i$  at all sites j where the cost of retrieval of applications is larger than the cost of update references to  $R_i$  from applications at any other site.  $B_{ij}$  is evaluated as the differences:

$$B_{ij} = \sum_{i=1}^n f_{kj} r_{kj} - C * \sum_k \sum_{P \neq j} f_{kp} u_{ki} \quad (4)$$

C is a constant, which measures the ratio between the cost of an update and retrieval access.  $R_i$  is allocated at site j such that  $B_{ij}$  is positive. When all  $B_{ij}$  are negative, a single copy of  $R_i$  is placed at the site such that  $B_{ij}$  is maximum.

- Secondly, we determined the solution of the non replicated problem and then progressively introduced replicated copied starting from the most beneficial. The process is terminated when no additional replication is beneficial. We can measure the benefit by placing a new copy of  $R_i$  in terms of increased reliability and accessibility of the system. Let  $d_i$  denotes the degree of redundancy of  $R_i$  and let  $F_i$  denotes the benefit of having  $R_i$  fully replicated at each site. The function  $B(d_i)$  measures the benefit.

$$B(d_i) = (1 - 2^{-d_i}) F_i \quad (5)$$

Then we evaluated the benefit of introducing a new copy of  $R_i$  at each site j by modifying the previous formula as follows,

$$B_{ij} = \sum_{i=1}^n f_{kj} r_{kj} - C * \sum_k \sum_{P \neq j} f_{kp} u_{ki} + B(d_i) \quad (6)$$

For banking operation, we need to store particular client's information together. It is unusual to store client's name in one place and account number in another place. Rather all the information is kept together for easier access. For this reason we have avoided vertical fragmentation.

### III. PROPOSED SOFTWARE ARCHITECTURE

Protocol can be defined as the rules governing the syntax and synchronization of communication. We have worked with two protocols: SS7 and SIGTRAN.

Signaling System no. 7 (SS7) is a set of telephony signaling protocols [1], which are used in the telephone network. SS7 provides a universal structure for telephony network signaling, messaging, interfacing and network maintenance [2]. SIGTRAN is the name given to an IETF working group that produced specifications [3] for a family of protocols that provide reliable datagram service and user layer adaptations for SS7 and ISDN communications protocols. Our developed software along with our designed hardware is capable of taking data from telephone network and convert that to SIGTRAN protocol data defined by ITU-T [1][2][4]. Our developed software takes SS7

protocol data from SS7 network through designed USB E1 card then converts it to SIGTRAN data and transmits that data through Ethernet port. The working principle of our developed hardware and software is like,

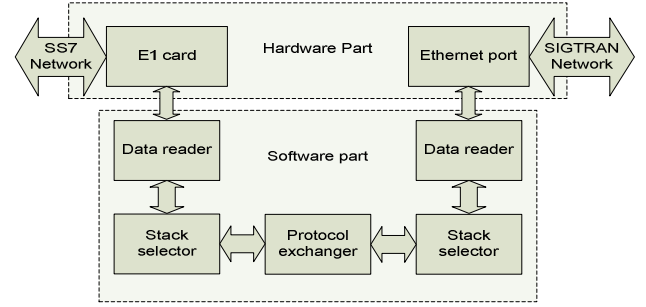


Fig. 3 Working principle of the proposed system

The first part of the development of our software was to develop the protocol stack of SS7 and SIGTRAN protocol.

**SS7 Stack:** The SS7 protocol stack borrows partially from the OSI Model of a packetized digital protocol stack. OSI layers 1 to 3 are provided by the Message Transfer Part (MTP) of the SS7 protocol; for circuit related signaling, such as the Telephone User Part (TUP) or the ISDN User Part (ISUP), the User Part provides layers 4 to 7, whereas for non-circuit related signaling the Signaling Connection and Control Part (SCCP) provides layer 4 capabilities to the SCCP user [6].

The MTP[9] covers the transport protocols including network interface, information transfer, message handling and routing to the higher levels. SCCP is a sub-part of other L4 protocols, together with MTP 3 it can be called the Network Service Part (NSP). ISUP provides a circuit-based protocol to establish connections for calls. TCAP is used to create database queries and invoke advanced network functionality mobile services (MAP), etc. [7]

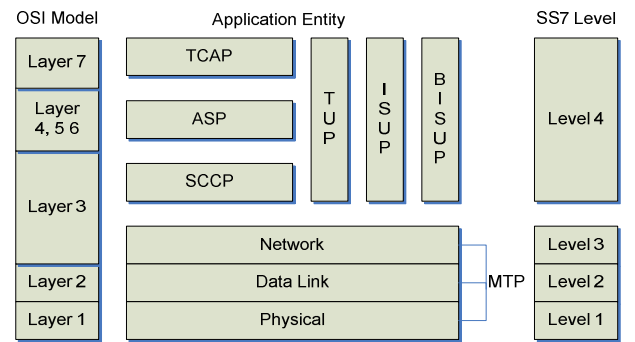


Fig. 4 SS7 protocol stack [8]

**SIGTRAN Stack:** The SIGTRAN protocols specify the means by which SS7 messages can be reliably transported over IP networks. The architecture identifies two components:

- A common transport protocol for the SS7 protocol layer being carried and
- An adaptation module to emulate lower layers of the protocol.

For example, if the native protocol is MTP (Message Transport Layer) Level 3, the sigtran protocols provide the equivalent functionality of MTP Level 2.

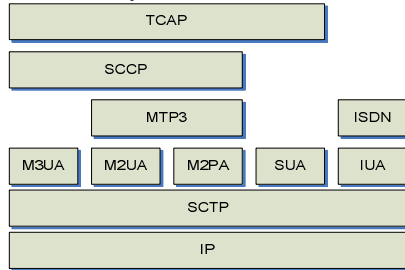


Fig. 5 SIGTRAN protocol stack [10]

The task of the data exchange between SS7 and SIGTRAN protocol is performed according to the following structure,

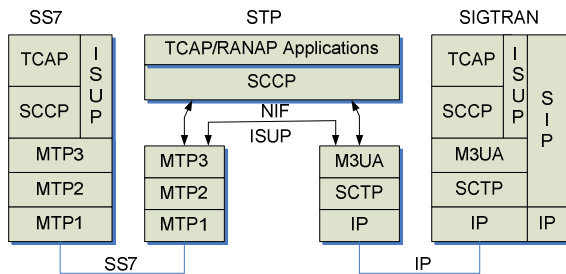


Fig. 6 SS7 and SIGTRAN protocol interconnection structure [11]

Our software exchanges SS7 protocol data via our designed USB E1 card and SIGTRAN protocol data from Ethernet port. This exchange of data between two different protocols takes place by analyzing the data [12] and maintaining the structure mentioned in figure 6. There are three part of the software which performs the data exchange operation. They are,

- Data reader
- Stack selector and
- Data exchanger

“Data Reader” window of our software is used to take SS7 data via E1 card or SIGTRAN data via Ethernet port. At first, the user has to specify the source to extract data for exchange. The selected source will be used as the base protocol which will be converted to the other protocol data. The software performing this task is shown bellow,

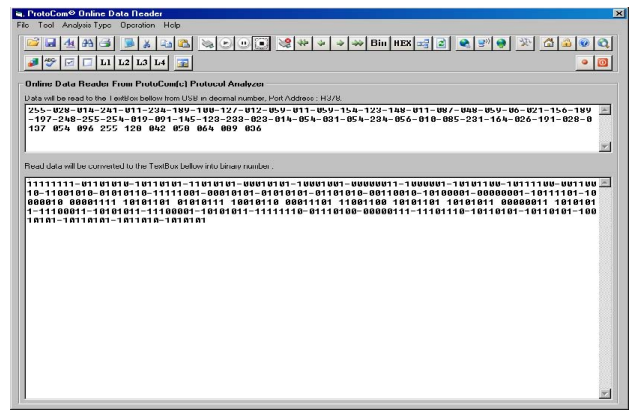


Fig.7 “Data reader” window of our developed software

“Stack Manager” window of our software lets the user to choose which message in the SS7 data is to be exchanged. It is very useful operation when the SS7 data length is very high. By choosing specific message, then user can reduce the number of exchanged messages so that only the desired message is transmitted. The software window of “Stack Manager” is shown bellow:

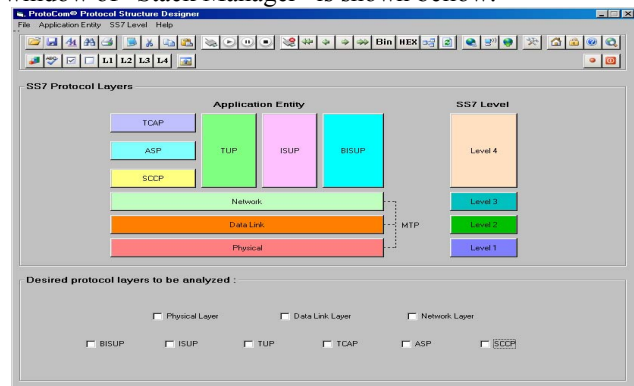


Fig.8 “Stack manager” window of our developed software

“Data Exchanger” is the last part of our software with which the selected portion of the base protocol is exchanged with other protocol. The software window performing data exchanging is shown bellow,

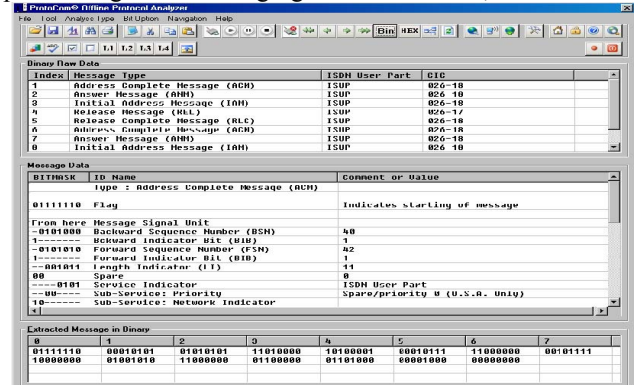


Fig.9 “Data exchanger” window of our developed software



#### IV. PROPOSED HARDWARE ARCHITECTURE

Our USB E1 card is an enhanced third-generation hardware that consolidates the essential pieces of industry-standard test equipment into a powerful, PC-Based USB E1 solution.

**PEF2256:** The PEF2256 is the latest addition to Infineon's FALC® family of sophisticated E1/T1/J1 framer and Line Interface Unit (LIU) transceivers which supports all standard E1/T1/J1 functions. [14]

**ATmega16:** ATmega16 is CMOS 8-bit microcontroller having 1 MIPS/MHz throughput. [15]

**FT245BM:** The FT245BM provides an easy cost-effective method of transferring data to / from a peripheral

and a host P.C. at up to 8 Million bits (1 Megabyte) per second. [16]

**74HC244:** These octal buffers are used to improve density of 3-state memory address drivers, clock drivers.

Using these ICs we have designed the USB E1 card. For avoiding complex PCB, we have designed the E1 card in two parts. They are,

1. Transceiver module: It includes transceiver IC, synchronizing clock, impedance matching transformer, transient voltage suppressor and connector for E1 link.

2. Controlling module: It includes microcontroller, USB interface, clock, serial and Parallel interface and power supply unit.

The schematics of the controlling and transceiver module are shown below:

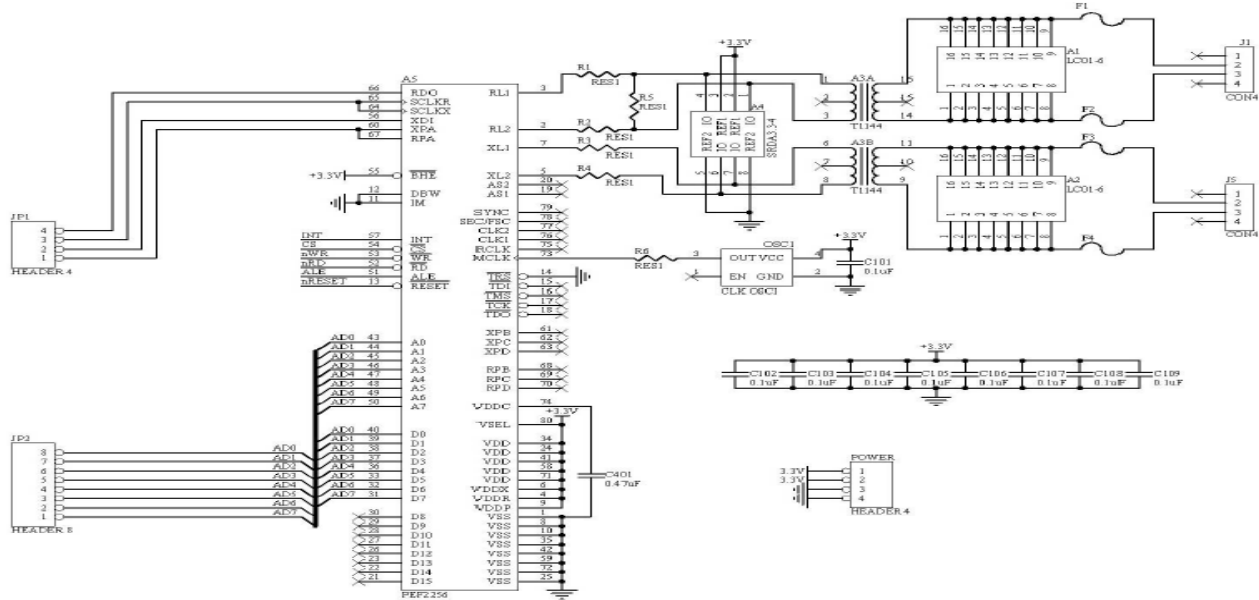


Fig. 10 Transceiver module of the E1 card.

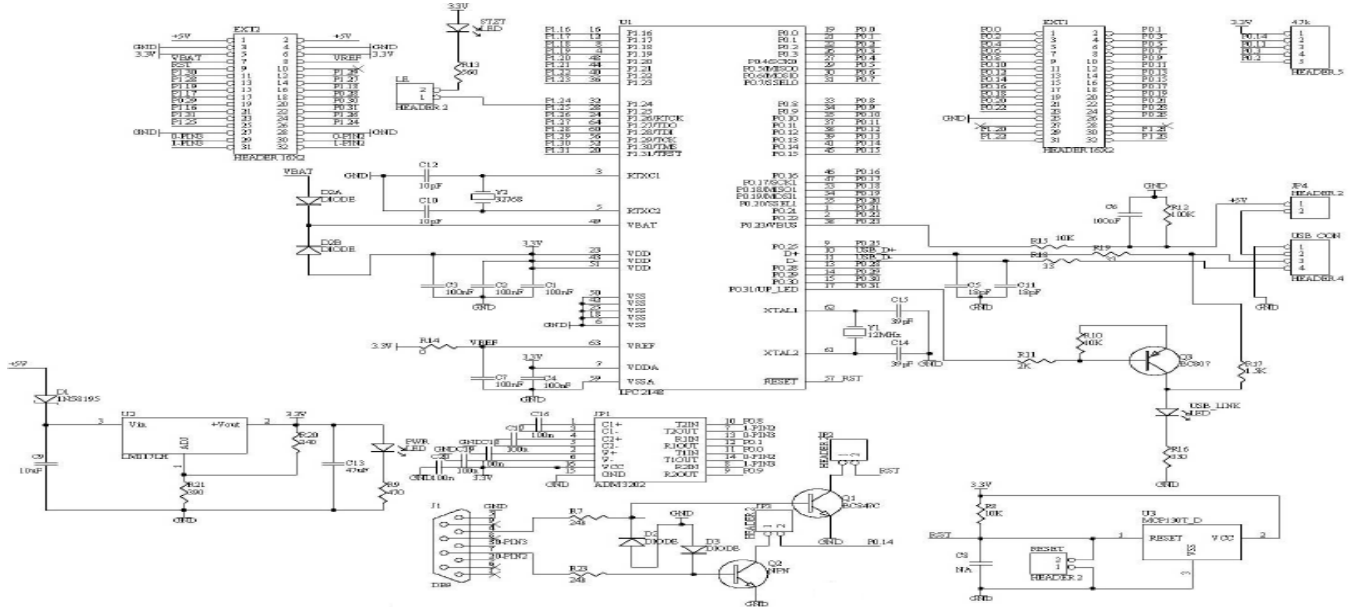


Fig. 11 Controlling part of the E1 card.



## V. RESULT

The hardware designing part is completed and currently it's working soundly.

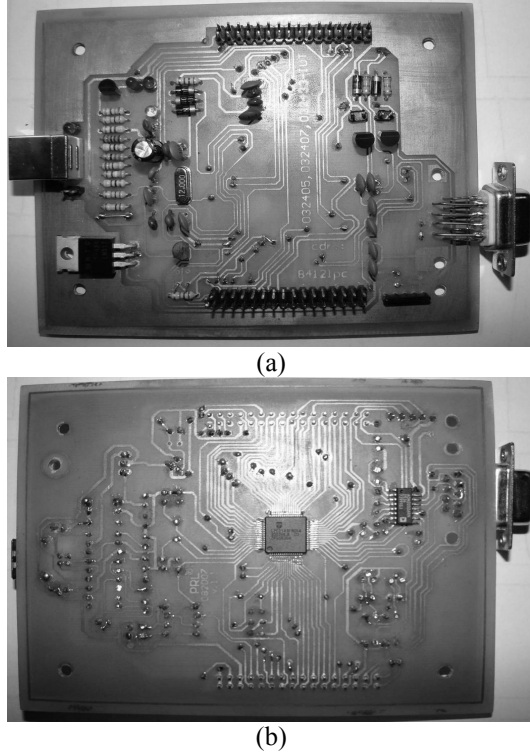


Fig. 12 (a) Top and (b) Bottom layer of the designed PCB of the E1 card

SS7 and SIGTRAN data can be read [Figure 7] according to the selected structure [Figure 8] and exchanged using our hardware and software [Figure 9] successfully. This shows that the interfacing of ATM network and IP network is done successfully. One of the major achievements of our work is the tremendous reduction of the cost. The manufacturing cost of the hardware is very minute compared to the market price of similar range hardware. A PCI slot E1 card costs about 500\$ in the current market whereas our hardware costs only 35\$.

## VI. CONCLUSION

This paper presents new SS7 and IP based SIGTRAN protocol interchanger software and related hardware to implement an improved distributed database based Automated Teller Machine (ATM) network where the countrywide existing IP network can be used. The current growing demand of ATM machines makes it essential to increase the number of ATM booth and to make the transaction faster and reliable. At this point we believe that

our software and related hardware will be a feasible solution of current demand of increasing the ATM booth number at a very low cost using the existing IP network. Also the distributed database based ATM machines will be a perfect solution of providing faster and reliable service to the customers.

## REFERENCES

- [01] ITU-T Recommendations, Q.700 to Q.716.
- [02] Fred Halsal, Multimedia communication.
- [03] IETF reference document "Framework Architecture for Signaling Transport", *RFC-2719*.
- [04] Performance Technologies protocol standard document.
- [05] Chukarin, A. Pershakov, N. Samouylov. 'Performance of Sigtran-based Signaling Links Deployed in Mobile Networks.' *Proc. of ConTel 2007*.
- [06] Dawis, E.P. 'Architecture of an SS7 protocol stack on a broadband switch platform using dualistic Petri nets.' *Communications, Computers and signal Processing*, 2001. PACRIM.
- [07] ITU-T reference document Q.701, Page 5.
- [08] "Tutorial on Signaling System 7 (SS7)" from Performance Technologies, Page 6.
- [09] ITU-T reference document Q.700 page 14.
- [10] SIGTRAN Protocol Suite by Jim Darroch, Page 8.
- [11] Reference document on "SS7 over IP signaling transport and SCCP" by *International Engineering Consortium*, Page 16.
- [12] Signaling System # 7 protocol analyzer software from GL communication.
- [13] GL communication protocol analyzer software standard
- [14] Falc2256 product brief of Infenion Technologies.
- [15] ATmega16L Datasheet from ATMEL Corporation.
- [16] FT245BM datasheet from Future technologies Ltd.

# Face Recognition using Gabor Filters

Md. Tajmilur Rahman and Md. Alamin Bhuiyan

Department of Computer Science and Engineering, Northern University Bangladesh and Department of Computer Science, Jahangir Nagar University, Shavar, Dhaka, Bangladesh.

**Abstract** — Gabor based face representation has achieved enormous success in face recognition. This paper addresses a novel algorithm for face recognition using neural networks trained by Gabor features. The system commences on convolving some morphed images of particular face with a series of Gabor filter co-efficient at different scales and orientations. Two novel contributions of this paper are: scaling of RMS contrast, and contribution of morphing as an advancement of image recognition perfection. The neural network employed for face recognition is based on the Multilayer Perceptron (MLP) architecture with back-propagation algorithm and incorporates the convolution filter response of Gabor jet. The effectiveness of the algorithm has been justified over a morphed facial image database with images captured in different illumination conditions.

**Index Terms** — Face Recognition, Gabor Filters, Morphing, RMS Contrast.

## I. INTRODUCTION

FACE representation using Gabor features has attracted considerable attention in computer vision, image processing, pattern recognition, and so on. The principal motivation to use Gabor filters is biological relevance that the receptive field profiles of neurons in the primary visual cortex of mammals are oriented and have characteristic spatial frequencies. Gabor filters can exploit salient visual properties such as spatial localization, orientation selectivity, and spatial frequency characteristics [1]-[2]. Considering these overwhelming capacities and its great success in face recognition, this paper addresses Gabor features to represent the face image and produces recognition task in tandem with neural network.

A fair amount of research works have been published in literature for Gabor based image recognition. Lades et al. developed a Gabor wavelet based face recognition system using dynamic link architecture (DLA) framework which recognizes faces by extracting Gabor jets at each node of a rectangular grid over the face image [3]. Wiskott et al. subsequently expanded on DLA and developed a Gabor wavelet-based elastic bunch graph matching (EBGM) method to label and recognize facial images [4]. In the EBGM algorithm, the face is represented as a graph, each node of which contains a group of coefficients, known as jets. However, both DLA and EBGM require extensive amounts of computational cost. Liu and Wechsler have developed a Gabor feature based classification protocol using the Fisher linear discriminate model for dimension reduction [5]. Shan et al. have developed an enhanced fisher

model using the AdaBoost strategy for face recognition [6]. Zhang et al proposed a face recognition method using histogram of Gabor phase pattern [7].

This paper proposes a Gabor filter coefficient based neural network approach for face recognition. Since rms contrast is sensible for image representation, attempts are focused on rms scaling. The scaling of rms contrast produces better recognition performance. Despite robustness, Gabor filter based feature selection methods are normally computationally expensive due to high dimensional Gabor features. To reduce feature dimension, this paper uses 15 Gabor filters; 3 for scaling and 5 for orientations.

The rest of the paper is organized as follows. Section II describes image pre-processing. Section III describes Gabor filter design. Section IV illustrates construction of the neural network with back-propagation algorithm. The experimental method and results are presented in Section V. Finally, in section VI, the results are discussed, conclusions are drawn and future works are proposed.

## II. IMAGE PRE-PROCESSING

The original images are first converted into gray scale images. Pointing the centers of two eyes on each face image, all images are properly rotated, translated, scaled and cropped into 100×100 pixels. Images are then subjected to some image pre-processing operations. The image pre-processing phase includes contrast and illumination equalization, histogram equalization, and morphing.

### A. Contrast and Illumination Equalization:

Contrast is a measure of the human visual system sensitivity. Although the role of contrast is significant in visual processing of computer displays, in almost all of the past literatures address the face recognition process in different lighting conditions with different illumination and contrast. To achieve an efficient and psychologically-meaningful representation, all images are processed with same illumination and RMS contrast. The RMS (root mean square) contrast metric, equivalent to the standard deviation of luminance, is given by [8]:

$$C_{rms} = \left[ \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad (1)$$

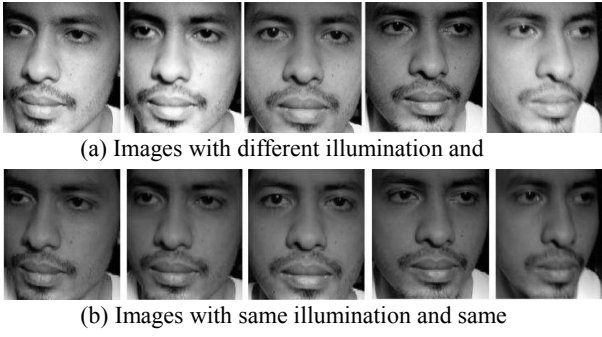


Fig. 1 Illumination and RMS contrast equalization. Images were captured at illumination angles of  $-38.4^\circ$ ,  $-21.6^\circ$ ,  $-0.2^\circ$ ,  $20.6^\circ$ ,  $37^\circ$ , respectively.

Where  $x_i$  is a normalized gray-level value such that  $0 < x_i < 1$  and  $x$  is the mean normalized gray level. With this definition, images of different human faces have the same contrast if their RMS contrast is equal. The RMS contrast does not depend on spatial frequency contrast of the image or the spatial distribution of contrast in the image. All images are maintained with the same illumination and same RMS contrast using the following equation:

$$\mathbf{g} = \alpha \mathbf{f} + \beta \quad (2)$$

Where  $\alpha$  is the contrast and  $\beta$  is the brightness to be increased or decreased from the original image  $\mathbf{f}$  to the new image  $\mathbf{g}$ . The values of  $\alpha$  and  $\beta$  are chosen empirically. The illumination and RMS contrast equalization process is illustrated in Fig. 1.

### B. Histogram Equalization:

The face images may be of poor contrast because of the limitations of the lighting conditions. So histogram equalization is used to compensate for the lighting conditions and to improve the contrast of the image [9].

### C. Morphing the images:

Morphing is a special effect in motion pictures and animations that changes (or morphs) one image into another through a seamless transition. Most often it is used to depict one figure turning into another through some magical or technological means or as part of a fantasy or surreal sequence. We are using the pixel based morphing [10] technique to enhance the perfection of facial image recognition more. Morphing gives us some intermediate stages between the source and target image which we can consider as the intermediate poses if the two images are of a same person.

When a face is in motion; turning left to right or right to left or some other else, the image captured at that moment might have motion blur. In order to recognize images at different poses, five different rotated images of a particular face at angle of  $-20^\circ$ ,  $-10^\circ$ ,  $0^\circ$ ,  $10^\circ$  and  $20^\circ$  generates 12

intermediate changing moments from the four sequential pairs with 3 images from each pair. So we are getting 17 different rotated facial images for a particular face image.

## III. DESIGNING GABOR FILTERS

Gabor filter works as a band pass filter for the local spatial frequency distribution, achieving an optimal resolution in both spatial and frequency domains. The 2D Gabor filter  $\psi_{f,\theta}(x,y)$  can be represented as a complex sinusoidal signal modulated by a Gaussian kernel function as follows [11]:

$$\psi_{f,\theta}(x,y) = \exp\left[-\frac{1}{2}\left\{\frac{x_{\theta_n}^2}{\sigma_x^2} + \frac{y_{\theta_n}^2}{\sigma_y^2}\right\}\right] \exp(2\pi f x_{\theta_n}), \quad (3)$$

$$\text{where, } \begin{bmatrix} x_{\theta_n} \\ y_{\theta_n} \end{bmatrix} = \begin{bmatrix} \sin \theta_n & \cos \theta_n \\ -\cos \theta_n & \sin \theta_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4)$$

$\sigma_x, \sigma_y$  are the standard deviations of the Gaussian envelope along the  $x$ - and  $y$ - directions,  $f$  is the central frequency of the sinusoidal plane wave, and  $\theta_n$  the orientation. The rotation of the  $x$ - $y$  plane by an angle  $\theta_n$  will result in a Gabor filter at the orientation  $\theta_n$ . The angle  $\theta_n$  is defined by:

$$\theta_n = \frac{\pi}{p}(n-1) \quad (5)$$

For  $n=1, 2, p$  and  $p \in \mathbf{N}$ , where  $p$  denotes the number of orientations.

Design of Gabor filters is accomplished by tuning the filter with a specific band of spatial frequency and orientation by appropriately selecting the filter parameters; the spread of the filter  $\sigma_x, \sigma_y$ , radial frequency  $f$ , and the orientation of the filter  $\theta_n$ . The important issue in the design of Gabor filters for face recognition is the choice of filter parameters. This research organizes 15 Gabor channels consisting of five orientation parameters  $\theta \in \left\{0, \frac{\pi}{5}, \frac{2\pi}{5}, \frac{3\pi}{5}, \frac{4\pi}{5}\right\}$  and three spatial frequencies,  $f \in \{0.6, 1.0, 1.4\}$  respectively.

The Gabor representation of a face image is computed by convolving the face image with the Gabor filters [12]. Let  $f(x,y)$  be the intensity at the coordinate  $(x,y)$  in a gray scale face image, its convolution with a Gabor filter  $\psi_{f,\theta}(x,y)$  is defined as:

$$g_{f,\theta}(x,y) = f(x,y) \otimes \psi_{f,\theta}(x,y) \quad (6)$$

where  $\otimes$  denotes the convolution operator. Fig. 2 illustrates the convolution result of a face image with a Gabor filter. The response to each Gabor kernel filter representation is a complex function with a real part  $\Re\{g_{f,\theta}(x,y)\}$  and an imaginary part  $\Im\{g_{f,\theta}(x,y)\}$  the magnitude response  $\|g_{f,\theta}(x,y)\|$  is expressed as:

$$\|g_{f,\theta}(x,y)\| = \sqrt{\Re^2\{g_{f,\theta}(x,y)\} + \Im^2\{g_{f,\theta}(x,y)\}} \quad (7)$$

This research uses the magnitude response  $\|g_{f,\theta}(x,y)\|$  to represent the features. To reduce the influence of the lighting conditions, the output of Gabor filter about each direction has been normalized. In the sequel, a transformation  $Q_{f,\theta}(x,y)$  to which  $\|g_{f,\theta}(x,y)\|$  is subjected is given by Fig. 3.

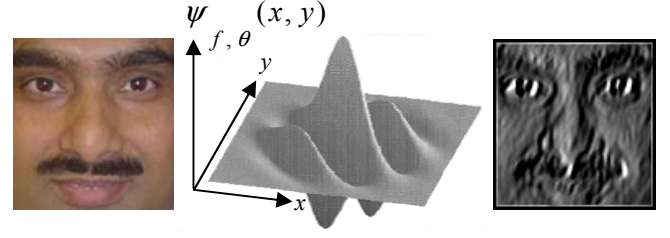


Fig. 3 Network architecture of Gabor based multi-layer Perceptron.

$$Q_{f,\theta}(x,y) = \frac{\|g_{f,\theta}(x,y)\|}{\sum_{\theta \in \Theta} \|g_{f,\theta}(x,y)\|} \quad (8)$$

The function  $Q_{f,\theta}(x,y)$  takes the value in the range  $[0, 1]$

and the parameter  $\theta$  is in  $\Theta \equiv \left\{0, \frac{\pi}{5}, \frac{2\pi}{5}, \frac{3\pi}{5}, \frac{4\pi}{5}\right\}$

#### IV. GABOR FILTER BASED NEURAL-NETWORK

Face recognition is achieved by employing a multilayer Perceptron with back-propagation algorithm. The architecture of the neural network is illustrated in Fig. 3.

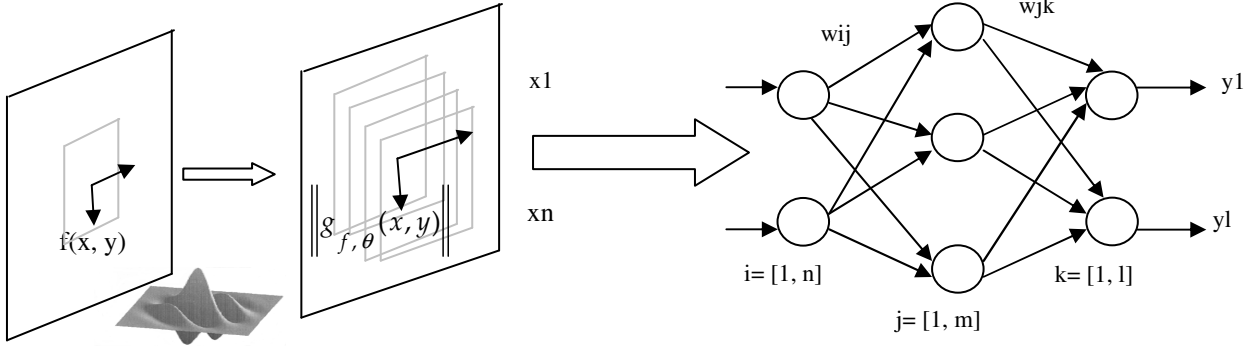


Fig. 3 Network architecture of Gabor based multi-layer Perceptron.

The first layer, called Gabor layer, receives the Gabor features. The number of nodes in this layer is, obviously equal to the dimension of the feature vector incorporating the Gabor features. The number of nodes in the output layer equals to the number of individual faces the network is required to recognize. The number of epochs for this experiment was 10,000 and the goal was to reduce the error margin up to 0.01. The recognition algorithm is illustrated in a flowchart as shown in Fig. 4.

In order to evaluate the effectiveness of the proposed method, experiments were carried out for real images at different illumination conditions and different orientations.

We used CMU Pose, Illumination, and Expression (PIE) database [13] and selected 200 images of 40 individual subjects with 2 different poses and 5 different illumination conditions. Half of the images in the database were taken for the morphing process. Three intermediate images of the morphing process for each of the images of were used as a training dataset and the remaining images were used as probe images in the recognition test.

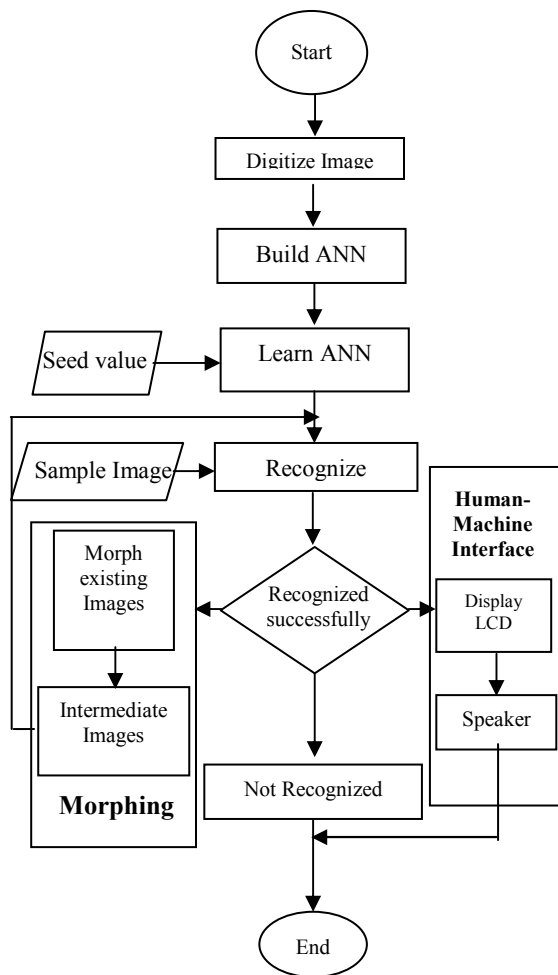


Fig. 4 Flow-chart of interactive image recognition for Human-Machine Interfacing.

All images were subjected to Gabor filters and were convolved with 15 Gabor filters. To each face image, the outputs were 15 images which record the magnitudes of the Gabor filter responses. Fig. 5 shows a typical face image being morphed and we took three different intermediate positions of morphing. Thus these three images (Fig. 5 i, ii, iii) were put into the database to train by Gabor filter. Fig. 6 shows the results for a typical face image.



Fig. 5 (a) Source image for morphing. (i) 30° pos (ii) 60° pos (iii) 90° pose of intermediate steps of morphing.

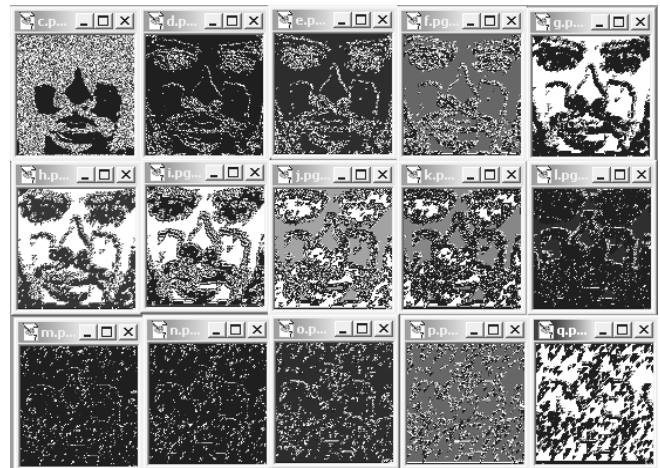


Fig. 6 Gabor filters response of a typical face image.

The output of the Gabor filters was used to train the neural network. The training curve, indicating the gradual reduction in error over several epochs due to back propagation learning algorithm, is shown in Fig. 7.

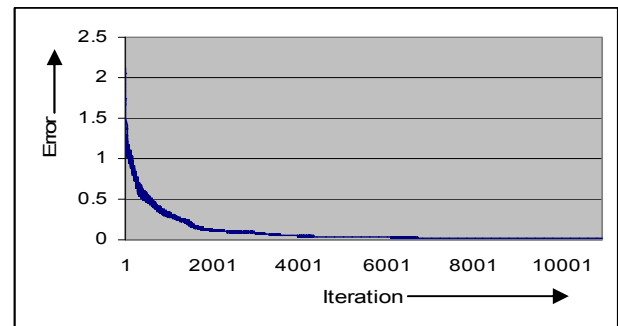


Fig. 7 Error versus iteration.

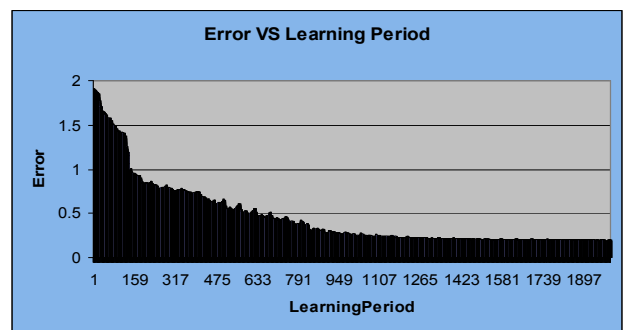


Fig. 8 Error Vs Iteration Graph for the image database without any morphed samples.

TABLE I  
RECOGNITION PERFORMANCE OF THE PROPOSED  
ALGORITHM

| Method                         | Correct Recognition (%) | Correct Rejection (%) |
|--------------------------------|-------------------------|-----------------------|
| EBGM                           | 75.29                   | 78.32                 |
| Log-polar                      | 77.38                   | 82.33                 |
| rms scaling Gabor              | 84.50                   | 87.75                 |
| Without any<br>Morphed Samples |                         |                       |
| rms scaling Gabor              | 96.07                   | 98.65                 |
| With Morphed<br>Samples        |                         |                       |

The graphical analysis implies that the training set of morphed images advances the learning and recognition process as well. This method can successfully trace the mutinous snapshots in between two different poses and increases the probability of recognition.

Finally, the effect of the proposed algorithm on the performance of face recognition has been analyzed. In order to evaluate the method for contrast equalization, we performed a comparison among our proposed system (rms scaling Gabor), the Elastic Bunch Graph Matching (EBGM), and the log-polar Gabor methods. The comparison results for the average recognition rates using different methods are furnished in Table I. The performance of the morphing process was evaluated for different images keeping the size of the source and target images same. To test the performance of using morphed intermediate poses for training set and image database we applied the gabor features to the image database not having any morphed samples. Two different image databases were created for the recognition of the face images. And trained for the process at a seed value of 0.5 and we got two different error vs. iteration graphs as shown in Fig. 7 and Fig. 8 respectively. For recognition we supplied same set of samples containing both images with motion blurs and without motion blurs to both the trained sets and the success rate of recognition was extensively high of our proposed method. Experimental results demonstrate that our method is capable of recognizing faces in gloomy environments also, with significant margin.

## VII. CONCLUSION

This paper presents a neural network based face recognition system using Gabor filter coefficients that can cope with illumination changes. The recognition performance has been improved substantially due to implication of contrast equalization using the rms value of the image pixels and adding the intermediate mutinous snapshots between two different poses by employing image morphing in the processing step.

A face recognition system should identify a face from a new image despite the variations between images of the same face. A common approach to surmount the impediments in image variations due to changes in the illumination conditions is to use image representations that are relatively insensitive to these variations. Examples of such representations are edge maps, image intensity derivatives, and images convolved with Gabor-like filters. This paper focuses on a pragmatic study that evaluates the effectiveness of the proposed method for facial images at different illuminations and different view angles. The processing time of Gabor transformation is reduced to less than one second with the image size of 100×100 pixels on a PC with a 1.2 GHz processor. From now on, the effectiveness of the method will be justified by testing it with face images of more persons and of some common databases. Since each pixel of the magnitude response of Gabor filter corresponds to a Gabor feature, the number of Gabor features for each sample is  $100 \times 100 \times 15 = 150,000$ . Therefore, our next step will be to improve the algorithm which would be able to employ more complex classifiers and distance measures to represent Gabor faces with spatial and frequency features.

## ACKNOWLEDGEMENT

This research is supported by Economic & Social Research Council RES-000-22-1759.

## REFERENCES

- [1] J.G. Daugman, "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters", *Journal of Optical Society America A*, Vol. 2, No. 7, 1985, pp. 1160 - 1169.
- [2] J. Buhmann, J. Lange, and C.V. Malsburg, "Distortion invariant object recognition by matching hierarchically labeled graphs", *Proceedings of International Conference on Neural Networks*, Washington DC, 1989, pp. 155-159.
- [3] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C.V. Malsburg, C. Wurtz and W. Konen, "Distortion invariant object recognition in the dynamic link architecture", *IEEE Trans. Computers*, Vol. 42, No. 3, 1993, pp. 300-311.
- [4] L. Wiskott, J.M. Fellous, N. Kruger, and C.V. Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Trans. Pattern Anal. Mach. Intel.*, Vol. 19, No. 7, 1997, pp. 775-779.
- [5] C. Liu, and K. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminate model for face recognition", *IEEE Trans. Image Process.*, Vol. 11, No. 4, 2002, pp. 467-476.
- [6] S. Shan, P. Yang, X. Chen and W. Gao, "AdaBoost Gabor Fisher classifier for face recognition", *Proc. IEEE Int. Workshop Analysis and Modeling of Faces and Gestures*, 2005, pp. 278-291.
- [7] B. Zhang, S. Shan, X. Chen, W. Gao, "Histogram of Gabor phase patterns: a novel object representation approach for face recognition", *IEEE Trans. Image Process.*, Vol. 11, No. 4, 2002, pp. 467-476.

- [8] E. Peli, "Contrast in Complex Images", Journal of Optical Society, Vol. 7, No. 10, 1990, pp. 2032-2040.
- [9] Y. Tae-Woong, and O. Il-Seok, "A fast algorithm for tracking human faces based on chromatic histograms", Pattern Recognition Letters, Vol. 20, No. 10, 1999, pp. 967-978.
- [10] Md. Tajmilur Rahman, Md. Alamin Bhuiyan, Ahsan Raja Choudhury, Alamin, Jobayer Bin Bakkre, "A Novel Approach of Image Morphing Based on Pixel Transformation", 10<sup>th</sup> International Conference on Computer and Information Technology (ICCIT 2007), UIU, Dhaka.
- [11] T. Andrysiak, and M. Choras, "Image retrieval based on hierarchical Gabor filters", International Journal of Mathematics and Computer Science, Vol. 15, No. 4, 2005, pp. 471-480.
- [12] M. Zhou, and H. Wei, "Face verification using Gabor Wavelets and AdaBoost", 18<sup>th</sup> International conference on Pattern Recognition, 2006, pp. 404-407.
- [13] T. Sim, S. Baker and M. Bsat, "The CMU Pose Illumination and Expression (PIE) Database", Proceedings of the 5th International Conference on Automatic Face and Gesture Recognition, 2002.

# High-speed distance measurement between moving vehicles with NIR-LED markers

M. Fukuzawa, H. Hama, N. Nakamori, and M. Yamada

Graduation School of Science and Technology, Kyoto Institute of Technology,  
Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan

**Abstract** — An image-based technique with high measurement rate has been proposed to measure the distance between two moving vehicles. By combining several near-infrared (NIR) LED markers mounted on the forward vehicle with an NIR-filtered CMOS camera mounted on the backward vehicle, one can clearly recognize the marker shape and positions without any adjusting the focus lens of camera, and hence he can extremely reduce the time of digital preprocessing such as noise elimination and marker segmentation. The distance may be easily calculated from the geometrical transformation of the marker positions obtained in sub-pixel resolution from the gravity center of the pixel value. The image-based technique proposed here is actually examined in the vehicles moving on linear and curved tracks. It is experimentally confirmed that the measurement rate attains to 1000 frames/sec in the windowed image of 752×34 pixels, with a small processor with the clock of 600MHz, including the capture time of the image, the calculation time of the gravity center, and the geometrical transformation time. The measurement error is less than 1 % for the distance range of 0.5 to 3 m.

**Index Terms** — distance measurement, image processing, infrared imaging, motion control, moving vehicles

## I. INTRODUCTION

Distance measurement is one of important applications in the vision system for industrial use such as nondestructive testing of products and motion control of vehicles and robots. The high-speed measurement in the wide field of vision is strongly demanded for the high-speed motion control of cargo vehicles which are moving on a common track in production facilities.

In order to cover the wide field of vision, several techniques of distance measurement were developed. In the case using active light source, several laser range-finding techniques such as optical triangulation, time-of-flight, phase shift measurements and frequency modulated continuous wave (FMCW) method were proposed [1]. In the case of passive measurement, several stereo vision techniques and the depth-of-focus techniques were proposed [2-4]. Those techniques mentioned above took a time for image capturing and digital preprocessing such as laser scanning, focus adjustment, noise elimination, marker segmentation and stereo matching, although there was an advantage that a priori knowledge of shape and position concerning on target was not needed. Therefore, a

new technique is desired to develop for the high-speed measurement of distance in the wide field of vision.

In this paper, we proposed an image-based technique to measure the distance between two vehicles at the high measurement rate in the wide field of vision. Here, we assume that the forward vehicle on which several luminous markers are mounted and the backward vehicle on which a single camera is mounted to take the image of markers are moving on a curved rail track. All the backgrounds except markers are effectively eliminated by combining NIR-LED markers and NIR-filtered camera and hence the time of digital preprocessing may be extremely reduced. Then, the distance can be numerically calculated from a straightforward technique of image processing: that is, the geometrical transformation of the marker positions obtained in sub-pixel resolution from the gravity center of the pixel value. The formulation of the geometrical transformation is given below. Also, experimental results are shown for confirming the performance of proposed technique.

## II. A DISTANCE MEASUREMENT TECHNIQUE BASED ON A PROJECTION OF LUMINOUS MARKERS

Figure 1 shows our arrangement of two vehicles moving on a curved track, in which the vehicles are drawn as cube by solid lines. The center of rear panel of the forward vehicle is defined as the origin of a local coordinate system  $(x, y, z)$ , and the forward, left and upper directions are represented as  $x$ -,  $y$ -, and  $z$ -axes, respectively. The other local coordinate system  $(x', y', z')$  is linked to the backward vehicle in the similar manner to the forward vehicle. A distance between the origins of two coordinates is defined as  $d$ . An angle between the  $x$ - and  $x'$ -axes is defined as  $\phi$ . The other angle between  $x'$ -axis and the direction of  $d$  is defined as  $\theta$ . The  $d$ ,  $\phi$ , and  $\theta$  are the values to be measured in our study. In the forward vehicle, eight luminous markers are attached at the corner and the center on the front, rear and side panels, which are noted as white circles in the Fig.1. In the backward vehicle, a single camera is mounted at the origin so that the optical axis of the camera should be similar to the  $x$ -axis.



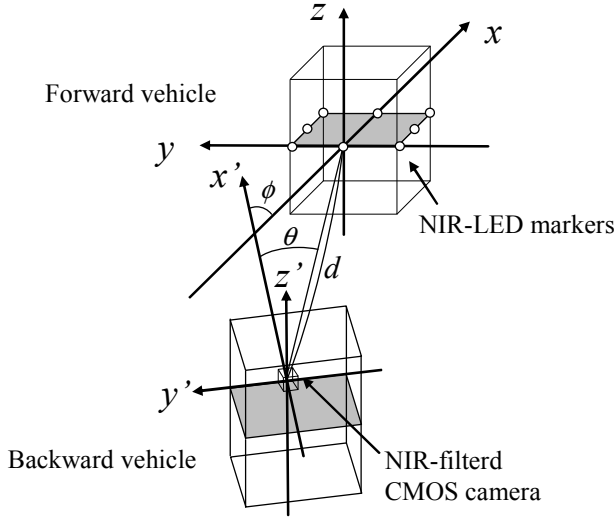


Fig. 1. Arrangement of two vehicles moving on a curved track.

In order to simplify the geometrical transformation from the marker positions to the vehicle distance, we assume that the  $x$ - $y$  plane is similar to the  $x'$ - $y'$  one. This assumption is reasonable when the track is flat without up and down slopes and the vehicles are not tilted. In this case, the markers are lined up along the  $y'$ -axis in the image surface of the camera. Then, the distance can be transformed only from the  $y'$ -coordinate of the marker positions in the image. Under the vehicle arrangement shown in the Fig. 1, the image reveals five markers on the rear and right-side panels because three markers on the front and left-side panels are hidden behind the vehicle itself. If the vehicles are moving on a linear track, the image reveals only three markers on the rear panel. In this study, we formulate the transformations only from the  $y'$ -coordinate of the marker positions in those two cases of three and five markers. However, it should be noted that the distance can also be transformed from the  $y'$ - and  $z'$ -coordinates of the marker positions even if the  $x$ - $y$  plane is not similar to the  $x'$ - $y'$  one, relating to the track slope or vehicle tilts.

Figure 2 shows  $x$ - $y$  in-plane arrangement of three markers on the rear panel of the forward vehicle and the camera on the backward vehicle. The markers are noted as  $p_0, p_1, p_2$ . The center of the camera lens with the focal length  $f$  is placed at  $p_c(x_c, y_c)$  and its vision field is assumed to be enough wide to cover the three markers. Strictly speaking, only three markers are revealed in the image if  $-S \leq y_c \leq S$  or  $-2S \leq x_c \leq 0$ . The  $y'$ -coordinate of the marker projected on the  $y'$ - $z'$  image surface is defined as  $p'_0, p'_1, p'_2$ . In this case,  $d, \phi$ , and  $\theta$  are obtained from  $p'_0, p'_1, p'_2, f$  and  $S$  by solving following equations:

$$\tan \phi = -\frac{(r_1 - r_0) + (r_1 - r_2)}{r_2(r_1 - r_0) + r_0(r_1 - r_2)},$$

$$r_k = \frac{p'_k}{f}, \quad (1)$$

$$d = \sqrt{x_c^2 + y_c^2}, \quad (2)$$

$$\begin{aligned} x_c &= -\frac{(r_1 \tan \phi + 1)y_c}{r_1 - \tan \phi} \\ &= -\frac{(r_0 \tan \phi + 1)(y_c - S)}{r_0 - \tan \phi} \\ &= -\frac{(r_2 \tan \phi + 1)(y_c + S)}{r_2 - \tan \phi} \end{aligned} \quad (3)$$

$$\begin{aligned} y_c &= \frac{(r_1 - \tan \phi)(r_2 \tan \phi + 1)S}{(r_2 - \tan \phi)(r_1 \tan \phi + 1) - (r_1 - \tan \phi)(r_2 \tan \phi + 1)}, \quad (4) \\ &= \frac{-(r_1 - \tan \phi)(r_0 \tan \phi + 1)S}{(r_0 - \tan \phi)(r_1 \tan \phi + 1) - (r_1 - \tan \phi)(r_0 \tan \phi + 1)} \end{aligned}$$

$$\theta = \tan^{-1}\left(\frac{y_c}{x_c}\right) - \phi. \quad (5)$$

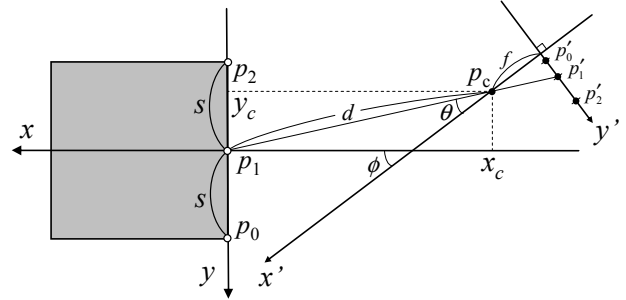


Fig. 2. In-plane arrangement of three markers on the rear panel of the forward vehicle and the camera on the backward vehicle.

Figure 3 shows  $x$ - $y$  in-plane arrangement of five markers on the rear and right-side panels of the forward vehicle and the camera on the backward vehicle. In addition to the  $p_0, p_1, p_2$ , the markers on the right-side panel are noted as  $p_3$  and  $p_4$  together with their projected positions as  $p'_3$  and  $p'_4$  in the Fig. 3. For convenience of the calculation, the marker  $p_2$  is placed at the origin and then, a distance between  $p_2$  and  $p_c$  and an angle between  $x'$ -axis and the  $p_2$ - $p_c$  direction are defined as  $d'$  and  $\theta'$ , respectively. In this case,  $d', \theta'$ , and  $\phi'$  are obtained from  $p'_0, p'_2, p'_4, f$  and  $S$  by solving following equations:

$$\tan \phi = \frac{r_4(r_2 - r_0) - (r_2 - r_4)}{(r_2 - r_0) + r_0(r_2 - r_4)},$$

$$r_k = \frac{p'_k}{f}, \quad (6)$$

$$d' = \sqrt{x_c^2 + y_c^2}, \quad (7)$$

$$\begin{aligned} x_c &= -\frac{(r_2 \tan \phi + 1)y_c}{r_2 - \tan \phi} \\ &= -\frac{(r_0 \tan \phi + 1)(y_c - 2S)}{r_0 - \tan \phi}, \quad (8) \\ &= -\frac{(r_4 \tan \phi + 1)y_c - 2S(r_4 - \tan \phi)}{r_4 - \tan \phi} \end{aligned}$$

$$\begin{aligned} y_c &= \frac{-2S(r_2 - \tan \phi)(r_4 - \tan \phi)}{(r_4 - \tan \phi)(r_2 \tan \phi + 1) - (r_2 - \tan \phi)(r_4 \tan \phi + 1)}, \quad (9) \\ &= \frac{-2S(r_2 - \tan \phi)(r_0 \tan \phi + 1)}{(r_0 - \tan \phi)(r_2 \tan \phi + 1) - (r_2 - \tan \phi)(r_0 \tan \phi + 1)} \end{aligned}$$

$$\theta' = \tan^{-1} \left( \frac{y_c}{x_c} \right) - \phi. \quad (10)$$

The  $d$ ,  $\phi$ , and  $\theta$  can also be derived from  $d'$ ,  $\theta'$ , and  $\phi$  by translation of the  $y$ -axis. Finally, the  $d$ ,  $\phi$ , and  $\theta$  for any arrangement of the vehicles are determined by using either eqs.(1) to (5) or eqs.(6) to (10) according to the number of markers revealed in the image.

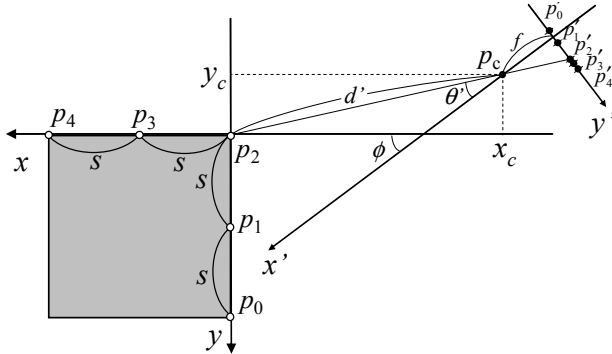


Fig. 3. In-plane arrangement of five markers on the rear and right-side panels of the forward vehicle and the camera on the backward vehicle.

### III. EXPERIMENTAL RESULTS

The distance measurement was performed with a horizontal flat track including linear and curved parts from

which two metal cages are vertically hanged as vehicles. The NIR-LED markers were attached to the forward vehicle and their image was taken by a NIR-filtered CMOS camera (752×480 pixels) mounted on the backward vehicle while varying the relative position of the vehicles. In order to reduce the processing time, a windowed image (752×34 pixels) including all the markers near the  $x'$ - $y'$  plane was captured and processed by a small signal processor. The windowed image was divided into several regions including a single marker based on their approximate position estimated from the profile along the  $y'$ -coordinate of the image. Consequently, the precise position of the marker  $p_k'$  was obtained by calculating the  $y'$ -coordinate of the gravity center in every region. In order to confirm the performance of our proposed technique, the distances and directions from the camera center to each marker were also measured as  $d_L$  or  $d'_L$  by a laser-distancemeter (Leica DISTO D3) with a precise rotary stage mounted on the backward vehicle.

Figure 4 shows the measurement results of the  $d$  on the linear track: (a) experimental arrangement, (b) windowed images, and (c) comparison between  $d$  and  $d_L$ . The windowed images clearly revealed three markers lined up along the  $y'$ -axis. The marker intervals in the image were decreased with the increase of the distance between the vehicles. The calculated values of  $d$  revealed very good agreement with those of  $d_L$ . The maximum error of  $d$  was less than 1 % for the distance range of 0.5 to 3 m.

Figure 5 shows the results of the  $d'$  on the curved track measured in the same manner as shown in the Fig. 4. The windowed images revealed that the marker positions were moved to the right and their intervals were decreased with the increase of the vehicle distance. The calculated values of  $d'$  also agreed with those of  $d'_L$ . The error of  $d'$  was similar extent to that in the linear track. Because of the fixed curvature of the track, the variable range of the  $\theta'$  was less than 4 degrees in this experiment. Further measurements of  $\theta'$  for error analysis is still in progress. However, the error of  $\phi'$  in the range of 20 to 80 degrees was less than 2 %, which is enough to identify the vehicle position on the curved track.

The measurement rate was examined by evaluating the execution time of image capture, calculation of the gravity center, and geometrical transformation. In the case of the signal processor with the clock of 600MHz, it attained to 1000 frames/sec for the windowed image of 752×32 pixels. It was very lower processing-cost than that of the versatile techniques with complex algorithms of digital preprocessing such as stereo visions. Therefore, we can conclude that the proposed technique is useful for high-speed motion control of the cargo vehicles.

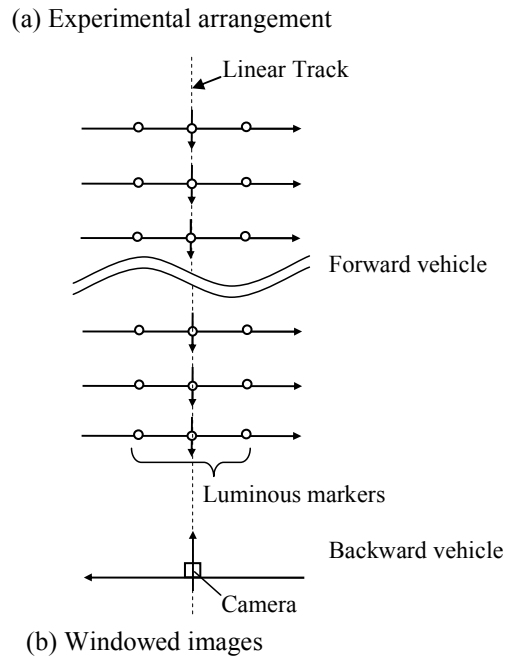


Fig. 4. Measurement results of  $d$  on the linear track: (a) experimental arrangement, (b) windowed images, and (c) comparison between  $d$  and  $d_L$ .

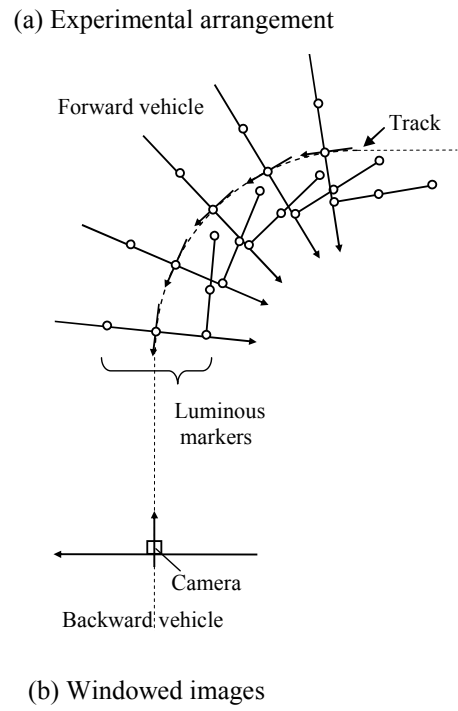


Fig. 5. Measurement results of  $d'$  on the curved track: (a) experimental arrangement, (b) windowed images, and (c) comparison between  $d'$  and  $d_L'$ .

## VI. CONCLUSION

High-speed distance measurement between two vehicles has been successfully performed by processing the image of the forward vehicle with NIR-LED markers taken by a NIR-filtered CMOS camera mounted on the backward vehicle. It was experimentally confirmed that the measurement rate attains to 1000 frames/sec in a windowed image of 752×34 pixels, with a small processor with the clock of 600MHz, including the capture time of the image, calculation time of the gravity center, and geometrical transformation time. The measurement error was less than 1 % for the distance range of 0.5 to 3 m. Therefore, we can conclude that the proposed technique is useful for high-speed motion control of the cargo vehicles.

## REFERENCES

- [1] M.-C. Amann, T. Bosch, M. Lescure, R. Myllylä, and M. Rioux, "Laser ranging: a critical review of usual techniques for distance measurement", *Opt. Eng.* vol. 40, no. 1, pp. 10-19, January 2001.
- [2] D.Nistér, O.Naroditsky, and J.Bergen, "Visual Odometry", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1, pp. I652-I659, 2004.
- [3] D.M. Gavrilu and V. Philomin, "Real-time object detection for 'smart' vehicles", *Proceedings of the IEEE International Conference on Computer Vision* 1, pp. 87-93, 1999.
- [4] John Ens and Peter Lawrence, "Investigation of methods for determining depth from focus", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 2, pp. 97-108, February 1993.

# SEI and SHI Representations for Human Movement Recognition

Mohiuddin Ahmad and Md. Zulfikar Hossain\*

Dept. of Electrical and Electronic Engineering, Khulna University of Engineering and Technology,  
Khulna-9203, Bangladesh

\*Dept. of Electronics and Communication Engineering, Khulna University of Engineering and  
Technology, Khulna-9203, Bangladesh

**Abstract** — In this paper, we propose spatio-temporal silhouette representations, called silhouette energy image (SEI) and silhouette history image (SHI) to characterize motion and shape properties for recognition of human movements such as human actions, activities in daily life. The SEI and SHI are constructed by using the silhouette image sequence of an action. The span or difference of the end time and start time is used to make the SHI. For addressing the human shape variability, we used the variation of the anthropometry of the person. We extract the features based on geometric shape moments. We tested our approach successfully in the indoor and outdoor environment. Our experimental results show that the proposed method of human action recognition is robust, flexible and efficient.

**Index Terms**—Anthropometry, human action recognition, shape variability, silhouette energy image, silhouette history image.

## I. INTRODUCTION

Recognition of human actions from multiple views by the classification of image sequences has the applications in video surveillance and monitoring, human-computer interactions, model-based compressions, video retrieval in various situations. Typical situations include scenes with moving or clutter backgrounds, stationary or non-stationary camera, scale variation, starting and ending state variation, individual variations in appearance and cloths of people, changes in light and view-point, and so on. These situations make the human action recognition a challenging task.

The standard approach for human action recognition is to extract a set of features from each image sequence frame and use these features to train classifiers and to perform recognition. Therefore, it is important to consider the appropriateness and robustness of features of action recognition in varying environment. Actually, there is no rigid syntax and well-defined structure for human action recognition available. Moreover, there are several sources of variability that can affect human action recognition, such as variation in speed, viewpoint, size and shape of performer, phase change of action, scaling of persons, and so on. In addition, the motion of the human body is non-rigid in nature. These characteristics make human action recognition a sophisticated task. Considering the above circumstances, we consider some issues that affect the development of models of actions and classifications, which are as follows: (i) An action can be characterized by the local motion of

human body parts, (ii) an action can be illustrated by the silhouette image sequence of the human body, which can be regarded as global motion flow, (iii) the trajectory of an action from different viewing directions is different; some of the body parts (part of hand, lower part of leg, part of body, etc) are occluded due to view changes, and (iv) human actions depend on several variability, such as anthropometry, method of performing the action, speed, phase variation (starting and ending time of the action), and camera view variations such as zooming, tilting, and rotating.

Among various features, the motion of the body parts and human body shape play the most significant roles for recognition. Motion based features can represent the approximation of the moving direction of the human body and human action can be effectively characterized by motion rather than other cues, such as color, depth, and spatial features. In the motion-based approach, the motion information of the human such as optic flows, affine variation, filters, gradients, spatial-temporal words, and motion blobs are used for recognizing actions. Motion-based action recognition had been performed by several researchers; a few of them are [3, 11, 12, 14]. However, motion-based techniques are not always robust in capturing velocity when motions of the actions are similar for the same body parts. On the other hand, the human body silhouette represents the pose of the human body at any instant in time, and a series of body silhouette images can be used to recognize human action correctly, regardless of the speed of movement. Different descriptors of shape information of motion regions such as points, boxes, silhouettes, and blobs are used for recognizing or classifying actions. Several researchers performed action recognition using shapes or silhouettes, such as [1, 2]. Bobick and Davis [1] proposed the motion energy image (MEI) and motion history image (MHI) for human movement representation and recognition and were constructed from the cumulative binary motion images. We propose silhouette energy image (SEI) and silhouette history image (SHI) which gives shape information with global motion information but MEI and MHI give only motion information.

In addition of shape and motion, several variabilities that occurred frequently are also responsible for human action recognition. Sheikh and Shah [15] explicitly identified three sources of variability in action recognition, such as viewpoint, execution rate, and anthropometry of actors and

they used the 3D space with thirteen anatomical landmarks for each image. In contrast to their work, we explicitly define and employ the anthropometry variation, camera observations (zooming of a person, slanting body, and rotation of human body), phase variation, speed variations and multiple views variation of the action. Related works have typically concentrated on the variability in viewpoint [13] by deriving view invariant features or proposing a view invariant algorithm.

During the action recognition of persons, we utilize the global shape motion features in addition to several variabilities for recognizing the periodic as well as non-periodic actions. The global shape motions are extracted from geometric shape of models. Therefore, based on the combined information of global motion, sources of variability's, and multiple views, human action recognition is more robust and flexible. We propose to recognize several actions of humans in the daily life from multiple views learning of local and global motion features using the multiclass support vector machine (MCSVM).

In our system, we assume that silhouettes of an image sequence are correctly captured. From the silhouette image sequence, we estimate the temporal boundary (i.e. period or duration) of each action. Depending on the temporal boundary, an action model (i.e. SHI) is constructed by the silhouette image sequence. Using the variability parameters and the SHI, variability models (adaptable models) are generated. The models are characterized by various geometric shape motions. We learn an action for multiple views global motion descriptors by using a MCSVM, and generate SVM models for specified actions. For recognizing actions, we classify (using the similarity of features) descriptions using SVM models. The actions modelling and classification in this work involve both the Korea university full body gesture database (FBGDB) [4] and the KTH database (KTHDB) [10]. Of particular interest is the detection method, which we use for the recognition of several daily actions of elderly people for human-robot interaction (HRI) or similar applications.

This paper is organized as follows: Section II presents action representation and variability generation in our system. Section III discusses global motion descriptors of combined models. Section IV presents experimental results and discussions of the selected approaches. Finally, conclusions are drawn in Section V.

## II. SEI, SHI, AND HUMAN SHAPE VARIABILITY

The silhouette history image represents the shape history information of an image sequence. This history information represents the time history for performing an action. SEI and SHI gives the idea of motion and action recognition by watching silhouettes.

### A. Human Action

Human action is the movement of humans for performing a task within a short period of time. The action may be simple or complex depending on the number of body limbs involved in the action. Many actions performed by humans have cyclic nature and they show periodicity of short duration. Besides, many actions show single occurrence or non-periodic with time frame of specific length (i.e. duration). We consider human actions daily performed which are almost cyclic in nature, either multiple cyclic (period= $nT$ ) or periodic actions (different types of walking, running, jogging, etc), and single occurrence (duration= $p$ ) or non-periodic actions (bowing, raising the hand, sitting on the floor, etc). Under the above circumstances, it is possible to transform a human action in the spatio-temporal space or 3D space, into a 2D spatial space, where the 2D space contains spatial information with temporal information.

### B. Silhouette Energy Image (SEI)

We define the silhouette history image as the image where the time variation of an action is represented by the body shape information. Let us assume  $x_t = f(x, y, t)$  is the silhouette image in a sequence at time  $t$ , which includes an action under duration or a period as shown in fig. 1. Therefore, SEI ( $SEI = S(x, y)$ ) is defined by Eq. (1): Moreover, the standard deviation image as well as motion variation image is shown in fig. 1(b) and 1(c) respectively.

$$A_{2D} = S(x, y) = \frac{1}{t_e - t_s} \int_{t_s}^{t_e} x_t dt \quad (1)$$

$$\sigma(x, y) = \sqrt{\frac{\int_{t_s}^{t_e} f^2(x, y, t) dt}{t_e - t_s} - \left( \frac{1}{t_e - t_s} \int_{t_s}^{t_e} f(x, y, t) dt \right)^2} \quad (2)$$

$$COV(x, y) = \frac{\sigma(x, y)}{S(x, y)} \quad (3)$$

Here,  $t_s$  and  $t_e$  are the starting and ending states of an action. Since the average 2D image stores the global motion distribution and orientation of the silhouette images, we can designate this as a SEI. The number of frames in the action depends on the person, time, and type of action. Since, we use the average of the time sequence silhouette images; the normalized variation affects are very low. Figure 1 shows the sample silhouette images with the SEI of the "bowing" action along with the variation of motion. This representation shows the shape as well as motion changes of an action.

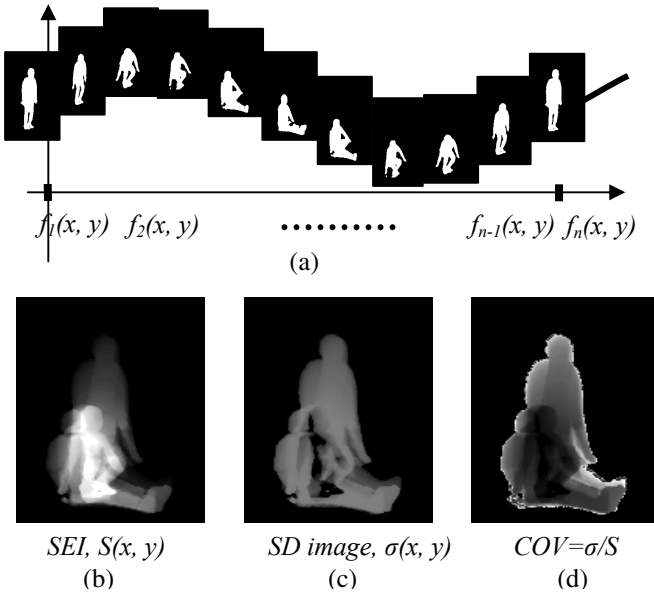


Fig. 1. Human action model representation. (a) Silhouette image sequence of sitting on the floor action. (b) SEI. (c) Standard deviation image. (d) Covariation image

The SEI represents an action model (AT), due to the following reason: (1) The energy of a pixel at every point is a result of an action formation; (2) Each silhouette represents the unit energy of a human action at any instance; (3) It determines the energy distribution of an action.

### C. Silhouette History Image (SHI)

Silhouette history image (SHI) refers to the shape variation of the image sequence in time. By SHI, we represent how the silhouette of an image sequence is moving. It does not show only the motion image representation as [1] but also it represents the global motion orientation of an action at any instant of time. The shape of the person at present state as well as the person's global motion orientation is visualized by the SHI.

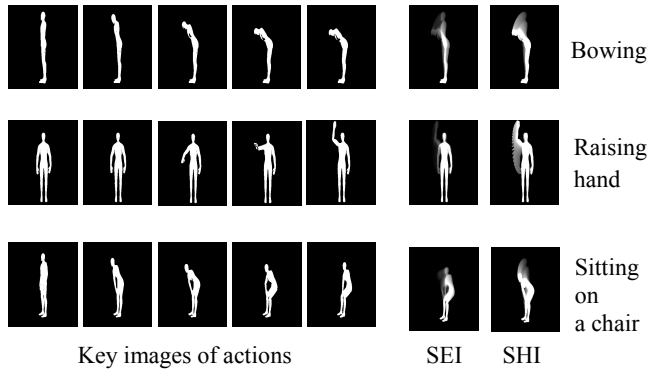


Fig. 2. Silhouette energy image (SEI) and silhouette history image (SHI) of human actions

From the image sequences of some actions, we show the SEI as well as SHI as shown in Fig. 2. Figure 2 shows several key frames of an action and the formation of both SEI and SHI. We construct these representations by using plastic model of human. From both SHI, we can show the action orientation or simply the direction of the action. From the MHI representation in [1], SHI represents both body shape and global motion change. The brighter region represents the recency of the human body shape.

### D. Variability of Body Shape

To consider the diversity of modeling (learning) and classifying actions, we consider multiple human body shape, and the produced models are known variability models (VTs). The anthropometry variation of an action represents the width and height variation of a person. Due to these variations, human action recognition should adapt anthropometry. We define basic eight sets of anthropometric variations by Eq. (4).

$$S(x, y)|_a = \begin{cases} S(x-a, y), & Hg \\ S(x, y-b), & Hh \\ S(x+a, y), & Lg \\ S(x, y+b), & Lh \\ S(x-a, y+b), & HgLh \\ S(x+a, y-b), & LgHh \\ S(x-a, y-b), & HgHh \\ S(x+a, y+b), & LgLh \end{cases} \quad (4)$$

Where,  $H$ =Higher  $g$ =girth,  $L$ =Lower, and  $h$ =height.

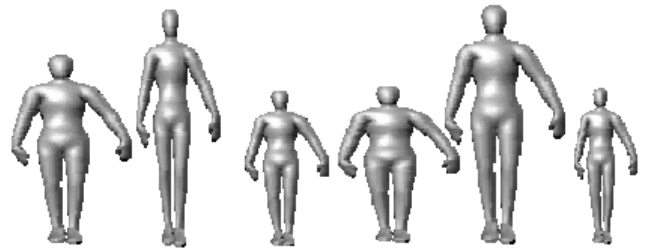


Fig. 3. Anthropometric variation images with different body width and height

where,  $a$  &  $b$  are width and height variation parameters with  $a \approx 0.067R \sim 0.1R$  and  $b \approx 0.067C \sim 0.1C$ . The variability models can be constructed by resizing the human body using bilinear interpolation method. We can generate a set of anthropometric images which are shown in Fig. 3.

### III. GLOBAL GEOMETRIC SHAPE & MOTION DESCRIPTION

We express the global shape motion description by the model orientation, span and elevation of motions, geometric and orthogonal moments. Therefore, we express the geometric shape motions by  $\{S_g, S_z, S_d\}$ . The symbols are defined in the following subsections.

#### A. Geometric Moments

Moments and function of moments have been utilized as pattern feature in pattern recognition applications. Such features capture global information about the image and do not require close boundaries as required by Fourier descriptors. By positioning the center of mass (COM), we can differentiate the motions for each action. Since this feature should be independent of the location of a person, then we consider the relative appropriate position for each action. Hu [6] introduced seven nonlinear functions,  $h_i$ , where  $i = 1, 2, \dots, 7$  defined on regular moments using central moments that are translation, scale, and rotation invariant. To achieve the consistent camera observations, we used non-orthogonal features, namely Hu moments,  $h = (h_1, h_2, \dots, h_7)$  [6] which are slightly modified to extract specific characteristics.

$$\eta_{pq} = \frac{\mu_{pq}}{m_{00}^\gamma} = \frac{1}{m_{00}^\gamma} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q S(x, y) \quad (5)$$

$$\text{where, } m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q S(x, y),$$

$(\bar{x}, \bar{y}) = (m_{10}/m_{00}, m_{01}/m_{00})$  and  $\gamma = (p + q + 2)/2$   $\forall (p + q) \geq 2$ . The advantage of a moment's methods is that they are mathematically concise and for the intensity image of action models, they reflect both shape and global motion distribution within it.

#### B. Zernike Moments

The geometric moment shows highly inaccurate results when the image is noisy. Zernike polynomials provide very useful moment kernels, present native rotational invariance and are far more robust to noise. Scale and translation invariance can be implemented using moment normalization. The magnitude of Zernike moments has been treated as shape features because they are rotation invariant. The two-dimensional Zernike moments of an image intensity function  $f(\rho, \theta)$  with order  $n$  and repetition  $m$  is expressed as follows [9]:

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 R_{nm}(\rho) e^{-jm\theta} f(\rho, \theta) \rho d\rho d\theta \quad (6)$$

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!((n+|m|)/2-s)!(n-|m|/2)!} \rho^{n-2s} \quad (7)$$

$$\text{where, } \rho = \frac{\sqrt{(2x - N + 1)^2 + (N - 1 - 2y)^2}}{N},$$

$$\theta = \tan^{-1} \frac{N - 1 - 2y}{2x - N + 1} \quad \text{and } 0 \leq \rho \leq 1. \quad \text{For each action}$$

model and one given value  $\rho$ , we obtain the Zernike moments value. We used the absolute Zernike moment, which is given by

$$Z_{nm} = \frac{n+1}{\lambda_N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) R_{nm}(\rho_{xy}) \exp(-jm\theta_{xy}) \quad (8)$$

as the global motion features.

#### C. Direction of Actions (Overall Width or Height)

The 2D orientation (direction of major axis, or minor axis) of the motion distribution for every action is different. Thus the relative differences in magnitude of the eigenvalues are an indication of the elongation of the image (action model). The global motion orientation is obtained from the eigenvalue  $\lambda_i$ , of the covariance matrix of the action models. The eigenvalue is given by

$$\lambda_i = \frac{\mu'_{20} + \mu'_{02}}{2} \pm \frac{\sqrt{4\mu'^2_{11} + (\mu'_{20} - \mu'_{02})^2}}{2} \quad (9)$$

where,  $\mu'_{pq} = \mu_{pq} / m_{00}$  and the covariance matrix is  $COV[S(x, y)] = \begin{pmatrix} \mu'_{20} & \mu'_{11} \\ \mu'_{11} & \mu'_{02} \end{pmatrix}$ . We have considered the projection of major and minor axis orientation and the direction of the major axis as global features for action models.

## IV. EXPERIMENTAL RESULTS & DISCUSSION

#### A. Databases

The FBGDB [4] contains 14 representative full body actions in the daily life of 20 performers. In the database, all the performers are elderly persons (both male and female) with ages ranging from 60 to 80. The database contains 3D motion data and 2D data. The 2D data consists of both video data and 2D silhouette data. As an example, the sample images for left view are shown in fig. 4.

The KTHDB is one of the largest databases with sequences of human actions taken over different scenarios [10]. The database contains six types of human actions, performed several times by 25 subjects in four scenarios: outdoors (s1), outdoors with scale variation (s2), outdoor with different cloths (s3), and indoor (s4). The database contains 2391 sequences. The sample images are shown in



Fig. 5. The image sequences have the spatial resolution of 160×120 pixels and have a length of four seconds in average.

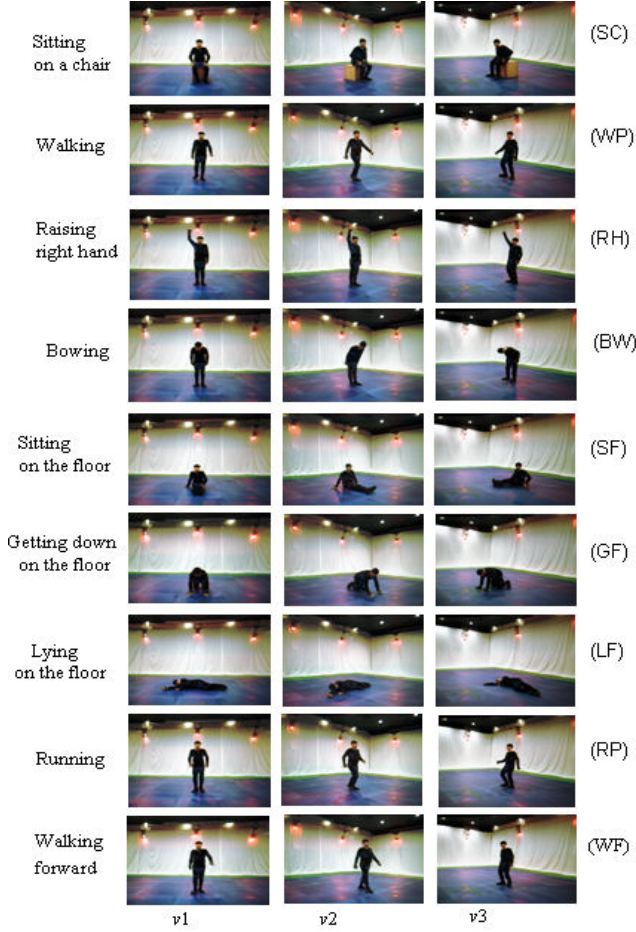


Fig. 4. Korea University Gesture database on all specified actions in three different views

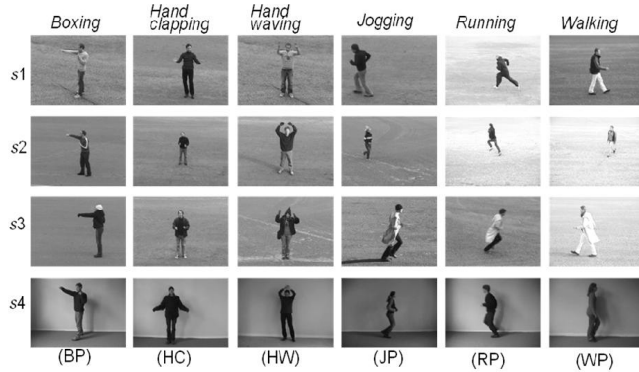


Fig. 5. KTH database on all specified actions in different scenarios

## B. Classification Results

Support vector machines (SVM) have high generalization capabilities in many tasks, especially in terms of object recognition. In applying MCSVM [5, 16], the training data of the global shape motions of the image sequences are divided into defined classes manually. The MCSVM predicts the class label information for arbitrary action. Before applying MCSVM, we normalized each component. Table I shows the action recognition results of FBGDB using MCSVM where we use components and global motions for each view. We use 9 subjects, 9 actions, and 4 views variation for testing (vA represents arbitrary view). The correct recognition rate (CRR) for any action is calculated as follows:

$$\text{CRR} (\%) = (N_c / N_a) \times 100 \quad (9)$$

Where,  $N_c$  is the total number of correct recognition sequences while  $N_a$  is the number of total action sequences. As can be seen, there is a clear separation among different kinds of actions. The overall CRRs of v1, v2, v3, and vA are 79.34, 84.12, 89.47 and 81.53 respectively, of FBGDB. We use combined models (AATs = ATs + VTs) to evaluate the performance.

TABEL I  
CRRS OF EACH ACTION AND VIEW OF FBGDB

|    | SC  | WP  | RH   | BW   | SF  | GF   | LF   | RF  | WF  |
|----|-----|-----|------|------|-----|------|------|-----|-----|
| v1 | .98 | .57 | 0.86 | 1.00 | .71 | 0.57 | 1.00 | .57 | .86 |
| v2 | .85 | .44 | 1.00 | 0.86 | .86 | 0.86 | 1.00 | .86 | .86 |
| v3 | .97 | .86 | 1.00 | 1.00 | .86 | 1.00 | 1.00 | .57 | .72 |
| vA | .94 | .57 | 0.94 | 0.99 | .76 | 0.76 | 0.90 | .76 | .76 |

We also have tested our approach by using the KTHDB, since it is one of the largest human action databases and several researchers used this database. We have tested 8 subjects, 6 actions, and 4 scenarios and each scenario contains 2 or 3 action sequences. Table II shows the recognition of each action in various scenarios for global shape motions. The CRRs of s1, s2, s3, s4, sA are 89.33, 83.17, 87.50, 88.3, 87.50 respectively for KTHDB, where sA is arbitrary scenario.

TABEL II  
RECOGNITION OF ACTIONS OF KTHDB

| Scenario | BP   | HC   | HW   | JP   | RP   | WP    |
|----------|------|------|------|------|------|-------|
| s1       | 0.98 | 0.97 | 0.97 | 0.91 | 0.74 | 0.82. |
| s2       | 0.98 | 0.88 | 0.94 | 0.83 | 0.61 | 0.78  |
| s3       | 1.00 | 0.96 | 0.97 | 0.74 | 0.81 | 0.81  |
| s4       | 0.95 | 0.94 | 0.93 | 0.81 | 0.67 | 0.80  |

We use 21 image sequences for classification of each action and each view in case of FBGDB. For arbitrary view recognition, we use more than 60 sequences for each action. It is important to mention that in some cases, the motion of

the elderly persons is similar. In our method, it is shown that by using the 2D action model with variability selection, the action recognition is more robust, since we use the natural actions of humans, with emphasis on elderly persons (FBGDB). The movement of elderly person's is significantly different than that of young people. For example, the speed and style of walking and running of elderly people are very similar.

We test the system performance without generating adaptable models, and we make a comparison of performance among AT, VT, and AAT. As an example, the performance (in CRR) of AT, VT, and AATs are 80%, 84.33%, and 89.33% respectively. The performance of AAT is significantly better than AT.

We compare our works with some state-of-art action recognition approaches by using the same database and similar test sequences but different methods. For example, we compare our method with [3, 7, 8, 12, 14] using KTHDB. Our results by global shape motions flow are compared with their results by spatio-temporal filters, volumetric features, spatio-temporal words, and local space time features. The overall comparison of different methods is listed in Table III. Compared to the mentioned researches, our approach yields better recognition results.

TABLE III  
COMPARISON RESULTS OF ACTION RECOGNITION

| Method              | CRR   | Scenario      |
|---------------------|-------|---------------|
| Dollar et al. [3]   | 81.17 | $sI+s2+s3+s4$ |
| Jiang et al. [7]    | 84.43 | $sI+s2+s3+s4$ |
| Ke et al. [8]       | 62.96 | $sI+s2+s3+s4$ |
| Niebles et al. [12] | 81.50 | $sI+s2+s3+s4$ |
| Schuldt et al. [14] | 71.72 | $sI+s2+s3+s4$ |
| Our method          | 87.50 | $sI+s2+s3+s4$ |

## V. CONCLUSION

We proposed a novel method for recognizing human action using the silhouette energy image and silhouette history image with the variation of anthropometric models. The SEI represented the energy content of the action model and SHI represented the energy history of the action model. The anthropometry variation provided a more natural and robust environment for human action recognition, using an advanced human-machine interface. Moreover, by adapting the shape variation, incomplete actions and partial occluded

actions were recognized successfully. We recognized human actions in individual and arbitrary views. With global motion features, the action recognition became sparse and flexible and it can be adapted to practical applications of human movement, human action recognition, and so on. Our future work will include the precise detection and recognition of action in complicated situations.

## REFERENCES

- [1] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," *IEEE Trans. on PAMI*, vol. 23 no. 3, pp. 257-267, 2001.
- [2] S. Carlsson and J. Sullivan, "Action recognition by shape matching to key frames", *IEEE Workshop on Models vs. Exemplars in CV*, pp. 263-270, 2002.
- [3] P. Dollar, G. C. V. Rabaud, and S. Belongie, "Behavior recognition via sparse spatio-temporal filters", *IEEE Workshop VS-PETS*, pp. 65-72, 2005.
- [4] FBGDB, <http://gesturedb.korea.ac.kr/>.
- [5] C. W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. on NN*, vol. 13, no. 2, pp. 415-425, 2002.
- [6] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. on Information Theory*, no. 8, pp. 179-187, 1962.
- [7] H. Jiang, M. S. Drew, and Z. N. Li, "Successive convex matching for action detection," *In Proc. CVPR*, vol. 2, pp. 1646-1653, 2006.
- [8] Y. Ke, R. Sukthankar, and M. Hebert, "Efficient visual event detection using volumetric features," *In Proc. ICCV*, pp. 166-173, 2005.
- [9] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. on PAMI*, vol. 12, no. 5, pp. 489-497, 1990.
- [10] KTHDB, <http://www.nada.kth.se/cvap/actions/>.
- [11] O. Masoud and N. Papanikolopoulos, "Recognizing human activities," *In Proc. AVSBS*, pp. 157-162, 2003.
- [12] J. C. Niebles, H. Wang, and L. Fei-Fei, "Unsupervised learning of human action categories using spatial-temporal words," *In Proc. BMVC*, vol. 3, pp. 1249-1258, 2006.
- [13] V. Parameswaran and R. Chellappa, "View invariants for human action recognition," *In Proc. CVPR*, vol. 2, pp. 613-619, 2003.
- [14] C. Schuldt, I. Laptev, and B. Caputo, "Recognizing human actions: a local svm approach," *In Proc. ICPR*, vol. 3, pp. 32-36, 2004.
- [15] Y. Sheikh, M. Shah, and M. Shah, "Exploring the space of a human action," *In Proc. ICCV*, pp. 144-149, 2005.
- [16] J. Westons and C. Wtkins, "Support vector machines for multiclass pattern recognition," *In Proc. European Symposium on ANN*, pp. 219-224, 1999.

# Generation of Mountain Ranges by Modifying a Controlled Terrain Generation Approach

K. Raiyan Kamal # and Dr. M. Kaykobad \*

Department of Computer Science and Engineering,  
Bangladesh University of Engineering and Technology  
#raiyan.kamal@gmail.com \*kaykobad@cse.buet.ac.bd

**Abstract**—The terrain generation approach proposed in ‘Parametrically Controlled Terrain Generation, (Proceedings of GRAPHITE 2007)’ focuses on generation of terrains having mountains of prespecified heights and spreads of base regions and peaks nearly at prespecified locations. However, inability to produce mountain ranges is a shortcoming of the algorithm. In this paper, we present a modification to the original algorithm that will enable generation of whole mountain ranges while preserving the similar parametric control over the resulting terrain. Experimental results of our algorithm have also been compared with results of existing algorithms.

**Index Terms**—Physically based modeling, procedural modeling, terrain generation.

## I. INTRODUCTION

Generation and rendering of artificial terrains has been extensively studied in computer graphics. The need for studying effective and efficient terrain generation and rendering techniques rises from the fact that artificially generated terrain models are useful across a broad range of application including geography, geology, climatology and other related fields. In recent years, the field of A.I. has also found artificially generated terrains useful in experiments due to growing interest in research for automatic trajectory generation for mobile robots [1]. Terrain models are also useful in military tactical decision making [2]. Besides scientific research and planning, artificial terrains are important part of many art and entertainment forms like computer or video games and animation movies.

Earlier terrain generation algorithms, consisting mostly of fractal techniques, are indeed able to generate sufficiently natural looking terrains but lack the control over the resulting terrains due to their unpredictable nature. Recent works have concentrated on controlled terrain generation providing significant control over the resulting terrain while preserving the randomness. Like any other procedural modeling, the randomness is a vital part in artificial terrain generation due to its responsibility to produce seemingly realistic yet unpredictable artifacts for the unspecified parts of the model.

The terrain generation approach presented in [3] offers control through a set of parameters, hence the name Parametrically Controlled Terrain Generation. A significant

drawback of the algorithm is its inability to produce mountain ranges or an entire landscape having several mountains. A possible solution to this problem would be to produce different terrain models with mountains at different location and features and then combine them to produce an entire landscape or mountain range. But the overhead of storing the intermediate results and combining them later on is significant. Thus the goal here for a better terrain generation approach is to provide means of producing mountain ranges or landscape with multiple mountains with low time and memory overhead while preserving the parametric controls provided by the original algorithm.

This paper presents Polyline Walk, a modified version of the Probing Algorithm. This extended version of the algorithm focuses on generation of mountain ranges according to prior specifications. Alongside the original set of features describing a single mountain peak, we introduce the following additional features which are vital for describing a mountain range.

- 1) Shape of the mountain range, e.g. linear shaped or zig-zag
- 2) Key mountain peaks, i.e. the significant mountain peaks characterizing the mountain range
- 3) Point of discontinuity, if any on the mountain range

The paper is organized as follows. Section 2 reviews in short the literature on terrain generation. Our proposed technique is presented in section 3. Section 4 depicts the observations and experimental results followed by future research directions and conclusion in Section 5.

## II. BACKGROUND AND PREVIOUS WORKS

### A. Data Structure

A widely used data structure for terrain representation is the height field. A height field is a mapping of two dimensional coordinates along X and Z axis with an elevation value along Y axis.

$$y = field(x, z) \begin{cases} min_x \leq x < max_x \\ min_z \leq z < max_z \end{cases}$$

This formulation leads to the implementation of height fields as a two dimensional rectangular grid of elevation

values. Triangular and hexagonal grid versions are also possible. Despite its drawbacks, height field is the most favorable data structure owing to ease of manipulation and other advantages including highly optimized rendering [4] and collision detection.

### B. Previous Works

Based on the degree of control offered, the known terrain generation algorithms can be classified into two categories. The ones providing explicit control over the result and the ones that do not. Most of the earlier works can be put into the second category while the recent works are of the first category.

1) *Early Terrain Generation Algorithms:* The Mid-Point Displacement Algorithm, a recursive subdivision algorithm introduced by Fournier et al [5], offers control over the resulting terrain through a roughness parameter  $H \in [0, 1]$ . Although roughness of the resulting terrain is controllable, the location and structure of resulting mountains and valleys are entirely unpredictable. Fault algorithm by Krten [6], basically mimicking the effect of earthquake, offers a similar type of control through specifying the amount by which one portion of the terrain to elevate at each iteration and thereby controlling the steepness of the resulting terrain. The erosion techniques introduced by Musgrave et al [7] simulate effects of thermal erosion, water-flow, air-flow, rain-fall or other geophysical phenomena on a supplied base terrain. Control over the result is achieved by tuning a set of parameters defining the nature and intensity of erosion. However, the influence of this action over the resulting terrain is indirect and the overall shape and structure of the result depend solely on the base terrain.

Results of these algorithms can be seen in figure 1.

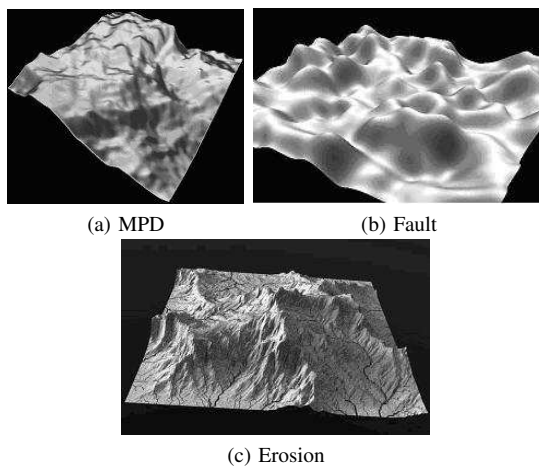


Fig. 1: results of early terrain generation algorithms

2) *Controlled Terrain Generation Algorithms:* Parry proposes a controlled terrain generation technique [8] offering control over the number of hills/mountains, their

average height with a certain standard deviation and slopes of the mountains. Based on the Modified Bivariate normal(MBVN) model [9], this Extended MBVN technique models each mountain as a bivariate normal distribution surface in the three dimensional space, possibly made unsymmetrical by tuning its parameters. Drawbacks of this technique are lack of reality and non-exact modeling.

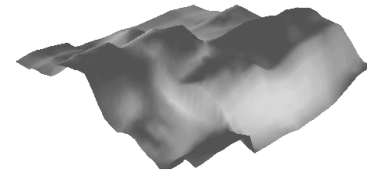


Fig. 2: Terrain generated by Genetic algorithm

Ong et al. proposed a method for controlled terrain generation [10] that uses genetic algorithm approach in two passes. The first pass refines an user provided rough sketch of terrain region boundaries using a genetic algorithm. Then this result is coupled with a database of given terrain data to generate an artificial terrain, which is again optimized using a second genetic algorithm. A limitation of the system is that in order to have realistic results the terrain database has to be populated with realistic terrain data, possibly GIS(Geographical Information System) information of real life locations, which are not always readily available. Terrain generated using this approach is seen in figure 2.

### C. The Original Probing Algorithm

The probing algorithm [3] is capable of generating a single mountain peak of given height at a given location of the terrain with spread of the base of the mountain also being controlled by algorithmic parameters. With advanced tuning of parameters, the algorithm can produce craters and other unusual artifacts as well. Result of this algorithm can be seen in figure 3.

The algorithmic parameters are

- $n$  : number of iterations to perform, integer
- $l$  : number of cutting lines drawn in each iteration, integer
- $r$  : probing depth, integer

The basic idea is to draw several random straight lines across the base terrain cutting the height-field into a number of polygons, hence the name *cutting line*. A set of polygons among the resulting ones is chosen by performing a probing through neighboring polygons starting from the polygon containing the prespecified location of the mountain peak and these polygonal regions are elevated by a constant amount. Two polygons are considered to be neighbors during the probing if they share a common side. This process is repeated several times.

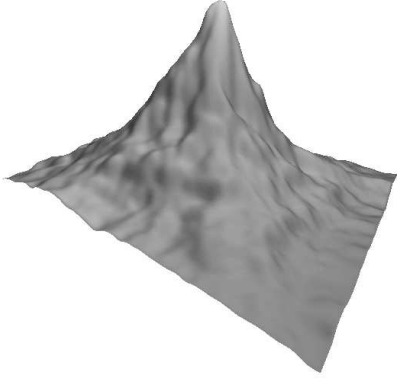


Fig. 3: Terrain generated by Probing algorithm

In advanced tuning of the algorithmic parameters, the elevation to be applied to the  $i$ th polygon found in course of probing is looked up from a deterministic function or from a vector initialized with precalculated values. This vector is referred to as *delta*,  $delta[i]$  holds a real value by which the  $i$ th polygon is to be elevated. Mountains of certain amount of steepness or flatness and unusual artifacts such as cavity and volcanic craters are also possible to produce using this technique.

### III. THE PROPOSED TECHNIQUE

#### A. Basic Notations

For convenience of discussion, let  $P$  denote a poly line, an ordered sequence of vertices. Let  $v_i = (x_i, z_i)$  be a vertex of the poly line where  $x_i$  and  $z_i$  are the coordinates of the vertex on the height field along  $X$  and  $Z$  axis respectively. Let  $a_i$  and  $r_i$  denote altitude values of the  $i$ th peak in the mountain range and span of its base respectively. A line segment connecting the vertices  $v_i$  and  $v_j$  is denoted by  $\langle v_i, v_j \rangle$ .

#### B. Modifications to the Original Algorithm

The following modifications and new features are introduced to the original algorithm in order to facilitate the generation of mountain ranges.

- Shape of the target mountain range is provided as an input parameter in the form of a polyline with  $N \geq 3$  vertices. Each vertex represent location of a key mountain peak.
- Instead of drawing  $l$  random lines,  $n - 1$  pairs of random lines are drawn as cutting lines. Thus the parameter  $l$  is made obsolete as it is not needed anymore. The details of drawing these lines are discussed later.
- Instead of randomly selecting the polygons to be elevated, only the polygons containing one of the

vertices of the polyline is elevated. Thus the parameter  $r$  is effectively replaced by  $N$ .

A more detailed discussion of the algorithm as a whole follows.

#### C. Generating a Mountain Range

Our algorithm generates a mountain range whose shape is specified as a poly-line having  $N \geq 3$  vertices. Each of the vertices represent location of a key mountain peak. The algorithm is of iterative nature. The number of iterations to be performed, denoted by  $n$ , is provided as an algorithmic parameter and the following operations are performed on each iteration. For each of the  $N - 1$  line segments that comprises the poly line, a random point  $p_i$  is chosen, which we choose to call pivot point, lying on the  $i$ th segment of the poly line. Through each of these pivot points, a pair of random straight lines is drawn in a fashion similar to drawing fault lines across the terrain in the fault algorithm [6]. These lines cut the terrain into several polygons from which, the set of polygons containing the key mountain peaks are chosen to be elevated by a constant amount, say 1. Figure 4 presents an illustration of this process. The resulting terrain after  $n$  iteration contains a mountain range with shape matching the given poly line and the key mountain peaks having the same altitude which is proportional to  $n$ . Pseudocode reflecting the procedure follows. Figure 5 depicts steps of generating a mountain range using this procedure.

#### **procedure** *MtRangeBasic*( $N, P, n$ )

```

1: for j=1 to n do
2: for i=2 to N do
3: $p_i \leftarrow \text{randomPointOn}(\langle v_{i-1}, v_i \rangle)$
4: $\text{draw2lines}(p_i)$
5: end for
6: for i=1 to N do
7: $\text{elevate}(v_i, 1)$
8: end for
9: end for
```

#### **procedure** *randomPointOn*( $\langle v_i, v_j \rangle$ )

```

1: returns a random point on the line segment $\langle v_i, v_j \rangle$
```

#### **procedure** *elevate*( $v, k$ )

```

1: elevates the polygon containing the vertex v by amount k using a flood-fill technique
```

#### **procedure** *draw2lines*( $v$ )

```

1: draws two random lines across the height-field through the vertex v
```

For drawing the cutting lines and erasing them at the end of the iteration, a procedure same as that of the original

algorithm is followed.

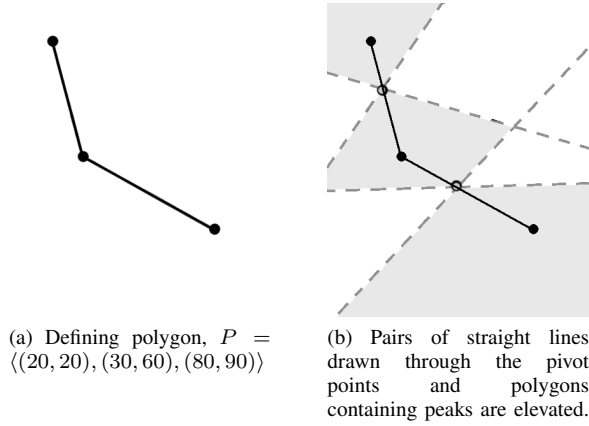


Fig. 4: Illustration of the process followed in one iteration of the algorithm. The filled circles, outlined circles, dashed lines and gray polygons denote locations of key mountain peaks, pivot points, cutting lines and elevated polygons respectively.

Two important parameters defining a mountain range are the altitude values of its key mountain peaks and span of the individual mountain bases. These parameters are considered next. Altitude of each of the key mountain peaks can be controlled by specifying their relative altitude  $a_i$  with respect to the highest peak in the range. Modifying the basic mountain range generating algorithm to elevate the polygon containing  $v_i$  by a value of  $a_i$  enables control over the altitude of each key mountain peak separately. A negative value for  $a_i$  produces a crater instead of mountain peak at the same location and negative value for all key mountain peaks produces a canyon instead of a mountain range. A farther modification in the algorithm to choose two pivot points within  $r_i$  distance from  $v_i$ , one from the line segments  $\langle v_{i-1}, v_i \rangle$  and  $\langle v_i, v_{i+1} \rangle$  for each vertex  $v_i$ ,  $2 \leq i \leq n-1$  of  $P$  instead of choosing one pivot point per line segment enables control over the span of the base of the mountains as well. The final algorithm enriched with these modifications is reflected in the following pseudocode.

```

procedure MtRagne(N, P, n)
1: for $j=1$ to n do
2: for $i=2$ to N do
3: $p_{ia} \leftarrow \text{RandomPointWithin}(\langle v_{i-1}, v_i \rangle, r_i)$
4: $p_{ib} \leftarrow \text{RandomPointWithin}(\langle v_i, v_{i+1} \rangle, r_i)$
5: $\text{draw2lines}(p_{ia})$
6: $\text{draw2lines}(p_{ib})$
7: end for
8: for $i=1$ to N do
9: $\text{elevate}(v_i, a_i)$
10: end for
11: end for

```

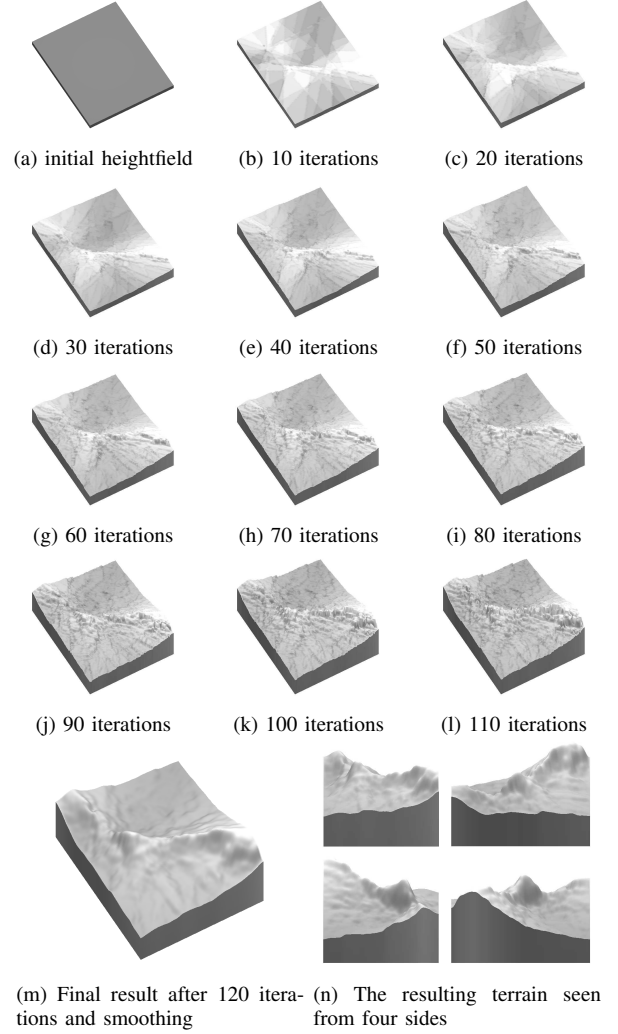


Fig. 5: Steps of generating a mountain range.

**procedure**  $\text{randomPointWithin}(\langle v_i, v_j \rangle)$

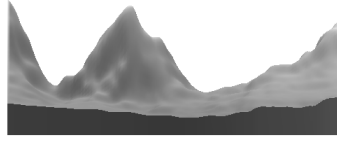
1: returns a random point on the line segment  $\langle v_i, v_j \rangle$  and within  $r$  distance from  $v_i$

Figure 6 demonstrates the effect of specifying maximum allowed base span,  $r_i$  on the resulting terrain. For a particular mountain peak, small value of  $r_i$  would cause that particular mountain to have a small base span making it somewhat detached from rest of the mountain range, creating a point of discontinuity on the mountain range. For larger values of  $r_i$ , the mountain base grows up to the point where it eventually merges with the rest of the range.

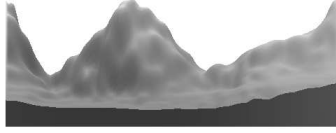
#### IV. EXPERIMENTAL RESULTS AND OBSERVATIONS

##### A. Generation of Simple Mountain Ranges

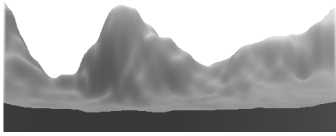
Figure 7 shows terrains containing simple mountain ranges generated using the basic procedure. The defining



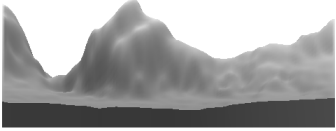
(a)  $r_2 = 30$



(b)  $r_2 = 40$



(c)  $r_2 = 50$



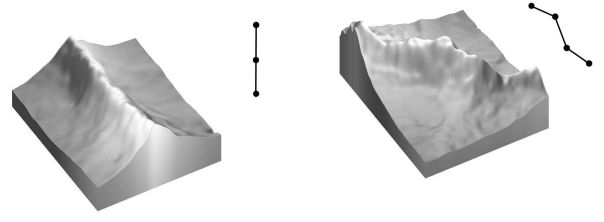
(d)  $r_2 = 60$

Fig. 6: Effect of maximum allowed base span,  $r_i$  on individual peaks. Base span of the mountain in the middle is being varied keeping all other parameters constant. For smaller values of  $r_i$ , smaller base span makes this particular mountain somewhat detached from rest of the range. With increasing allowed base span, the mountain base grows larger and eventually merges with the rest of the range.

polygons of the two mountain ranges are linear and zig-zag shaped respectively and so are the resulting mountain ranges.

### B. Generation of Complex Terrains

Further control of the resulting terrain is required to produce mountain ranges with complex attributes. Such a terrain consisting of a peak surrounded by two craters on both sides can be seen in figure 8. The defining polygon is a simple linear one and relative altitude values for the three peaks here are  $-0.8$ ,  $1.0$  and  $-0.7$  respectively. Note that the negative value for altitude has turned the first and third peak into craters effectively.



(a) Simple linear shaped mountain range.

(b) Zig-zag shaped mountain range.

Fig. 7: Simple terrains generated by the basic procedure.

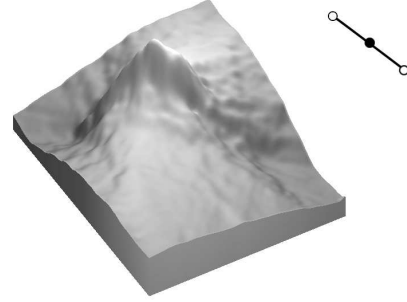


Fig. 8: Generating mountain range with unusual characteristics. White dots on the poly line indicates vertices representing spots of negative altitude, i.e. craters.

Figure 6 demonstrates the effect of specifying  $r_i$ , maximum allowed base span on individual peaks, as discussed earlier.

### C. Experimentation with MPD Algorithm

MPD(Mid-Point Displacement) algorithm has the property that if the initial height field is biased with some seed value at a certain location before running the algorithm, the seeded location retains that height in the resulting terrain as well. Apparently this property is favorable in generating mountain ranges by seeding the coordinates of vertices from the defining poly line with relative altitude values. In effect, the resulting terrain indeed contains desired altitude at the desired points but the remaining part of the terrain stays still beyond control obscuring the distinctive elongated high altitude region of the mountain range if any be created at all. Results of such experimentation can be seen in figure 9. In this experiment, attempts were taken to produce a simple linear shaped mountain range, but as can be seen from the experimental results, the generated terrain does not show any convincing mountain range with distinctive features. Moreover, the results differ widely from one run of the algorithm to another making this approach inappropriate for serving this purpose.



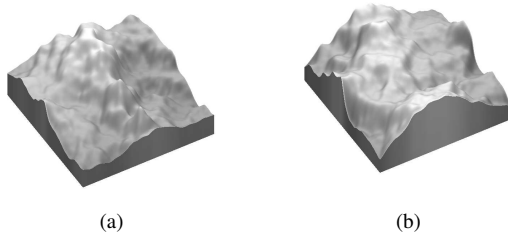


Fig. 9: Limitation of MPD algorithm at producing simple linear shaped mountain ranges by biasing.

#### D. Experimentation with Fault Algorithm

Crude height fields resembling mountain ranges of intended shape (linear in this case) and altitude attributes were provided as seed terrains to the Fault algorithm. As can be seen in figure 10, the resulting terrain indeed contains a linear shaped mountain range. But the lack of realistic features and uncontrolled altitude along the length of the range is evident. Results of another experiment in figure 11, aiming a mountain range with two distinct peaks of same height. Altitude of the same mountain peak differs greatly between two runs of the procedure.

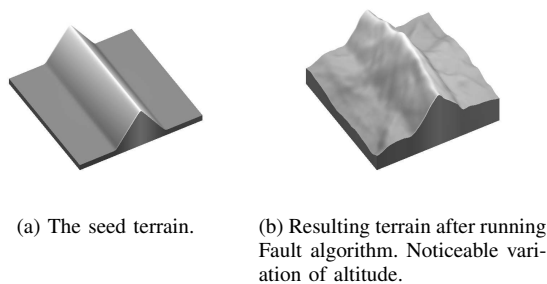
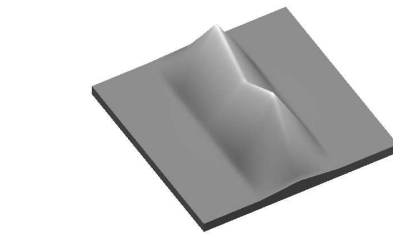


Fig. 10: Experimenting with Fault algorithm. Attempting to produce a simple linear shaped mountain range.

#### V. CONCLUSION AND FUTURE WORKS

The presented modification to the algorithm described in [3] effectively eliminates the drawback discussed earlier in this paper. Additional advantages of the proposed technique are low additional overhead and reduction of user inputs. Currently the algorithm can work with mountain ranges without branching. Future works would focus on adapting the algorithm to accept more complex mountain ranges with branching and loops, i.e. trees or planar graph structures with cycles as well.



(a) The seed terrain of a mountain range containing two distinct peaks.

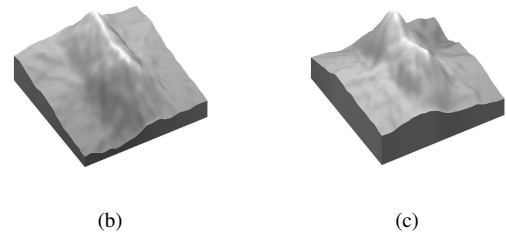


Fig. 11: Drawback of Fault algorithm on a seed terrain method. Mountain peaks on the resulting terrain does not always have same altitude and shape.

#### REFERENCES

- [1] THOMAS M. HOWARD, ALONZO KELLY 2005. Terrain-adaptive generation of optimal continuous trajectories for mobile robots. *Proceedings of ISAIRAS 2005 Conference, Munich, Germany*.
- [2] MARK ST. JOHN, HARVEY S. SMALLMAN, TIMOTHY E. BANK, MICHAEL B. COWEN 2001. Tactical routing using two-dimensional and three-dimensional views of terrain *Proceedings of the human factors and ergonomics society 45th annual meeting- 2001*.
- [3] KAMAL, K. RAIYAN AND SARWAR UDDIN, MD. YUSUF. 2007. Parametrically Controlled Terrain Generation *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia (GRAPHITE 2007)*, 2007, Perth, Western Australia, 1-4 December 2007.
- [4] DUCHAINEAU, MARK ET AL. 1997. ROAMing Terrain: Real-time Optimally Adapting Meshes. *proceedings of IEEE Visualization '97. IEEE Computer Society Press, 1997*.
- [5] FOURNIER, A., FUSSELL, D. AND CARPENTER, L. 1982. Computer Rendering of Stochastic Models. *Communication of the ACM*, 25, 6, 371-384.
- [6] KRTEEN, R. 1994. Generating Realistic Terrain. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 19,7, 26-28.
- [7] MUSGRAVE, F. K., KOLB, C. E. AND MACE, R. S. 1989. The Synthesis and Rendering of Eroded Fractal Terrains. *Computer Graphics*, 23, 3, 41-50.
- [8] PARRY, SAM H. 1977. The generation and use of parameterized terrain in land combat simulation *Proceedings of the 9th conference on Winter simulation - Volume 1*.
- [9] NEEDELS, CHRISTOPHER J. 1976. Parameterization of Terrain in Army Combat Analysis. *Naval Postgraduate School, 1976*.
- [10] TEONG JOO ONG, RYAN SAUNDERS, JOHN KEYSER AND JOHN J. LEGGET. 2005. Terrain generation using genetic algorithms. *Proceedings of Genetic and Evolutionary Computation Conference, GECCO 2005, Washington DC, USA, June 25-29, 2005*



# A Machine Vision Based Automatic System for Real Time Recognition and Sorting of Bangladeshi Bank Notes.

Raihan Ferdous Sajal, Mohammed Kamruzzaman, Faruq Ahmed Jewel.

Visual Magic Corporation Limited, House # 13/1, Road # 12, Baridhara, Dhaka-1212.

**Abstract** — This paper presents an efficient machine vision algorithm for real time image analysis and recognition of different features of Bangladeshi bank notes by using an automatic banknotes sorting system. The features recognized by this algorithm include denominations, orientations and sides of the bank notes. In a mechatronic system the Bangladeshi bank notes are fed together. The system draws the notes one by one and gets their images using a CCD sensor from a specific side. Then the system determines the denomination, orientation and side of the bank notes by analyzing the images grabbed by the CCD sensor. The average recognition speed is 8 to 9 bank notes per second and the rate of success is 100% for the banknotes having no extremely fatal damage.

**Index Terms** — Machine vision, Banknotes sorting, Image processing, Neural networks, Simplified Bi-directional associative memory.

## I. INTRODUCTION

Automatic banknotes recognition and sorting systems have become mandatory in the current fast going world. In the banks where a large transaction of money is an everyday reality, the officials find it tedious to count and sort the notes manually.

This paper presents a real-time machine vision based algorithm for recognizing and sorting the Bangladeshi banknotes. The algorithm is implemented and dumped in a banknotes sorting system, in which the notes are placed upon a feeding tray and the system draws them one by one and grabs the images of a specific side of the banknotes with a CCD sensor. The images are then analyzed with a microcontroller, LPC-H2106, where the implementation of this algorithm is housed. The image analysis and recognition is performed in real time and the notes are dispatched to two different terminal pockets according to their denominations, orientations and side.

## II. BACKGROUND

Number of researches have been conducted throughout the world for paper currency identification, sorting and recognizing the counterfeit notes. William J. Jones et al [1] developed an automatic currency processing system that uses the digital signal processing method for identifying the denomination of the notes. Danek and

Menelly [2] patented an idea for currency paper verification and denomination finding device that uses a linear array of photo-emitters and photo-detectors in one side of the banknotes and determines the authenticity of the banknotes from the reflected and transmitted light. Paul A. Jones et al [3] contributed another document processing system that analyzed the full image.

Joshi et al [4] developed a method for determination of the validity of the bills by visual and reflective spectral response. Rhodes and Gustafson [5] developed a system for checking the genuineness of the notes by verifying the watermarks in them.

## III. METHODOLOGY

### A) Image Acquisition:

When the banknotes are placed upon the system tray, an optical sensor issues an interrupt to the system notifying the banknote's arrival. The system then draws the note with its first roller. When the note moves inside the system, it slides on a long and slender CCD sensor array. The sensor array is one dimensional, and it can take the image of a single line of the note at any particular instance. But when the entire note slides on the sensor, the complete image of the note is grabbed.

### B) Determination of the denomination, orientation and side:

The banknotes images are obtained in grayscale. The monochrome images are then filtered by using a high-pass filter mask. These filtered images are then converted to binary images by using a histogram based thresholding technique. Since the threshold value is determined by analyzing the histogram of the image, the binary images remain almost the same irrespective to the new, old, blurred or faded notes.

These binary images are divided into equal sized rectangles and the average gray-levels of each rectangle is computed. These average values constitute the feature vectors of the banknotes. Since the notes cannot rotate inside the system, and the images are taken from a fixed distance, the size and direction of the notes in their images remain the same in all cases. It facilitates the design of the algorithm, because the effects of scaling and rotation of

the notes images are almost completely absent. The feature vectors obtained from the banknotes are first used to train an artificial neural network, and then it is used to classify the notes.



Figure 1: The grayscale image of a 50 taka note.

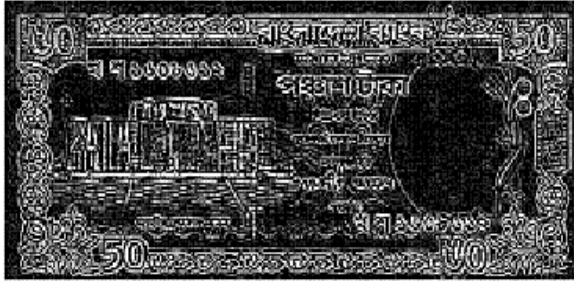


Figure 2: The filtered image.

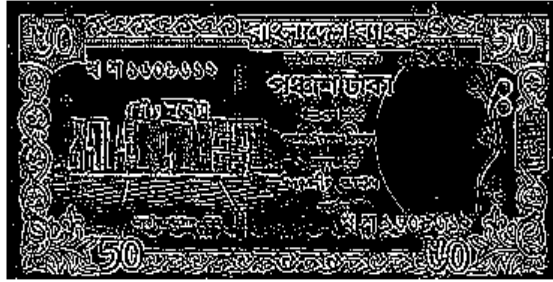


Figure 3: The binary image after thresholding.

The neural network used in this research for classifying the notes is known as Simplified Bi-Directional Associative Memory (sBAM) [7]. First proposed by S. Rajasekaran and V. Pai in 1998, this associative memory can manipulate real-number data while the conventional associative memory can manipulate the binary data only. But since it is an associative memory, it has a major drawback: it must associate any input pattern with an output pattern although the input pattern is absolutely absurd. That means, an associative memory cannot classify a pattern as indeterminate while other sorts of neural network like the multilayer feed-forward network has the capability to do it. For example, if an associative memory is trained with the patterns of Bangladeshi 500, and 100 taka images, and then if any other arbitrary image is given as input, the associative memory cannot classify it as an indeterminate pattern. Rather it classifies it as one of the aforementioned input patterns of 500 and 100

Bangladeshi taka. It implies that, if an associative memory is trained with the Bangladeshi banknotes images, and if any arbitrary piece of paper is placed on the system's input tray, the program will definitely recognize that vague piece of paper as one of the Bangladeshi banknotes.

In order to get around this problem, first the feature vector extracted from an input image is given to sBAM. The sBAM then classifies that as a Bangladeshi banknote with a particular denomination, orientation and side. Then a particular portion of the testing banknote is cropped and compared with the portion in the same location of the respective banknote image. If both the cropped portions are found substantially similar, the banknote is considered to be classified successfully. Otherwise, the algorithm classifies the input image as an indeterminate one and the therefore the microcontroller notifies the system about this problem.

C) The Simplified Bi-Directional Associative Memory (sBAM):

For both the training and recognition of the patterns, the Simplified Bi-Directional Associative Memory (sBAM) follows a twofold approach: The input normalization and evaluation.

Given a set of N training pairs  $(A_i, B_i)$ ,  $i=1,2,3,\dots,N$  where  $A_i = (a_{i_1}, a_{i_2}, \dots, a_{i_n})$  and  $B_i = (b_{i_1}, b_{i_2}, \dots, b_{i_p})$ ,  $a_{i_j}, b_{i_j} \in \mathbb{R}$ , the normalized input pattern  $(\hat{A}_i)$  is given by

$$(\hat{A}_i) = \frac{A_i}{\|A_i\|},$$

$$\text{where } \|A_i\| = \sqrt{\sum_{j=1}^n (a_{ij})^2}$$

Clearly, the vector  $(\hat{A}_i)$  is a unit vector.

Now, let,  $\alpha$  be the vector to be submitted to retrieve the associated pair  $\beta$ . Now,  $\alpha$  could represent a stored pattern or a noisy pattern. Irrespective of the case, the general system of equation is given as:

Frame the co-relation vector

$$M_{n \times 1} = [m_j], [m_j] = \phi(\langle \alpha, \hat{A}_j \rangle), j = 1, 2, \dots, N.$$

where,  $\phi(x) = 1$  when  $x = \max(\langle \alpha, \hat{A}_j \rangle)$  and 0 otherwise.

Now,  $\beta$  can be retrieved as,

$$\beta_k = \sum_{i=1}^N (\hat{a}_{ik} m_k), k = 1, 2, \dots, p$$

Where  $\beta_k$  is the k-th component of the vector  $\beta$  and  $\hat{a}_{ik}$  is the k-th component of  $\hat{A}_i$ . It is to be observed that the computation of  $\beta_k$  is unnecessary since for some  $d$ , when

TABLE I  
THE RESULT OF DENOMINATION AND ORIENTATION RECOGNITION

| Denomination | Orientation | Number of notes | Successful recognition | Wrong recognition | Recognition as indeterminate | Average time per note (in milliseconds) |
|--------------|-------------|-----------------|------------------------|-------------------|------------------------------|-----------------------------------------|
| 500          | Forward     | 20              | 20                     | 0                 | 0                            | 115                                     |
| 500          | Reverse     | 20              | 20                     | 0                 | 0                            | 115                                     |
| 100          | Forward     | 50              | 48                     | 0                 | 2                            | 105                                     |
| 100          | Reverse     | 50              | 47                     | 0                 | 3                            | 105                                     |
| 50           | Forward     | 100             | 97                     | 0                 | 3                            | 97                                      |
| 50           | Reverse     | 100             | 97                     | 0                 | 3                            | 97                                      |
| 20           | Forward     | 100             | 95                     | 0                 | 5                            | 100                                     |
| 20           | Reverse     | 100             | 97                     | 0                 | 3                            | 101                                     |
| 10           | Forward     | 100             | 96                     | 0                 | 4                            | 80                                      |
| 10           | Forward     | 100             | 94                     | 0                 | 6                            | 84                                      |

$m_d = \phi(\langle \alpha, \hat{A}_d \rangle) = 1$ , the corresponding  $\hat{B}_d$  is the vector to be retrieved.

#### IV. EXPERIMENTAL RESULTS

The overall system has been tested using numbers of banknotes ranging from Bangladeshi 10 taka to 500 taka. While recognizing the denomination and the orientation or side of the banknotes, the system could determine the appropriate values for 96.08% of the given banknotes images. In the remaining 3.92% cases the system could declare the banknotes as the indeterminate ones because they were too much old, irregularly torn, or violently corrupted by different noises. But there was no such case that the system mistakenly determined a wrong denomination or orientation. The result is shown in the table 1.

#### V. CONCLUSION

The automatic banknotes recognition and sorting systems are used allover the developed countries of the world. But in Bangladesh it is not available yet. However, in order to make the monetary transactions more dynamic such systems capable to recognize and sort the Bangladeshi banknotes must be introduced in the banking sectors and different business centers properly. This research is probably the first attempt to develop machine vision based practical banknotes manipulation systems in Bangladesh. The primary results shown in this paper are certainly promising. Nevertheless the authors believe that there is still a long way to go to make the methods more accurate, fast and efficient.

#### ACKNOWLEDGEMENT

The authors wish to acknowledge the strong enthusiasm and support of Dr. Kenji Yoshida, the chairman of Visual Magic Corporation Limited, and professor of Hosei University, Japan.

#### REFERENCES

- [1] William J. Jones et al, "Automatic Currency Processing System", patent description in details available at the site: <http://www.patentstorm.us/patents/5905810.html>
- [2] Danek and Menelly, "Currency Paper Verification and Denomination Device," The patent description available at: <http://www.patentstorm.us/patents/5416307.html>
- [3] Paul A. Jones et al, "Document Processing System Using Full Image Scanning" The detail description is available at: <http://www.patentstorm.us/patents/7187795.html>
- [4] Joshi et al, "Fake Currency Detection using Visual and Spectral Response", patent description is available at: <http://www.freepatentsonline.com/20060115139.html>
- [5] Rhodes and Gustafson, "Digital Watermarking Employing both Frail and Robust Watermarks", available at: <http://www.patentstorm.us/patents/6574350-description.html>
- [6] Gonzalez and Woods, "Digital Image Processing, 2<sup>nd</sup> edition", Pearson Education, 2004.
- [7] S. Rajasekaran and V. Pai, "Neural Network, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications", Prentice-Hall of India Private Limited, New Delhi, 2003.
- [8] Theodoridis and Koutrombas, "Pattern Recognition".

# An Efficient Implementation Scheme for Multidimensional Online Analytical Processing

Nilufar Easmin, K. M. Azharul Hasan, Nilufa Parvin

Department of Computer Science and Engineering, Khulna University of Engineering and Technology, Khulna 9203, Bangladesh

**Abstract**— Multidimensional online analytical processing (MOLAP) processes data that is already stored in a multidimensional array in which all possible combinations of data are reflected. It employs multidimensional array as their basic data structure. The intent of this paper is to develop a MOLAP model based on the scheme called Extended Karnaugh Map Representation (EKMR) and evaluate it. The basic idea of EKMR scheme is to present multidimensional array by a set of two dimensional arrays. This reduces data access time from secondary storage if the tuple of subscripts are known. Effectiveness of this scheme over traditional multidimensional array is shown with sufficient analysis and experimental results.

**Index Terms**— MOLAP, Dice, Slice, Multidimensional Array, EKMR.

## I. INTRODUCTION

Decision support systems are rapidly becoming a key to gaining competitive advantage for businesses. Many corporations are building decision-support databases, called *data warehouses*, from operational databases. Users of data warehouses typically carry out on-line analytical processing (OLAP) for decision making. Multidimensional aggregation plays an important role in OLAP technology in providing decision support environment for the users. In order to present this multidimensional view, the data is usually stored in the form of “summary tables”. This approach of storing data cube is known as ROLAP (relational OLAP) [1][2]. The other approach which the data cube is stored as a multidimensional array is known as MOLAP [3].

ROLAPs are built on top of standard relational database systems. MOLAPs are based on multidimensional database systems. A MOLAP is a set of *multidimensional datasets* and is designed to allow for the efficient and convenient storage and retrieval of large volumes of data that is closely related, viewed and analyzed from different perspectives. These perspectives are called *dimensions*. It is realized in the database technology that there is no single best data structure for all applications within an enterprise [4]. In a simple MOLAP model, a multidimensional dataset consists of *dimensions* and *measures*, represented by  $R(D_1, D_2, \dots, D_n; M_1, M_2, \dots, M_k)$ , where  $D_i$ 's are dimensions and  $M_j$ 's are measures. The data structures in which ROLAPs and MOLAPs store datasets

are fundamentally different[5]. ROLAPs use relational tables as their basic data structure on the hand MOLAPs store their datasets as multidimensional arrays. MOLAPs only store the values of measures in a multidimensional array. The position of the measure values within the array can be calculated by the dimension values. The latest step in the evolution of databases is the MOLAP database and this technology has matured into the database engine for data analysis applications. The MOLAPs are designed for analyzing large groups of records whereas relational databases make it easy to work with individual records. The advantages of wide acceptance of the MOLAP includes ease of data presentation and navigation, ease of maintenance and high performance for storage and retrieval costs [4] [6].

From the data model point of view, the multidimensional data model is one of the paradigms which can assist the data analysis and data summarization from multiple attribute values of database [11]. Many multidimensional data structures have been proposed [8]-[11]. Most of the MOLAP implementation use traditional multidimensional array as their basic data structure[1]-[3]. [12][13] Used multidimensional extendible array [8] as their basic data structure for MOLAP and multidimensional database implementation. Here in this paper a new representation scheme for multidimensional array is used for the implementation. The representation scheme is named as Extended Karnaugh map Representation[14][5]. In EKMR, an  $n$  dimensional array is represented by a set of 2 dimensional arrays. The superiority of the scheme over the traditional array for matrix operations is shown in [14].

## II. THE NOTION OF EKMR

In the EKMR scheme, a multidimensional array is represented by a set of two-dimensional arrays. The idea of the EKMR scheme is based on the Karnaugh map. Consider the 3 input K-map and its corresponding EKMR(3) in Fig. 1. Let  $A[k][i][j]$  denote a three-dimensional array based on Traditional Multidimensional Array TMA(3) with a size of  $3 \times 4 \times 5$ . The corresponding EKMR(3) of array  $A[3][4][5]$  is shown in Fig.1. The analogy between the EKMR(3) and the 3-input Karnaugh map is that the index variables  $i$ ,  $j$ , and  $k$  are

corresponding to the variables X, Y, and Z, respectively (Fig.1(a) and Fig.1 (b) ). According to the 3-input Karnaugh map, a three dimensional array based on the TMA(3) can be presented by a two-dimensional array based on the EKMR(3). The EKMR(3) is represented by a two-dimensional array with the size of  $4 \times (3 \times 5)$ . Here index variable  $i$  is used to indicate the row direction and the index variable  $j$  is used to indicate the column direction. The index variable  $i$  is the same as the index variable  $i$  and the index variable  $j$  is a combination of the index variables  $j$  and  $k$ .

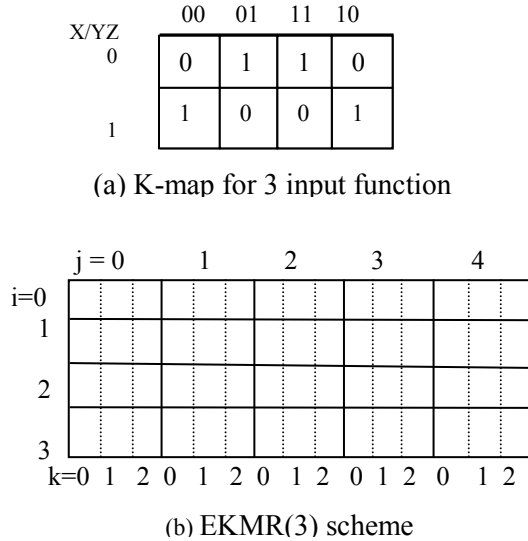


Fig. 1: 3-input K-map and its corresponding EKMR(3)

### III. EMPLOYING MULTIDIMENSIONAL ARRAY

The records stored in the relational table are the gathering of the column values to each dimension of the multidimensional array. The example representing in Fig. 2 is a representation of a 3 column relation in a 3 dimensional array. Making each column each dimension of 3 dimensional array each column value is being corresponded with the subscript of each dimension. Also, each cube of array shows the position inside the array of each record. The subscript of the dimension is assigned name to 1, location to 2, and age to 3. Due to this, each record can stand for as an element of 3 dimensional array. For example, the record whose *Name* column value is “Yesmin”, *Location* column value is “Kushtia” and *Age* column value is 22 can be expressed as the subscript values (2,2,1) of the three dimensional array; these subscript values(i.e., coordinate) stand for the corresponding position in the array. The column values

need to be transformed into the subscript of the array to express the relation with the multidimensional array.

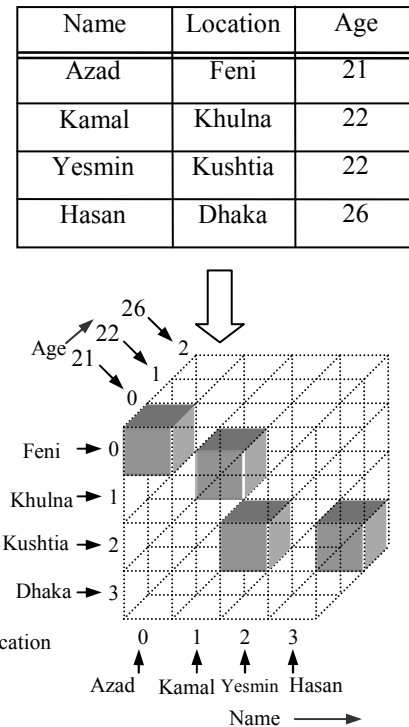


Fig. 2: Relational table implementations using multidimensional array

### IV. AN EFFICIENT MOLAP IMPLEMENTATION MODEL

In this Section the proposed MOLAP Implementation model and its operations are described. In the following, the EKMR based implementation is called as EI and the Traditional Multidimensional (TMA) based implementation is called as TI.

In Traditional Multidimensional Array (TMA) based implementation of MOLAP scheme each of the  $k$ th column of a  $n$  column relational table is mapped to the  $k$ th dimension of the multidimensional array. Each column value is mapped to a unique subscript and the measure value (i.e. sales value) of the relational table is inserted into the corresponding cell in the multidimensional array. Figure 3(b) represents the TI based MOLAP model for the relational table of the Figure 3(a). Hence for a  $n$  column relational table the corresponding multidimensional array is  $n-1$  dimensional array.

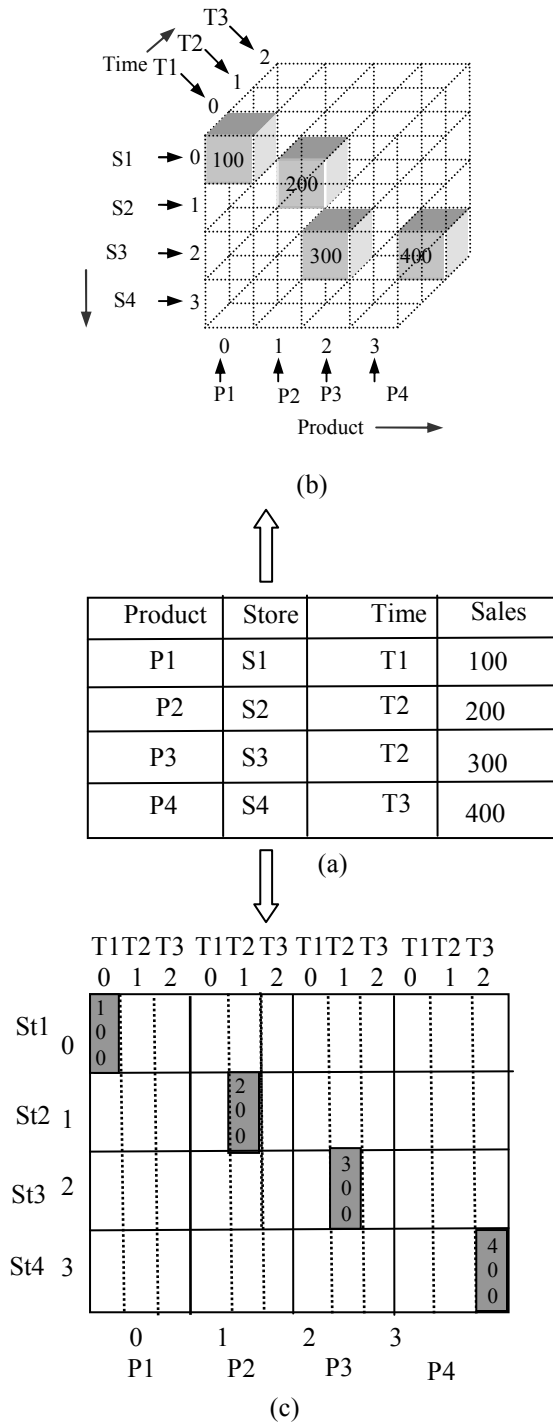


Fig. 3: MOLAP implementation scheme by TMA and EKMR.

Consider a simple multidimensional model in Figure 3, in which we have the column names product, store, time, and sales ( Figure 3(a)). There are three dimensions in Figure 3(b) product, store, and time respectively. The sales value

is stored in the corresponding cell of the array as fact data. For example the sales value 400 in TI (Figure 3(b)(c) ) indicates the fact that this is the sales value for the product P4 in store S4 at the time T3 which corresponds to the coordinate  $\langle 4,4,3 \rangle$ .

The similar approach of mapping is also considered for EI approach as shown in Figure 3(c) for MOLAP implementation. From Figure 3(b) and 3(c) we see that the representation scheme for MOLAP is fundamentally different in EI and TI. The 4 dimensional array is replaced by a 2 dimensional array. Hence the location of the sales value or array cell number will be different for TI and EI. And for the same reason, the same relational table implementation for MOLAP will be different.

## V. MOLAP OPERATIONS

The slice and dice operation for different situation is described in this section. We are taking a 5 column relational table as an example.

### A. Slice Operation

The slice operation for a 4 dimensional array implemented as MOLAP is the basis for answering the following questions

- Q1) Find the sales value of product P4
- Q2) Find the sales value of store S3
- Q3) Find the sales value at time T2
- Q2) Find the sales value for model M1

In slice operation the number of known dimension is only one and on the basis of the value of this dimension expected sales value is found. The physical organization of these situations is explained in Figure 4 using row major order of array organization. In Figure 4(a), the candidate array cells are shown for the case Q1. We can see all the candidate values are linear and can be accessed very easily if the array is stored in the secondary storage. All the candidate sales values are collected and added to produce the result for Q1. For this case the retrieval for slice operation will be advantageous for EI.

In Figure 4(b), we see that, the candidate array cells are organized in three different streams having 9 values are in the same stream. It will be necessary to access the streams to collect and produce the result for Q2. In Figure 4(c), (for Q3) the candidate values are organized in 9 different streams 3 values for each stream. And finally, In Figure 4(d), all the values are organized different stream. These values are collected and added to get the result of the slice operation.

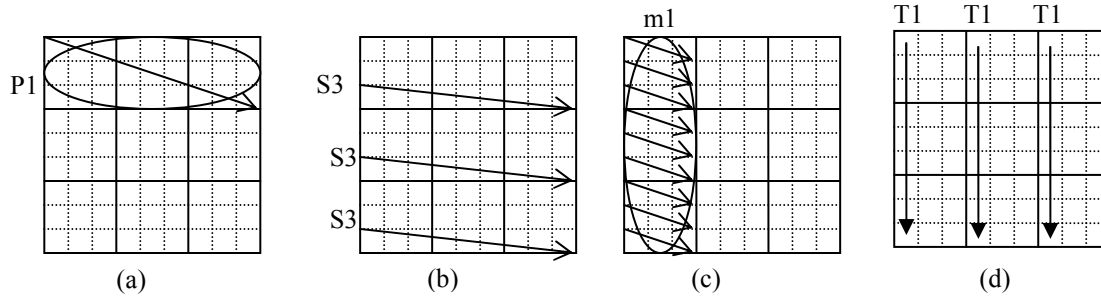


Fig. 4: Different situation for Slice operation for 4 dimension case ( $\rightarrow$  indicates those values which can be accessed linearly because of row major order).

### B. Dice Operation

The dice operation is similar to that of the slice operation. Dice operation is the basis for answering of the questions like

- Q1) Find the sales value of product P2 at store S2  
Q2) Find the sales value of product P2 for store S3 at time T2, and so on.

The number of known dimension in slice operation is only one on the other hand number of known dimension for dice operation is more than one. For example in the Q1 above the known dimension is product dimension and

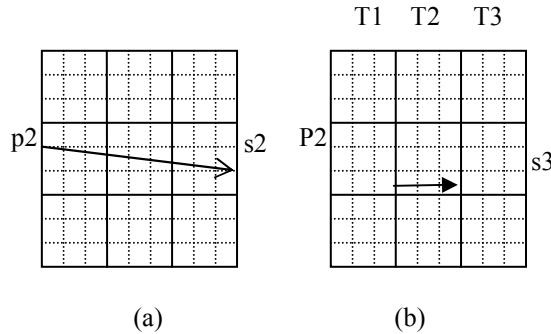


Fig. 5: Different situations for Dice operation( $\rightarrow$  indicates those values which can be accessed linearly)

store dimension. Figure 5(a) shows the situation for the case Q1 and Figure 5(b) shows the situation for case Q2. All the values are collected and added to get the sales value for dice operation.

## VI. EXPERIMENTAL RESULTS AND ANALYSIS

In this Section the experimental result for both EI and TI are described and compared. The parameters that are considered are as follows.

### A. Parameters

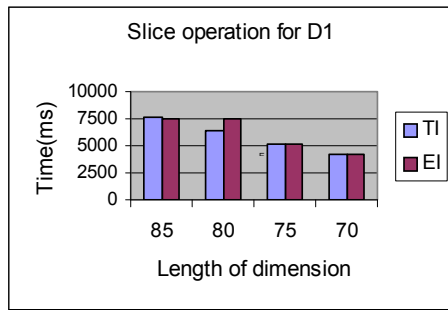
| Parameter | Description                                                                                                                                                                                           |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $N$       | Total number of dimensions                                                                                                                                                                            |
| NR        | Total no. of records in the relational table.                                                                                                                                                         |
| $D_i$     | Length of dimension of dimensions $i(i \leq N)$ . We assumed all the lengths of dimension $D_i$ is equal for all $i$ ( $1 < i < n$ ). Hence we write $D$ to indicate the lengths of dimension $D_i$ . |
| $\rho$    | Density of records in the Array $\rho = NR / IID_i$ . We assumed the records are uniformly distributed in EI and TI.                                                                                  |

### B. Comparison between EI and TI for Slice Operation

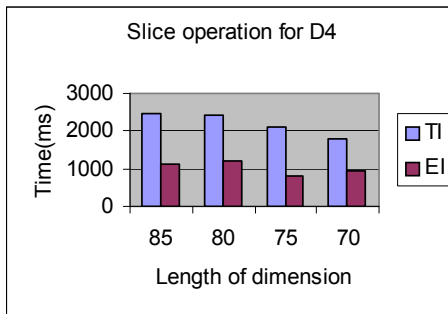
We have taken executing time for dice and slice operation for MOLAP at different length of dimension and at different density of records for each dimension for the case of EI and TI scheme. The retrieval time for slice and dice operation at  $\rho=0.3$  for  $N=65$  are shown in Table 1. In Table1  $D_1$ ,  $D_2$ ,  $D_3$  and  $D_4$  are dimension 1, 2, 3 and 4 respectively for which the slice operation is done. From table1 we see that for slice operation if dimension 4 is known then executing time is 90 ms this is the situation pointed out in Figure 4(a). This is because when

Table1: Processing time for slice and dice operation

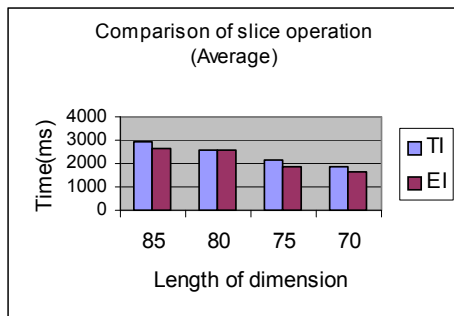
|                 |          |           |             |        |
|-----------------|----------|-----------|-------------|--------|
| Slice operation | D1       | D2        | D3          | D4     |
|                 | 6990 ms  | 199627 ms | 15703 ms    | 90 ms  |
| Dice operation  | D1,D4    | D1,D3     | D1,D2       | D2,D3  |
|                 | 3974 ms  | 210 ms    | 81 ms       | 250 ms |
|                 | D1,D2,D3 | D2,D3,D4  | D1,D2,D3,D4 |        |
|                 | 20 ms    | 20 ms     | 0 ms        |        |



(a)



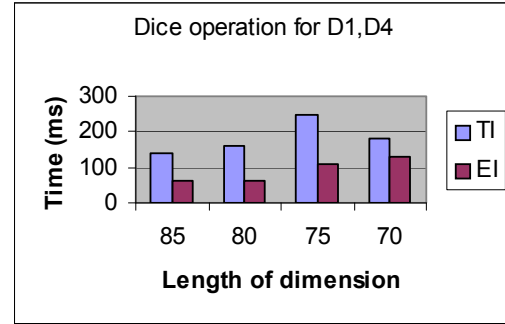
(b)



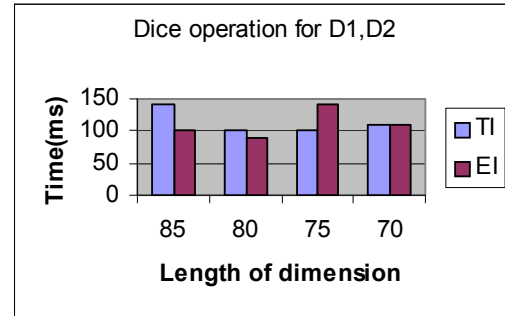
(c)

Fig. 6: Slice operation comparison for EI and TI for different values of known dimension

dimension 4 is known then more data are linear and can be collected easily. If dimension 1 is known then executing time is 6990ms this corresponds to the situation pointed out in Figure 4(b), if dimension 3 and dimension 2 corresponds to the situation described in Figure 4(c) and 4(d) respectively.



(a)



(b)

Fig. 7: Dice operation comparison for EI and TI for different values of known dimension

Comparison between EI and TI for slice operation at  $\rho = 0.25$  density for length of dimension 70, 75, 80 and 85 are shown Fig. 6 which shows the retrieval time for slice operation for different known value of dimension for varying length of dimension for EI and TI having  $\rho = 0.25$ . We see that, the time required to perform slice operation EI scheme is superior over TI for different known value of dimension. Some of the results have unexpected value as shown in the figure. This is because the uniform distribution that is used to generate the data automatically we put one dimension as reference dimension. So when the query that is adopted is not reference dimension it might cause less data because the data placement idea in EI is completely different than that of TI. One more important point is that the same array location for TI and EI does not refer to same record in a relational table. Hence some unexpected result for a query may occur. The superiority of EI scheme will be well understood by the



Figure 6(c) which shows the average execution time for TI and EI. We can see all the length of dimension values the EI scheme has advantages over TI. The average is calculated using the values resulted for the Q1 to Q4 (Section V).

### C. Comparison between EI and TI for Dice Operation

Fig. 7(a) and (b) show the dice operation comparison for different combination of known dimensions for EI and TI. EI has advantages in most of the cases of the experiment in EI over TI. Finally it can be concluded that the EI scheme proposed in this paper has advantages over TI. Although both are multidimensional arrays but due to the organization difference the EI scheme is superior over TI scheme for large volume MOLAP design and implementation.

## VI. CONCLUSION AND FUTURE WORKS

We propose and evaluate an efficient MOLAP implementation scheme using a new array representation scheme as opposed to the conventional multidimensional arrays. Our performance result shows that the proposed scheme is superior in most of the cases of slice and dice operations. We believe our scheme can also be applied to the multidimensional database implementations effectively specially for data warehousing applications for multidimensional analysis.

The future direction of this paper is to proceed for compressed data warehouse design and implementation using this scheme. The large number of array cells that are empty should be ignored and that's why compression is important. One more direction is to apply for parallel data warehouse design because many operations described in this paper can be done independently that will make efficient the structure to be applied in parallel environment.

### REFERENCES

- [1] S. Agarwal, R. Agrawal, P.M. Deshpande, A. Gupta, J. F. Naughton, and S. Sarawagi, "On the Computation of Multidimensional Aggregates", In Proceedings of the International Conference on Very Large Databases, pp. 506-521, 1996.
- [2] P. M. Deshpande, S. Agarwal, J. F. Naughton and R. Ramakrishnan, "Computation of Multidimensional Aggregates", Technical Report 1314. University of Wisconsin, Madison, 1996.
- [3] Y. Zhao, P. M. Deshpande and J. F. Naughton, "An Array-Based Algorithm for Simultaneous Multidimensional Aggregates", In Proceedings of the ACM SIGMOD Conference on Management of Data, pp.159-170, 1997.
- [4] Kenan Technologies, "An Introduction to multidimensional database technology", Available from <http://www.kenan.com/>, 1995.
- [5] Colliat, G., "OLAP, Relational and Multidimensional Databases Systems", SIGMOD Record, Vol.25, No.3, 1996.
- [6] J. Li and J. Srivasava, Efficient aggregation algorithms for compressed data warehouses. IEEE transaction of knowledge and data engineering, Vol.14, No.3, June 2002. pp. 515-529.
- [7] D. Taniar, and C. H. C. Leung, "The impact of load balancing to object-oriented query execution scheduling in parallel machine environment", Information Sciences, Vol. 157, pp. 33-71, 2003.
- [8] E. J. Otoo and T. H. Merrett, "A storage scheme for extendible arrays", Computing, Vol. 31, pp. 1-9, 1983.
- [9] J. Nievergelt, H. Hinterberger, and K. C. Sevcik, "The grid file: An adaptable, symmetric multikey file structure", ACM Transactions on Database Systems, 9(1), pp. 38-71, 1984.
- [10] K. Y. Whang and R. Krishnamurthy, "The multilevel grid file: A dynamic hierarchical multidimensional file structure", Proc. of DASFAA, pp. 449-459, 1991.
- [11] R. Bayer, "The universal B-tree for multidimensional indexing: General concepts", Proceedings of the International Conference on Worldwide Computing and Its Applications (WWCA '97), pp. 198-209, 1997.
- [12] K. M. Azharul Hasan, M. Kuroda, N. Azuma, T. Tsuji, K. Higuchi, "An Extendible Array Based Implementation of Relational Tables for Multidimensional Databases", Proc. of Data warehousing and Knowledge Discovery (DaWak'05), pp 233-242, 2005.
- [13] K.M. Azharul Hasan, Tatsuo Tsuji, and Ken Higuchi, An Efficient Implementation for MOLAP Basic Data Structure and Its Evaluation. DASFAA 2007, LNCS 4443, pp. 288 – 299, 2007.
- [14] Chun-Yuan Lin, Yeh Ching Chung, Jen-Shiuh Liu, "Efficient Representation Scheme for Multidimensional Array Operations.", IEEE transactions on computers, VOL 51, NO.3, pp. 327-345, March 2002.
- [15] Chun-Yuan Lin, Yeh-Ching Chung, Jen-Shiuh Liu, "Efficient Data Parallel Algorithms for Multidimensional Array Operations Based on the EKMR Scheme for Distributed Memory Multicomputers", IEEE Transactions on parallel and distributed systems, VOL. 14, NO. 7, pp. 625-639, JULY 2003.

# Decision Tree Based Routine Generation (DRG) Algorithm: A Data Mining Advancement to Generate Academic Routine for Open Credit System

<sup>1</sup>Md. Ashiqur Rahman, <sup>2</sup>Shafaat Sheik Giasuddin, <sup>3</sup>Rashedur Rahman

<sup>1,2,3</sup>Department of Computer Science and Engineering,

North South University, Banani, Dhaka – 1213, Bangladesh.

<sup>1,3</sup>{ashiq\_rahman, rashedurrahman}@yahoo.com, <sup>2</sup>rshafaat@gmail.com

**Abstract** — Mining data from a database using classification algorithm is not an innovative technique. This paper presents a technique namely Decision Tree based Routine Generation (DRG) algorithm to create an academic routine. The concept of open credit course registration system (after completion of pre-requisite courses any student may choose any course in any semester) makes the research more challenging and complex to accomplish. This is an NP-Hard Problem and hence unsolvable to satisfy all students and teachers at the same time. A level of tolerance has to be added to make the evaluation efficient and effective to achieve the goal, a conflict free routine. OLAP representation helps to classify the courses along with the proposed algorithm to eliminate some constraints. Day-based pattern, minimum Manhattan Distance between courses of same teacher, minimum conflicted course distribution has been stage-managed to classify the courses. To alleviate the algorithm, decision tree based models and sequential search methods are espoused with computational results.

**Index Terms** — OLAP, Crosstable, Conflict List, Favorite Slot, Faculty Choice, Course Color, Day-time slot pattern.

## I. INTRODUCTION

In this paper the proposed Decision Tree based Routine Generation (DRG) algorithm describes how to generate a university class routine within a tolerable range of some constraints. A decision-tree based classification algorithm has been introduced to solve this NP-Hard Problem. CPL (constraint logic programming) is a respected technology for solving hard problems which include many (non-linear) constraints [1]. Constraint propagation technique has been applied to overcome the preferential requirements for slots of teachers, courses from pre-advising by students and class room allocation. Versatile choices for courses may lead to a deadlock situation. Golz used priorities heuristic ordering [2] where Abdennadher introduced an optimized cost-based rule mining [3,4] to solve these type of problems. On the other hand, knowledge based in a hyper heuristic course scheduling using case based reasoning is used to maximize the rule covering area [5]. DRG presents key features of generating class routine with minimum computational time. Heterogeneous distribution of courses is classified with maximum satisfaction of all

constraints. Section II describes about previous related works and preliminaries in details. Section III illustrates the problem dimensions with the data filtering technique used in the paper to summarize for further manipulation; pursued by the proposed algorithm, DRG, and the classification procedure to find the feasible solution in Section IV. Section V explains the computational results and the performance analysis.

## II. PRELIMINARIES AND RELATED WORK

A university class routine generation problem – as considered in this work – consists of assigning each course in a set of slots (classes) in a limited class rooms within the teachers' favorite time slots. This highly constrained problem is optimized by simulated annealing and genetic algorithm [6,7]. Seven different major and minor objectives are discovered and deadlock situation is overcome by randomly exploring the composite neighborhood [8]. The most closely related attempt with this work appears to be the constraint programming approach used by Boizumault [9] and the simulated annealing approaches explored by Dowsland and Thompson [10,11]. The principal innovation in DRG is the sequential use of these two methods. DRG may select some poor slots, with respect to teachers or students, under a tolerable conflict range. A similar sequential approach has been taken in DRG on other problems: White and Zhang [12] used constraint programming to second a starting point for tabu search in solving course timetabling problems. For a high school timetabling Yoshikawa tested several combinations of two stage algorithms, including a greedy algorithm followed by simulated annealing and a constraint programming phase followed by a randomized greedy hill climbing algorithm (which is deemed to be the best combination of those used). In a similar vein, Burke, Newell and Weare [13] used their work on sequential construction heuristics [14] to generate initial solutions for their memetic algorithm [15].

### III. PROBLEM DESCRIPTION AND DATA FILTERING

The routine maps a set of courses chosen by students and teachers to a specific room and time-slot. A major objective, in developing an automated system, is to minimize the hassle of separating conflicted courses from choices by students. In this paper the major identified problems are (a) Number of lectures per week for each course are fixed, (b) Room overlapping is prohibited, (c) Fitting the routine with teacher's favorite timeslots, and (d) trying to assert different timeslots to same level of courses. On the other hand, the minor objectives are day-timeslots pattern for the course, room capacity, avoiding gaps between classes of same teacher, if possible, single class for student per day and ensuring compactness of interclass time difference.

The required scattered data contains total courses (course choices from pre-advising by students)  $C = \{c_1, c_2, c_3, \dots, c_n\}$  where the dependencies between courses are also maintained. Here course dependency can be defined as  $\exists C_i, C_i \rightarrow C_j$  where  $C_i, C_j \subset C$ . For this paper the students',  $S = \{S_1, S_2, S_3, \dots, S_z\}$ , course choices can be derived as  $S_j = \{c_i\}$  where  $\exists i, c_i \in C$  and  $1 \leq i \leq n$  and  $|S_j| = \max\_course\_choice$  for the student as shown in Fig. 1. Teachers' favorite timeslots are grouped according to day-timeslots pattern. Here group A and B is formed for the teacher  $t_k$ , where  $T = \{t_1, t_2, t_3, \dots, t_m\}$ ,  $A(t_k) = \{\text{favorite time\_slots of } t_k \mid \text{sequential time slots for Saturday, Monday and Wednesday}\}$  and  $B(t_k) = \{\text{favorite time slots of } t_k \mid \text{sequential time slots for Sunday, Tuesday and Thursday}\}$ , whereas  $\text{time\_slots} = \{1, 2, \dots, 30\}$  contains 5 sequential slots per day starting from Saturday. Here Friday is considered as an off-day. Priority of the teacher has been introduced by  $P(t_k) = \{1, 2, \dots, 10\}$  where a higher value represents higher priority. An exceptional priority is also introduced as 11 reflecting part-time faculty, whose projected time cannot be changed. Fig. 2. and Fig. 3. shows the teachers' wish-list and the course distribution among the teachers'. Target of this work is to find the values of the "class slot routine" field of the Fig. 3. The resultant routine vector,  $V = \{\{c_i\}_q\} \forall i, c_i \in C$  and  $1 \leq i \leq n$  and  $1 \leq q \leq 30$ , consists of the course classification as per day required for each course and class room availability.

In this paper, this huge dimensionality of dataset is reduced by initiating an OLAP (On-Line Analytical Processing) representation [16]. Here a *Crosstable* (Cr) of  $(n \times n)$  courses are initialized.  $Cr(n \times n) = \{\text{conflict}_{i,j}\}$  where  $1 \leq i,j \leq n$  and 'n' is the total number of courses requested by the student (or is 'n' number of courses offered by the department). Here the conflict of  $Cr_{i,j}$  is a positive integer that reflects the common students between

$c_i$  and  $c_j$ . The diagonal values of  $Cr_{i,i}$  show the total number of students requesting for the course  $c_i$ .

| Student ID     | Course ID      | Semester ID |
|----------------|----------------|-------------|
| S <sub>1</sub> | c <sub>1</sub> | 1           |
| S <sub>1</sub> | c <sub>3</sub> | 1           |
| S <sub>2</sub> | c <sub>1</sub> | 1           |
| S <sub>2</sub> | c <sub>3</sub> | 1           |
| S <sub>3</sub> | c <sub>2</sub> | 1           |
| S <sub>3</sub> | c <sub>4</sub> | 1           |
| S <sub>3</sub> | c <sub>5</sub> | 1           |
| S <sub>4</sub> | c <sub>4</sub> | 1           |

Fig. 1. Courses Pre-Advised by the Students.

$\sum Cr_{i,j}$ ,  $[1 \leq j \leq n]$  is the total conflict for the course  $c_i$   $[\forall i, i \neq j]$ . Maximum "chaos" (conflict) courses can be easily sorted out from this two dimensional conflict distribution (*Crosstable*).

| Teacher ID     | Favorite Slots             | Priority |
|----------------|----------------------------|----------|
| t <sub>1</sub> | 7,8,9,13,17,18,19,22,23,24 | 7        |
| t <sub>2</sub> | 2,12,22,29                 | 11       |
| t <sub>3</sub> | 5,10,15,20,25,30           | 9        |
| t <sub>4</sub> | 5,8,9,12,15,18,22,23,24,25 | 10       |

Fig. 2. Teachers' Slots Preferences along with Priority.

| Teacher ID     | Course ID      | Class Slot Routine | Sem. ID | Class per Week |
|----------------|----------------|--------------------|---------|----------------|
| t <sub>1</sub> | c <sub>1</sub> | 7,17               | 1       | 2              |
| t <sub>1</sub> | c <sub>4</sub> | 8,18,24            | 1       | 3              |
| t <sub>2</sub> | c <sub>2</sub> | 2,12               | 1       | 2              |
| t <sub>3</sub> | c <sub>3</sub> | 5,15               | 1       | 2              |
| t <sub>4</sub> | c <sub>5</sub> | 5,15,25            | 1       | 3              |

Fig. 3. Courses and Classes Distribution for the Teachers.

To minimize the potential for time conflict, an admissible heuristic ( $h$ ) is imposed to regroup the courses according to their dissimilarity. Graph Coloring algorithm is used to cluster where the conflict in *Crosstable* confirms the weights of the edges in the graph. Here the admissible heuristic is applied as the maximum number of colors needed to find the minimum number of groups of courses. Here faculty redundancy is also considered as weight, that is, same faculty of different courses cannot be in the same group. The output of Graph Coloring Algorithm assigns a color to all courses individually (same colored courses are treated as a group). Fig. 4. shows the resultant *Crosstable*. Each group may consist more than the threshold limit

members with tolerable conflict range. Here the number of the rooms is considered as the threshold value. In this work the tolerable conflict range is set to 0.

| Course ID      | c <sub>1</sub> | c <sub>2</sub> | c <sub>3</sub> | c <sub>4</sub> | Total Conflict |
|----------------|----------------|----------------|----------------|----------------|----------------|
| c <sub>1</sub> | 40             | 5              | 7              | 0              | 12             |
| c <sub>2</sub> | 5              | 50             | 0              | 2              | 7              |
| c <sub>3</sub> | 7              | 0              | 35             | 5              | 12             |
| c <sub>4</sub> | 0              | 2              | 5              | 40             | 7              |
|                | t <sub>1</sub> | t <sub>2</sub> | t <sub>3</sub> | t <sub>1</sub> |                |

Fig. 4. Crosstable for  $n \times n$  Courses.

This easy formation of coloring may lead to a measure of the undesirability of having classes overlapped in the routine. It will be effective to try to fit the most “chaos” courses of high priority teachers in the routine first. Random selection may be used to select teacher having same priority. The data used in this work to test the algorithm is real.

#### IV. THE DECISION TREE BASED ROUTINE GENERATION (DRG) ALGORITHM

The aim of the proposed DRG algorithm is to classify all the courses with a degree of satisfaction. Enormous permutation of courses may lead to a time consuming process. So a standardized branch & bound condition may be applied to reduce the problem surface area. The DRG sequentially follows 4 sets of cascading decision trees. Depending upon the emergence and success rate, the result of one tree is propagated to another tree. These transitions may lead to a solution but also may degrade the satisfaction threshold. A control portion helps to justify the problem solution needed to be more explored or not.

##### A. Priority regulated DRG (PDRG)

The first decision tree accumulates the high prioritized teacher  $t_k$  and most conflicted course  $c_i$  of  $t_k$  to the routine vector with maximum fitness value. The *max\_fit* function tries to discover the day-time pattern for the course. If the consequential day-time slots are already occupied by other courses, it checks the corresponding course color. If the color of the courses is same it classifies the course  $c_i$  with a degree. Here the fitness value of a course is referred as degree. The highest returned degree of a day-time slot, as maximum fitness value, is selected as the course class for  $c_i$ . This operation provides a pattern-based course distribution,  $A(t_k)$  or  $B(t_k)$ , in the routine although some courses may be placed partially as per their *class\_per\_week* and *max\_class\_per\_slot* constraints. In this manner, the low priority teachers may suffer by not getting the classes in a sequential manner. But in practice

26% of the total courses can be placed with zero conflict and with a high level of satisfaction. The level of satisfaction is a quantitative measure of placement for a course with respect to the teachers’ favorite slots. The definitive PDRG tree is shown in Fig. 5. The partially placed and not yet placed courses then elected as cascaded input to second level of exploration. The computational time of this operation requires  $O(n \times m)$  where the maximum number of courses per teacher is ‘n’ and the number of faculties is ‘m’.

##### B. Chaos eradication DRG (CDRG)

In this decision tree, the less demanded time slots (with respect to number of rooms) are labeled as “cold” whereas high demanded ones are labeled as “hot”. Among the remaining most conflicted courses (may not or partially placed) with low frequent time slots of the corresponding teacher are chosen for the second decision tree. If the considerable slot is empty then the course is placed in that

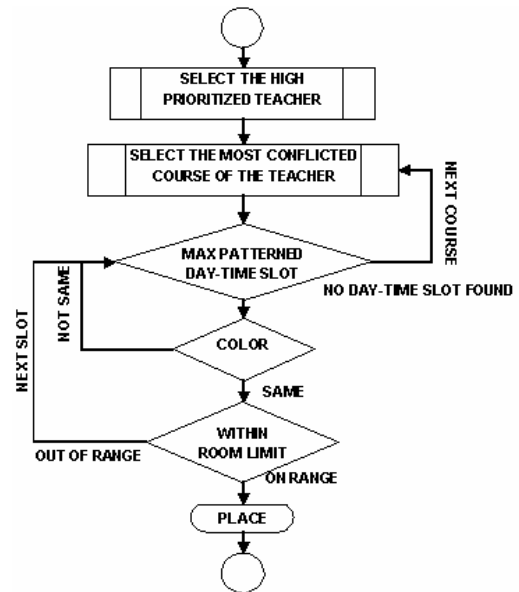


Fig. 5. Decision Tree of PDRG

slot, otherwise the colors have to be matched. If the color of the courses placed in the slot matches with the concerning course, the latter course is classified, and if not other options are taken into consideration for the teacher. The considerable courses in CDRG are the overlooked courses from PDRG, where PDRG confirms the day-time pattern is not possible for the courses due to color and the scattered choice of slots by the teachers. So day-time slots pattern are ignored in this operation. After using this second decision tree few courses may remain unchanged. Nevertheless this time the number of remaining courses is much less than the previous one. Around 32% of the remaining courses are placed successfully by CDRG. The

combined effort of decision trees still provides high confidence. Fig. 6. demonstrates the Chaos eradication DRG. Now this tree holds time complexity of  $O((n - d) \times (m - l))$  where the maximum number of courses per teacher is 'n', the number of faculties is 'm' and 'd' is the already classified courses of each teacher by the PDRG and 'l' is the number teacher whose all courses were placed in the routine by PDRG.

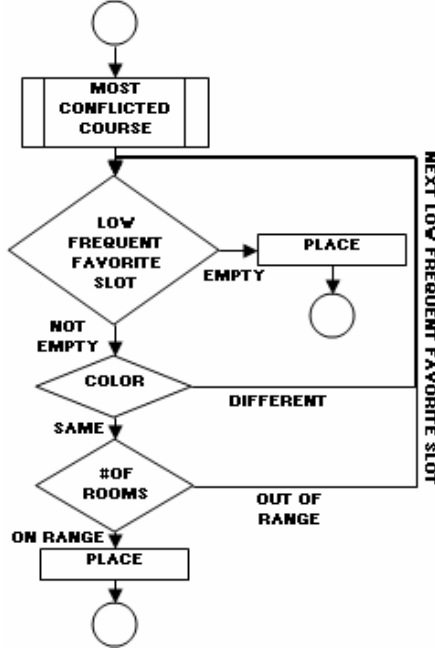


Fig. 6. Decision Tree of CDRG

### C. Tolerable DRG (TDRG)

The first two decision trees are aimed at automated generations of a better assignment. The second approach seeks to find an assignment of vector which may be more difficult to locate in the search space using the already assigned vector. The third decision tree allows the remaining courses according to the priority of the teachers to find a place into the routine within a tolerable conflict range of the subsequent teachers' favorite slots. Here the course color is overlooked. This classification now introduces errors into the system by considering the tolerable students conflicts only. Important issue is that, this manipulation may iterate several times to include as many courses possible to place in to the routine. 22% of the unclassified and partially classified courses are labeled with a tolerable error. The flowchart presentation of this decision tree is shown in Fig. 7. The average time complexity of TDRG is approximately  $O(t \times (n - (d + o)) \times (m - (l + p)))$  where the maximum number of courses per teacher is 'n', the number of faculties is 'm' and 'd' is the already classified courses of each teacher by the PDRG and 'l' is the number teacher whose all courses were placed in the routine by PDRG. 'o' is the number of

courses placed by CDRG and 'p' is the number of teachers whose courses are completely placed by CDRG. Here 't' is the number of TDRG iterates.

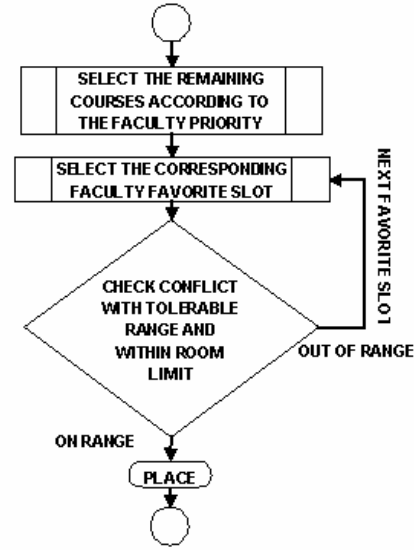


Fig. 7. Decision Tree of TDRG

### D. Neighboring Tour DRG (NTDRG)

The fourth possibility search allows the courses to look over the contour of the teachers' favorite slots to find the least conflicted slots. Although the students' scattered choices is overlooked in TDRG but the placed courses do not displease teachers' preference. The final decision tree is modeled in such a manner so that the rest of unplaced courses are going to be graded into the routine vector within a minimum distance of the teachers' choice. Manhattan Distance is calculated for this placement of courses. Manhattan Distance for a teacher is a vital performance measuring tool to find the slot gaps per day for an individual teacher. The optimization can be done by keeping cumulative Manhattan Distance for the teachers as low as possible. It is assured that the courses were not placed on that day. From the remaining courses, a course is selected according to the priority of the teacher. Complement of the intersection between the classification value of that course and the corresponding teachers' used slots are considered as new host slots. The neighboring slots of the new hosts are the most likely candidates. Among the candidates the most "cold" (less desired slots) slots are considered as candidates. Considerable issues in this placement are tolerable conflict range, allocable number of rooms and day-time misjudgment. Fig. 8. demonstrates the overall scenario of Neighboring Tour DRG (NTDRG). The approximate time complexity is  $O(2 \times (n - (d + o + u)) \times (m - (l + p + v))) \approx O((n - (d + o + u)) \times (m - (l + p + v)))$  where the maximum number of courses per teacher is 'n', the number of faculties is 'm'

and 'd' is the already classified courses of each teacher by the PDRG and 'l' is the number of teacher whose all courses were placed in the routine by PDRG. 'o' is the number of courses placed by CDRG and 'p' is the number of teachers whose courses are completely placed by CDRG. 'u' is the number of courses placed by TDRG and 'v' is the number of teachers whose courses are completely placed by TDRG.

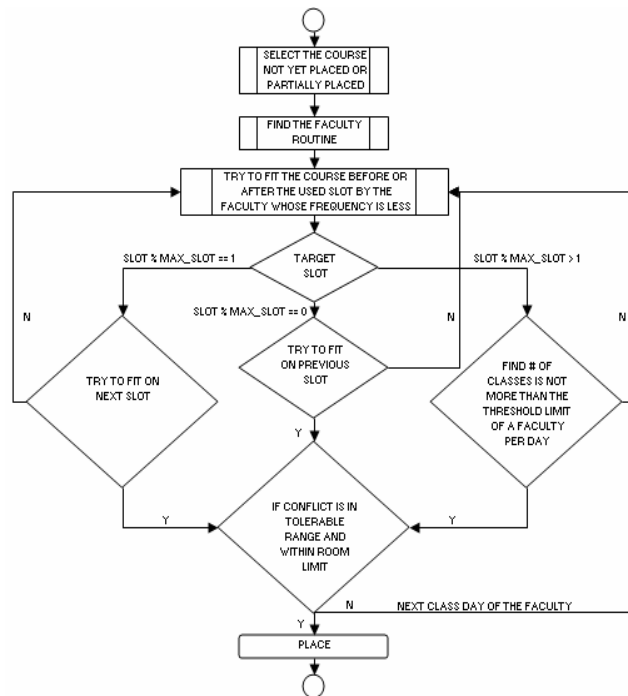


Fig. 8. Decision Tree of NTDRG

These four sequential decision trees feed forward to an acceptable solution. The cumulative time complexity of DRG is  $O(A.mn + B.m + C.n)$ , where A, B and C are constants depending upon the number of iterations of TDRG algorithm and number of courses placed by each and every algorithm. The algorithm degrades due to the fact that the students have a freedom to choose any course (assuming that the prerequisite course is completed). First search PDRG faces problem if same prioritized teacher focus into same favorite time slots. By using CDRG this situation may prevail over considering a level of discontinuity into the day-time pattern for the courses.

In second run, if the low “chaos” course holds high prioritized teacher then the classification may dissatisfy the teacher. For the third rotation the conflict may arise for the students for not considering the color. For NTDRG if the host slot is elected as the first and last class of the day then the previous and the next consecutive slot are from

different days. So, this day jump increases huge distance for the teacher which may lead to an unfeasible classification.

## V. EXPERIMENTAL RESULTS

Test results for DRG algorithm are carried out on a PC with Pentium IV/1.6 GHz processor and 256 MB of memory. Table I. shows the computational results for semester 1 and semester 2 with 66 and 61 courses correspondingly. Where for semester 1, around 24 teachers, with minimum 10 classes per week and at least 3 courses and 120 students with 430 combinations of choices of courses are considered as input. For semester 2, around 23 teachers, with minimum 10 classes per week and at least 3 courses and 106 students with 414 combinations of choices of courses are considered as input.

In Table I. *unclassified courses* refers to the number of courses that were not classified by any decision trees, *partially classified courses* gives the number of courses partially classified (the number classes already placed into the routine is less then the required classes). *Day-time pattern* shows the numbers of courses that followed the day-time pattern. *student conflict* estimates the percentage of unsatisfied requirements for courses by students. *Unsatisfied time* refers the number of time slots automatically generated beyond teachers' favorite choice. Randomization or prediction on classification is not used. So the output of this  $O(A.mn + B.m + C.n)$  based deterministic algorithm strictly depends upon the input.

## VI. CONCLUSION

This paper presents an algorithm that generates a promising result. It should be noted that the proposed Decision Tree based Routine Generation (DRG) algorithm using OLAP representation, to construct a university class routine, brings the complexity to a considerable level. This solution classifies 96% - 97% of the courses as well 93% - 95% satisfaction for teacher. For this data set, students' satisfaction is 100% but in general 90% - 93% satisfaction may be achievable by using DRG. Preferential requirements on time variables are met around 93%.

## ACKNOWLEDGEMENT

The proposed algorithm is employed to construct a class routine for a reputed university in Bangladesh. Author(s) of the paper express gratitude to the university authority for providing indispensable information.

TABLE I  
COMPUTATIONAL RESULTS OF DRG

|                   |                                      | PDRG | CDRG | TDRG | NTDRG | Final Finding |
|-------------------|--------------------------------------|------|------|------|-------|---------------|
| <b>Semester-1</b> | <b>Unclassified courses</b>          | 5    | 3    | 1    | 0     | 0             |
|                   | <b>Partially classified courses</b>  | 45   | 32   | 2    | 2     | 2             |
|                   | <b>Day-time patterned classified</b> | 16   | 31   | 63   | 64    | 64            |
|                   | <b>Student conflict</b>              | 0    | 0    | 0    | 0     | 0             |
|                   | <b>Unsatisfied time</b>              | 0    | 0    | 3    | 7     | 9             |
| <b>Semester-2</b> | <b>Unclassified courses</b>          | 2    | 2    | 1    | 0     | 0             |
|                   | <b>Partially classified courses</b>  | 41   | 26   | 2    | 2     | 2             |
|                   | <b>Day-time patterned classified</b> | 18   | 33   | 58   | 59    | 59            |
|                   | <b>Student conflict</b>              | 0    | 0    | 0    | 0     | 0             |
|                   | <b>Unsatisfied time</b>              | 0    | 0    | 5    | 8     | 11            |

#### REFERENCES

- [1] Christelle Guert, Narendra Jussien, Patrice Boizumault and Christain Prins, "Building university timetable using constraint logic programming," Springer-Verlag LNCS 1153, pp. 130-145, 1996.
- [2] Hans-Joachim Goltz, Georg Küchler and Dirk Matzkae, "Constraint-based timetabling for universities," INAP'98, pp. 75-80, 1998.
- [3] Slim Abdennadher and Michael Marte, "University course timetabling using constrain handling rules," Journal of Applied Artificial Intelligence, vol. 14, no. 4, pp. 311-326, 2000.
- [4] Thom Früwirth, "Constrains handling rules," Constrain Programming: Basic and Trandes, LNCS 910, Springer, 1995.
- [5] E.K. Burke, B.L. MacCarthy, S. Petrovic and R. Qu, "Knowledge Discovery in a Hyper-Heuristic for Course Timetabling Using Case-Based Reasoning," PATAT 2002, 4<sup>th</sup> International Conference, pp. 90-103, August 2002.
- [6] D. Abramson, "Constructing school timetables using simulated annealing: Sequential and parallel algorithms," Management Science, vol. 37, pp. 98-113, 1991.
- [7] A. Schaert, "Tabu search techniques for large high-school timetabling problems," 13<sup>th</sup> National Conference on Artificial Intelligence AAAI'96, pp. 363-368, 1996.
- [8] Luca Di Gaspero and Andra Schaerf, "Multi-Neighbour Local Search for Course Timetabling," PATAT 2002, 4<sup>th</sup> International Conference, pp. 128-132, August 2002.
- [9] P. Boizumault, Y. Delon, and L. Peridy, "Constrain logic programming for examination timetabling," Journal of Logic programming, vol. 26, pp. 217-233, 1996.
- [10] K. Dowslan, "Using simulated annealing for efficient allocation of students to practical classes," Applied Simulated Annealing – Lecture Notes in Economics and Mathematical System, Springer-Verlag, vol. 396, pp. 125-150, 1993.
- [11] K. Dowslan, "Off-the peg or made-to-measure? Timebaling and scheduling with sa and ts," PATAT'97, Springer – Verlag, pp. 37-52, 1998.
- [12] G.M. White and J. Zhang, "Generating complete university timetables by combining tabu search with constraint logic," PATAT'97, Springer – Verlag, pp. 187-198, 1998.
- [13] E.K. Burke, J. Newall, and R.F. Weare, "Initialization strategies and diversity in evolutionary timetabling," Evolutionary Computation Journal, vol. 6.1, pp. 81-103, 1998.
- [14] E.K. Burke, J. Newall, and R.F. Weare, "A simple heuristically guided search for the timetable problem," Proceeding of the International ICSC Symposium on Engineering of Intelligent System, pp. 574-579, 1998.
- [15] E.K. Burke, J. Newall, and R.F. Weare, "A memetic algorithm for university exam timetabling," Lecture Notes in Computer Science, Springer-Verlag, vol. 1153, pp. 241-250, 1996.
- [16] Tan, Steinbach and Kumar, Introduction to Data Mining, pp. 130-139, 2004

# Extended Algorithm for Spatial Characterization and Discrimination Rules

Md. Rashidul Hasan, Md. Zakir Hossain, Fahim Md. Chowdhury, Mohammad Khalad Hasan

Department of Computer Science & Engineering, Shah Jalal University of Science & Technology, Sylhet, Bangladesh. {shaikot\_cse, zakir\_01913, tanim\_csesust}@yahoo.com, khalad-cse@sust.edu.

**Abstract** — Spatial data mining, i.e., mining knowledge from large amounts of spatial data, is a demanding field since huge amounts of spatial data have been collected in various applications, ranging from remote sensing to geographical information systems (GIS), computer cartography, environmental assessment and planning. The collected data far exceed people's ability to analyze it. The number and the size of spatial databases are rapidly growing which results in an increasing need for spatial data mining. In this paper, we have presented a new spatial data mining algorithm for spatial characterization and discrimination rules. For spatial data mining algorithm, it is important that class membership of a database object is not only determined by its non-spatial attributes but also by the attributes of objects in its neighborhood. We have implemented the algorithm within a general framework for spatial data mining providing a small set of database primitives on top of a commercial spatial database management system. At the end, performance analysis using a real geographic database demonstrates the effectiveness of the proposed algorithms.

**Index Terms** — Characterization, data mining algorithm, discrimination, spatial database primitives, spatial data mining.

## I. INTRODUCTION

Knowledge discovery in databases (KDD) has been defined as the non-trivial process of discovering valid, novel, potentially useful and ultimately understandable patterns from data; a pattern is an expression in some language describing a subset of the data or a model applicable to that subset [1]. The process of KDD is interactive and iterative, involving several steps such as data selection, data reduction, data mining, and the evaluation of the data mining results [2]. The heart of the process, however, is the data mining step which consists of the application of data analysis and discovery algorithms that, under acceptable computational efficiency limitations, produce a particular enumeration of patterns over the data [1].

Spatial Database Systems (SDBS) [3] are database systems contain spatial objects representing natural objects like mountain or rivers, infrastructure like railroad, location, highways and/or extension as well as by several non-spatial attributes. Figure 1 illustrates a spatial database named Branches where the location of the branch has stored as spatial object.

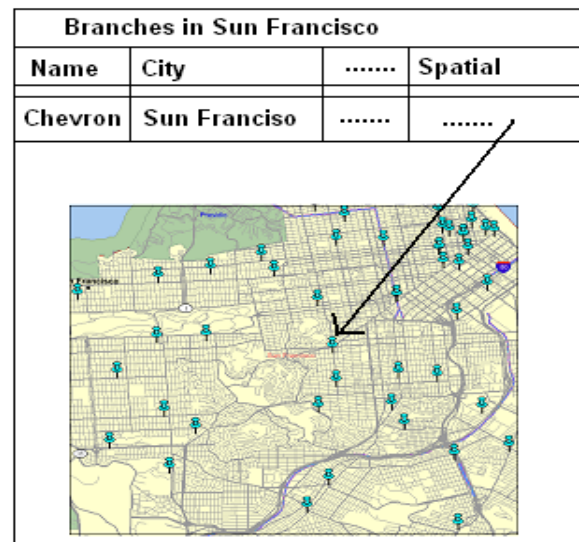


Fig. 1. Spatial and Non-spatial attribute of a data set.

Spatial data mining algorithms are used to find implicit regularities, rules or patterns hidden in large spatial databases, e.g. for geo-marketing, traffic control or environmental studies. A variety of data mining algorithms for mining in relational as well as spatial databases have been proposed in the literature [4][5][6].

The discovery process for spatial data is more complex than for relational data. This applies to both the efficiency of algorithms as well to the complexity of possible patterns that can be found in a spatial database. The reason is that, in contrast to mining in relational databases, spatial data mining algorithms have to consider the neighbors of objects in order to extract useful knowledge. This is necessary because the attributes of the neighbors of some object of interest may have a significant influence on the object itself [2].

In this paper, we present a new spatial data mining algorithm for both characterization and discrimination rules which extends current spatial characterization algorithm [4]. These tasks, especially characterization and discrimination rules in spatial databases were also studied in [6, 7, 8] and [12] respectively. A characterization rule is an assertion which characterizes a concept satisfied by all



or a majority number of the examples in the class undergoing learning (called the target class). For example, the symptoms of a specific disease can be summarized by a characterization rule. A discrimination rule is an assertion which discriminates a concept of the class being learned (the target class) from other classes (called contrasting classes) [12]. For example, to distinguish one disease from others, a discrimination rule should summarize the symptoms that discriminate this disease from others. We have developed an algorithm that can characterize target class and find out the concepts by which it is discriminated from other classes. For example, the algorithm can summarize the symptoms of a specific disease as well as it also can summarize the symptoms that discriminate this disease from others. Our algorithm for spatial characterization and discrimination are presented within a general framework based on database primitives [9] for spatial data mining.

## II. DATABASE PRIMITIVES FOR SPATIAL DATA MINING

For the efficient support of the spatial data mining algorithms one should analyze typical operations and data structures that can support different types of knowledge discovery queries. Martin Ester, Alexander Frommelt, Hans-Peter Kriegel and Jörg Sander have presented the concepts of neighborhood graphs, neighborhood paths and neighborhood indices [9]. The neighborhood concept covers different type of spatial relations including topological relations (meet, overlap etc.), distance relations (close to), and direction relations (north, west, under, etc.). The neighborhood graph for a relation neighbor is a graph where nodes are objects in the database and edges between nodes  $n_1$  and  $n_2$  represent the fact that the relationship neighbor ( $n_1, n_2$ ) holds. The neighborhood path in the neighborhood graph  $G$  is a set of nodes directly connected through the edges of the graph. The authors presented basic operations for the construction of graphs and for browsing and expansion of paths. Some important definition and basic operation has been described which are used in our algorithm.

### A. Neighborhood graphs and paths

Let *neighbor* be a neighborhood relation and  $DB$  be a database of spatial objects. A *neighborhood* graph  $G_{neighbor}^{DB} = (N, E)$  is a graph with nodes  $N = DB$  and edges  $E \subseteq N \times N$  where an edge  $e = (n_1, n_2)$  exists iff *neighbor*( $n_1, n_2$ ) holds. A *neighborhood* path of length  $k$  is defined as a sequence of nodes  $[n_1, n_2, \dots, n_k]$  where *neighbor* holds for all  $(n_i, n_{i+1})$ . The following basic operations are briefly described.

- *neighbors*:  $Graphs \times Objects \times Predicate \rightarrow Sets\_of\_objects$
- *paths*:  $Sets\_of\_objects \rightarrow Sets\_of\_paths$
- *extensions*:  $Graphs \times Sets\_of\_paths \times Integer \times Predicates \rightarrow Sets\_of\_paths$

The operation *neighbors*(*graph*, *object*, *predicate*) returns the set of all objects connected to object in graph satisfying the conditions expressed by the predicate *predicate*. The operation *paths*(*objects*) creates all paths of length 1 formed by a single element of objects and the operation *extensions*(*graph*, *paths*, *length*, *predicate*) returns the set of all paths of the specified length in graph extending one of the elements of paths. The extended paths must satisfy the predicate *predicate*. The elements of paths are not contained in the result implying that an empty result indicates that none of the elements of paths could be extended.

Because the number of neighborhood paths may become very large, the argument *predicate* in the operations *neighbors* and *extensions* acts as a filter to restrict the number of neighbors and paths to certain types of neighbors or paths. The definition of predicate may use spatial as well as non-spatial attributes of the objects or paths.

The techniques of implementing these basic operations (which is described above) in commercial spatial database system have been described here [2]. Typically, spatial index structures, e.g. R-trees, are used in a spatial DBMS (such as oracle 11g spatial) to speed up the processing of spatial queries such as region queries or nearest neighbor queries. Using the spatial index structures, a neighbor algorithm (which has signature *neighbors* (*Graph*, *object*, *predicate*)) [2] have presented which returns neighbor of a spatial object for some predicates in a graph. We have used the algorithm in the implementation of our proposed algorithm.

## III. CONCEPTS OF THE ALGORITHM

Spatial characterization of a set of target objects with respect to the database containing them as a description of the spatial and non-spatial properties which are typical for the target objects but not for the whole database. To obtain a spatial characterization, not only the properties of the target objects should be considered, but also the properties of their neighbors up to a given maximum number of edges in the neighborhood graph [4].

A spatial discrimination is a general description of the features discriminating or contrasting a class of spatial data (target objects) from other class (contrast objects),

like the comparison of price ranges of houses in different geographic regions. The relative frequencies of non-spatial attributes and the relative frequencies of different object types can be used as interesting properties for describing spatial characterization and spatial discrimination rules [2]. Figure 2 illustrates relative frequencies in the database as well as in the target regions and the ratio of these frequencies in comparison with the specified level of significance.

A *property* of some spatial object by a tuple of one of the two following formats: either *(attribute, value)* or

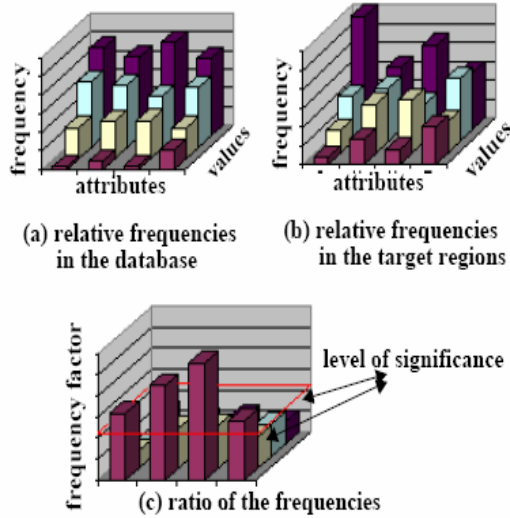


Fig. 2. Sample frequencies and differences [2].

*(“type”, type)* where “type” denotes the special “attribute” representing the object type such as community, mountain or river [2].

The task of spatial characterization is to discover the set of all properties for which the relative frequency in a set of target objects, extended by its neighbors, is significantly different from the relative frequency in DB [4]. The task of spatial discrimination is to discover the values of that set of properties in a set of contrast objects, extended by its neighbors and compare them with the values of the set of the properties for the target region. For spatial discrimination, it is important to find out properties which are significant with the properties of target objects and their neighbors, not with total databases because only same property values can be comparable.

So, to obtain spatial discrimination, the following steps are required:

1) First find the properties of target class which are significant from the properties of total database with expansion to maximum neighbors.

2) Find the values of the properties of the characterized target class in contrast objects with expansion to maximum neighbors. The values of these properties should have the maximum frequency in contrast region.

3) Compare the target and contrast objects using these properties.

#### IV. PROPOSED ALGORITHM

The current spatial characterization algorithm [2] is suitable for characterizing some spatial objects like finding characteristics of weather pattern in a certain region. We think the comparison of characteristics of weather pattern among two or more different regions is also necessary for business and real life purposes. In comparison, it is important to ensure that the relevant properties of two classes have been compared so that comparison makes valid sense. So, we have extended the current spatial characterization algorithm and make it compatible for finding both spatial characterization and discrimination rule. Using our proposed algorithm, the characteristics of some spatial objects (selected by some predicates) can be found as well as what the characteristics of that spatial objects discriminate from other contrast spatial objects can also be found. We have used in our algorithm same basic notations of current existing spatial characterization algorithm [2].

Let  $G_{neighbor}^{DB}$  be a neighborhood graph and *targets*, *contrasts* are subset of DB. Let frequency  $^s(prop)$  denote the number of occurrences of the property *prop* in the set *s* and let  $card(s)$  denote the cardinality of *s*. The *frequency factor* of *prop* with respect to set of spatial objects *objects* and *DB*, denoted by, is defined as follows [4]:

$$f_{objects}^{DB}(prop) \Rightarrow \frac{freq_{objects}(prop)}{card(objects)} \bigg/ \frac{freq^{DB}(prop)}{card(DB)}$$

Let *significance* and *proportion* be real numbers and let *max-neighbors* be a natural number. Let  $Neighbors_G^i(s)$  denote the set of all objects reachable from one of the elements of *s* by traversing at most *i* of the edges of the neighborhood graph *G*. Then, the task of *spatial characterization* is to discover each property *prop* and each natural number  $n \leq max-neighbors$  such that (1) the set  $objects = Neighbors_G^n(targets)$  as well as (2) the sets  $objects = Neighbors_G^n(\{t\})$  for at least *proportion* many  $t \in targets$  satisfy the condition:

**Discover-Spatial-Characterization-Discrimination** (graph  $Gr_r^{DB}$ , set of objects *targets*, set of objects *contrasts*, real *significance*, real *proportion*, integer *max-neighbors*)

initialize the set of (*targets\_characterizations*, *contrast\_characterizations*, *final\_targets\_characterizations*, *final\_contrast\_characterizations*, *final\_contrast\_characterizations\_nonspatial*, *final\_contrast\_characterizations\_spatial*) as empty.

Initialize the set of *target\_regions* to *targets*, the set of *contrast\_regions* to *contrasts*.

initialize integer *n* to 0, real *contrast\_region\_frequency\_factor* to 0, property *maximum\_frequency\_property* as empty ;

calculate  $frequency^{DB}(prop)$  for all properties *prop* = (attribute, value) in DB;

**while**  $n \leq \text{max-neighbors}$  **do**

**for** each attribute of DB and for the spatial attribute object type **do**

*contrast\_region\_frequency\_factor* = 0;

**for** each value of attribute **do**

            calculate  $frequency^{target\_regions}(prop)$  for property *prop* = (attribute, value);

            calculate  $frequency^{contrast\_regions}(prop)$  for property *prop* = (attribute, value);

**if**  $frequency_{target\_regions}^{DB}(prop) \geq \text{significance}$  or  $frequency_{target\_regions}^{DB}(prop) < \frac{1}{\text{significance}}$  **then**

**add** (*prop*, *n*,  $frequency_{target\_regions}^{DB}(prop)$ ) to the set of *targets\_characterization*;

**if**  $frequency_{contrast\_regions}^{DB}(prop) > \text{contrast\_region\_frequency\_factor}$

*contrast\_region\_frequency\_factor* =  $frequency_{contrast\_regions}^{DB}(prop)$  ;

                    replace *maximum\_frequency\_property* by *prop*;

**end of for.**

**add** (*maximum\_frequency\_property*, *n*,  $frequency_{contrast\_regions}^{DB}(prop)$ ) to the set of *contrast\_characterizations*;

**end of for**

**if**  $n < \text{max\_neighbors}$  **then**

**for** each object in *target\_regions* **do**

**add** neighbors( $Gr_r^{DB}$ , object, true) to *target\_regions*;

**for** each object in *contrast\_regions* **do**

**add** neighbors( $Gr_r^{DB}$ , object, true) to *contrast\_regions*;

    increment *n* by 1;

**end of while.**

*final\_targets\_characterizations* = extract all tuples (*prop*, *n*, *factor*) from *targets\_characterizations* which are significant in at least proportion of the regions with *n* extensions;

**for** each non-spatial attribute in *final\_targets\_characterizations* **do**

*final\_characterizations-nonspatial* = extract all tuples (*prop*(non spatial attribute, value), *n*, *factor*) from *contrast\_characterizations* set for that nonspatial-attribute which has maximum frequency-factor with *n* extensions;

**end of for.**

*final\_contrast\_characterizations\_spatial* = extract all tuples (*prop* (spatial-attribute, value)) from *contrast\_characterizations* set which are significant in at least proportion of the regions with *n* extensions;

**add** (*final\_contrast-characterizations-nonspatial*, *final\_contrast-characterizations-spatial*) to *final contrast-characterizations*.

**return** *final\_target\_characterization* set and the rule generated from these discriminations;

Fig. 3. Algorithm for discovering spatial characterization and discrimination rules

$$frequency_{Objects}^{DB}(prop) \leq \text{significance}$$

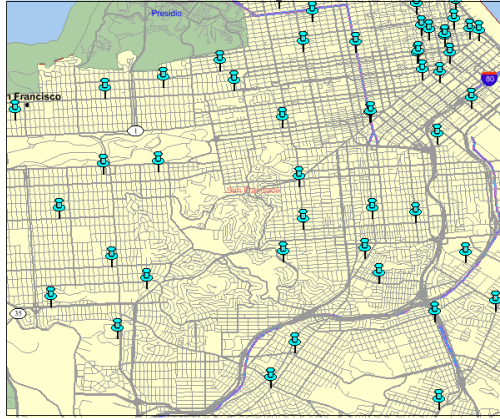
$$\text{Or } frequency_{Objects}^{DB}(prop) < \frac{1}{\text{significance}}$$

In point (1) the union of the neighborhood of all target objects is considered simultaneously, whereas in point (2) the neighborhood of each target is considered separately.

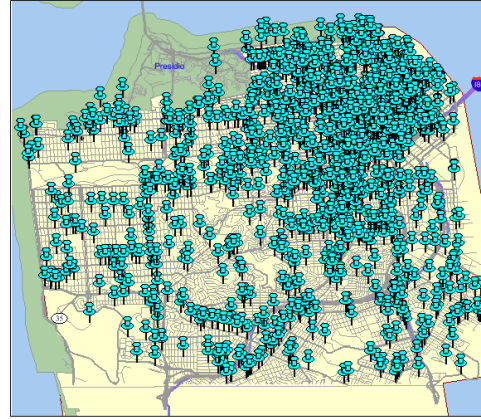
The parameter *proportion* specifies the minimum confidence required for the characterization rules and the frequency factors of the properties provide a measure of their interestingness with respect to the target objects. The parameter *proportion* is relevant only for the last step of the algorithm, i.e. for the generation of a rule. Figure 3 presents the algorithm for discovering spatial characterizations and discriminations.

In this algorithm, If the relative frequency of the property (attribute, value) in target region differ significantly (either grater or smaller) from relative

[11], an application server for Oracle spatial database for the visualization purpose of those spatial data.



Target Objects



Maximally Expanded Region

Fig. 4. Target objects and after expansion with maximum neighbors.

frequency of the property (attribute, value) in DB then the property is added to set of target characterization. If frequency factor of a property for current value of the attribute is grater than the previous value of that property in the contrast region, then the property is stored and added into the set of contrast characterization. After traversing up to maximum neighbor, for each non-spatial attribute in the actual set of target characterization, we extract same non-spatial attribute tuples from the set of contrast characterization and for spatial attributes, extract all tuples of spatial attribute which are significant at least proportion in the contrast regions with n extensions. Then we evaluate the discrimination rules by using by comparing the characterization sets of two regions.

## V. EXPERIMENTAL SETUP

We have implemented the database primitives on commercial spatial database Oracle 11g using its 2D spatial data which offers R trees in the spatial index structure. A geographical datasets was collected from sample spatial datasets [10] which contain three datasets (Customers, Branches and Competitor Branches). Each datasets contains several non-spatial attributes (such as name, category, city etc.) and one spatial object named Location which specifies their locations in the geodetic co-ordinate system. The total amount of spatial data among all the datasets is 3944. We have used Map Viewer

## VI. EXPERIMENTAL RESULT

The algorithm usually starts with small set of target and contrast objects, selected for instance by a predicate on some non-spatial attribute(s) such as 'city = Sun Francisco' for target objects and 'city = Washington DC' contrast objects in a dataset (such as Branches). You can also select target and contrast objects from different datasets (such as, target objects from Branches datasets and contrast objects from Competitors Branches). Here, we describe a testing result from our algorithm where target objects are selected from Branches datasets and contrast objects are selected from the Competitor Branches using predicate for both objects "city = Sun Francisco " so that these two types of branch (Branches and Competitors Branches) characteristics and their customers behavior can be analyzed. Then, the algorithms expands region around the target objects, simultaneously selecting those attributes of the regions for which the distribution of values differs significantly from the distribution in the whole database [Figure 4]. In the same way, it also expands region around contrast objects and selects those attributes for contrast region.

In the last step of the algorithm, characterization rules are generated for describing the target regions and discrimination rules are generated by comparing both target and contrast objects characteristics.

TABLE I. CHARACTERISTICS VALUES FOR TARGET AND CONTRAST OBJECTS.

| Attribute Name    | Target Attribute Value | Target Value Frequency Factor | Contrast Attribute Value | Contrast Value Frequency Factor |
|-------------------|------------------------|-------------------------------|--------------------------|---------------------------------|
| Customer Category | Hotel                  | 1.6962                        | Hotel                    | 1.3374                          |
| Street Number     | 301                    | 1.602                         | 502                      | 1.421                           |
| Street Name       | Mission ST             | 2.0007                        | Montgomery ST            | 1.801                           |
| City              | Sun Francisco          | 1.744                         | Sun Francisco            | 1.7105                          |
| Customer Grade    | Silver                 | 0.8488                        | Gold                     | 0.5432                          |

TABLE II. PERFORMANCE OF SPATIAL CHARACTERIZATION AND DISCRIMINATION ALGORITHM

| No. of Test Case | No. of objects in target region | No. of objects in contrast region | Run time for the algorithm (seconds) |
|------------------|---------------------------------|-----------------------------------|--------------------------------------|
| 1                | 1214                            | 627                               | 86.312                               |
| 2                | 1475                            | 516                               | 88.714                               |
| 3                | 1752                            | 547                               | 94.183                               |
| 4                | 1832                            | 841                               | 101.714                              |

Table 1 show the results of algorithm and specifies the characteristics of target objects which are most significant from the total database and values for the same properties in the contrast objects. We can informally say the characterization rules for target objects “Maximum branches in Sun Francisco are situated besides in the Mission Street and their maximum customers are normally hotels and silver graded”. The discrimination rule can be “Maximum branches in the city Sun Francisco are situated besides in the Mission Street and their customers are silver graded whether their competitor branches are situated in the Montgomery Street and their customers are gold graded”.

Table 2 reports the efficiency of our spatial characterization and discrimination algorithm. As the algorithm efficiency of the algorithm depends on the neighbors operation, so we provide here number of spatial objects have been manipulated in each test case after expansion neighbors in both target and contrast objects. Target and contrast objects have selected using various non-spatial predicates in different datasets.

## VII. CONCLUSION

In this paper, we have presented a new spatial data mining algorithm for both spatial characterization and

discrimination rules. In this algorithm, we consider not only the properties of target and contrast objects but also the properties of their neighbors as they impact on the characteristics of them. The algorithms were implemented successfully within a general framework for spatial data mining providing a small set of database primitives on top of a commercial spatial database management system Oracle 11g and performance analysis using a real geographic data show the effectiveness of the proposed algorithm.

## REFERENCES

- [1] Fayyad U. M., Piatetsky-Shapiro G., and Smyth P. 1996 “Knowledge Discovery and Data Mining: Towards a Unifying Framework”, Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining. Portland, Oregon, AAAI Press, Menlo Park, California, pp. 82 - 88.
- [2] ‘Algorithms and Applications of Spatial Data Mining.’ Martin Ester, Stefan Gundlach, Hans-Peter Kriegel, Jörg Sander.
- [3] Gueting, R.H. 1994. An Introduction to Spatial Database Systems. VLDB Journal 3(4).
- [4] Ester M., Frommelt A., Kriegel H.-P., and Sander J. 1998 “Algorithms for Characterization and Trend Detection in Spatial Databases”, Proc. 4th Int. Conf. on Knowledge Discovery and Data Mining, New York City, NY, pp. 44-50.
- [5] Chen M.-S., Han J., and Yu P. S. 1996 “Data Mining: An Overview from a Database Perspective”, IEEE Trans. on Knowledge and Data Engineering, Vol. 8, No. 6, pp. 866-883.
- [6] Koperski K., Adhikary J., Han J. 1996 “Knowledge Discovery in Spatial Databases: Progress and Challenges”, Proc. SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, Technical Report 96-08, University of British Columbia, Vancouver, Canada.
- [7] Lu W., Han J., Ooi B.C. 1993. Discovery of General Knowledge in Large Spatial Databases. Proc. Far East Workshop on Geographic Information Systems, Singapore, 275-289.
- [8] Knorr E.M. & Ng R.T. 1996. Finding Aggregate Proximity Relationships and Commonalities in Spatial Data Mining. IEEE Transactions on Knowledge and Data Engineering, 8(6):884-897.
- [9] Ester M., Kriegel H.-P. & Sander J. 1997. ‘Spatial Data Mining: A Database Approach’. Proc. 5th Int. Symp. on Large Spatial Databases, 47-66, Berlin: Springer.
- [10] Sample spatial datasets (<http://www.apress.com>)
- [11] <http://www.otn.oracle.com/software/products/mapviewer>
- [12] ‘Knowledge Discovery in Databases: An Attribute-Oriented Approach’- Jiawei Han, Yandong Cai, and Nick Cercone, School of Computing Science, Simon Fraser University Burnaby, British Columbia, Canada V5A 1S6.

# Isolating Significant Phrases in Common Natural Language Queries to Databases

Mohammad Moinul Hoque, Md. Shahriar Mahbub, and S.M. Abdullah Al-Mamun

Department of Computer Science and Engineering, Ahsanullah University of Science and Technology, Tejgaon, Dhaka, 1208, Bangladesh

**Abstract** — This paper describes a novel methodology for extracting significant phrases from common natural language queries to databases. We show the usability of the methodology for developing database front ends, and propose it as an acceptable alternative to developers of databases that are supposed to be subject to massive interaction by non-technical people. Here we localize the matching procedures by isolating the phrases that are important for mapping the queries to database details, and propose to use heuristics and soft computing techniques for inexact matching of the phrases.

**Index Terms** — Natural Language Query, Database Interface, Attribute Value Phrase, Object Identification.

## I. INTRODUCTION

To face the multifaceted problem of ambiguity present in natural language texts, usual NL (Natural Language) query processing limits its scope within finding key words or their synonyms in the queries [1]. Others like Microsoft English Query, much talked about over the Internet, propose extensive use of English grammar. We rather opted to go for an approach that emphasizes structural analysis of queries. This one supposes cashing on the facts that most common queries addressed to databases have a comparatively small number of strictly formulated structures. These structures can be outlined by marking the boundaries or delimiters of phrases significant for formulating expressions in a form that is widely used, like SQL (Structured Query Language). Our basic assumptions are as follows.

- The problem of creating a Natural Language Interface (NLI) to database remains open as far as accuracy, ease of implementation and generality of a solution is concerned.
- Unlike an arbitrary NL sentence, an NL query remains, as always, provocative to search for a technique to map it automatically to a SQL query that demands no effort for a commoner once placed for execution in a computer system.
- Finally, database developers, who are expected to know the domain as well as the users of the database they are going to develop, can incorporate an extension to the data dictionary along with a few algorithms for providing a good NLI to the database without much effort.

Very common aspects of Natural language processing like dealing with phrase structures are emphasized in the proposed methodology and it gives, we think, the fragrance of some sort of generality. Although the earlier works claimed to be portable [3], those were, as correctly observed in [1], basically not so natural language oriented but mostly check-box based. Later works claimed achieving reliability [8], but at the cost of scope of functionality. Recent works claim of no generality at all [10], [11]. With this background we propose our methodology, and as of today, we restrict our investigation to common queries, but we believe that the methodology will turn out to be significantly accommodative due to the generality of approach. Our work is closer in principle to the theory proposed in [8], where exhaustive tokenization of a limited type of queries and mapping tokens to database features are proposed, and closer to some ideas like semantic structure extraction referred to in [2], [5], [6], [7], [9].

Experimental encounters of the proposed methodology were designed keeping in front three of a very much common type but heavy duty database systems. The first one was the Result Processing System (RPS) of the University we work at and the other two, Air Travel Information System (ATIS) and Employee Information Management System (EIMS), had typical design and data that many can think of at ease. We analyzed queries that are common, but most significant in number like ‘What is the CGPA of the student 04020032?’, ‘List the students who are from Dhaka.’, ‘Show the total number of female employees who joined before 2007.’, ‘What is the type of the Thai airways plane that flies from Dhaka on Sunday evening for London?’

And for this type of queries we have found common phrase structures belonging to the set of a very limited number of predefined ones that are very simply traced down to SQL queries.

## II. SIGNIFICANT PHRASES

As our assumptions go, significant phrases are those that help to map an NL query to various components of an information repository for effective retrieval of the

information addressed to. We restrict our investigation by considering a relational database as a possible repository of information on one hand and restricted types of English Interrogative and Imperative sentences on the other. This implies that significant phrases are related to different aspects of object descriptions that are explicitly or implicitly presented in a relational database. It also implies that the phrases can be isolated using a limited number of syntactic and semantic rules.

First of all we can take into account Attribute Phrases (APs). By an AP we mean a phrase that denotes an 'attribute' in a 'relation' or 'table' which is most common, but obvious of the parameters used in a formal query language for information retrieval. 'student ID number' or 'roll' may simply denote the attribute 'id' in a possible 'student' relation.

Next come Attribute Value Phrases (AVPs). Specifying the values associated to the attributes described within a query is very much likely. This leads to identification of phrases that contain concrete values of the attributes specified in a query. We eventually come to the concept of Attribute Value Phrases or AVPs. In the query 'What is the CGPA of the student bearing identification number 050204003?', the phrase 'identification number 050204003' is a simple AVP, while 'newly enrolled' in 'Display the rolls and names of newly enrolled students in the department of CSE.' may pose to be a difficult one.

The concept of Object Identifying Phrases or OIPs is to be introduced for coping up with those nuances of NL queries that essentially include pointers to specific relations of the database in terms of typical descriptors common to the multiple relations of the same database. Distinguishing 'male patient' and 'male doctor' in a Hospital Information System is a must. More simple use of OIPs can be demonstrated in the query 'List the students of third year 1st semester', where no AP is specified. Here records of students are to be fetched although it is unlikely that an ordinary Data Dictionary allows a relation to be accessed bypassing the attributes. In a better situation when 'all' is added after 'List' it becomes easier to match to a predefined structure.

We have to be aware of the presence of Aggregate Function Phrases (AFPs) in common queries as well. There is no doubt that we are going to have phrases containing words like average, total, best, last, sum, lowest, etc. It also comes out interesting to isolate AFPs and recognize them in combination with APs, AVPs and OIPs present in a query.

Let's consider the query, Q1 given below.

Q1: What are the names and resident addresses of the students bearing id numbers 050402001 and 050204009?

Analysis of this query can be depicted using the diagram at Fig. 1 and it turns into a grammar rule like the one below.

Query  $\rightarrow$  HW1 AP and AP of OIP [oip\_del1] AVP avp\_del1 AVP

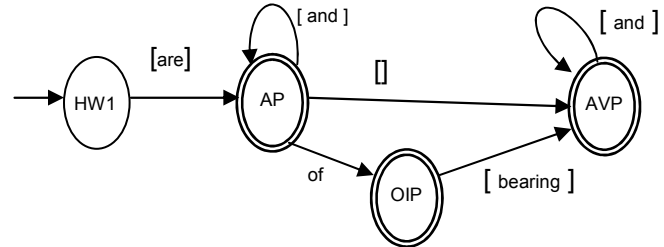


Fig. 1. Structure of the query, Q1

Here HW1 may stand for What, Who, etc. and oip\_del1 is a delimiter that may possibly appear after an OIP, like bearing, whose etc., and so on. Let's look into another example that begins with List.

Q2: List the students who got A+ in CSE203.

The pictorial representation of the sentence is shown in Fig. 2 followed by the grammatical representation.

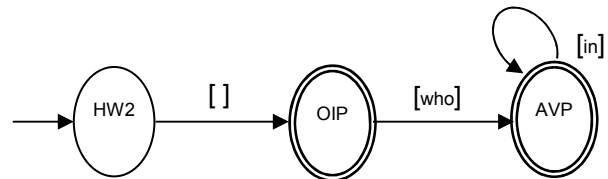


Fig. 2. Structure of the query, Q2

Query  $\rightarrow$  HW2 OIP [oip\_del2] AVP

Here HW2 may stand for List, Display, etc. and oip\_del2 is a delimiter that may possibly appear after an OIP, like of, bearing, who, etc. We present another example to illustrate aggregate function queries.

Q3: Show the average grade points of 1st year 1st semester CSE students in CSE103.

In case of this query, the representations look like the ones below.

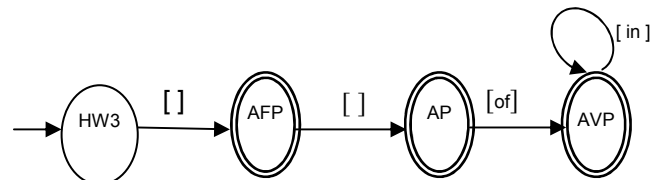


Fig. 3. Structure of the query, Q3

Query  $\rightarrow$  HW3 AFP AP [of] AVP [in] AVP

Here HW3 may stand for average, total, sum, etc.



### III. MARKING PHRASE BOUNDARIES

For marking the phrase boundaries, we propose that the NL query be preprocessed (filtered) first. The preprocessed NL query (PNLQ) is then used as input to the Algorithm1, the output of which is a phrase structure (PS). A PS is a sequence of pairs containing Non-delimiter phrases (NDPs) and NDP classes. Important features of the algorithm are that it uses a set of DFAs to recognize the delimiter structure and a set of Moore Machines (MMs) to generate the structure consisting of the NDP-NDP class pairs. The machines ensure efficient search by providing a strict formalism.

In the typical queries we specified in the scope of our investigation we have observed that a number of words do not play any role in formation of an executable query. It is in line with the idea of ‘words not semantically important’ in [8]. We have marked them as ‘insignificant’. The words ‘is’ and ‘the’ may be marked insignificant in the NLQ ‘What is the CGPA of the 3rd year student Abdul Karim?’. The sequence is supposed to undergo further refinement that frees it from words that may come into conflict with predefined phrase delimiters. While ‘of’ is a very important phrase delimiter in our objects of investigation, this very ‘of’ in the phrase ‘department of CSE’ will be a misleading delimiter. We have included the filtering processes above in the ‘Preprocessing’ part of our proposed system, for which correctly written specific NL queries are inputs and PNLQs are outputs.

Further processing for demarcation is performed using an algorithm that is specified below in Algorithm1. A set of rules is assumed that exhaustively describes the possible structures of queries each of which is numbered. A DFA with the same index recognizes the sequence of delimiters of each structure.

A possible DFA that can be used to recognize the delimiters of the extended version of Q1, Q1’: ‘What are the roll numbers, names, resident addresses and CGPAs of the students bearing id numbers 050102001, 060204034 and 050204009?’ is depicted below (Fig. 4).

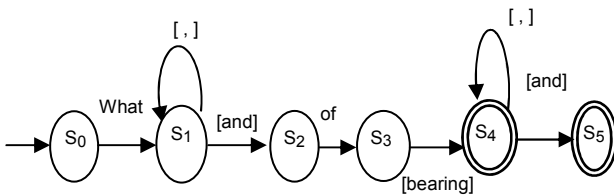


Fig. 4. DFA that recognizes the delimiters of Q1’

Then the relevant Moore Machine from the predefined set classifies the phrases isolated from the query may look like the following (Fig. 5).

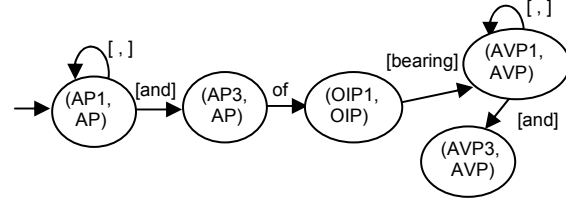


Fig. 5. Moore Machine that generates PS of Q1’

*Major steps of Algorithm 1:*

- Take a PNLQ as input.
- Isolate Delimiter Structure (DS) from the PNLQ and store the NDPs in the Buffer.
- If any DFA, DFA<sub>i</sub> of the DFA Set accepts the DS, then return the corresponding Rule number, R<sub>i</sub>, else return ‘Query unrecognized’ and terminate with ‘Unsuccessful’ notice.
- Find the Moore Machine, MM<sub>i</sub> from the MM Set using R<sub>i</sub>.
- Generate the structure consisting of the sequence of NDP-NDP class pairs, and the rule number.

Final processing supposes understanding the phrases in the PS returned by Algorithm 1, that is to say, mapping each phrase isolated in the PS to appropriate database details. It will include the Extended Data Dictionary (EDD). Algorithm 2 described in the next section is assigned the job.

### IV. ASSOCIATING PHRASES WITH DATABASE DETAILS

We propose an algorithm, Algorithm 2 to associate database details to the previously isolated phrases. The algorithm takes as input the set of phrase structures (PS) returned by the Algorithm 1 and itself returns the PS associated to the database object descriptors. The output has been abbreviated as PSO.

An extension of Data Dictionary in line with the idea temporal phrase affixes associated to concrete presented in [4], is referred here which has been proposed to have a number of relations (tables), such as a table with known APs associated to concrete attributes of the database tables, a table with known OIPs associated to concrete database tables and so on. A special table called Typical Value Table (TVT) is also proposed that would contain frequently used values of certain attributes, patterns of such values and common attributes of database tables.

*Major steps of Algorithm 2:*

**Begin Procedure**

For each element of the sequence of NDP-NDP class from the PS do



IF the NDP class is AP then  
 Try to match the NDP with the APs of the EDD.  
 In case the NDP sufficiently matches to an AP,  
 place the corresponding object-attribute details  
 along with the NDP class in the PSO accordingly;  
 Otherwise place the NDP marked 'Unrecognized'  
 along with the NDP class.  
 ELSE IF the NDP class is OIP then  
 Try to match the NDP with the OIPs of the EDD.  
 In case the NDP sufficiently matches to an OIP,  
 place the corresponding object details along  
 with the NDP class in the PSO accordingly;  
 Otherwise place the NDP marked  
 'Unrecognized' along with the NDP class.  
 ELSE IF the NDP class is AVP then  
 Try to match the NDP with the APs of the EDD.  
 IF the NDP sufficiently matches to an AP  
 Separate the AP portion of the NDP and take  
 the remaining part as the value of the AP,  
 incase this is a suffix or prefix of the NDP,  
 and place the corresponding object-attribute  
 details along with the value portion and NDP  
 class in the PSO accordingly; In case it is not,  
 place the NDP marked 'Unrecognized' along  
 with the NDP class.  
 ELSE  
 Try to establish that value + OIP = AVP in a  
 similar way.  
 ELSE  
 Try to match the NDP with the values in  
 TVT (Typical Value Table).  
 IF a match is found then  
 Place the corresponding object  
 attribute details along with the NDP as its  
 value in the PSO accordingly.  
 ELSE  
 Analyze the element inserted in the PSO  
 immediately before. In case it is an  
 AVP, treat the current NDP as an AVP  
 also; In case not, place the NDP

marked 'Unrecognized' along with the  
 NDP class.  
 END IF  
 END IF  
 END IF  
 End Procedure

## V. EXPERIMENTAL VERIFICATION

We have verified our algorithms over three different databases named earlier. For experimentation, commonly used database systems like RPS, ATIS and EIMS were chosen to diversely test our system using different level of queries in terms of complexity. Various types of possible queries were formulated based on these systems. We have manually extracted different phrases from these queries and compared these phrases with the recorded output of the algorithms. Although the system showed near 100% accuracy in some cases, it also showed some average performance for some complex and not well formulated queries. From a sample set of NLQs, 15 versatile queries were selected for the experimental purpose and the following table (Table 1) depicts a comparison between actual significant phrases and detected significant phrases presented in those NLQs. The resulting graphs for detecting APs, AVPs and OIPs are shown in Fig. 6, Fig.7 and in Fig. 8.

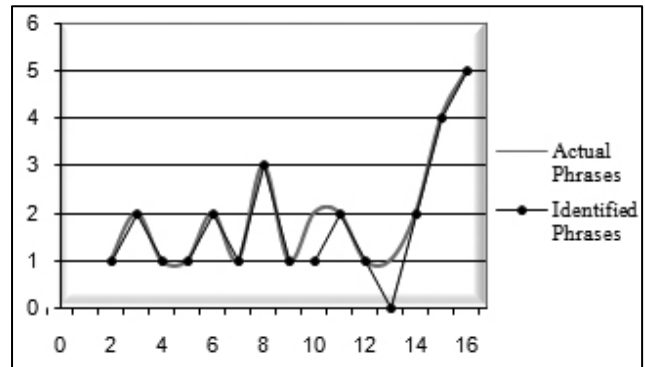


Fig. 6. Identification of APs.

TABLE 1  
 COMPARISON OF ACTUAL SIGNIFICANT PHRASES AND DETECTED SIGNIFICANT PHRASES IN NLQS

| Query Number    | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|-----------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| Actual APs      | 1 | 2 | 1 | 1 | 2 | 1 | 3 | 1 | 2 | 2  | 1  | 2  | 2  | 4  | 5  |
| Identified APs  | 1 | 2 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 2  | 1  | 1  | 2  | 4  | 5  |
| Actual AVPs     | 3 | 2 | 5 | 2 | 2 | 1 | 2 | 6 | 2 | 3  | 1  | 2  | 1  | 2  | 3  |
| Identified AVPs | 2 | 2 | 4 | 3 | 2 | 1 | 2 | 5 | 2 | 3  | 1  | 2  | 1  | 2  | 3  |
| Actual OIPs     | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2  | 2  | 1  | 1  | 1  | 1  |
| Identified OIPs | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2  | 2  | 1  | 0  | 1  | 1  |

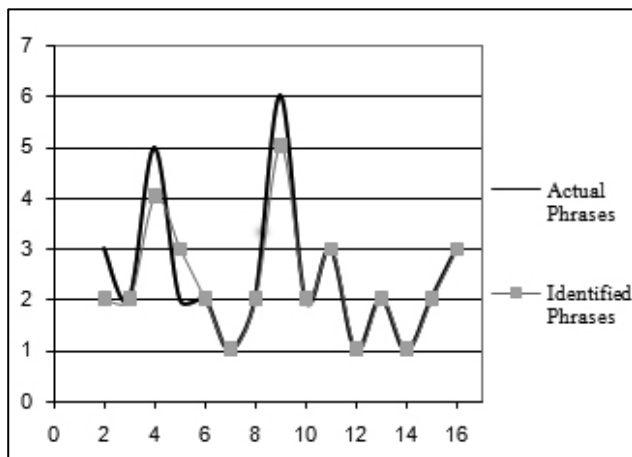


Fig. 7. Identification of AVPs.

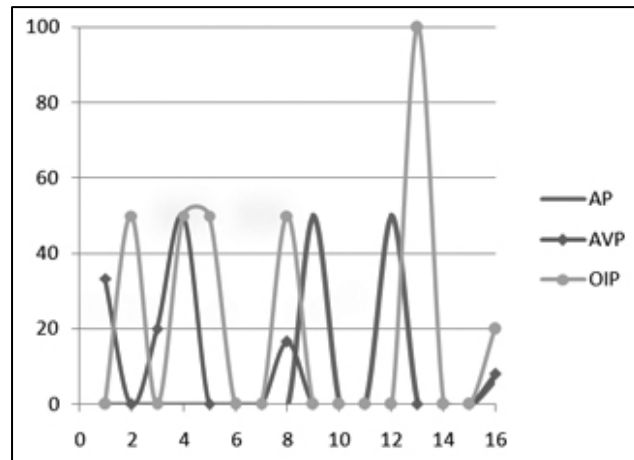


Fig. 9. Error detection rate in different NLQs.

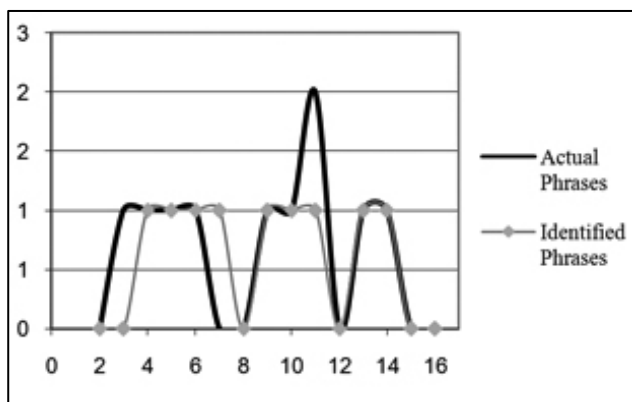


Fig. 8. Identification of OIPs.

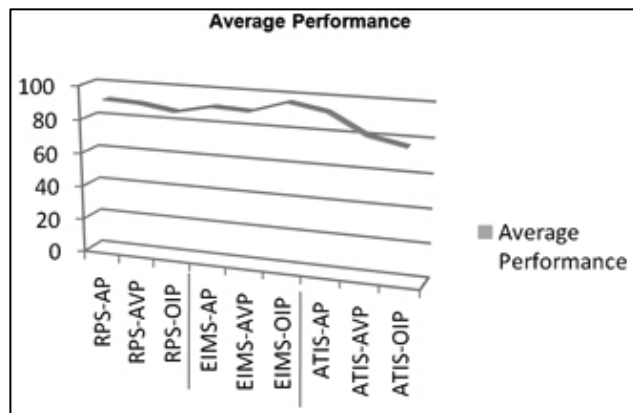


Fig. 10. Overall performance of the system.

The rate of errors that occurred in various queries while detecting APs, AVPs and OIPs are given in Table 2 and the pictorial representation is also shown in Fig. 9.

Comparing over three different database systems, we have found that, on an average, a maximum of 20% error occurred in identifying OIPs and the error rate in indentifying APs and AVPs were below 10% which was quite satisfactory.

We have used over two hundred of possible NLQs taken from three different database systems and on an average the system showed around 88.5% success rate in indentifying significant phrases as depicted in Fig. 10.

## VI. DISCUSSION AND FUTURE WORKS

Evidences present in the related publications urge to believe a few things which we tried to address in our work. First of all, it is evident that a relatively small number of basic structures involving significant phrases are there in NL queries. And this directed us to make one of our basic assumptions for the current investigation.

Secondly, we come to the fact that extensive use of NL grammars may be avoided. That means we need to handle less ambiguity, which is much easier to implement in contrast. Further, portability across natural languages

TABLE 2  
ERROR RATE IN IDENTIFYING DIFFERENT SIGNIFICANT PHRASES

|     | 1    | 2  | 3  | 4  | 5  | 6 | 7 | 8    | 9  | 10 | 11 | 12 | 13  | 14 | 15 | Average |
|-----|------|----|----|----|----|---|---|------|----|----|----|----|-----|----|----|---------|
| AP  | 0    | 0  | 0  | 0  | 0  | 0 | 0 | 0    | 50 | 0  | 0  | 50 | 0   | 0  | 0  | 6.7%    |
| AVP | 33.3 | 0  | 20 | 50 | 0  | 0 | 0 | 16.7 | 0  | 0  | 0  | 0  | 0   | 0  | 0  | 8%      |
| OIP | 0    | 50 | 0  | 50 | 50 | 0 | 0 | 50   | 0  | 0  | 0  | 0  | 100 | 0  | 0  | 20%     |

besides database domains can be achieved more easily by dealing with phrase structures in sentences avoiding grammatical nuances.

Next, it is also evident that we can avoid extensive token matching by isolating the significant phrases first. Extensive token matching becomes intractable in wider range. Soft computing and other nontraditional techniques like partial matching, heuristics in the form of studying prefixes and suffixes, etc. come to the rescue to certain extent in this regard.

Lastly, we find that a kind of generality may be achieved by accommodating techniques that find phrase structures from the types of sentences other than those we investigated. We think, it would be a good addition to NL understanding.

And we would like to present our beliefs as our reasons in favor of our proposed methodology. Besides efficiently constructing intelligent interfaces for massive databases, we think, it can be used as a tool to shed an insight into understanding sentences in general. It would be more appropriate if the representation of the world is assumed to be a collection of semantic networks or ontology of various things. The paper is presented here as a report of an ongoing research, and we hope to go on adding efforts in this line. Our future plan also includes extensive use of soft computing techniques for matching phrases to DB components as well as enhancing NL understanding study using proposed structures.

#### REFERENCES

- [1] M. Dittenbach, D. Merkl and H. Berger., "A Natural Language Query Interface for Tourism Information", In Proceedings of 10th International Conference on Information Technologies in Tourism (ENTER'03)., pp.152-162., Helsinki, Finland, January 29-31, 2003.
- [2] H. Ishikawa et al., "Designing a knowledge-based natural language interface", IEEE Expert, 2(2):57-71, 1987.
- [3] S. Jerrold Kaplan., "Designing a Portable Natural Language Database Query System", ACM Transactions on Database Systems (TODS), vol.9, pp. 1-19., March 1984.
- [4] Fred J. Damerau., "Problems and some solutions in customization of natural language database front ends", ACM Transactions on Information Systems (TOIS), vol.3., n.2, pp.165-184, April 1985.
- [5] Beck, H.W. Navathe, S., "Integrating natural language, query processing, and semantic data models", In Digest of Papers, Thirty-Fifth IEEE Computer Society International Conference, Compcon Spring'90. pp. 538-543, 1990.
- [6] H.H. Mengg, K.C. Siu., "Semiautomatic Acquisition of Semantic Structures for Understanding Domain-Specific Natural Language Queries", IEEE Transactions on Knowledge and Data Engineering, pp. 172-181, Jan.– Feb. 2002.
- [7] D. G. Shin., "Semantics modeling issues for processing natural language database queries", In Proceedings of the 1990 ACM annual conference on Cooperation ,Washington, D.C., United States, pp. 8 – 14, 1990.
- [8] Ana-Maria Popescu, Oren Etzioni, Henry Kautz., "Towards a Theory of Natural Language Interfaces to Databases", In Proceedings of the 8<sup>th</sup> international conference on Intelligent user interfaces, Miami, Florida, USA., pp. 149 – 157, 2003.
- [9] Jürgen M Janas., "The semantics-based natural language interface to relational databases, Cooperative interfaces to information systems", Springer Verlag, New York, 1986.
- [10] Majdi Owda, Zuhair Bandar, Keeley Crockett, "Conversation-Based Natural Language Interface to Relational Databases", In proceedings of ACM International Conferences on Web Intelligence and Intelligent Agent Technology Workshop, pp. 363-367, 2007.
- [11] Onur Küçüktunç, Ugur Gündükbay, Özgür Ulusoy., "A Natural Language-Based Interface for Querying a Video Database", IEEE Multimedia, Vol. 14, No. 1, pp. 83-89, January-March, 2007.

# Meliorated Approach for Extracting Bilingual Terminology from Wikipedia

Anand Gupta, Akhil Goyal, Aman Bindal and Ankuj Gupta

Department of Computer Science

Netaji Subhas Institute of Technology, Delhi

Emails - anand@coe.nsit.ac.in, akhil587@gmail.com, aman.bindal@nsitlounge.in,  
ankuj.gupta@nsitonline.in

**Abstract**— With the demand of accurate and domain specific bilingual dictionaries, research in the field of automatic dictionary extraction has become popular. Due to lack of domain specific terminology in parallel corpora, extraction of bilingual terminology from Wikipedia (a corpus for knowledge extraction having a huge amount of articles, links within different languages, a dense link structure and a number of redirect pages) has taken up a new research in the field of bilingual dictionary creation. Our method not only analyzes interlanguage links along with redirect page titles and linktext titles but also filters out inaccurate translation candidates using pattern matching. Score of each translation candidate is calculated using page parameters and then setting an appropriate threshold as compared to previous approach, which was solely, based on backward links. In our experiment, we proved the advantages of our approach compared to the traditional approach.

**Index Terms**— Wikipedia, Bilingual Dictionary, Wikipedia Mining, Automated translation.

## I. INTRODUCTION

Dictionary extraction from large amounts of bilingual text corpora is an emerging research area. A bilingual dictionary is a dictionary that is usually used to translate words or phrases from one language to another. There can be two approaches for extracting bilingual terminology—manual or automatic creation. But as stated in [1], the manual creation of bilingual dictionaries is not efficient, as they are time consuming and new or highly specialized domain specific words are difficult to cover. Moreover, linguistic knowledge is expensive. The problem proliferates when we consider two entirely different languages like English and Japanese as sufficiently large parallel corpora is not available for both languages.

Therefore, in order to provide a higher degree of accuracy and coverage, [1] propose the extraction of bilingual terminology from multilingual encyclopedias like Wikipedia, as it is continuously expanding encyclopedia containing millions of pages in hundreds of languages with a broad variety of topics. For extracting bilingual terminology from Wikipedia structure an approach is given in [1]. The previous research work was based only on backward links of each translation candidate neglecting page parameters. As a result, the dictionary [4] proposed

by [1] lags behind in many aspects, namely - fewer results, lesser coverage of words/ phrases, case sensitivity.

Thus, in order to cover all the aspects that were missing in the previous approach, we performed pattern matching which eliminated most of the wrong translations. For all the translation candidates, we calculated a score based on page parameters like page length and page random (which are present in the Wikipedia database), so that we can eliminate the translations with low scores and to make a better relativity with the searched string. In order to evaluate our method, we extracted Japanese translations for 200 English words and compared them for the accuracy and coverage with that of the previous dictionary [4].

The remainder of this paper is organized in the following way, related work in the field of bilingual terminology extraction, followed by shortcomings which we discovered in earlier approaches and the process exercised by us to amend them. Subsequent sections explains the Basic terminology, System design, Algorithm implemented by us. Further, the results obtained from both approaches are compared. In the last section, we have concluded the paper by mentioning certain features that can be worked upon.

## II. RELATED WORK

In 2008, Maïke Erdmann et al.[1] collected bilingual terminology based on the dense linked structure of the Wikipedia. They used number of backward links as the criteria for generating score of a translation candidate. But a score calculated without considering the page parameters defined by Wikipedia is inappropriate.

For translations from English to Japanese, one of the most commonly used dictionaries is, freely available online dictionary **EDICT** [2]. It comprises of more than 99,300 terms as on 2004. However, even with the aid of a large community, the manual creation of such a dictionary is a time-consuming process as mentioned in [1].

Besides this, the coverage of domain specific words and their related terminology was less. Moreover, while for high frequency terms usually good results can be achieved, the accuracy decreases drastically when the terms to be translated are domain specific.

### III. MOTIVATION AND CONTRIBUTION

The research described in this paper is motivated by the larger goal of improving the previous work [1]. The dictionary [4] based on [1] had some grave problems which are stated as under :-

- 1) Wrong translations for some words came into picture while trying. For instance, take "India" whose one of the translation candidate came out to be "heaven".
- 2) The translation candidate's score was solely calculated based on the number of backward links not taking into consideration the page parameters.
- 3) Another problem was case sensitivity. There were words like "eu" (abbreviation for European Union) which did not have any results but "Eu" had results.
- 4) The dictionary didn't have common words used in the day to day life like evening, etc. even though they were present in Wikipedia.
- 5) The number of translation candidates or related topics was pretty less and the score given for these related topics was not appreciable.

In the view of above mentioned points we analyze the dense link structure of Wikipedia and found a scope of improvement in them.

We briefly describe our approach to amend these points. For incorporating the first point, we performed pattern matching of the translations using the interlanguage link title and redirect page title which eliminated most of the incorrect results.

We explored some new parameters present in Wikipedia i.e. page length and page random instead of 'backward links' which was used by the previous research [1]. This concept helped the score to become more flexible and authentic. The coding was structured in such a way so as to make dictionary case insensitive. The search string was searched in the page table of the Wikipedia database which has all the titles of Wikipedia articles, so that each and every word of Wikipedia can be covered (including common words used in the daily life).

We improved the quality of the score by assigning proper weights to the parameters based on experimental results. This gave scores varying from 0-1 which made the dictionary more versatile and effective.

In continuation with the above mentioned points we enlarged our scope to other languages also. We downloaded the SQL tables from the Wikipedia dump [3] for three languages: English, Japanese and French and use them to make the following language conversions:-

- English to Japanese
- English to French
- Japanese to French
- Japanese to English
- French to English and
- French to Japanese

### IV. WIKIPEDIA LINK STRUCTURE

Our idea is to use a multilingual Web-based encyclopedia such as Wikipedia for extracting bilingual terminology. Wikipedia currently contains more than 10 million articles. It covers general topics, domain specific topics as well as proper nouns, containing even latest terminology since Wikipedia is being updated all the time. Moreover, Wikipedia contains a lot of links among its articles, not only within the articles of one language but also between articles of different languages. Before describing our method we would like to explain the different types of links and pages that are present in Wikipedia.

**Interlanguage Link:-** An interlanguage link in Wikipedia is a link between two articles in different languages. We assume that in most cases, the titles of two articles connected by an interlanguage link are translations of each other.

**Redirect Pages:-** Redirect pages in Wikipedia are pages containing no content but a link to another article (target page) in order to facilitate the access to Wikipedia content. When a user accesses a redirect page, he will automatically be redirected to the target page. Redirect pages are usually strongly related to the concept of the target page. They often indicate synonym terms, but can also be abbreviations, more scientific or more common terms, frequent misspellings or alternative spellings etc.

**Link Texts:-** A link text, is the text part of a link, i.e. the text that is presented to the user in the browser where he clicks to reach the target page. In Wikipedia, the title of the target article is displayed as the link text by default. The link texts are usually strongly related to the target page title. In many cases, they differ only in capitalization, but sometimes they are changed in other ways to fit in the sentence structure of the linking article.

**Page Parameters:-** Wikipedia provides a number of parameters for each page that exists within Wikipedia. Each page has an entry within its databases which identifies it by title and contains some essential metadata.

- `page_is_redirect`: Indicates whether a page is redirect or not.
- `page_random`: Random value between 0 and 1.
- `page_len`: Uncompressed length of page in bytes.

Before moving further we have defined few terms and notations.

### V. TERMS AND NOTATION

- 1) **Source Page(*sp*)**: It is the page corresponding to search string in input language.
- 2) **Target Page (*tp*)**: It denotes the target page in the desired language.
- 3) **Target Page's Title (*ill*)**: It represents the target page's title.
- 4) **Redirect Page Titles (*RP{tp}*)**: It denotes the titles extracted from the redirect pages of *tp*.

- 5) **LinkTexts Titles ( $LT\{tp\}$ )**: It signifies the titles of the linktexts of  $tp$ .
- 6) **Translation Candidates ( $TCs$ )**: This parameter gives the value of the total number of translation candidates after extraction and pattern matching.

$$TCs = \sum \{ill\}URP\{tp\}ULT\{tp\}$$

- 7) **Type**: It signifies two ways to get the result set. One is basic which covers only  $ill$  and other enhanced which extends to  $RP\{tp\}$  and  $LT\{tp\}$ .
- 8) **Filtered  $TCs$ ( $FTCs$ ) and Removed  $TCs$ ( $RTCs$ )**: The undesired translation candidates are filtered out using pattern matching. Let the filtered out translation candidates be denoted by  $FTCs$  and the removed translation candidates be denoted by  $RTCs$ . Thus  $FTCs$  is equal to

$$FTCs = TCs - RTCs$$

- 9) **Result Set( $RS$ )**: These are  $FTCs$  after applying threshold.
- 10) **False Positive Ratio( $FPR$ )**: It is the count of false positives appeared in the results per total number of results obtained

$$FPR = \frac{Count(false\ positives)}{Count(RS)}$$

- 11) **Zerom**: It is the count of the total number of results whose translations could not be found in the dictionary.
- 12) **Sensitivity**: This term checks for the case sensitivity in searching of translations.

In the remainder of the paper we will be using the above mentioned abbreviations and terminology.

## VI. SYSTEM DESIGN

The blocks in Figure 1 are explained in the following text.

- **User Interface**: It provides a basic interface for user to enter the string that the user wants to search,  $type$  and the choice of languages. It also includes Form Validator, which checks for the correctness of the data entered. If the entries are correct the values are passed to server for further processing.
- **Language Code Generator**: It generates code for input and output languages based on the Wikipedia codes.
- **Dname and Tablename Generator**: It takes the source language code and desired language code and decided which databases need to be used.
- **Query Generator**: The function of the query generator is to generate SQL queries to be used by four processes namely: *Source Page finder*, *Target Page Finder*, *Redirect Page Finder* and *LinkText Finder*. It

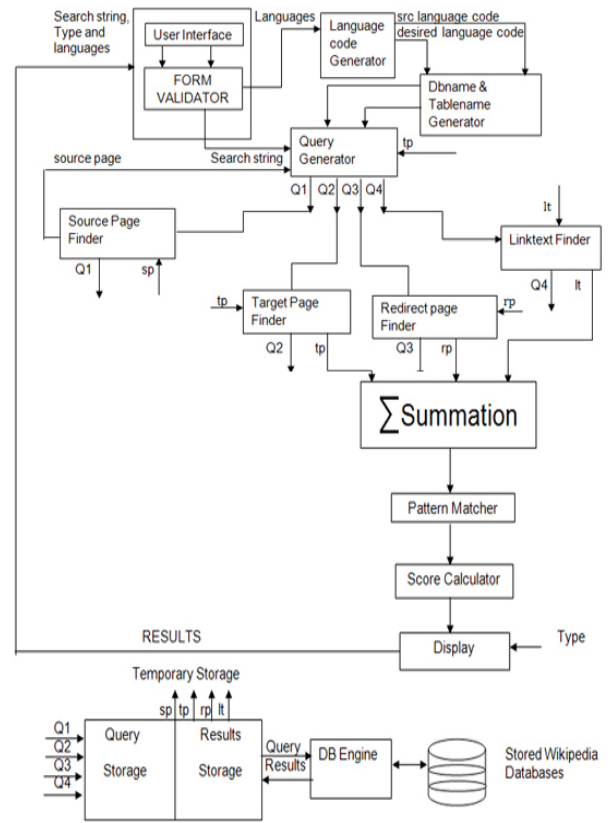


Fig. 1. System Design: Meliorated Approach for Extracting Bilingual Terminology from Wikipedia

generates four type of queries namely  $Q_1$ ,  $Q_2$ ,  $Q_3$  and  $Q_4$  which are passed into process in earlier mentioned order.

- **Temporary Storage**: Temporary storage is partitioned in two segments one each to store the incoming query from the processes and one to store the incoming resultsets from the DBengine.
- **DBengine**: A database engine is the underlying software component that a Database Management System (DBMS) uses to create, retrieve, update and delete data from a database.
- **Source Page Finder**: It searches for the source page in the source language. It takes query  $Q_1$  from Query Generator passes it to Temporary Storage which passes the query to DBengine which finally fires the query to the stored Wikipedia databases. The tables generated are stored again in the temporary storage from where they are returned back to the calling process i.e. Source Page Finder.
- **Target Page Finder**: After the source page is found its interlanguage link's page is found in the desired language. It is known as target page. The title  $ill$  of  $tp$  becomes the translation candidate. It takes query

$Q_2$  from Query generator and performs the same operation as mentioned for *Source Page Finder*.

- **Enhancement by Redirect Pages:** Once the target page is found, its redirect pages are searched. The title of these redirect pages becomes the Translation candidate (TCs). The process is called *Redirect Page Finder*. It takes query  $Q_3$  and performs the earlier mentioned process.
- **Enhancement by LinkTexts:** The TCs is further improved by including the titles of all the linktexts within the target page. The title set of linktexts is denoted by  $LT\{tp\}$ . The process is known as *LinkText Finder* and it takes query  $Q_4$  and performs the above mentioned process.
- **Summation:** Now all the outputs from three processes form the part of Translation Candidates. Thus, all the three outputs are summed to give the final set of Translation Candidates. Now TCs becomes

$$TCs = \sum \{ill\}URP\{tp\}ULT\{tp\}$$

- **Pattern Matcher:** Since all the translation candidate are not suitable pattern matching is performed based on the target page  $tp$  title  $ill$  and the redirect page titles  $RP\{tp\}$ . The undesired translation candidates are removed. It gives FTCs as output
- **Score Calculator:** Even after pattern matching some of the TCs are still not suitable translations of the source page title. Experimental results confirmed that the page parameters of the target page can be used to estimate the accuracy of a TC. We calculated the score of every TC by assigning weightages based on observation to these parameters. Moreover within the translation candidates that were filtered out, if it is a redirect page title then it has more priority as compared to linktext page titles, but less than that of the target page's title  $ill$ . This is due to the fact that the linked texts are just titles related to the target page title and not the exact translations. After score calculation a threshold was set to filter out the undesired results.
- **Displayer:** It returns the resultset back to User Interface.

## VII. SYSTEM WORKING

First, user inputs the search string, *type*(basic or enhanced), and the input & output languages. The languages are passed to the *Language Code Generator* which generates the code of the languages as used in the Wikipedia. These codes are processed by the *Dname and Tablename Generator*. Output of this along with search string,  $tp$  and  $sp$  is fed into the *Query Generator* which generates SQL queries based on the set templates which are  $Q_1$ ,  $Q_2$ ,  $Q_3$  and  $Q_4$ . The output queries are compared to the set templates and are then sent to the

respective processes namely *Source Page Finder*, *Target Page Finder*, *Redirect Page Finder* and *LinkText Finder*. From these processes queries are passed to Temporary Storage which calls the DBEngine. The returned results are stored in Temporary storage, from where they are transferred to the calling process. Now, these results from above mentioned processes is passed to *Summation*. It generates TCs which act as input to *Patternmatcher*. The output of *Patternmatcher*, which is FTCs is then passed to score calculator which eliminates undesired FTCs using a threshold. Finally, the RS thus obtained is sent to user interface to be displayed based on the choice of type. With the aim to implement the above mentioned system design we have proposed the following methodology.

---

**Algorithm 1** *Fundamental variables to be used in our algorithms*

---

- Require:** query: search string entered by user  
**Require:** queries: An array storing SQL queries  
**Require:** string: Search string entered by user  
**Require:** type : basic or enhanced  
**Require:** input language: Language for input  
**Require:** output language: Language for RS  
**Require:** srclang : input language code  
**Require:** deslang : output language code  
**Require:**  $Q_1$ ,  $Q_2$ ,  $Q_3$  and  $Q_4$ : Predefined SQL queries template  
**Require:** dbname : Name of database on which query is to executed  
**Require:** tablename : Name of table on which query is to executed
- 

---

**Algorithm 2** *PatternMatcher( $ill, RP\{tp\}, LT\{tp\})(n)$*

---

- Require:** An integer  $i = 0$   
**Require:** An integer  $count\_ill = count(ill)$   
**Require:** An integer  $count\_rp = count(RP\{tp\})$   
**Require:** An integer  $count\_lt = count(LT\{tp\})$   
**while**  $i \neq count\_ill$  **do**  
     $j \leftarrow 0$   
     $k \leftarrow 0$   
    **while**  $j \neq count\_rp$  **do**  
        **while**  $k \neq count\_lt$  **do**  
            **if**  $eregi(ill[i], LT\{tp\}[k])$  or  $eregi(RP\{tp\}[j], LT\{tp\}[k])$  **then**  
                 $results[] = LT\{tp\}[k]$   
            **end if**  
            **end while**  
             $k++$   
        **end while**  
         $j++$   
    **end while**  
     $i++$   
**end while**
- 

## VIII. EXPERIMENTS AND RESULTS

We implemented the Bilingual Dictionary extraction algorithm, proposed by us, using PHP (version 5.2.4) and

---

**Algorithm 3** *Language Code Generator*(input language, output language, srclanguage, deslanguage)

---

```

if input language == NULL or output language
==NULL then
 exit
else
 for $i = 0$ to $\text{length}(\text{languages})$ do
 if inputlanguage == languages[i] then
 srclanguage = languages[i][code]
 break
 end if
 end for
 for $i = 0$ to $\text{length}(\text{languages})$ do
 if outputlanguage == languages[i] then
 deslanguage = languages[i][code]
 break
 end if
 end for
end if

```

---



---

**Algorithm 4** *Bilingual Dictionary*(input language, output language, srclanguage, deslanguage)

---

```

LanguageCodeGenerator(inputlanguage, output
language, srclang, deslang)
queries = QueryGenerator(srclang, deslang,
sp, tp, string, dbname, tablename)
for $i = 0$ to $\text{length}(\text{queries})$ do
 if queries[i] LIKE Q_1 then
 $sp \leftarrow \text{SearchpageSearcher}(\text{query})$
 end if
 if queries[i] LIKE Q_2 then
 $tp \leftarrow \text{TargetpageSearcher}(\text{query})$
 end if
 if queries[i] LIKE Q_3 then
 $RP\{tp\} \leftarrow \text{RedirectSearcher}(\text{query})$
 end if
 if queries[i] LIKE Q_4 then
 $LT\{tp\} \leftarrow \text{LinkTextSearcher}(\text{query})$
 end if
end for
 $TCs \leftarrow \text{Summation}(\text{ill}, RP\{tp\}, LT\{tp\})$
 $FTCs \leftarrow \text{Patternmatcher}(\text{ill}, RP\{tp\}, LT\{tp\})$
 $RS \leftarrow \text{Scorecalculator}(FTCs)$
Displayer(RS, type)

```

---

MySQL (5.0.45 community-NT server) on an Intel Celeron 1.7 GHz personal computer with 1GB DDR RAM and MS Windows Xp operating system.

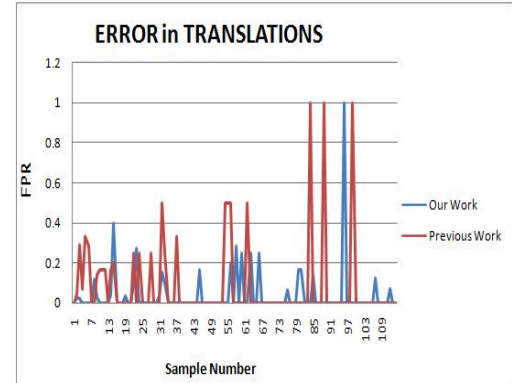
First, we implemented the Wikipedia MySQL databases available at [3]. After that we designed a list L, which consisted of a sample set of 200 words covering domain

specific (egs. methane, fauna, usb, etc.) and frequently used (egs. shirt, industry, etc.) terms from various fields like history, biology, accounts, etc. In order to prove effectiveness of our approach, we compared the resultsets for list L, returned by both methods(ours and [4]). To calculate the number of false positives RS were converted back to English and were checked for the relatedness or closeness to the searched term. Results which were synonyms or were related to searched string have been considered desired cases. The resultsets were compared on the parameters namely: Number of RS and FPR for each word in list L and graphs were plotted. Besides this a graph pertaining to words which were only present in our dictionary was also drawn for the above mentioned parameters. Experimental proof of our observations is given by the accompanying screenshots.

Search String is:India

| Page Title  | Page score     |
|-------------|----------------|
| インド         | 1              |
| インド国        | 0.872583896527 |
| バールト        | 0.895130856602 |
| インド連邦       | 0.978923122844 |
| インド連邦共和国    | 0.944674366588 |
| インド人        | 0.903096968805 |
| インド共和国      | 0.848908743471 |
| イギリス東インド会社  | 0.671200618274 |
| インド (曖昧さ回避) | 0.266633955238 |
| インドの世界遺産    | 0.664510624769 |
| インドの国旗      | 0.428246602592 |
| インドの大統領     | 0.400636432055 |
| インドの数学      | 0.487063788081 |
| インドの歴代首相    | 0.607263241115 |
| インドの歴史      | 0.601823959128 |
| インドの観光地の一覧  | 0.722859421068 |
| インドの言語      | 0.673416590205 |

(a) Our screenshot



(b) Graph 1

Fig. 2. Comparison of FPR

Graph 1 in Figure 2 (b) depicts the FPR for each word. High peaks of graph shows uncorrelation among the resultset and the searched string. It is noticeable that the FPR in our case is less with respect to previous approach. Figure 2 (a) denotes the screenshots for the word india in our dictionary whose one of the translation candidate was wrong in the previous approach [1].

The previous dictionary did not show the results for "eu" but "EU" showed results. But in our approach we



Search String is:Eu

| Page Title  | Page score     |
|-------------|----------------|
| 欧州連合        | 0.96           |
| EU          | 0.924858025737 |
| 欧州連合の外交     | 0.74675944736  |
| 欧州連合の拡大     | 0.380487255233 |
| 欧州連合の政治     | 0.586333028407 |
| 欧州連合の旗      | 0.457226341786 |
| 欧州連合の経済     | 0.364985639739 |
| 欧州連合の経済通貨統合 | 0.346595028837 |
| 欧州連合の言語     | 0.741963621676 |
| 欧州連合加盟国     | 0.686081874302 |
| 欧州連合基本条約    | 0.515305327236 |
| 欧州連合理事会     | 0.294556942787 |

Fig. 3. Example - Case Sensitivity is removed in our approach

got results for both the cases. This is apparent from the accompanying screenshot in Figure 3.

Search String is:Evening

| Page Title | Page score     |
|------------|----------------|
| 夕          | 1              |
| 夕方         | 0.855304140329 |
| 夕方         | 0.573479298354 |
| 夕焼け        | 0.456217881615 |
| 夕食         | 0.404546815299 |

Fig. 4. Zerom - Example of new word extracted

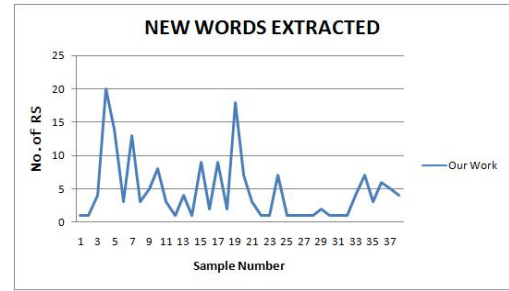
There were a lot of commonly used words which were not present in the dictionary available at [4]. But whose RS existed in ours. These terms are referred to as **Zerom**. As cited in our screenshot in Figure 4 we have results for the word "evening" but the previous work did not incorporate it. Further graphs for the RS of such words and their FPR were plotted in Figure 5.

Graph 4 in Figure 6 shows the number of RS for each word in List L. It is clearly evident from the graph that number of RS is much more in our approach as compared to previous approach.

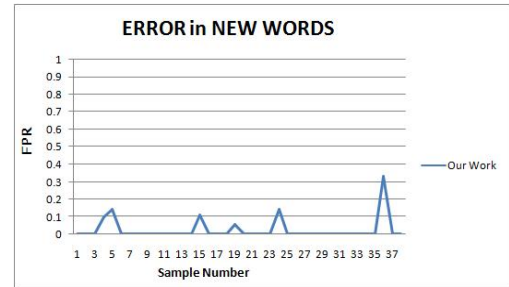
## IX. CONCLUSION AND FUTURE WORK

In this paper we have presented an approach for extracting bilingual terminology from Wikipedia. We have proposed the use of page parameters to provide a relativity between the searched string and the translation candidates. Our method is useful for specialized domain-specific terms because accuracy and coverage are better than that of the bilingual text corpus approach. We believe that Wikipedia will become more comprehensive in near future which will result in a better coverage.

We are planning to include Indian languages which are not present in the online dictionaries. It is also promising



(a) Graph2



(b) Graph3

Fig. 5. Zerom

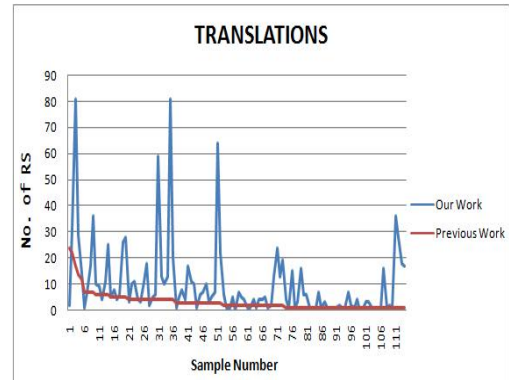


Fig. 6. Graph 4

to analyze and find some more parameters to enhance the dictionary. Moreover, a Web Bot may be implemented to automatically update the databases from the website of Wikipedia.

## REFERENCES

- [1] M. Erdmann, K. Nakayama, T. Hara, S. Nishio, "An Approach for Extracting Bilingual Terminology from Wikipedia", *Conference on Database Systems for Advanced Applications, DASFAA*, DASFAA 2008, LNCS, vol. 4947, pp. 380-392, Springer, March, 2008, New Delhi.
- [2] James Breen, "Jmdict: a Japanese-multilingual dictionary", *COLING 2004 Multilingual Linguistic Resources*, August 2004, pp. 65-72, COLING, 2004, Geneva, Switzerland.
- [3] "Wikimedia Foundation: Wikimedia downloads", <http://download.wikimedia.org/>.
- [4] "Dictionary proposed by Nakayama et al", <http://wikipedia-lab.org:8080/WikipediaBilingualDictionary>.

# Comparing FLV and MPEG4 (H.264/AVC) Multimedia File Format with Wireless Network Parameters

<sup>1</sup>Mehedi Mahmud, <sup>1</sup>Sadia Monir, <sup>1</sup>Tanjila Khan, <sup>2</sup>Mohammad Abdur Rahim Mustafa and <sup>3</sup>Hasan Sarwar

<sup>1</sup>Department of CSE, MIST, Dhaka, Bangladesh, <sup>2</sup>Department of CSE, BUET, Dhaka, Bangladesh, <sup>3</sup>Department of CSE, UIU, Dhaka, Bangladesh.

**Abstract** — Accommodating multimedia resources in the wireless media is a long cherished service. In this paper, we compared the FLV and MPEG4 (H.264/AVC) format of multimedia content using frame rates, sizes and byte rates. Thus we showed that MPEG4 performs better than FLV.

**Index Terms** — FLV, MPEG4, CoQV, MBMS, pixel, macroblock, screenshare.

## I. INTRODUCTION

As wireless communication has lower bandwidth support, many researchers have been trying to transmit multimedia resources using different file format such as MPEG4 part2, MPEG4 part10 (H.264/AVC), 3GP, FLV (V6) etc. So we compared MPEG4 and FLV video file format as to ensure this purpose.

The intention of using H.264/AVC for the sake of its capability to provide good video quality at substantially lower bit rates than previous standards (e.g. half or less the bit rate of MPEG2, MPEG4 part2) without increasing the complexity of design so much that it would be impractical or excessively expensive to implement.

Flash Video is used because it is a file format used to deliver video over the low bandwidth as it is one of the upcoming features in short times.

In order to fulfill our purpose we used different tools to parse video format and measured the variations.

## II. RELATED WORKS

The frame rate of MPEG4 has been examined in [1]; video standard, video sequences and the comparison have been measured to generate measure the coefficient of quality variance (CoQV). The lower the variance, the quality is better.

Prevailing studies of MPEG4 part 2, H.264/AVC (MPEG4 part 10), MPEG4/SVC codec [1], [2] generally illustrates the rate distortion performance, video quality function (PSNR) and CoQV. The study of the bit rate variability or frame rate variability as a function of video quality demonstrates bit-rate-variability-distortion (VD) curve.

Comparison based studies like [3] expresses the comparison of actual average numbers of texture bits per frame with corresponding models, comparison of actual variances of numbers of texture bits per frame with corresponding models, comparison of actual average numbers of motion bits per frame with corresponding models etc.

It is quite obvious that the gradual migration of transcoding from MPEG2 to H.264 is at the expense of higher computational complexity due to its improved inter-and intra-prediction modes. So a novel macroblock (MB) mode decision algorithm for P-frame prediction based on machine learning techniques [4] and based on data mining techniques [10] to be used as part of a very low complexity MPEG-2 to H.264 video transcoder.

Implementing MPEG2 to H.264 transcoding schemes in the pixel domain introduces a high degree of computational complexity. In order to solve this algorithms and equations are proposed that compensate for the luminance half-pixel and chrominance quarter/three-quarter pixel interpolation errors are derived. Thus, the compensation scheme includes a closed-loop H.264 motion compensation process, which is implemented in the pixel domain and our scheme is compared with two different cascaded pixel-domain transcoding structures [5]. The experimental results show that the proposed compensation algorithms achieve 5-dB quality improvement over the open-loop transform-domain-based transcoding. Again in order to motion re-estimation in video transcoding and to improve the searching speed a new transcoding scheme that uses the MPEG2 DCT coefficients [8] to predict the block size partitioning for H.264 is introduced.

H.264 has expand its arena in Medical video compression also and in this process it improves the H.264 rate control algorithm [6] with better perceptual quality. H.264 can achieve a significant average peak signal-to-noise ratio (PSNR) gain of up to 4.35 dB for the test medical video sequences, and thus is much more effective when applied in medical video compression.

The Achilles' heel of video over wireless services continues to be the limited bandwidth of wireless

connections and short battery life of end-user handheld devices such as cell-phones and PDAs. In [7], strategies are proposed to reduce the complexity of various H.264 operations. Using the knowledge of the context of the scenes in the video sequences, unimportant regions in the frames are isolated and unnecessary processing is avoided.

On the basis of architectural features of architectural features of TMS320DM642 and the computational complexity analysis of H.264 encoder [11], for the video sequences with CIF format, the optimized H.264 encoder can achieve the encoding speed of more than 24 frames per second, which can meet the real-time requirements of the applications. Now rate control algorithm is developed [12] for the H.264 baseline-profile encoder under the constant bit rate constraint, to improve performance.

Recently, three Rate Distortion (RD) optimization algorithms (a graph-based algorithm) are put forward [13] for near optimal Soft Decision Quantization (SDQ) in H.264 baseline encoding given motion estimation and quantization step sizes, an algorithm for near optimal residual coding in H.264 baseline encoding given motion estimation, and an iterative overall algorithm to optimize H.264 baseline encoding for each individual frame given previously coded reference frames-with them embedded in the indicated order.

Research on the statistical analysis of video traffic and video traffic models has also received significant interest [1]. The existing works generally focuses on the statistical report of traffic and bit rate variability as a function of quality level or quantization scale.

In contrast, in this paper we comprehensively study and compare the frame size [byte] per 0.8 second and the byte rate of MPEG 4 with FLV (Flash Video) and in addition we find out the variations plotting those points to decide which on is of better quality between these two.

### III. BASIC DEFINITIONS

#### A. H.264/AVC

H.264 (MPEG 4 part 10) is a latest video coding standard of the ITU-T Video Coding Expert Group (VCEG) and the ISO/IEC Moving Picture Expert Group (MPEG). H.264/MPEG4-AVC has recently become the most widely accepted video coding standard and it covers all common video conferencing to IPTV, HDTV and HD video storage.

H.264 has fringe benefits [9] and H.264/AVC's intra-frame coder is so efficient that we can make it operate as a still image coder that outperforms the existing state-of-the-art still image compression standard (JPEG-2000). MB coding mode decisions in H.264 video have a correlation with the distribution of the motion compensated residual in the decoded video. Data mining

tools can be used to exploit the correlation and derive decision trees [10] to classify the incoming decoded MBs into one of the several coding modes in H.264.

Keeping full compatibility with the H.264 standard syntax and decoding behavior in case of single description reception based on the concept of redundant slices [14], An analytical setup is employed in order to optimally select the amount of redundancy to be inserted in each frame, taking into account both the transmission condition and the video decoder error propagation.

#### B. Flash Video (FLV V6)

Flash Video files contain video bit streams which are a variant of the H.263 video standard, under the name of Sorenson Spark. The Flash Video file format supports two versions of a so called 'screenshare' codec which is an encoding format designed for screencasts. These formats are bitmap tile based can be lossy by reducing color depths and are compressed using zlib.

#### C. MBMS

Multimedia Broadcast/Multicast Service (MBMS), as specified in Release 6 of 3GPP For such mobile TV services, improved video compression [17] performance, in conjunction with appropriate mechanisms for error robustness, is key — a fact that is well reflected by the use of H.264/MPEG4-AVC together with forward error correction schemes in all of those mobile-broadcasting systems.

#### D. Coefficient of Quality Variance (CoQV)

The coefficient of variation of frame sizes [unit free] is defined as

$$\text{CoQV} = \frac{S_x}{\bar{X}}$$

Where  $S_x$  is the root of variance and  $\bar{X}$  is the mean frame size. CoQV is widely employed as a measure of the variability of the frame sizes, i.e., the bit rate variability of the encoded video. Plotting the CoQV as a function of the quantization scale (equivalently, the PSNR video quality) gives the rate variability-distortion (VD) curve [1].

#### E. Mode Decision Algorithm

The algorithm of mode decision for inter-frame coding is introduced and the complexity of the H.264 encoder is analyzed from the side of transform and quantization. In order to speed up the H.264 encoder without causing a

great decrease in encoding efficiency, a fast intra-mode decision method for inter-frame coding is also introduced. The experimental results show that the inter-frame coding with fast intra mode decision can reduce the computational complexity of the H.264 encoder remarkably while incurring little, if any, loss in quality [18].

#### IV. DATA EXTRACTION AND ANALYSIS

##### A. Parsing MPEG4 (H.264)

The comparative study mainly focuses on the frame size per second between MPEG4 and FLV. In order to do so, we use several tools such as MPEG4 Parser, FLV Converter, Xilisoft video converter, flvmdigui (XML generator for FLV). The metadata of MPEG4 is generated using the tool “MPEG4 PARSER” developed by the video trace research group at Arizona State University. This material is based upon work supported by the National Science Foundation, USA. As we have compared on the basis of frame size per second we generated frame size and frame type using MPEG4 Parser. We used several video files of MPEG format and we parsed these MPEG format video files into CSV file.

A lot of values of Frame size [byte] and frame type [P type or I type] have been generated. Then those are parsed to calculate the linearity and byte rate. Total 266 points of byte rate have been used to calculate the average.

##### B. Parsing FLV

We converted the video into FLV format using FLV converter and by the help of FLV parser (flvmdigui) we generated the same data set for FLV using FLV parser and FLV parser generates the accumulated value for frame size in every 0.8 second in XML. Then we parsed to calculate the byte rate. Total 302 points of byte rate have been used to calculate the average.

##### C. Environment

For this comparison we worked on HP laptop, core2duo processor (centrio), 2GB RAM, Windows Vista. For plotting points we used MatLab7.1.

#### V. EXPERIMENTAL RESULT

Using MPEG4 Parser we generated the following frame type and frame size:

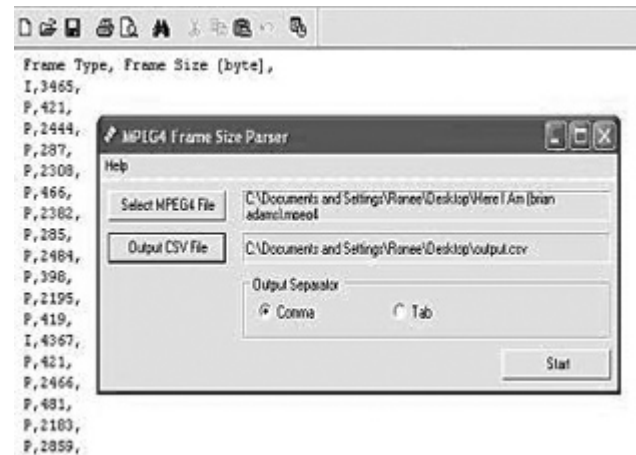


Fig. 1. Parsing MPEG4.

Using parser we generated the following frame size per 0.8 second:

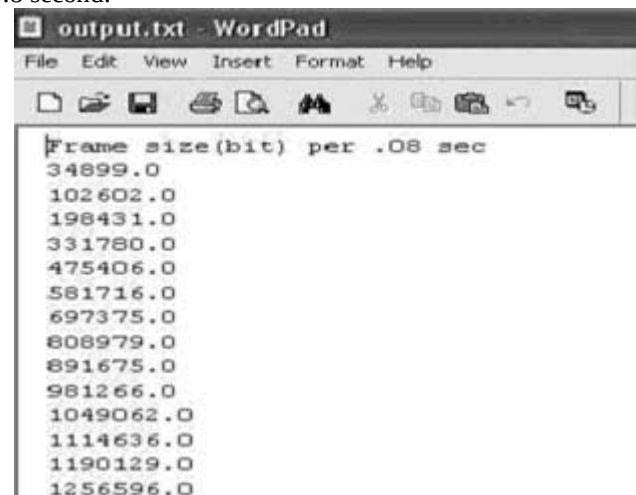


Fig. 2. Snapshot of the output data of frame size [byte] per 0.8 second

Using flvmdigui following XML data has been generated:

```
<?xml version="1.0" encoding="UTF-8" ?>
<root>
 <duration>252.472</duration>
 <width>352</width>
 <height>288</height>
 <videodatarate>0</videodatarate>
 <framerate>15</framerate>
 <videocodecid>2</videocodecid>
 <audiosamplerate>22050</audiosamplerate>
 <audiosamplesize>16</audiosamplesize>
 <stereo>true</stereo>
 <audiocodecid>2</audiocodecid>
 <filesize>8190320</filesize>
 <lasttimestamp>252.447</lasttimestamp>
 <lastkeyframetimestamp>252.4</lastkeyframetimestamp>
 <audiodelay>0</audiodelay>
 <canSeekToEnd>true</canSeekToEnd>
 <audiodatarate>0</audiodatarate>
 <creationdate>unknown</creationdate>
```

Fig. 3. Snapshot of root of XML data for FLV

```

- <keyframes>
- <filepositions>
 <value1>6306</value1>
 <value2>27137</value2>
 <value3>49998</value3>
 <value4>74262</value4>
 <value5>98718</value5>
 <value6>123906</value6>
 <value7>145314</value7>
 <value8>166795</value8>
 <value9>187850</value9>
 <value10>212679</value10>
 <value11>233019</value11>
 <value12>252791</value12>
 <value13>272036</value13>
 <value14>294768</value14>
 <value15>312783</value15>
 <value16>330871</value16>
 <value17>348496</value17>
 <value18>366862</value18>
 <value19>385411</value19>
 <value20>403694</value20>

```

Fig. 4. Snapshot of frame size of XML data for FLV

Fig 1 shows the MPEG4 parser and generated CSV file. Fig 2 depicts the incremental output data for frame size [byte] at every 0.8 second. Fig 3 shows the XML root data that demonstrates different data-type for FLV file format. Fig 4 is the continuation of Fig 3 and represents the frame size per 0.8 second.

Graphical representation of the experimental result of average byte rate for both MPEG4 and FLV format is depicted below:

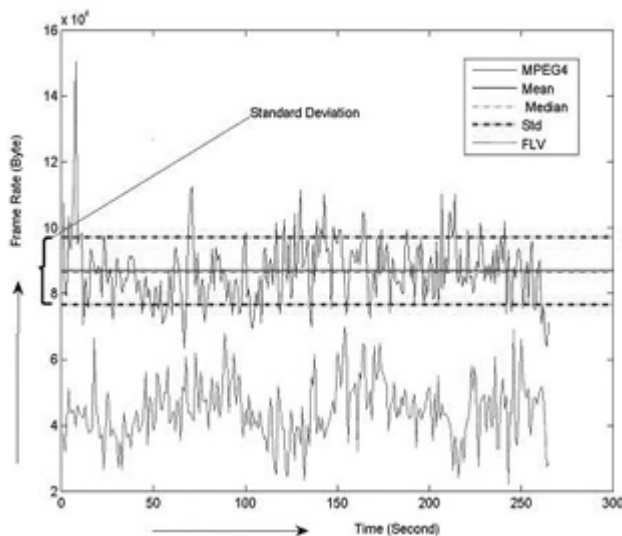


Fig. 5. Comparison of FLV and MPEG4 data file with respect to byte rate and variations showing mean, median and standard deviation of MPEG4.

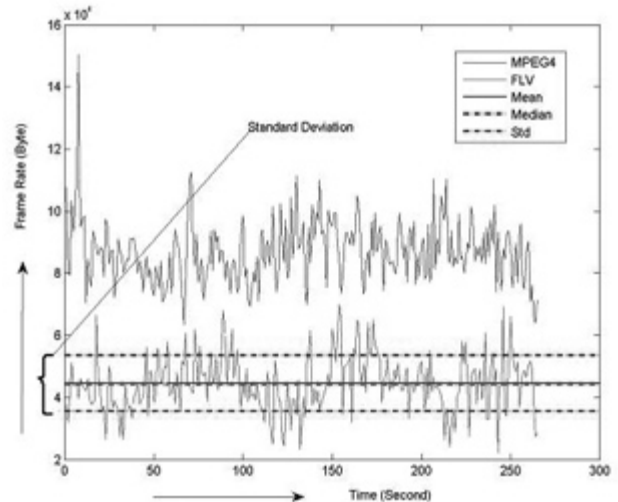


Fig. 6. Comparison of FLV and MPEG4 data file with respect to byte rate and variations showing mean, median and standard deviation of FLV.

Fig 5 & 6 shows the experimental result of FLV and MPEG4 format. In Fig 5 the X axis represents the time and the Y axis represents the frame rate. It is quite evident that FLV has better byte rate.

From Figure 5 we found, the MPEG4 has mean 86900 byte, median 86310 byte and standard deviation 10190 byte. We also found the variation and fluctuation of MPEG4 is high and MPEG4 has lack of constantability.

From Figure 6 we found, the FLV has mean 44540 byte, median 44320 byte and standard deviation 9067 byte. We also found the variation and fluctuation of FLV is lower than MPEG4 and FLV is more stable than MPEG4.

## VI. CONCLUSION

In this paper we examined in detail the variations that occur due to variable frame size per second. From the study, the characteristics that we can convey are: Firstly, FLV generates lower variations than MPEG4 format. Secondly, FLV has better linearity. Thirdly, from the comparison the spikes and fluctuation of frame rate of MPEG4 is more than FLV. Finally, FLV is more competent introducing it in wireless multimedia features.

## ACKNOWLEDGEMENT

I would like to take this opportunity to thank all people who have been instrumental in the preparation of the work. First of all, appreciation goes to Mohammad Abdur Rahim Mustafa for his in depth guidance and support during the course of this thesis. I am extremely grateful to him for providing assistance in literature survey and

suggesting improvements to bring the thesis to its final form. I would also like to thank Dr. Hasan Sarwar for helping me in many helpful technical discussions. Finally I would thank MIST CSE department for giving me facilities to carry on this work.

## REFERENCES

- [1] Geert Van der Auwera, Prasanth T. David, Martin Reisslein, and Lina J. Karam, "Traffic and Quality Characterization of Scalable Video Streams Encoded with the H.264/AVC Scalable Video Coding Extension", May 2008.
- [2] Geert Van der Auwera and Martin Reisslein, "Implications of Smoothing on Statistical Multiplexing of H.264/AVC and SVC Video Streams", May 2008.
- [3] Geert Van der Auwera, Martin Reisslein, and Lina J. Karam, "Video Texture and Motion Based Modeling of Rate Variability-Distortion (VD) Curves", IEEE Transactions on Broadcasting, Vol. 53, No. 3, pp 637-648, September 2007.
- [4] Fernandez-Escribano, G. Kalva, H. Cuenca, P. Orozco-Barbosa, L. Garrido, A. Univ. de Castilla-La Mancha, Albacete., "A Fast MB Mode Decision Algorithm for MPEG-2 to H.264 P-Frame Transcoding", "Circuits and Systems for Video Technology", IEEE Transactions on Publication Date: Feb. 2008, Volume: 18, Issue: 2, pp: 172 – 185.
- [5] Qiang Tang Nasiopoulos, P. Ward, R.K. Univ. of British Columbia, Vancouver, "Compensation of Requantization and Interpolation Errors in MPEG-2 to H.264 Transcoding", "Circuits and Systems for Video Technology", IEEE Transactions on Publication Date: March 2008, Volume: 18, Issue: 3, on page(s): 314 – 325, pp: 314 – 325.
- [6] Hongtao Yu Zhiping Lin Feng Pan Sch. of Electr. & Electron. Eng., Nanyang Technol. Univ., Singapore, "Applications and improvement of H.264 in medical video compression", Circuits and Systems I: Regular Papers, IEEE Transactions on Publication Date: Dec. 2005, Volume: 52, Issue: 12, on page(s): 2707 – 2716.
- [7] Sahafi, L. Randhawa, T.S. Hardy, R.H.S. Sch. of Eng. Sci., Simon Fraser Univ., Burnaby, BC, Canada, "Context-based complexity reduction of H.264 in video over wireless applications", Multimedia Signal Processing, 2004 IEEE 6th Workshop on , Publication Date: 29 Sept.-1 Oct. 2004, On page(s): 23 - 26 .
- [8] Qiang Tang Nasiopoulos, P. Ward, R. Univ. of British Columbia, Vancouver, BC, "Compensation of Requantization and Interpolation Errors in MPEG-2 to H.264 Transcoding", Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on Publication Date: March 31 2008-April 4 2008, On page(s): 1029 - 1032 , Location: Las Vegas, NV.
- [9] de Queiroz, R.L. Ortis, R.S. Zaghetto, A.Fonseca, T.A, Univ. de Brasilia, Brasilia, "Fringe benefits of the H.264/AVC", Telecommunications Symposium, 2006 International, Publication Date: 3-6 Sept. 2006, on page(s): 166 – 170, Location: Fortaleza, Ceara.
- [10] Fernandez-Escribano, G. Bialkowski, J.Gamez, J.A. Kalva, H. Cuenca, P. Orozco-Barbosa, L. Kaup, A. Univ.de Castilla-La Mancha, Albacete. Low-Complexity Heterogeneous Video Transcoding Using Data Mining, Multimedia, IEEE Transactions on Publication Date: Feb. 2008, Volume: 10 , Issue: 2, On page(s): 286 – 299.
- [11] Li Zhuo Qiang Wang Feng, D.D. Lansun Shen, Univ. of Sydney, Sydney, "Optimization and Implementation of H.264 Encoder on DSP Platform", Multimedia and Expo, 2007 IEEE International Conference on Publication Date: 2-5 July 2007, on page(s): 232 – 235.
- [12] Kwon, D.-K. Shen, M.-Y. Kuo, C.-C. J. Univ. of Southern California, Los Angeles, "Rate Control for H.264 Video With Enhanced Rate and Distortion Models", CA. Circuits and Systems for Video Technology, IEEE Transactions on Publication Date: May 2007, Volume: 17 , Issue: 5 , On page(s): 517 – 529.
- [13] En-Hui Yang Xiang Yu, Univ. of Waterloo, Waterloo, "Rate Distortion Optimization for H.264 Interframe Coding: A General Framework and Algorithms", Image Processing, IEEE Transactions on Publication Date: July 2007, Volume: 16, Issue: 7 On page(s): 1774 – 1784.
- [14] Tillo, T. Grangetto, M. Olmo, G. Politecnico di Torino, Torino, "Redundant Slice Optimal Allocation for H.264 Multiple Description Coding", Circuits and Systems for Video Technology, IEEE Transactions on Publication Date: Jan. 2008, Volume: 18 , Issue: 1, On page(s): 59 – 70.
- [15] Geert Van der Auwera, Prasanth T. David, Martin Reisslein, "Video Traffic Analysis of H.264/AVC and Extensions: Single-Layer Statistics", July 2007.
- [16] Geert Van der Auwera, Martin Reisslein, and Lina J. Karam, "Video Texture and Motion Based Modeling of Rate Variability-Distortion (VD) Curves of I, P, and B Frames", IEEE International Conference on Multimedia & Expo (ICME), pp 1405-1408, Toronto, Canada, July 2006
- [17] Detlev Marpe and Thomas Wiegand, Heinrich Hertz Institute (HHI), Gary J. Sullivan, "The H.264/MPEG4 Advanced Video Coding Standard and its Applications", Standard Report by Microsoft Corporation 2007.
- [18] Yun Cheng Zhiying Wang Jianjun Guo Kui Dai Dept. of Comput. Sci., Coll. of Hunan Humanities, Sci. & Technol., Loudi, China, "Research on intra modes for inter-frame coding in H.264", Computer Supported Cooperative Work in Design, 2005. Proceedings of the Ninth International Conference on Publication Date: 24-26 May 2005, Volume: 2, On page(s): 740 - 744 Vol.2.
- [19] "Computer Graphics– Principals & Practice" (2nd edition) by Foley, VanDam, Feiner, Hughes.
- [20] "Computer Graphics - Using Open GL" (2nd edition) by F. S. Hill, Jr.

# Direction of Arrival Algorithms for Adaptive Beamforming in Next Generation Wireless Systems

Mohammad Ismat Kadir<sup>1</sup>, Md. Shadiul Hoque<sup>2</sup>, Saiful Islam<sup>3</sup>

Electronics and Communication Engineering Discipline

Khulna University, Khulna – 920800, Bangladesh

E-mail: ismat\_kadir@yahoo.com<sup>1</sup>, sadiul\_ece04@yahoo.com<sup>2</sup>, sunny\_khulna@yahoo.com<sup>3</sup>

**Abstract** - Different beamforming algorithms like Side-lobe Cancellers, Linearly Constrained Minimum Variance (LCMV), Least Mean Squares (LMS), Recursive LMS, and Direction of Arrival (DOA) exist in literature. Among the Direction of Arrival (DOA) algorithms, MUSIC and ESPRIT play the most important role. These two algorithms were implemented and their performances were compared. The algorithms were simulated for different signal levels and the DOAs were computed for use in next generation wireless. ESPRIT was found to be a better DOA technique for uncorrelated source used in beamforming.

**Keywords** - Antenna array, Beamforming, DOA estimation, Eigenvalues, ESPRIT algorithm, MUSIC algorithm.

## 1. INTRODUCTION

The performance of next generation wireless can be greatly improved by using adaptive beamforming algorithms [3], [6].

Beamforming can meet the challenge of increasing spectral efficiency and improving wireless communication system performance by significantly increasing the reception and transmission ranges and reducing the probability of interception of secure transmission.

Adaptive Beamforming is a technique in which an array of antennas is exploited to achieve maximum reception in a specified direction by estimating the signal arrival from a desired direction (in the presence of noise) while signals of the same frequency from other directions are rejected. This is achieved by varying the weights of each of the sensors (antennas) used in the array.

## 2. BEAMFORMING BACKGROUND

Beamforming is a form of spatial filtering used to distinguish the spatial properties between a SOI and the noise and the interfering signals. Beamforming principles apply to both the transmission and reception of signals.

Beamforming is accomplished through the use of an array of sensors such as antenna, hydrophones and so on. In order to proceed with the discussion of beamforming, it is important to note some basic assumptions. First, a signal originating far away from the sensor array can be modeled as a plane wave. Next the signal received by each sensor element is a time-delayed (phase shift) version of the signal

received by other sensor elements. Finally, an N-element beamforming system is capable of forming up to N beams.

For the beamformer to steer the radiation in a particular direction and to place the nulls in the interfering directions the direction of arrival has to be known beforehand. The Direction of arrival algorithms does exactly the same; they work on the signal received at the output of the array and compute the direction of arrivals of all the incoming signals. Once the angle information is known it is fed into the beamforming network to compute the complex weight vectors required for beam steering.

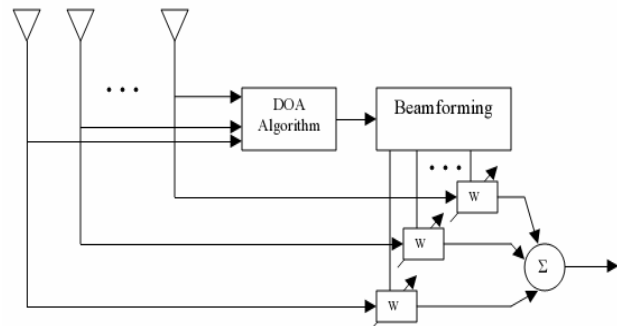


Fig. 1: Beamforming setup with DOA Estimation

## 3. DIRECTION OF ARRIVAL

### A. Signal Model

Let a uniform linear array be composed of  $N$  sensors, and let it receive  $M$  narrow band source signals  $S_m(t)$  from , as shown desired users arriving at directions  $\theta_1, \theta_2, \dots, \theta_M$  as shown in Figure-2. The array also receives  $I$  narrow band source signals  $s_i(t)$  from undesired (or interference) users arriving at directions  $\theta_1, \theta_2, \dots, \theta_I$ . At a particular instant of time  $t = 1, 2, \dots, K$ , where  $K$  is the total number of snapshots taken. The desired users signal vector  $x_s(t)$  can be defined as

$$x_M(t) = \sum_{m=1}^M a(\theta_m) s_m(t) \quad (1)$$

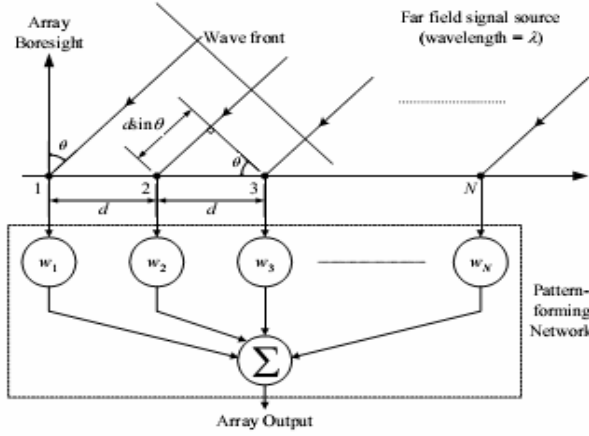


Fig. 2: Geometry of a uniform linear array.

where  $a(\theta_m)$  is the  $N \times 1$  array steering vector which represents the array response at direction  $\theta_m$  and is given by

$$a(\theta_m) = [\exp\{j(n-1)\psi_m\}]^T : 1 \leq n \leq N \quad (2)$$

where  $[\cdot]^T$  is the transposition operator, and  $\Psi_m$  represents the electrical phase shift from element to element along the array. This can be defined by

$$\Psi_m = 2\pi \left( \frac{d}{\lambda} \right) \sin(\theta_m) \quad (3)$$

where  $d$  is the inter-element spacing and  $\lambda$  is the wavelength of the received signal. The desired users signal vector  $x_M(t)$  of (1) can be written as

$$x_M(t) = A_M s(t) \quad (4)$$

where  $A_M$  is the  $N \times M$  matrix of the desired users signal direction vectors and is given by

$$A_M = [a(\theta_1), a(\theta_2), \dots, a(\theta_M)] \quad (5)$$

and  $s(t)$  is the  $M \times 1$  desired users source waveform vector defined as

$$s(t) = [s_1(t), s_2(t), \dots, s_M(t)]^T \quad (6)$$

We also define the undesired (or interference) users signal vector  $x_I(t)$  as

$$x_I(t) = A_I i(t) \quad (7)$$

where  $A_I$  is the  $N \times I$  matrix of the undesired users signal direction vectors and is given by

$$A_I = [a(\theta_1), a(\theta_2), \dots, a(\theta_I)] \quad (8)$$

and  $i(t)$  is the  $I \times 1$  undesired (or interference) users source wave form vector defined as

$$i(t) = [i_1(t), i_2(t), \dots, i_I(t)]^T \quad (9)$$

The overall received signal vector  $x_M(t)$  is given by the superposition of the desired users signal vector  $x_M(t)$ , undesired (or interference) users signal vector  $x_I(t)$ , and an  $N \times 1$  vector  $n(t)$  which represents white sensor noise. Hence,  $x(t)$  can be written as

$$x(t) = x_M(t) + n(t) + x_I(t) \quad (10)$$

where  $n(t)$  represents white Gaussian noise. The conventional (forward-only) estimate of the covariance matrix defined as

$$R = E\{x(t)x^H(t)\} \quad (11)$$

Where  $E\{\cdot\}$  represents the ensemble average; and  $(\cdot)^H$  is the Hermitian transposition operator. Equation (11) can be approximated by applying temporal averaging over  $K$  snapshots (or samples) taken from the signals incident on the sensor array. This averaging process leads to forming a spatial correlation (or covariance) matrix  $R$  given by [8]:

$$R = \frac{1}{K} \sum_{k=1}^K x(k)x^H(k) \quad (12)$$

Substituting for  $x(t)$  from (10) in (12) yields

$$R = A_M R_{ss} A_M^H + n(k)n(k)^H + A_I R_{ii} A_I^H \quad (13)$$

where  $R_{ss} = E\{s(t)s^H(t)\}$  is an  $M \times M$  desired users source waveform covariance matrix;  $R_{ii} = E\{i(t)i^H(t)\}$  is an  $I \times I$  undesired users source waveform covariance matrix.



### B. Algorithm for MUSIC estimation

Consider an  $N$ -element linear array that detects  $M$  signals impinging on it whose directions of arrival need to be known. From the previous discussion we know that the received signals at the output of the array have the following form.

$$x(t) = A(\theta)s(t) + n(t)$$

Or in matrix notation it can be represented as,

$$X = AS + n \quad (14)$$

Where  $S$  is the signal vector,  $A$  is the array propagation vector and  $n$  is the noise vector with zero mean and variance.

The covariance matrix is given by,

$$\begin{aligned} R_x &= E\{XX^*\} = AE\{SS^*\}A^* + E\{nn^*\} \\ &= AP^*A + \sigma^2 I \end{aligned} \quad (15)$$

Where  $P = E\{SS^*\}$

$$AP^*A = R_x - \sigma^2 I$$

When the number of signals  $M$  is less than  $N$  then  $AP^*A$  is singular and has a rank less than  $N$ . The eigenvalues of can be found by,

$$|AP^*A| = |R_x - \sigma^2 I| = 0$$

The eigenvectors of  $R_x$  must satisfy

$$R_x e_i = \sigma^2 I e_i \quad (16)$$

$$(R_x - \sigma^2 I) e_i = 0 \quad (17)$$

where  $e$  is the eigenvector and  $i$  varies from 1 to  $N-M$ .

Let the noise eigenvector be defined as  $E_N$  such that,

$$(R_x - \sigma^2 I) E_N = 0 \quad (18)$$

Or,

$$AP^*A E_N = 0 \quad (19)$$

Based upon this approach, the pseudospectrum  $P(\theta)$  is given by

$$P(\theta) = \frac{1}{A(\theta)^* E_N E_N^* A(\theta)} \quad (20)$$

When the pseudospectrum  $P(\theta)$  is plotted, peaks appear at the angles of arrival of the incident signals.

### C. Algorithm for ESPRIT estimation

ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique) is one of the most efficient and robust methods for DOA estimation. It uses two arrays in the sense that the second element of each pair is displaced by the same distance in the same direction relative to the first element. It is not required to have two separate arrays but can be realized using a single array by being able to select a subset of elements.

Let the array signals received by the two arrays be denoted by  $x(t)$  and  $y(t)$  such that

$$x(t) = As(t) + n_x(t) \quad (21)$$

$$y(t) = A\theta s(t) + n_y(t) \quad (22)$$

$A$  is a  $K \times M$  matrix; where  $M$  is the number of steering vectors produced by  $N$  elements of the array.  $n_x(t)$  and  $n_y(t)$  denotes the noise induced at the elements of the two arrays. Now, by using the available methods, the numbers of directional sources  $M$ , are estimated based on principles such as Akaike's information criterion (AIC) and Minimum description length (MDL). Two matrices  $U_x$  and  $U_y$  are formed which denote the  $M$  eigenvectors corresponding to the largest eigenvalues of the two array correlation matrices  $R_{xx}$  &  $R_{yy}$  (Array correlation matrices). The eigenvectors of the following  $2M$  by  $2M$  matrix are obtained and are denoted by

$$V = \begin{bmatrix} H & H \\ U_x & U_y \end{bmatrix} \begin{bmatrix} U_x & U_y \end{bmatrix} \quad (23)$$

Once the eigenvector  $V$  is obtained its eigenvalues  $\lambda_m, m = 1, \dots, M$  can be computed.

Now the DOA is given by

$$\theta_m = \cos^{-1} \left\{ \frac{\text{Arg}(\lambda_m)}{2\pi \Delta_0} \right\}, m = 1, \dots, M \quad (24)$$

$\Delta_0$  is the element separation in terms of wavelength. Other variations of ESPRIT include beam-space ESPRIT, resolution-enhanced ESPRIT multiple invariance ESPRIT and higher order ESPRIT.

#### 4. IMPLEMENTING MUSIC AND ESPRIT

The above algorithms were implemented and the directions of arrival were computed. It is obvious from the analysis that MUSIC gives the pseudospectrum from which we can compute the direction of arrival as the pseudospectrum peaks. But in case of ESPRIT, we can get the direct polar plot.

Two sets of amplitudes have been used to be the signals at the elements of a linear array. Using the algorithms, the DOAs are calculated. The arrays are simulated to have three elements. There can be the provision for variable number of elements too. The distance between the elements is  $\lambda/2$ . The amplitudes of three element signals for first case are considered to be [1 2 3] and for second case [1.5 2.5 3.5] in voltage units.

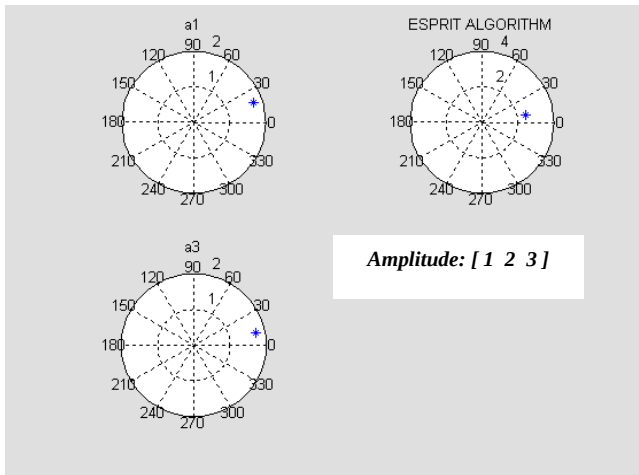


Fig. 3: ESPRIT for first signal

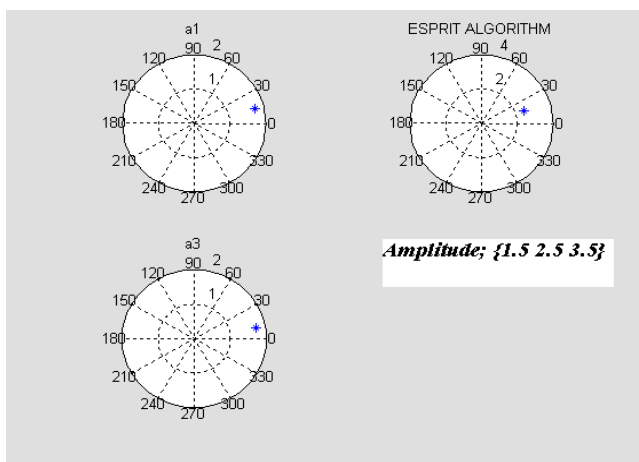


Fig. 4: ESPRIT for second signal

The directions of arrival for different element signals are shown for the first amplitude set in Fig. 5 and those for the second set in Fig. 6. The pseudospectrum  $P(\theta)$  s for the two cases are plotted in Fig.3 and Fig.4 respectively, from where the spectrum peaks can be taken to be the DOAs.

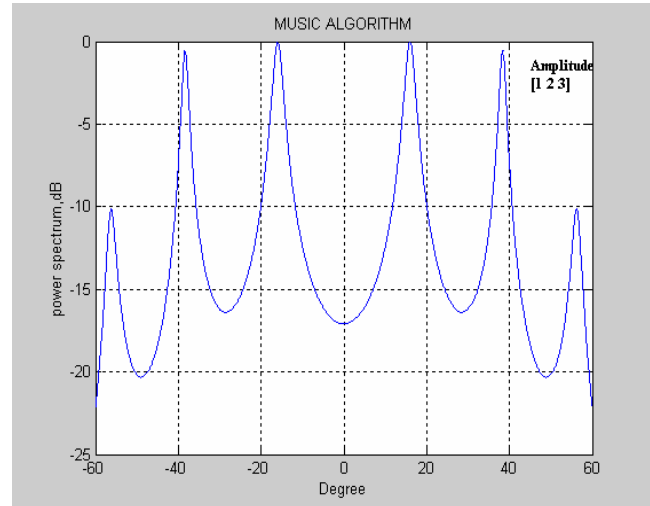


Fig. 5: MUSIC for first signal

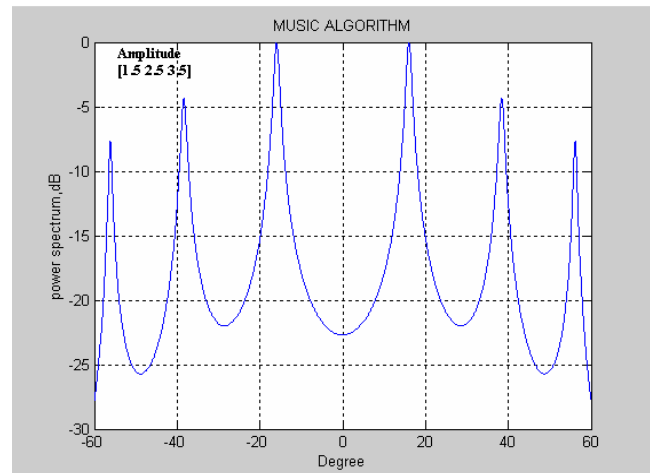


Fig. 6: MUSIC for second signal

Table 1 summarizes the results found in the corresponding experimentations. We can have a comparative view of the two algorithms for beamforming from this result.

Amplitude	MUSIC			ESPRIT		
[1 2 3]	17 <sup>0</sup>	38 <sup>0</sup>	57 <sup>0</sup>	20 <sup>0</sup>	10 <sup>0</sup>	13 <sup>0</sup>
[1.5 2.5 3.5]	15 <sup>0</sup>	38 <sup>0</sup>	55 <sup>0</sup>	15 <sup>0</sup>	19 <sup>0</sup>	10 <sup>0</sup>

Table 1: Comparative DOA

## 5. CONCLUSION

For first signal in MUSIC Estimation we get the peak of power spectrum at 17, 38, 57 degrees and for second signal at 15, 38, 55 degrees respectively. This gives an idea about the process of DOA estimation. The element angles exhibit greater variation, although in case of signal amplitude variation, the angles of peak power spectrum does not differ that much. In case of ESPRIT estimation we get 20, 10, 13 degrees and 15, 19, 10 degrees for the first and the second signals for the three elements respectively. No doubt, the variations are not that much pronounced and the estimated angles for different elements remains to be almost same.

So, the ESPRIT DOA algorithm can be treated as a more robust and faster estimation technique as compared to MUSIC. The computation is also less complex comparatively. However, ESPRIT has the disadvantage of not being capable of handling correlated sources. Finally, this accurate direction will help to obtain better beamforming which is very essential for next generation wireless systems.

## REFERENCES

- [1]. E. M. Al Ardi, R. M. Shubair and M. E. Al Mualla, "Computationally Efficient DOA Estimation in a Multipath Environment," IEEE Electronics Letters, Volume 40, Issue 14, July 2004, Pages 908-909.
- [2]. E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Direction of Arrival Estimation in a Multipath Environment: An Overview and a New Contribution," Applied Computational Electromagnetics Society Journal: Special Issue on Phased and Adaptive Array Antennas, Volume 21, Issue 3, November 2006, Pages 226-239.
- [3]. [http://etd.lob.fsu.edu/thesis/available/etd-04092004-143712/unrestricted/ch\\_5adaptivbeamforming.pdf](http://etd.lob.fsu.edu/thesis/available/etd-04092004-143712/unrestricted/ch_5adaptivbeamforming.pdf), 7AM, 20<sup>th</sup> April, 2008,
- [4]. H. L. Van Trees, "Detection, Estimation and Modulation Theory", Part IV: Optimum Array Processing. John Wiley & Sons, 2002.
- [5]. John D. Kraus, "Antennas: For All Applications", 3<sup>rd</sup> edition, Tata Mc-GrawHill, 2002
- [6]. L. C. Godara, "Application of Antenna Arrays to Mobile communications. II. Beamforming and Direction-of-

Arrival Considerations," Proceedings of IEEE, Volume 85, Issue 8, August 1997, Pages 1195-1245.

- [7]. Raed M. Shubair, Mahmoud A. Al-Qutayri, and Jassim M. Samhan, "A Setup for the Evaluation of MUSIC and LMS Algorithms for a Smart Antenna System", Journal of communications, VOL. 2 NO. 4, June 2007.
- [8]. Simon Haykin, "Adaptive Filter Theory". Prentice-Hall Inc., 4th Ed., 2002.
- [9]. John G. Proakis, "Digital Signal Processing: Principles, Applications and Algorithms". 4<sup>th</sup> edition, Prentice-hall India Ltd., 2007.
- [10]. Theodore S. Rappaport, "Wireless Communications: principles and practice." 2<sup>nd</sup> edition, Prentice-Hall India Ltd.

# The Channel Capacity of a GSM Frequency Hopping Network with Intelligent Underlay-Overlay

Md. Majharul Islam Rajib, and Wahida Nasrin

Department of Electronics & Communication Engineering

Khulna University of Engineering & Technology

Khulna-9203, Bangladesh

E-mail: majhar951@yahoo.com, jums\_24@yahoo.com

**Abstract** — In this paper a complete step by step method is described to increase the capacity of Global System for Mobile Communication (GSM) using the Intelligent Underlay Overlay(IUO) with frequency hopping and Discontinuous Transmission (DTX). The IUO splits the available frequencies into two layers, super and regular. Super layer frequencies have tighter reuse factor and can be accessed by only mobile stations (MS) having good C/I ratio while the regular frequencies can be used by MS's throughout the cell. The most promising work of this paper is that C/I ratio (hence reuse factor) is obtained by considering the first tier co-channel cells are also using IUO network structure. Analysis is performed by taking some default parameters into account shows that capacity enhancement up to 31 percent can be obtained with respect to a normal frequency hopping network.

**Index Terms** — Carrier to Interference Ratio (C/I), Discontinuous Transmission (DTX), Global System for Mobile Communication (GSM), Grade of Service (GOS) and Intelligent Underlay Overlay (IUO).

## I. INTRODUCTION

At present, increasing the capacity of the cellular network with limited resources (like available spectrum or number of base stations) is one the most crucial problem for the network operators. The general approach towards increasing the capacity with limited spectrum is cell splitting, sectoring, hierarchical cell structure like micro and pico cells [1, 2], or multiple reuse patterns with frequency hopping [3]. But most of these involve additional network elements like base stations or transmission network-which costs a lot. The most desired solution is to increase the capacity with minimal investment while maintaining the quality of service (QoS). There are two possible solutions may be mentioned here such as Frequency hopping (FH), and Intelligent underlay overlay (IUO) which was first proposed by Nokia Telecommunications [4]. The idea behind IUO is to partition the channels into two groups with different frequency reuse factors. The MS's having good C/I ratio

(better than a predefined threshold) will be assigned to the frequencies with tighter reuse factor which eventually enhances the capacity. The other frequency group with normal reuse factor provides the coverage. Several design methods have been proposed for the IUO. Some are well suited for networks without frequency hopping [4, 5, 7, 8]. In some papers, analysis was done by taking predefined value of reuse partitioning and channel assignment strategy [7, 9, 10]. Though the C/I ratio was considered in [5, 8, 10, 11] it was not a function of co-channel cells super area coverage factor. In this paper, an analytical approach has been taken considering that the C/I ratio is a function of super area coverage factor and carried traffic in co-channel cells. Hence, dynamic channel assignment strategy for a given Grade of Service (GOS) can be obtained. In analysis, it is also assumed that the network is large, homogeneous with symmetric traffic and using frequency hopping. At last the IUO model is compared with a normal cell considering some default parameters for performance evolution.

## II. PRINCIPLE OF INTELLIGENT UNDERLAY OVERLAY (IUO) SCHEME

The principle of IUO is that, the available radio frequencies are split in two groups, a super layer and a regular layer frequency group. The super frequencies are meant for the mobile stations with a good C/I, while the regular frequencies can be used by all MS's. A mobile station always starts on a regular frequency, because the C/I ratio of the call is not yet known. If this calculated C/I ratio is higher than a predefined C/I good threshold, the mobile is allowed to enter one of the super frequencies. If there is no available super frequency, the MS stays in the regular group but continue trying to enter the super layer. In the same way while the mobile station is already using a super frequency, it is continuously evaluated whether the C/I ratio is less than so called good C/I ratio, in that case

the MS is handover to the regular frequencies. The IUO algorithm doesn't distinguish moving or non moving vehicles. It considers that a non moving vehicle will not try for handover while the moving vehicles will. And there will not be any direct handover from super to a neighbor cell. The handover rate from super to regular or regular to super are calculated in [7] considering the speed of MS, the coverage factor of the super layer, the radius of the cell, the probability that the MS will move in a direction to increase or decrease the C/I ratio.

### III. MODELING ASSUMPTIONS

A model is considered for the analysis. The assumptions taken for modeling is described here,

- A single cell is taken into considerations for the sake of simplicity. It is also considered that the amount of incoming handover is equal the amount of outgoing traffic from the cell. However the model kept the opportunities to consider this. It is advised that in the practical case it should be considered.
- The source of traffic is uniformly distributed throughout the cell.
- For a particular coverage factor same reuse factor is considered for all the super layer frequencies.
- Number of super channel is SC and regular channel is RC. Any time the amount of occupied channel in regular layer is  $R_{reg}$  and occupied channel of regular but residing in the super is  $R_{sup}$ . The number of super channel occupied is  $Sup$ .
- The ratio of moving MS is  $Rmv$ .
- The moving MS are assumed to move uniformly in a random direction with a velocity  $Vms$
- New call arrival rate is assumed to be a Poisson process with arrival rate .
- The total incoming handover rate is  $\lambda_h$  and outgoing is  $R_{reg} \cdot \mu_{hout}$
- The time needed to calculate the C/I ratio is exponentially distributed with mean  $\frac{1}{\mu_{cal}}$ .
- The call holding time is assumed exponentially distributed with mean  $\frac{1}{\mu_t}$ .
- The channel dwell time is exponentially distributed with mean  $\frac{1}{\mu_{hout}}$ .

- The time needed for a moving vehicle to move from a super to regular area or regular to super is exponentially distributed random variable with mean value  $\frac{1}{\mu_{sup-reg}}$  and  $\frac{1}{\mu_{reg-sup}}$  which is given in [7]

### IV. EQUILIBRIUM MODEL

The model described here utilizes the fact that for certain amount of traffic in the super and regular layer, the sum of call termination rate and call blocking rate (GOS given by the operator) will be equal to call arrival rate. Hence at maximum traffic carrying condition there will be an equilibrium state. Channel allocation and suitable coverage factor will be calculated from there. From fig 1, we find the maximum carried traffic for blocked call rate B,

$$\begin{aligned} & \lambda \cdot cov + R_{reg} \cdot \mu_{reg-sup} - R_{sup} \cdot \mu_t - R_{sup} \cdot \mu_{cal} \\ & - B \cdot \lambda \cdot cov - R_{sup} \cdot \mu_{sup-reg} \\ & = 0 \end{aligned} \quad (1)$$

$$\begin{aligned} & \lambda \cdot (1 - cov) + \lambda_h + sup \cdot \mu_{sup-reg} + R_{sup} \cdot \mu_{sup-reg} \\ & - R_{reg} \cdot \mu_{reg-sup} - R_{reg} \mu_t \\ & - R_{reg} \mu_{hout} - B \lambda (1 - cov) \\ & = 0 \end{aligned} \quad (2)$$

$$R_{sup} \cdot \mu_{cal} - sup \cdot \mu_{sup-reg} - sup \cdot \mu_t = 0 \quad (3)$$

Solving equation 1, 2, and 3  $R_{reg}$ ,  $R_{sup}$  and  $sup$  can be found.

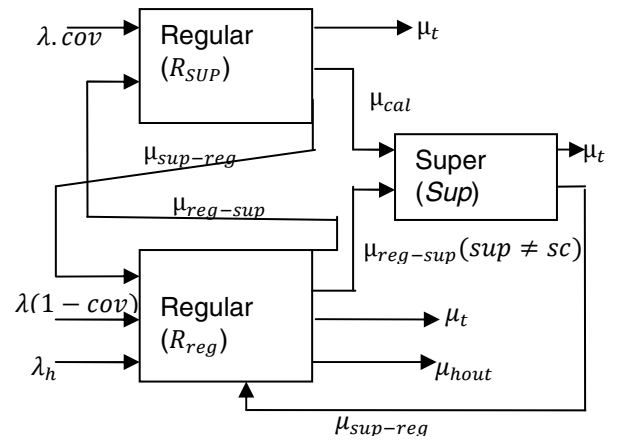


Fig 1. Equilibrium model of IUO cell

TABLE I  
TRANSITION RATE [7]

Event	Rate
New Call in regular layer	$(1 - cov). \lambda$
New call in super layer	$cov. \lambda$
Incoming handoff	$\lambda_h$
Good SIR calls	$R_{sup} \mu_{cal}$
Super to regular	$sup. \mu_{sup-reg}$ $R_{sup} \cdot \mu_{sup-reg}$
Soft blocking	$sup. \mu_{sup-reg}$
Regular to super	$R_{reg} \cdot \mu_{reg-sup}$
Call terminating	$R_{reg} \cdot \mu_t$ $R_{sup} \cdot \mu_t$ $sup. \mu_t$
Outgoing handoff	$R_{reg} \cdot \mu_{hout}$

#### V. DESIGN PROCEDURE

The design procedure is described with the flow chart in fig 2.

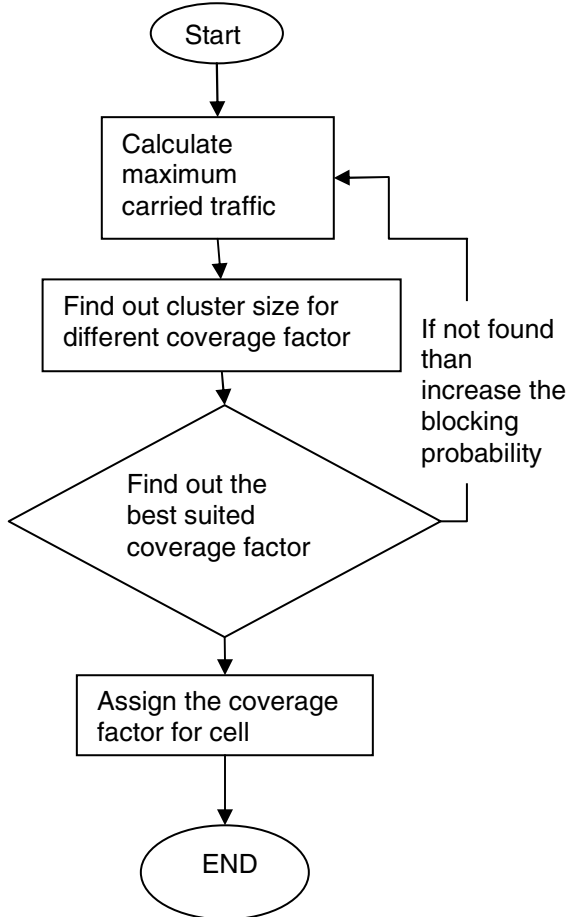


Fig. 2. Design procedure with flow chart

#### A. Calculation of traffic

The maximum traffic can be calculated from section IV. Here it is noticeable that  $sup = SC$ . All the other values are different for different cells and can be estimated by field measurement.

#### B. Cluster size

It is assumed that the users are uniformly distributed, hence the probability that a user using a slot in a carrier in a frequency hopping network is uniformly distributed [12]. In a simple cell this distribution of user is extended over the area with radius 0 to R. In a model like IUO and coverage factor  $cov$  distribution will take place in two different areas. For super the users are distributed from 0 to  $cov.R$ . In case of regular layer a fraction of regular layer user  $R_{sup}$  resides in super layer area. And the other are distributed in regular layer area  $R_{reg}$ . So the average interference power present in BTS due to super frequency user  $sup$  in j th cell will be [11].

$$E[Y_j] = \frac{S}{\pi(cov.R)^2} \int_0^{cov.R} \int_0^{2\pi} \left(\frac{r}{d_j}\right)^\alpha E\left[10^{\frac{\zeta}{10}}\right] r dr d\phi$$

And for regular frequency user ( $R_{sup} + R_{reg}$ ) this will be a conditional case as the frequency hopping takes place. So the average interference power will be

$$E[Y_j] = \left(\frac{R_{sup}}{R_{reg} + R_{sup}}\right) \cdot \frac{S}{\pi(cov.R)^2} \cdot \int_0^{cov.R} \int_0^{2\pi} \left(\frac{r}{d_j}\right)^\alpha \cdot E\left[10^{\frac{\zeta}{10}}\right] \cdot r dr d\phi + \left(\frac{R_{reg}}{R_{reg} + R_{sup}}\right) \cdot \frac{S}{\pi(R - cov.R)^2} \int_{cov.R}^R \int_0^{2\pi} \left(\frac{r}{d_j}\right)^\alpha \cdot E\left[10^{\frac{\zeta}{10}}\right] r dr d\phi$$

Here  $d_j$  is the distance between MS in j th cell and BTS. It

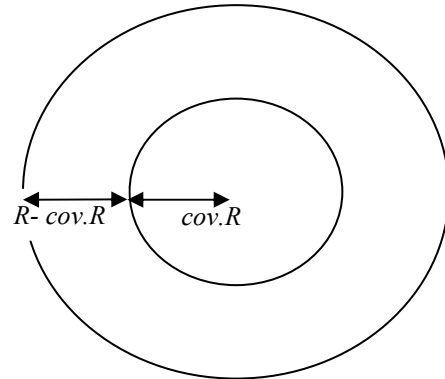


Fig. 3. IUO cell

is a function of cell radius, coverage factor  $cov$  and cluster size. Now, the C/I ratio at the BTS for uplink transmission (assuming reciprocity) will be,

$$C/I = \frac{S}{\mu \sum_{j=1}^6 E[\gamma_j]}$$

Where  $\mu$  is the voice activity factor (VAF), which allows discontinuous transmission (DTX) to include in the model. To model the IUO for the worst case we considered here VAF=1. For various coverage factors  $cov$  and to maintain a minimum level of C/I ratio which is predefined by operator, we will get different cluster size. However, in the case of co-located overlay and underlay cell, Broadcast Control Channel (BCCH) is carried by a regular frequency [4], hence, this BCCH should be kept clean as much as possible. It is the next problem to allocate frequency to super and regular layer so that none them is allocated more or less with respect to expected carried traffic.

### C. Optimization

First the required number of frequency is estimated considering for a given GOS. The frequencies are allocated considering that regular layer frequencies consist of 7 TCH, and super layer consist of 8 [7]. Then a graph is plotted for required total frequency vs. offered traffic. From the curve a minimum point can be achieved which best fulfills the offered traffic and power management.

## VI. ANALYTICAL RESULT

To observe the practical outcome of the proposed model, we analyzed a cell with some default configurations. The objective is to provide adequate frequency for super and regular layer. If there is not enough frequency, to the GOS has to be increased or the available frequency should be increased which is not desired or cost effective. The default parameters are presented in Table 2 and the results are analyzed in subsequent graphs.

TABLE 2  
ANALYSIS ASSUMPTION PARAMETERS

Parameters	Values
Call holding time	80sec.
Cell radius	3km
Ratio of moving MS	50%
Speed of moving MS	50km/hr
The path loss exponent,	4
The standard deviation of shadow fading,	8
Voice activity factor, $\mu$	1
Blocked call rate, $b$	5%
Threshold level	12db

The resultant graphs show that for different values of  $cov$  the C/I ratio for MS in super frequencies decreases as the  $cov$  increases.

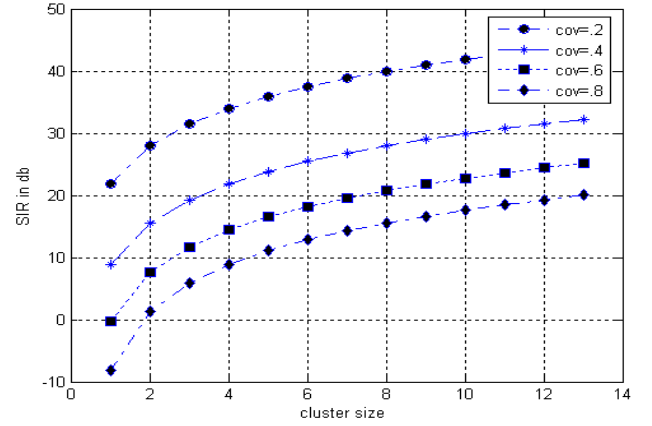


Fig. 4. Uplink SIR versus super layer cluster size at different super area coverage factor

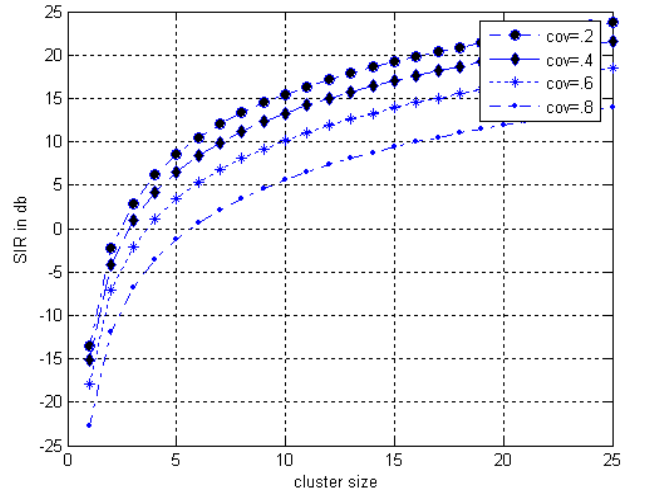


Fig. 5. Uplink SIR versus regular layer cluster size at different super area coverage factor

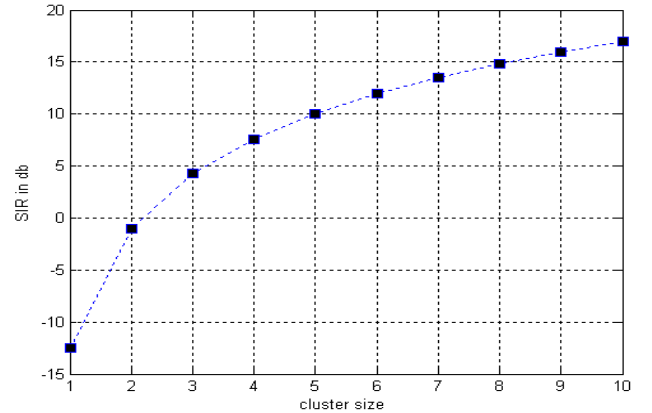


Fig. 6. Uplink SIR versus cluster size for a normal cell

## ACKNOWLEDGEMENT

The authors wish to acknowledge the assistance and support of the IEEE Book and Information Services and the IMS2004 Publication Committee.

## REFERENCES

- [1] Theodore S. Rappaport "Wireless Communications Principles and Practice", Second edition, Prentice Hall India, 2004.
- [2] Timo Halonen, Risto Aalto, Esa Tuovien "Increasing the Capacity of a Cellular Radio Network", *US Patent 6985736* Issued on January 10, 2006.
- [3] Pedro Assunção, Rui Estevinho and Luis M. Correia. "Assessment of Cellular Planning Methods for GSM" *Proc. of the 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2001
- [4] Nokia Telecommunications Web Site: [www.nokia.com](http://www.nokia.com).
- [5] J. Zander, "Generalized Reuse Partitioning in Cellular Mobile Radio" *Proc. of the 43<sup>rd</sup> IEEE Vehicular Technology Conference*, 1993
- [6] Y. K. Ling, J. Elling, T. Toftgrad Nielson, "Capacity of Intelligent Underlay and Overlay Network", *Proc. of the IEE Colloquium on Advanced TDMA Techniques and Applications*, pp. 23-31, oc. 1996.
- [7] Khalid Begain, Gergo Istvan Rozsa, Andras Pfening, Miklos Telek, "Performance Analysis of GSM Networks with Intelligent Underlay-Overlay", *Proceedings of the Seventh International Symposium on Computers and Communications*, 2002.
- [8] Kumaran, Krishnan, Whitling, P. Alfred "Method for Designing Underlay-Overlay Networks for Mobile Wireless Communication." *European Patent EP 0993210B1* issued on January 30, 2002
- [9] T. T. Nielsen, J. Wigard, P. Morgensen, "On The Capacity of a GSM Frequency Hopping Network with Intelligent Underlay-Overlay", *Proc. of the VTC Phoenix, Arizona, USA, May 1997*
- [10] J. Wigard, T. T. Nielson, P. H. Michaelsen, P. Mogensen "Improved Intelligent Underlay-Overlay Combined With Frequency Hopping" *Proc. of the 8th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 1997
- [11] S. Papavassiliou, L. Tassiulas, P. Tandon, "Meeting QoS Requirements in Cellular Network with Reuse Partitioning" *IEEE Journal on selected areas in Communications*, 1994
- [12] Raymond Steele, Chin-Chun Lee, Peter Gould "GSM, CDMA1, and 3G Systems", Jhon Wiley & Sons Ltd, 2001.
- [13] Sheldon M. Ross "Introduction to Probability and Statistics for Engineers and Scientists" Third Edition, Elsevier Academic Press, 2004.

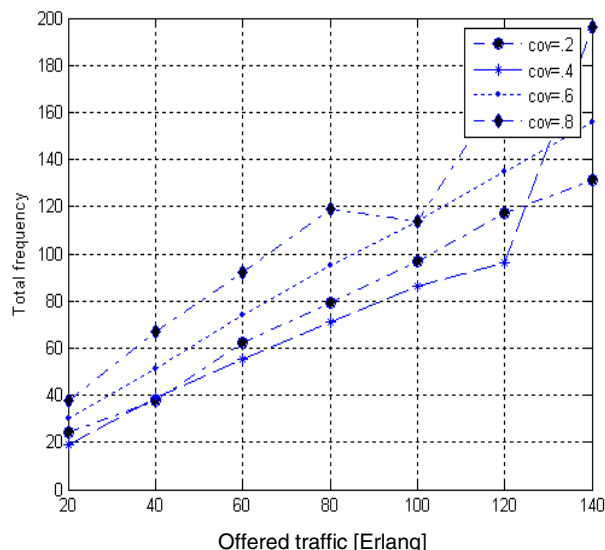


Fig. 7. Total channel versus offered traffic for different super area coverage factor

The same result for regular is obtained. Hence the cluster size increases as *cov* increases. This was expected because as the *cov* increases more and more MS spread throughout the cell. Hence the MS comes into more vicinity to neighbor co-channel BTS, hence the interference power increases thus the C/I decreases. Another thing is that different number of frequencies is required to meet the GOS for different *cov*. A careful observation is necessary before allocating frequencies. A handoff between the amounts of simultaneous subscriber served in a cell and the number of frequencies available to operator is to be considered. Performance improvement will depend completely upon the efficient partitioning of available frequency and maintaining the minimal threshold of C/I at the same time. By assuming lost call cleared system, it is seen that the capacity is increased up to 31% with respect to the FH GSM network. If the cluster size is not taken some discrete values like 3,7,12 ( $i^2 + ij + j^2$ ) etc, more enhancement can be found.

## VII. CONCLUSION

This paper focused on designing Intelligent Underlay Overlay (IUO) scheme and optimization of capacity without compromising the Quality of Service of a cellular network. The analysis results characterize the effect of IUO on the performance of a cell as a function of the coverage factor, offered traffic, the mobility of MSs, QoS and GOS. The proposed model indicates that significant capacity enhancement can be achieved with increased complexities of frequency partitioning and network optimization.



# Cell Database Based Time Specific Resource Reservation Scheme

S. M Didar-Al- Alam, Md. Nazmul Hasan, Md. Mahfuzul Islam, Akramul Azim

Department of Computer science and Information Technology,  
Islamic University of Technology, Gazipur, Bangladesh  
didar522@yahoo.com, bappy\_iut@yahoo.com, adip\_iut@yahoo.com, bijoy456@gmail.com

## ABSTRACT

**Abstract** — Increasing cellular network offers services with diverse quality-of-service supports users to access telephony, paging, instant messaging, and trivial web-browsing on the same device. The mobility of communication devices and channel imperfection makes the service offerings much more challenging with limited resources. Resource reservation plays a very important role in cellular communication in terms of availability of services. There are various schemes present to manage the resources in an efficient way. In this paper, we have tried to analysis several resource management schemes and their drawbacks. We have proposed a time dependent resource reservation scheme which has significant advantages over other schemes.

**Index Terms** — Mobility Management, handoff, Call Admission Control (CAC), Call dropping probability (CDP), Call Blocking Probability (CBP), Cell Visiting Probability (CVP).

## I. INTRODUCTION

In a cellular network MU is connected with base station (BS), which uses a dedicated frequency spectrum. The BS is connected with the BSC by microwave antennas that are again connected with MSCs for controlling purpose. MSCs are interconnected using the optical fiber backbone in most of the cases except some conditions where wireless connectivity is needed that acts as a gateway to the external links.

Bandwidth [8] reuse principle is used on the limited available radio frequency spectrum that divides it into smaller spectrums for maximum uses. Smaller sized cells required by the next generation cellular networks to support higher transmission rates of multimedia applications that create increasing handoff problems in the cellular network. That's why resource management becomes a major concern in the perspective field. There are three major aspects to deal with. These are Call Admission Control (CAC), Resource Reservation, and Resource Allocation.

The paper is organized as follows. Introduction, Different past policies and difficulties with them, Our contribution, Our proposed scheme, Mathematical Relations, Conceptual backend database design, Comparison with old schemes, Conclusion and Future works.

## II. DIFFERENT POLICIES FOR RESOURCE RESERVATION

### A. Elementary guard channel (GC)

This GC policy [3] reserves fixed portions of the cell capacity causing huge wastage of resources & higher call blocking probability. Instead of resource reservation, controlling new service admission random walk model decreases CDP below a threshold [1, 3, 5] level. Shadow cluster is formed with neighbouring cells likely to be visited by an MU. Resources are reserved estimating the CVP for an MU relying on historic information rather than current mobility information. Initiating a new service is done on basis of resource limitation by giving priority to the existing call. Most of the CAC [1,5] & resource reservation schemes are designed by reserving a certain bandwidth in each cell or by keeping CDP less than threshold value.

### B. Non-Mobility Based Resource Reservation Schemes

A fixed portion of bandwidth is reserved in each cell for handoff services in conventional resource management schemes (Ex: Guard channel scheme). Unused channels are shared by handoff and new services. If we say  $CH_i^T$  is the number of total channels,  $CH_i^u$  number of unused channels,  $CH_i^a$  allocated channels,  $CH_i^r$  reserved channels,  $CH_i^f$  and number of free channels, then according to this scheme

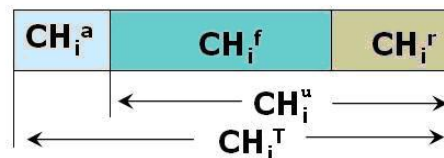


Fig. 1. Cell Capacity Distribution

a new service is admitted if and only if  $CH_i^f = CH_i^u - CH_i^r$ , the free channels can support it. Distribution is shown in Fig.1. Handoff is functional when it is supported by unused channels. Handoff service queuing (HQ) model on top of the GC scheme was introduced by Hong and Rappaport.

In this model they assume that a MU can simultaneously communicate with more than one cell;

Received Signal Strength Indicator (RSSI) level is used for handoff decision. For a MU the handoff procedure is implemented if RSSI level is less than the handoff threshold signal strength level and also less than the RSSI of any neighbouring cell's BS. For unavailable channels of the next cell service are queued until the RSSI of the current BS drops below an acceptable received threshold. Tekinay and Jabbari introduced in HQ model measured based priority-queue scheme (MBPS) to maintain priority in the queue, higher priority is given to the services with lower RSSI level. Performance of this HQ model has been studied both experimentally and analytically using the birth-death queuing model [4]. This scheme though can detect handoff situation perfectly but resource reservation part has not been given much concern. Fixed amount of resource is reserved in all the neighbouring cells costing a huge amount of resource wastage.

### C. Symmetric Random Walk Model Based Schemes

SRWM [7] estimates CDP and/or CBP using different statistical measures & keep them below some threshold values for new/handoff services admission decision.

#### C.1. Threshold Bounded New Service Admission Control Models

In threshold bounded new service admission control schemes [1, 2, 5] new services are accepted if the number of new services or CDP in a cell is below a given threshold. Handoff service is admitted by the remaining unused channels in the cell.

#### C.2. Diffusion Equation Based Models

The properties of the diffusion equation are used to predict the mobility of MU. The influences of MU from the neighbouring cells are taken into consideration. Diffusion equation based models takes the CAC decision that expected CDP is kept below  $P_i^D$  (a certain limit). No resource is reserved for this purpose. Principles of the diffusion process cannot explain free-will motion of MUs in a cellular environment guided by geographical features.

### D. Handoff Profiling Based Schemes

This schema predicts future mobility of an individual MU based on the mobility history stored in profiled databases. Path segment or mobility, route & timing information of the visiting MU is recorded in the database.

#### D.1. Path Segment Based Schemes

For dynamic location tracking Tabbane used user specific handoff mobility profiling database based on the quasi-deterministic mobility behavior. Liu and Maguire further pursued the model by tracking repetitions of movement patterns. Location of the MU estimated based on these movement patterns, using pattern matching/recognition-based mobile motion prediction (MMP). Random movements cannot be

classified by the simple mobility patterns. MMP algorithm is high sensitive to this is its main drawback. To solve the problem Liu et al. proposed inter-cell mobility predictor. Here routine routes of the user are identified from mobility profile database by using an approximate pattern matching technique. For each user's mobility patterns (UMPs) are recorded and indexed by the occurrence time. It is much less sensitive to small deviation from the user's actual path (UAP). In UMP it is assumed that the direction of MU is not changed within a path segment like within a cell. For a cellular network with micro-cells or Pico-cells using UMP is meaningless. It has high computational complexity and huge storage requirements that is not suitable for real-world situations.

#### D.2. Mobility Route Based Schemes

Levine et al. proposed the scheme where route based user mobility information is profiled at regular intervals. That estimates handoff, cell residence time for each cell and cell-independent call duration. It gives us the time-dependent CVP for the MU considering all possible routes to visit neighbouring cells.

This technique is very expensive as it requires constant updating for each MU in profiled databases. Again the number of possible routes could be infinite. Recording individual or aggregated mobility history is time and space consuming task. Reducing complexity by distributing database demands extra memory and processing power. As aggregated mobility history must be kept centrally, distributed database cannot resolve the problem too.

#### D.3. Route Independent Traffic Flow Based Schemes

This scheme profiles up handoff histories in the form of quadruplets. It has the departing time from the current cell, the index of the previous cell the MU resided in before entering the current cell, the index of the cell the MU entered after departing the current cell, and the current cell dwelling time. CVP is then estimated. Profiling each handoff for a large period of time (at least 24 hours) is space consuming. Here from all the handoff profiling based schemes we found that a major common problem is cost. These models cost huge time and money. Different models records different topics like Movement patterns, Route based mobility information, Handoff histories etc for taking resource reservation decision. But among this Route Independent Traffic Flow Based Schemes are better as they records handoff, though it is very much costly.

### E. Cell Topography Based Schemes

Sleem et al. extended cell topography knowledge to predict CVP directly from the location road-map. From its current position an MU takes the next path with equal probability from the paths ahead. At location A, MU can turn left or turn right or take a U-turn. CVP for cells 1, 3, and 5 are 1/3 from cell 4. It is shown in Fig 2. Information of road networks is taken into account

ignoring mobility information. It is highly complicated in a city with complex road-networks.

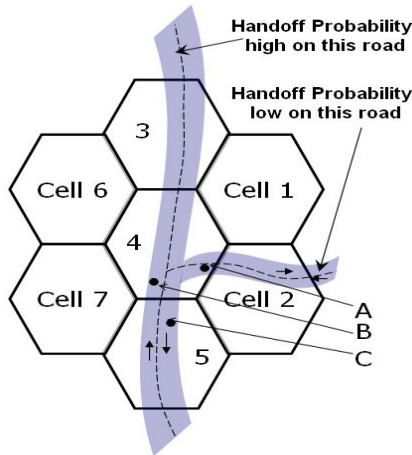


Fig. 2. Cell Topography based on CVP estimation

In this model road networks and environment are taken under consideration. Surely this knowledge helps to handle handoff situation perfectly.

#### F. Parametric Mobility Support Schemes

In this scheme key mobility parameters are used. They are position, direction, and speed. This information is used for CVP estimation, shadow cluster formation technique, solitary resource reservation policy, complete resource reservation and CAC policies.

##### F.1. Shadow Cluster Formation without Explicit CVP Estimation

Using the direction and speed information of the MU this scheme proposed an elliptical shaped (Fig 3) shadow cluster formation technique [6, 10]. It does not estimate any CVP. MU is the focus of the ellipse and direction is the axis.

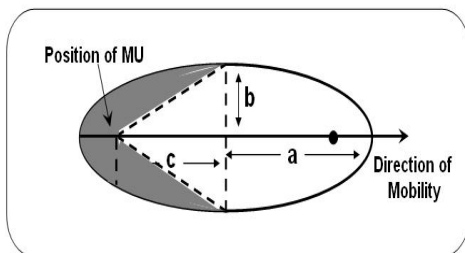


Fig. 3. An elliptical shaded shadow cluster is formed.

In the shaded region which is the backside of the MU's current position, visiting probability is extremely low. Elliptical cluster reserves resources unnecessarily for the area. From the position of the MU the radial span of the cluster should be equal. This should be irrespective of the direction. Next, CVP has no bearing. So the shape and size of the cluster remains same irrespective of the MU current cell. As there is no CVP estimation [11, 12], so resource reservation policy is not effective one. Direction of MU has no influence on the amount of reservation.

### III. OUR CONTRIBUTION

As far as we know the database based schemes are mostly not applicable because of their high cost and maintenance problems. Again these database based schemes can not give an optimally efficient performance for the whole network at a time. They actually focus on giving optimal performance for most of the cells.

We have proposed a new database based scheme that is able to reserve resources much more effectively. We tried to find out the causes behind performance problems and proposed the new idea with some specific improvements. In this paper we also tried to show how we can overcome the problems with the help of this theory. This scheme would take decision for each and every cell separately. This would be effective for having a constant efficient performance for whole the network. Details for these topics are given below.

### IV. PROPOSED SCHEME

#### A. Time Specific Cell Database Scheme:

In the proposed time specific cell database scheme a database is maintained for each and every cell. Now the cell database is created by watching the MUs' movement for a period of time. It then takes the decision depending on the database. It updates its database after a defined period of time to notice the change in the cell area.

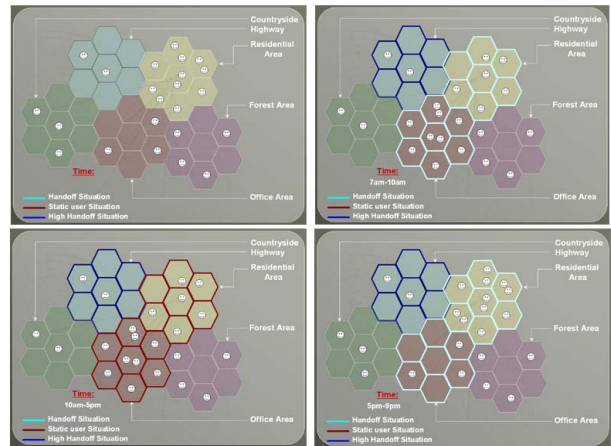


Fig.4. Time specific cell database scheme

In the Fig. 4 we can see how the cells are using the database to take the reservation decision. Suppose, a city is divided into two parts which are Residential area and office area. Now the cell observer finds out that during the 7am-10am period of time the handoff is happening very frequently in these cells because of the movement of people from home to office. So, for this period the handoff probability is very high in these cells. So the larger portion is reserved for handoff managing to control the handoff rush properly.

Then from 10am-5pm cell observer finds a static situation within the cell and that's why use larger portion of its resources for call initiating.

Again from 5pm-9pm again rush on road causing same situation during 7-10am because of rush on the road. On the highway, as all the vehicles are moving fast and handoff probability is much higher so always handoff portion should be more for highway cells.

Again for weekends we should have different strategies as there is no rush for office area in that period. For amusement & market areas the rush would be higher & handoff should be managed just as before.

### B. Mathematical relationship

Here we are actually trying to give a basic idea how different parameters are related with the handoff probability and how they change its value. To give the relations mathematically we assume-

$H_p$  = Handoff Probability for the cell

$R_h$  = Resource reserved for handoff

$R_i$  = Resource reserved for call initiating

Then the relation would be:

$$H_p \propto R_h \quad (1)$$

$$H_p \propto 1/R_i \quad (2)$$

As considering only equation 1 we can take our threshold value decision we are only focusing on this-

$$H_p \propto R_h \Rightarrow H_p = k \cdot R_h \quad (3)$$

Here K depends on the factors like,

U= Number of user resident at the moment.

C= Number of call initiated

N= Number of users in neighboring cells

P= Probability of users to move in this cell

T= Average time period for the initiated calls.

And K is dependent on this factor in the following way-

$$\frac{NP}{UCT}$$

So equation 1 can be written as-

$$H_p \propto \frac{NP}{UCT} R_h \quad (4)$$

### C. Conceptual database design for Cell Database Scheme

Suppose we store data in the database for a 2 hour time gap for all 7 days in a week. Observer observes the cell for a specific period of time and stores the handoff percentage for the cell happened in each slot of time. From database if we find handoff probability for a cell more than 40 % then we reserve larger portion of the resource so that we can minimize the CDP. Again if handoff probability is less than 40 % then resource portion is smaller to minimize the CBP.

	6-8 am	8-10 am	10-12 am	12-2 pm	2-4 pm	4-6 pm	6-8 pm	8-10 pm	10-12 pm	12-2 am	2-6 am
Fri	20	25	30	40	40	70	60	30	20	10	10
Sat	20	50	40	30	30	40	60	30	20	10	10
Sun	20	70	55	30	40	55	60	30	20	10	10
Mon	20	70	55	30	40	55	60	30	20	10	10
Tue	20	70	55	30	40	55	60	30	20	10	10
Wed	20	70	55	30	40	55	60	30	20	10	10
Thu	20	70	55	30	40	55	60	30	20	10	10

Fig. 5 Database model view for the proposed scheme

So, for the database schema design we have to have one database in the MSC and one in the BSC because it is difficult to maintain bulky data at BSC. This costs much. To reduce the cost and to gain speed in retrieving we use a distributed database. A sample view of database for a Time Specific Cell Database Scheme is shown in Fig 5. The schema design for the backend database is shown.

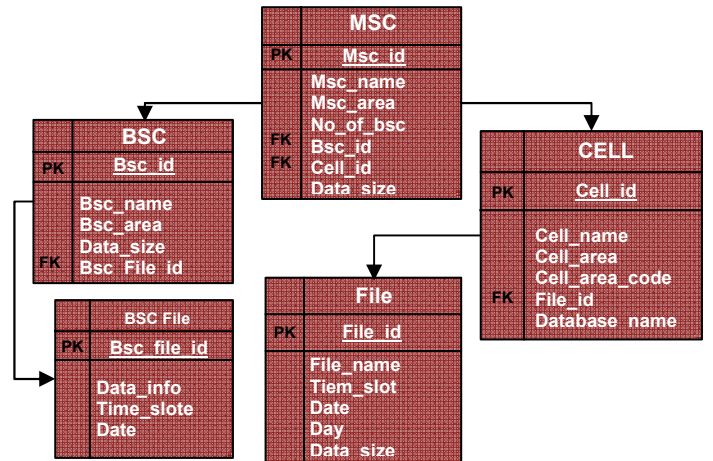


Fig. 6. Schema design of backend database for Cell Database Scheme

### D. Assumptions & Specifications

#### D.1 High data retrieving rate

Data retrieving rate is high in this scheme as we are storing data in the BSC database. So the data can be retrieved in a much shorter time than other schemes where others are retrieving data from the OMC.

From the graph (Fig. 7) we can see that data size with respect to time is growing up. Though the rate of it low in other schemes in spite of that the time limit is higher. On the other hand the growing scale takes a high jump after 2GB as we consider more than 2GB of data should



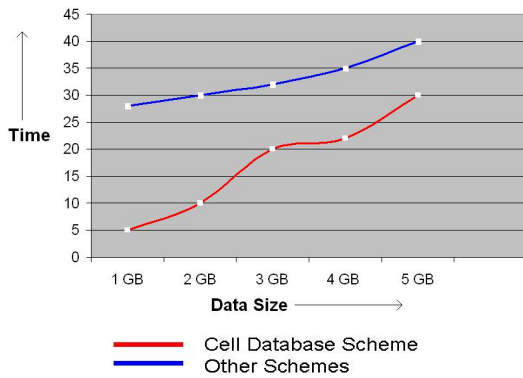


Fig.7. Data retrieving rate graph

be retrieved from the MSC. But the graph clearly shows us that high growing rate does not make it ineffective but still it is effective in time savings and high rate data retrieving.

#### D.2 Constant & Efficient CBP & CDP Management

As in other schemes frequency is managed for whole network at a time depending on historic movement of MU. Though this is applicable for most of the cells but for some cells it may cause hazardous situation.

From the graph (Fig. 8) given we can see the effect of this situation. From 7am-10am MUs' much movement causing heavy handoff rushes. For this reason, larger portion of frequency is reserved for managing this handoff situation in the whole network. But it is not efficient for all of the cells. Cells like residential area may need higher portion of frequency for call initiating while much frequency is reserved for managing handoff situation. This causes high CBP. Again from 10am-5pm frequency is reduced for handling handoff situation now which is given for call initiating. But this also causes problems for highway as handoff is occurred frequently in highway. Some other situations are also given in the graph.

In our schema as we maintain cell database, it helps us to take decision on a specific cell that what level of resource it should have for managing the handoff. So here we can ensure that while we are giving high reservation in most of the areas that may not cause high CBP in the residential area. As we are serving it with a low resource reserved, this gives us a constant efficiency Fig. 8) all the time in all the cells. From the graph we get some more situations where CDP & CBP is managed efficiently.

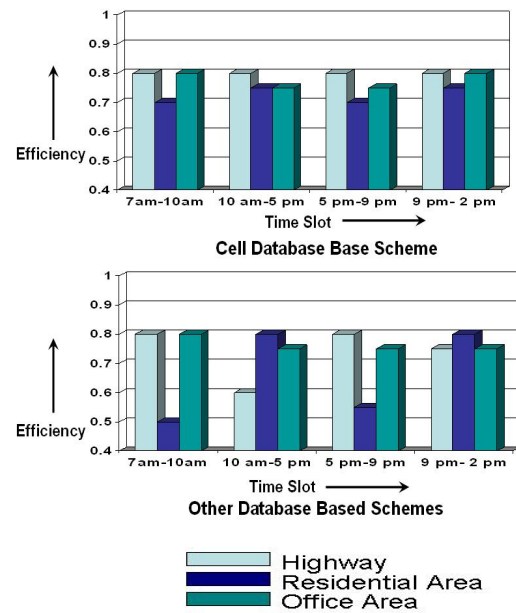


Fig. 8. Efficient management of CBP & CDP

#### D.3 Cost Effective data storage

In Mobility Based schema path movement for each MU is taken into consideration & movement of each MU is stored in the database. So here database is huge which causes high cost for managing as well as for storing & retrieving.

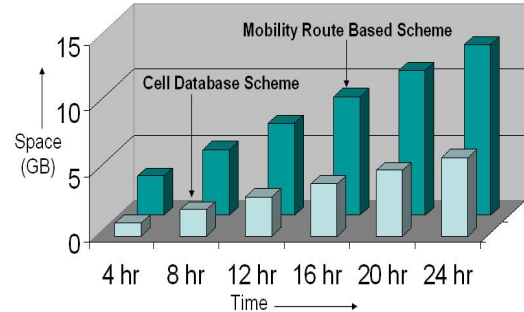


Fig. 9. Cost effective data storage

But in Cell database schema number of handoff per cell is calculated with respect of time. As here we don't calculate each MU's movement other then handoff condition. So, database is small. It is also easy to manage the database as it is distributed. So, our scheme has less costly data storage than Mobility route based scheme (Fig. 9).

#### E. Working procedure for the Cell Database Scheme:

This flow chart (Fig. 10) is given for showing the working procedure of Cell Database Scheme. From the flow chart it is visible that each time it is taking the decision that it is now going to get new data to store or it is going to take decision on stored data. If it needs to take decision it then checks for the data size. If it is

smaller than a fixed amount it looks for the data in the BSC database else it looks in the MSC database which is more time consuming.

After gathering of data it takes the decision of the amount of resources to be reserved. According to the decision it reserves the resources. Or it finds out that it is going to gather new data to store. Then it starts by checking out when last database was updated. If it is more than four months ago then the observer gathers new data and stores it. If it was updated within four months then it takes the decision that data is valid for use.

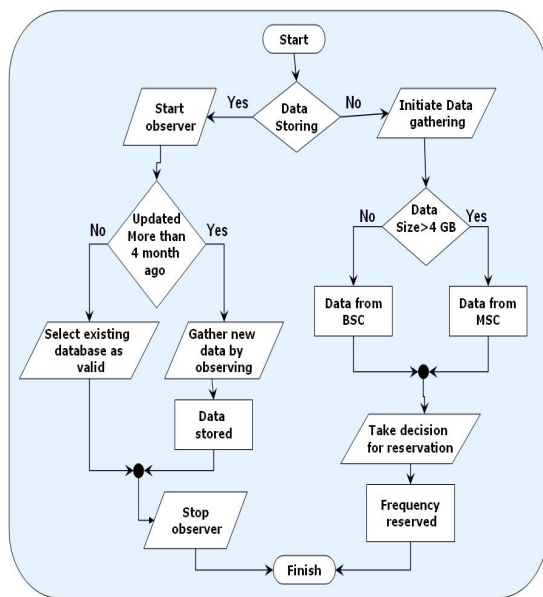


Fig. 10. Cell Database Scheme Flow chart

### III. CONCLUSION & FURTHER WORK

Above we have discussed some related schemes that contributed a lot in the telecommunication evaluation. But none of them are the optimal one. We have proposed a new concept of resource reservation, where we have taken under consideration the problems of the previous schemes. But we have some further improvements to do in it like:

1. Call duration should be taken into account. That means a database should be maintained for ongoing calls for each cell to limit its excessive holding of a channel. If a call duration limit exceeds then its CDP would be higher and new initiating calls will be given much priority.

2. Mobility and Non mobility based issues should be merged. Previous schemes discussed only with one of those aspects completely ignoring the other one in most of the cases. Schemes those considers both of the aspects actually focuses only one aspect in practical use. Schemes concerned about both the aspects practically will be much more effective.

3. As cell database is used, the database would be a distributed one, which makes it hard to manage. So database management should be taken under consideration.

### REFERENCES

- [1] Fang Y. and Y. Zhang, "Call admission control schemes and performances analysis in wireless networks," IEEE Transactions on Vehicular Technology, vol. 51, no. 2, pp. 371-381 March 2002.
- [2] Hou J. and Y. Fang, "Mobility-based call admission control schemes for wireless mobile networks," Wireless Communications and Mobile Computing, vol. 1, no. 3, pp. 269-282, 2001.
- [3] Posner E. C. and R. Guerin, "Traffic policies in cellular radio that minimize blocking of Handoff calls", in Proc. 11th ITC, pp. 294-298, Kyoto, Japan, Sept. 1985.
- [4] Van V. P. and S. Glisic, "Performance analysis of queuing schemes for priority handoff and call admission control in mobile cellular radio networks", International Conference on Communications (ICC), vol. 7, pp. 2291-2295, 2001
- [5] Wu S., K. Y. M. Wong, and B. Li, "A dynamic call admission policy with precision QoS guarantees using stochastic control for mobile wireless networks," IEEE Transactions on Networking, vol. 10, no.2, pp. 257-271, April 2002.
- [6] Zhou X., R. Chen, and C. Gao, "An elliptical shadow algorithm for motion prediction and resource reservation in wireless cellular networks," International Conference on Computer Communications and Networks (ICCN), pp. 234-239, 15-17 October 2001.
- [7] Oliviera C., J. Kim, and T. Suda, "An adaptive bandwidth reservation scheme for high speed multimedia wireless networks," IEEE Journal on Selected Areas in Communications, vol. 12, pp. 858-874, 1998.
- [8] Gaasvik P., M. Cornefjord, and V. Svensson, "Different methods of giving priority to sthandoff traffic in a mobile telephone system with directed retry," 41 IEEE Vehicular Technology Conference 'Gateway to the future Technology in Motion, pp. 549-553, 1991.
- [9] Sahinoglu Z. and A. Vetro, "Mobility characteristics for multimedia service adaptation," Signal Processing: Image Communications, vol. 18, pp. 699-719, 2003.
- [10] Chen R., X. Zhou, F. Zheng, and C. Gao, "A distributed algorithm enhancing the connection quality and bandwidth utilization in wireless cellular networks," International Conference on Computer Networks and Mobile Computing, pp. 79-84, 2001.
- [11] Liu T., P. Bahl, and I. Chlamtac, "Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks," IEEE Journal on Selected Areas in Communications, vol. 16, no. 6, pp. 922-936, August 1998.
- [12] Liu T., P. Bahl, and I. Chlamtac, "An optimal self-learning estimator for predicting inter-cell user trajectory in wireless radio networks," IEEE International Conference Universal Personal Communications, vol. 2, pp. 438-442, 12-16 October 1997.

# Developing a Model of E-governance for Urban and Rural Areas of Bangladesh

Gazi Zahirul Islam<sup>1</sup>, Md. Ibrahim Khan<sup>2</sup>, Md. Anas Boksh Mazady<sup>3</sup>

<sup>1</sup>CSE Department, Primeasia University, Banani, Dhaka, Bangladesh.

<sup>2</sup>CSE Department, Chittagong University of Engineering & Technology, Chittagong, Bangladesh.

<sup>3</sup>Core Network Development, Aktel, Gulshan-1, Dhaka, Bangladesh.

**Abstract** — Several developing countries including Bangladesh are realizing the role ICT (Information & Communication Technology) can play in the governance sector, and are putting into practice innovative e-governance models that may be technologically simple but are drastically changing the way information is distributed in the society. This paper depicts and implements a model of e-governance for urban and rural areas in Bangladesh. The project deeply considered existing ICT infrastructure in Bangladesh and put effort to implement the model in a cost effective and flexible manner. A clear, comprehensive layout and network architecture for rural and city areas is given separately. We also described complementary tasks such as database implementation, system development, web related task etc. for e-governance. The main objective of the project is to implement e-governance in an easy and cost effective way to suit with existing ICT infrastructure in Bangladesh to provide better information and service delivery of government to the people.

**Index Terms** — ICT, VPN, Public network, Private network, Baud rate, Dial-up network

## I. INTRODUCTION

E-governance is the public sector's use of information and communication technologies with the aim of improving information and service delivery, encouraging citizen participation in the decision-making process and making government more accountable, transparent and effective. Bangladesh and many other developing countries are apprehending the importance of information and communication technology (ICT) in the governance arena, and are putting into practice innovative e-governance models that might be technologically simple but are abruptly changing the way information is distributed to the public [15], [11]. In this paper we delineate a model of e-governance through rural and city areas in Bangladesh.

The project deeply considered current ICT infrastructure in Bangladesh and tried to implement the project in a cost effective manner. A clear, comprehensive layout for urban and rural areas is drawn separately and is shown in section III of this paper. Since, in Bangladesh government telephone line is the most widely spread data

communication line; we have chosen a VPN (virtual private network) model for implementing e-governance and is explained in section IV. In section VI we described complementary tasks such as database implementation, system development, web related task etc. to establish a full-fledged model of e-governance. We have developed sample database, forms and web pages which views are included in this paper. The goal of this paper is to implement e-governance in a flexible and cost effective manner to suit with current infrastructure in Bangladesh to provide better information and service delivery of the government to the public; and hence promote democratic practices through public participation and consultation.

## II. OBJECTIVE OF E-GOVERNANCE

E-government refers to the delivery of government services and information to the public, using electronic means to transform the way it serves its constituents [3], [9]. In this case, in general, citizens are seen as passive recipients of the digital information and services. But the trend has been changing drastically and citizens are now demanding more user-friendly programs.

However, most importantly, e-government must be on governance and not just IT; on citizens and their needs and not just on computers; on transforming the process not on translating the process; and on software not on just hardware and focus on people but not on systems. Thus, e-government is not just 'e' but about facilitating the system by 'e' and getting it to work for citizens to benefit them in their daily activities of practical life [10].

### Goals of e-governance

- Improve the internal organizational processes of governments
- Provide better information and service delivery
- Increase government transparency in order to reduce corruption
- Reinforce political credibility and accountability

- Promote democratic practices through public participation and consultation
- Focus on disadvantaged communities, who otherwise will be excluded
- Provide that information or service which otherwise will not be provided
- Create an outcome which in absence of ICT will not be produced efficiently or timely

### III. LAYOUT DESIGN

Like other developing countries, the infrastructure of ICT between rural and urban areas in Bangladesh is different [1]. So, separate layout should be considered for cities and villages. Design of an e-governance scheme for a rural area involves many constraints such as connectivity, bandwidth and available infrastructure [12], [5]. In city areas the structure is stronger and easier to implement e-governance. By various scheme initiated by Government of Bangladesh, rural telephony penetrates into most of the remote villages of the country. This paper describes the design issues of a practical e-governance system for rural and urban areas. Emphasis was on giving a cost effective solution with minimal changes to the existing communication infrastructure and user-friendly interface to the system using English and Bengali Language. Also the issues like getting active network connection through dial-up modem have been taken into consideration.

#### A. Layout for rural areas

In rural areas District Commission Office (DCO) is the centre of control, where database server, web and application server, network hub etc. will be kept. Other areas are connected to District Commission Office. Kiosk or booth should be set up in different places for public to communicate with DCO office. We may choose Upazela Offices and Union Offices as the place to set up the kiosks. Without Upazela Office and Union Office a remote populated or important place may be choose to set up kiosk for ease of public. All kiosks should be equipped with at least a computer, a modem, telephone line and other equipments to communicate with DCO office. Figure 1 shows the layout of the model for rural areas. Upazela Office, Union Office and distant kiosk are directly connected to District Commission Office (DCO).

#### B. Layout for city areas

The infrastructure of ICT in cities are substantial than rural areas. It is easier to implement e-governance in city areas than rural areas. We can easily divide a city area by wards. In every Ward Commission Office a kiosk has to be setup. Here, City Corporation Office (CCO) will take the control of governance [2]. Every kiosk is connected to the CCO. The equipments in the kiosk are more equipped and

sophisticated in cities than rural villages. Figure 2 depicts the model.

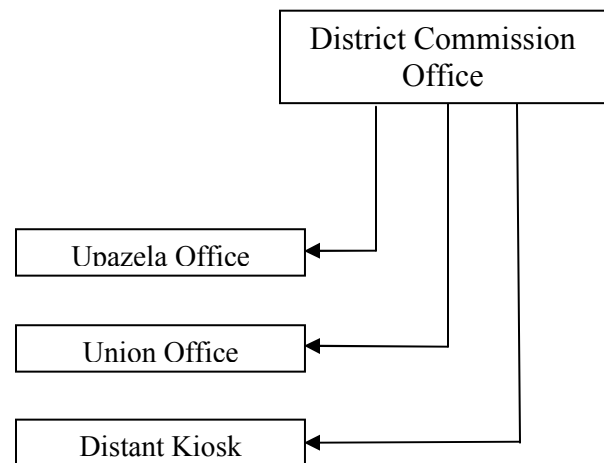


Fig. 1. Layout for rural areas

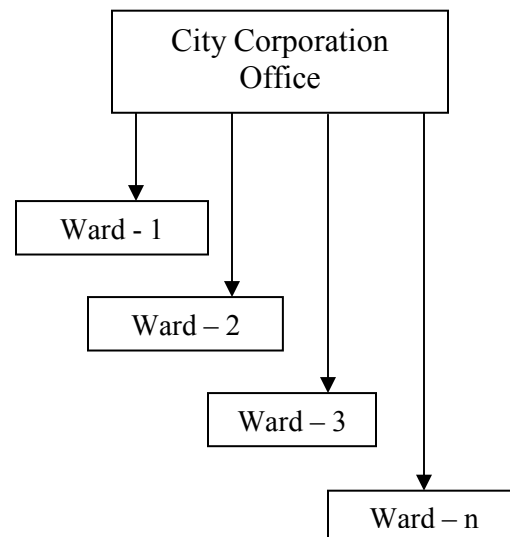


Fig. 2. Layout for city areas

### IV. NETWORK ARCHITECTURE

There is a big gap between infrastructure needed for a pervasive e-governance and what is in-fact present in rural and city areas [14]. So, information and communication network designer needs to measure the success rate in terms of technological breakthrough of the system under design [13]. The concept of e-governance can be translated into practice by creating required ICT infrastructure and introduction of appropriate operational strategies. Following are the few ICT network scheme:

- Public Network
- Private Network
- Virtual Private Network (VPN)



A public network is owned and operated by a service provider similar to a telephone service. The chief advantages of a public network are flexibility and the ability to use state-of-art networking without maintaining technical expertise. A public network is flexible because an arbitrary subscriber at an arbitrary location can connect to the network at any time. But the system can not provide sufficient security and it is not suitable for particular decisions or policies. A network is said to be private if use of the network is restricted to the corporate or individual owner. The chief advantage of a private network is that the owner has complete control over both the technical decisions and policies. But a private network can be expensive to install and maintain. A Virtual Private Network (VPN) combines the advantages of private and public networks by allowing a company with multiple sites to have the illusion of a completely private network and to use a public network to carry traffic between sites [6].

We will choose a VPN model for implementing e-governance. The model will use government telephone line for data transmission. In our country government telephone line is the most widely spread data communication line. Actually, telephone network is the most suitable means for rural areas that can be used for implementing e-governance. So, we need to establish a dial-up communication line through government telephone line. Figure 3 shows the network structure for rural areas, where Upazela Office, Union Office and Distant Kiosk are connected to District Commission Office through the telephone line. It is not hard or complex to implement a Virtual Private Network (VPN) that could provide sufficient security for such a project. VPN can be implemented with a combination of easily acquirable hardware and software [4], [7].

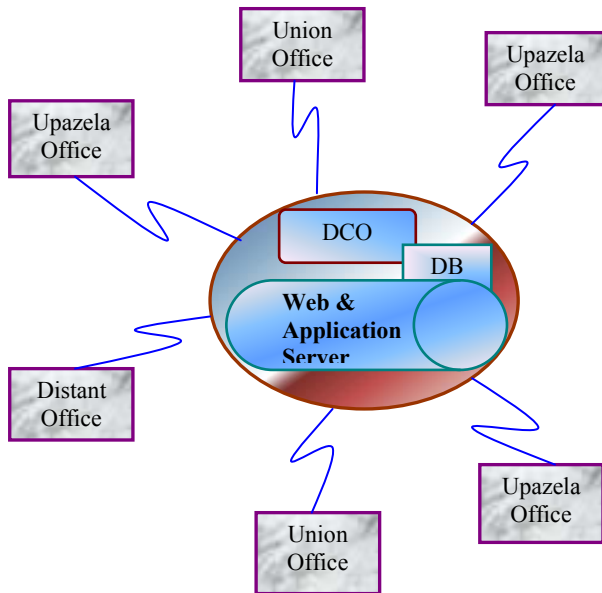


Fig. 3. Dial-up networks of rural areas

Table 1. Data rate and dial-up speed

Forms	Characters	Required data rate	Worst case dial-up connection
Minimum field content	300	2700 bps	26100 bps
Maximum field content	2800	25200 bps	26100 bps

In city areas every ward offices are connected to City Corporation Office. Figure 4 shows the network structure for city areas. The communication medium of rural areas and city areas is same that is telephone line. Actually, more sophisticated medium and equipment may be implied in city area as the demand of city people is more than rural people. Since telephone line is available in every place in city and we consider a minimized cost, we designed the model with dial up telephone line as villages. Figure 4 shows the structure of city areas.

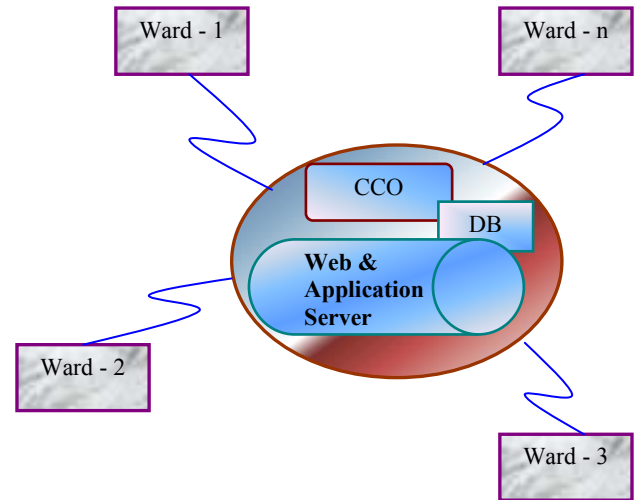


Fig. 4. Dial-up networks of city areas

## V. DATA RATE SPECIFICATION

The bandwidth of the telephone line should be considered to implement the model. To reduce the bandwidth, only data fields are transmitted to store it in the main database. Data rate is the number of bits per second and Baud rate (B) is the number of characters per second. A character frame consists of 8 data bits and 1 stop bit. So, character frame length is 9 bits. If modem speed is set at 33600 bps, then Baud rate,

$$B = 33600 / 9 = 3733 \text{ characters per second}$$

Therefore, 3733 characters per second can be transmitted through the telephone line. We consider that the worst case dial-up speed at distant places is at 26100 bps. Data will be transmitted successfully at this rate as described in table 1.

## VI. DATABASE AND SOFTWARE

To implement e-governance a huge class of software is needed in various developing stages. Many commercial softwares are available which should be purchased to implement the system. Of course, some software should be developed to provide desired services to the citizen. An important task is to create database. A large class of database is a significant criterion for such a project. One thing should be kept in mind during software development is that software should be flexible to use and be simple to suit with the existing communication infrastructure.

We have created sample databases, web pages and forms for our e-governance model [8], [16]. For convenience, we developed a virtual web site for Dhaka City Corporation. The real implementation would need rigorous analysis and planning to meet the people needs and corporation needs. We developed database, web site, and some forms for online job application. Figure 5 shows the view of the web site. The web pages are developed by html code and with the help of Microsoft Front Page.



Fig. 5. Web site of Dhaka City Corporation

There are several links of the main page of DCC website. The button “Apply for Job” is the link to the job application form. Clicking the button job form will appear. Users who are interested of the job they fill the form and submit online, then the data of the form is stored in the main database of the city corporation. We created a database named ‘JOB\_APPLICATION’ with php and mysql software. Figure 6 shows first portion of the job application form. This portion required personal information and career & application information about the intended candidate.

After submitting first portion successfully then second portion of the form will come. Figure 7 shows second portion which required information about academic and training information of the candidate. After filling the fields of second portion of the form, candidate has to submit it. If it is successfully submitted then last portion experience information will come.

PERSONAL INFORMATION			
User ID	00034		
Full Name	Md. Hasan	Sur Name	Md.
Father's Name	Md. Rahim	Sex	MALE
Date of Birth	1980-5-6	E-Mail	hasan@yahoo
Mailing Address	new eskaot, agraba		Marital Status
Home Phone no	03284543	Mobile No	018232322
Religion	Islam	Nationality	Bangladeshi
CAREER & APPLICATION INFORMATION			
Your Preferred Level of Position	<input checked="" type="radio"/> Entry Level <input type="radio"/> Mid Level <input type="radio"/> Top Level		
Available For	<input checked="" type="radio"/> Full Time <input type="radio"/> Part Time <input type="radio"/> Contract Job		
Present Salary	8000	Expected Salary	12000
Ready to Join	2006-1-1	Have you applied before?	No
Source of application	<input checked="" type="radio"/> Newspaper Add <input type="radio"/> Web Add <input type="radio"/> Friends <input type="radio"/> Others		
Main Reason for applying	It looks better for me		
Future Plan	Taking Masters degree		
Submit			

Fig. 6. First portion of the of the job\_applicant form

ACADEMIC INFORMATION					
User ID	3245555				
Name of Degree	Educational Institute	Year From	Year To	Class/CGPA	Major/Subject
SSC (SSC or Equivalent)	R. F. High S	1996	1998	First	Scienc
HSC (HSC or Equivalent)	R. F. High S	1998	2000	First	Scienc
BSC in e (Graduation)	Chittagong C	2001	2005	Secor	Econor
 (Post Graduation)					
 (Other)					

Submit

Fig. 7. Second portion of the of the job\_applicant form

At last, final and third portion of the form which required information about work experience will appear, if previous portion submitted accurately. Figure 8 shows the experience information portion. If this portion is completed and submitted successfully then all required information of the job application form would be gathered and stored into the main database of DCC.

## VII. SYSTEM IMPLEMENTATION

We focus to develop a flexible system as possible for developing our e-governance model. A cost effective and efficient system should be implemented to suit existing infrastructure. Special concentration should be given for rural areas due to weaker communication infrastructure in villages. Client and user software should be developed as simple as possible. Software interfaces and forms should provide information also in Bengali, so that rural people can easily understand the information and communicate with the government. Data can be submitted either in online or offline. Offline data submission reduces operational cost and also very flexible for a noisy and low bandwidth network. Data can be uploaded to central database 2 or 3 times in a day or more than that as required and availability of good connection. People can fill up the application forms

Work Experience			
User ID	565676		
Company Name	PHP	Position Held	Executive
Served From	2000	Served Till	2003
Address	ffdfs rrtrtt	Phone	03443233
Contact Person	Md. adfrfe hor	E-Mail	hasan@yah
Major Responsibilities	dfdf		
Achievement/Contributes	dfdffff		

Submit

Fig. 8. Last portion of the of the job\_applicant form

at their nearest kiosk. After submission of the form he would be given a unique token number from DCO or CCO for identification of the submitted application. Local kiosks are also capable of delivering status information for online queries from central databases. Kiosks should also provide all of the facilities as people required specially for cities such as scanner, printer, web camera etc.

## VIII. CONSTRAINTS AND RECOMMENDATIONS

There are some infrastructural constraints in rural areas which should be carefully considered before implementing e-governance. So, project designer needs to measure the success rate in terms of technological breakthrough of the system under design [13].

In rural areas often the telephone lines are idle due to some technical failure or too noisy for computer data transmission. Network bandwidth, data speed and a noise free network connection are desired for on-line data transmission. To reduce the cost of operation of dial-up ICT network and getting this dial-up connectivity active the form submission process should made offline also. To reduce the bandwidth, only data fields could be transmitted to store it

in the central database. These types of communication problems may not arise in city areas.

To establish e-governance following principles should be obeyed:

- Consider rural and urban infrastructure separately to implement the project
- Focus on disadvantaged communities, who otherwise will be excluded
- Provide that information or service which otherwise will not be provided
- Focus on utilizing and where possible building upon what is existing rather than thrusting a new intervention
- Create an outcome which in absence of ICT will not be produced efficiently or timely

## IX. CONCLUSION

The term e-governance is of recent origin and has become the buzzword in all developing countries. In fact, it is the use of Information and Communication Technology (ICT) for redesigning the traditional mode of governance, through better coordination and sharing of power among different components of government; thus taking care of the traditional problems of governance. The lack of information and communication infrastructure results in people having to waste time and money [5], [12]. Lack of clarity in processes, and corruption and mismanagement in systems and operations, is rampant. In this context, Information and Communication Technologies (ICTs) can play a significant role in making information available at a reasonable cost. ICTs promise to provide innovative solutions to the problems of poverty and inequality by accelerating development and introducing transparency into systems and operations.

At last, establishing e-governance is utmost important for our country. In this paper, we proposed a comprehensive and cost effective model of e-governance considering existing ICT infrastructure of Bangladesh. We hope that this paper could provide important guidelines and information to government and others who deal with e-governance.

## REFERENCES

- [1] Md. Shahidul Islam, M. Ismail Jabiullah, M. Lutfar Rahman, "M- Commerce Infrastructure Design and Development: Bangladesh Perspective", Paper No- 074, ICCIT 2004.
- [2] Dhaka City Corporation, <http://www.dhakacity.org/>
- [3] DigitalGovernance.org Initiative, <http://216.197.119.113/artman/publish/index1.shtml>
- [4] Cisco Systems, Inc <http://www.cisco.com/>
- [5] Sanjib Das, L. N. Sharma, A. K. Gogoi, "Enabling E-Governance for Low Bit Rate Network through Rural ICT Infrastructures in Indian Villages", Paper No-075, ICCIT 2004.
- [6] Douglas E. Comer, "Computer Networks and Internets", 4th edition.
- [7] Andrew S. Tanenbaum, "Computer Networks", 4th edition.
- [8] Jeffrey A. Hoffer, Mary B. Prescott, Fred R. McFadden, "Modern Database Management", 6th edition.
- [9] DG Communities, <http://topics.developmentgateway.org/egovernment?goo=1554>
- [10] E-Democracy, E-Governance and Public Network, <http://www.publicus.net/articles/edempubpublic network.html>
- [11] Global Information Infrastructure Commission, <http://www.giic.org>
- [12] UNESCO Communication and Information, [http://portal.unesco.org/ci/en/ev.phpURL\\_ID=1657&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.phpURL_ID=1657&URL_DO=DO_TOPIC&URL_SECTION=201.html)
- [13] V. Rajaraman, "Analysis and Design of Information Systems", 2nd edition.
- [14] Pankaj Sharma, "E-Governance", 2nd edition.
- [15] the-south-asian.com, <http://www.the-south-asian.com/Oct 2001/E-Governance.htm>
- [16] Dave W. Mercer, Allan Kent, Steven D. Nowicki, David Mercer, Dan Squier, Wankyu Choi, "Beginning PHP5", 2004 edition.

# SECURE ONLINE SEALED BID AUCTION

*Md. Nazrul Islam, Mohammad Zahidur Rahman*

Department of Computer Science Engineering,  
Jahangirnagar University, Savar, Dhaka.  
Email-: rmzahid@juniv.edu,  
mdnislam2006@yahoo.co.in

## Abstract

**Security and privacy have become crucial factors in auction design. Various schemes to ensure the safe conduction of sealed bid auctions have been proposed recently. We propose a secure sealed bid auction protocol. We introduce a new standard of privacy for auctions, that prevents extraction of bid information despite any collusion of participants. Present cryptographic protocol which is the realization of an electronic auction being the component of the e-government or e-commerce system. Our protocol provide non-reputation and correctness.**

**Keywords:** Sealed bid auction, Auction security, auction corruption, auction protocols.

## 1 Introduction

The Internet has been providing a commercial environment, where secure auctions are in demand. A usual type of them is a sealed bid auction. In the sealed bid auction, bidders submit their own bid prices to the auctioneer secretly, until the deadline. After the deadline, the auctioneer announces the winner and the clearing price. In the highest-price auction, the bidder who submitted the highest bid is the winner, and clearing price is the highest bid.

In the sealed-bid auction, two important properties are required. One is the bid privacy. The minimum is

the privacy until the deadline. Additionally, it is desired that the privacy continues even after the deadline, as pointed out in [6], since the information on the bids (i.e., prices) indicates bidders' preferences or powers. Furthermore, the statistics of bids in an auction help the auctioneer to learn the strategy of the bidders, which is useful for managing other auctions. The important property of privacy is Anonymity.

This paper is organised as follows: Section-2: discusses types of auction, Section-3: describe related works, Section-4: describes auction notation and terminology, Section-5: provides basic electronic auction protocol, section-6: discusses Security goals for sealed bid auction and section-8: provides conclusions.

## 2 Types of Auction

Auctions are means to transfer from one entity to another where the price of the goods is not easily determined and depends on the perception of the customer [23].

Auction can be classified according to the number of sellers and the manner by which bidders submit bids. An auction that has only one seller and many buyers is referred to a single auction. An auction with many sellers and many buyer's (i.e a securities market) is referred to as a double auction. The classification of Auctions are shown in bellow.

Bidding in an auction can typically be two types: sealed and open. In a sealed bid auction, secret bid (not known to other bidders ) are submitted to the

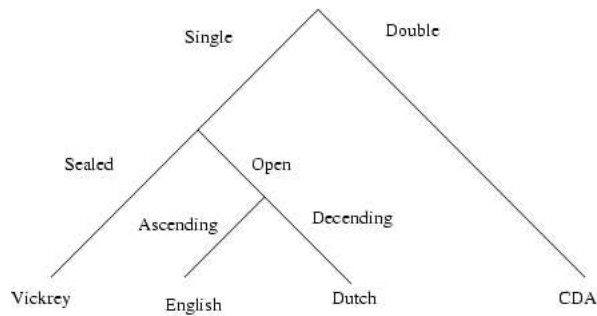


Figure 1: Taxonomy of Electronic Auction Schemes.

Auctioneer. Once the period closes, the bids are opened and the winner is determined according to some publicly known rule (e. g. ,the highest bidder wins) [6].

We look at different auction mechanisms for a sealed bid auction . We take the case of single - item auctions and study First Price and Second Price mechanisms, in this context. We note that both these mechanisms lead to an efficient allocation and under some simplifying assumptions, fetch the same revenue to the auctioneer.

First Price Auction mechanism, sealed bids are invited from all the bidders. The item is then allocated to the bidder with the highest bid, at his/her bid price

An English auction is a highest price, open bid auction. The bid amounts are known to all parties and bidders attempts to outbid each other. The highest bidder wins and must pay an amount equal to the wining bid. An English auction is the most well known type of auction and is widely used for real estate. An English auction is known as physical Auction. In physical auctions the responses must be received within seconds, while in cyber auctions it is conceivable that several minutes or hours will be allowed for the responses. English and vickrey auctions are considered as ascending due to the bidders bidding up the settlement price.

Alternatively an auction can be decreasing. This decreasing price (Dutch auction) of auction is similar to the English auction in that the bidding price varies over time; however, in this case, the price decreases

and at a time  $i$  is  $k \cdot i^* \nabla$ .

In a continuous Double Auction (CDA) there are many buyers and sellers who continuously trade a single commodity. CDAs are typically open bid. The best known example of CDA is a securities market

### 3 RELATED WORKS

Many papers consider secure sealed bid auction protocols. We shall now review selected proposals from the literature. The first cryptographic auction scheme was proposed by Franklin and Reiter [7]. The protocol uses a set of distributed auctioneers to evaluate bids and is based on secret sharing. The protocol prevents any single auctioneer from cheating during the auction, but reveals every bid to all the auctioneers once the winner is determined. so this scheme is not fully private, in the sense, that it only ensures that confidentiality of bids until the end of the protocol. Kikuchi, Harkavy and Tygar presented an anonymous sealed bid auction that uses encrypted vectors to represent bidding prices [8]. Harkavy, Tygar and Kikuchi proposed a Vickrey auction, where the bidding price is represented by polynomials that are shared by the auctioneers [15]. Kudo used a time server to realize sealed bid auctions [14]. Cachin proposed a sealed bid auction using homomorphic encryption and an oblivious third [4]. Sakurai and Miyazaki proposed a sealed bid auction in which a bid is represented by the bidders undeniable signature of his bidding price [13]. Naor, Pinkas and Sumner realized a sealed bid auction by combining Yao,s secure computation with oblivious transfer [16]. Juels and Szydlo improved this scheme [2]. Sako proposed a sealed bid auction in which a bid is represented by an encryption message with a public key that corresponds to his bidding price [21]. Kobayashi, Morita and Suzuki proposed a sealed bid auction that uses only hash chains [12][9] . Omote and Miyaji proposed a sealed bid auction that is efficient [18]. Kikuchi proposed an  $M+1$ -st price auction, where the bidding price is represented by the degree of a polynomial shared by auctioneers [11]. Baudron and Stern proposed a sealed bid auction based on circuit evaluation using homomorphic encryption [3]. Chida, Kokayashi

and Morita proposed a sealed bid auction with low round complexity [10]. Abe and Suzuki proposed a M+1-st price auction using homomorphic encryption [1]. Lipmaa, Asokan and Niemi proposed a M+1-st price with out threshold trust [17]. Omote and Miyaji proposed a M+1-st price auction using p-th residue problem [19]. For instant the electronic auction protocols have to guarantee confidentiality, privacy, integrity, non reputation anonymity of bidders. These security requirements are solved by the conception of the protocols using some cryptographic primitives. Achieved these properties correctly is not obvious and an automatic verification is useful to avoid some flows. There are many tools to check formally and automatically the cryptographic protocols [20].

Felix Brandt introduce a new standard of privacy for auctions ( full privacy) that prevents extraction of bid information despite any collusion of participants. His protocol is a verifiable bidder-resolved auction protocol. Shuo Zheng, Luke McAven and Yi Mu proposed another form of bidder resolved auction protocol [22]. Both protocol distributes their trust onto bidders divide their bids into n shares, keep one and send one share to each other bidder [20][5]. Felix Brandt states that the bidders are less likely to share their secrets since they are less likely to share their secrets since they are competing with other. Our protocol is based on the basic idea of n parts of a auction house resolved auction.

## 4 NOTATION AND TERMINOLOGY

- i) Auctioneer -:  $A_M$
- ii) Bidders -:  $B_i$  (  $i = 1, 2, 3, \dots, n$ )
- iii) Sub auctioneer -:  $A_i$  (  $i = 1, 2, 3, \dots, m$ )
- iv) Registration number -:  $ID_i$
- v) Public key -:  $PK$
- vi) Private key -:  $SK$ .
- vii) Digital signature -:  $DSig_i()$
- **Bid close time:** A time after which no new bids are allowed. In an English auction the bid

close time may be fixed or may be defined by a period of inactivity.

## 5 THE FIRST PRICE SEALED BID CRYPTOGRAPHIC AUCTION PROTOCOL

Our proposed sealed bid auction protocol contains four phases: Notification, Bidding, Opening and Winner selection. There are n bidders, one auctioneer and m sub-auctioneer. The role of auctioneer is to organize each auction run and announces the bid result at the end of the protocol run. He receives bids from bidders in a predetermined form and distributes the bids to m sub- auctioneers for the selection of the winning price and winner, then publishes the result. The channels between auctioneer and sub-auctioneers are supposed to be secure and reliable which means all messages transferred between two entities finally reach to the communicating party without compromising. We describe the protocol in details as follows.

### 5.1 Notification

The auctioneer publishes the details of the item for auction. A price vector,  $V = (v_1, v_2, v_3, \dots, v_n)$  is also published. The n prices satisfy  $v_i < v_j$ . The security parameter m,  $2 \leq m \leq n$  is also published. Each party interested in the item downloads the public certificate for the  $A_M$ , which includes the public key  $PK_A$  ( auctioneer public key) and the price vector V.

Every participant is supposed to have a particular identification say registration number to attend auction. Now here we describe a sub-protocol for registration. Each bidder generates a pair of keys ( $SK_{B_i}, PK_{B_i}$ ). To register they submit a note of interest with their own public key certificates, including their encryption and verification public keys. For each bidder  $B_i$ , the seller publishes the certificates and a unique identity  $ID_i$ , simply a non negative integer representing the order in which participants registered. Let N be the number of sub-auctioneers.

## 5.2 Bidding

Each registered  $B_i$  performs following operations:

- Selects a bidding price  $v_j$  from the vector  $V$  and determines  $R_i = H(v_j + Seq)$ .
- $A_M \leftarrow C_i = \{DSig_i(ID_i || R_i)\}_{PK_{A_M}}$  as a commitment of bid.
- Uses the rules below to generate a vector  $V_i = (H(v_j + Seq), H(v_j + Seq), \dots, R_i, v_{j+1}, \dots, v_{n-1}, v_n)$ . If  $B_i$  selects price  $v_j$  then the  $j^{th}$  value in  $V_i$  equals  $R_i$ . Higher indexed positions have the same value as the corresponding position in  $v$ , while lower indexed positions have values  $H(v_j + Seq)$ .
- Divides  $V_i$  into  $m$  shares  $V_{ik}$ ,  $1 \leq k \leq m$  such that  $V_i = V_{i1} + V_{i2} + V_{i3} + \dots + V_{im}$ .
- Each bidder  $V_i$  is send to the auctioneer with identity  $(ID_i + k)(mod N)$ . Then auctioneer each  $V_{ik}$ ,  $1 \leq k \leq m$  is send to the sub-auctioneer with  $(ID_i + k)(mod N)$ .
- Each participant sums all  $m$  parts and submits the result to the auctioneer as  $V'_i$ .

## 5.3 Opening

The auctioneer perform the following-:

- Adds all incoming shares together to recover 
$$V' = \sum_{i=1}^N V'_i.$$
- Evaluates the vector  $W = V' - NV$  and finds the index  $w$  of the first non zero value in  $W$ , reading from right. The highest bidding price is the value  $(V_w)$  from  $V$ .
- Calculates  $B_{id} = W_w - (N - 1)v_w$ .
- Opens all the commitments  $C_i$  and locates the winner who has the same commitment as  $B_{id}$ .

## 5.4 Winner selection

Let  $w$  is the index of the winner. In the full privacy model the  $A_M$  publishes the winners signed commitment  $Dsigs(R_w)$ , so every one can verify whether the revealed  $R$  is the same as that committed to the seller.

In the partial privacy model the seller publishes the winner's price  $V_w$  and also  $Dsigs(ID_w || R_w)$ . The winner is asked to publish the random  $r_w$  chosen to create  $R_w$ , so that each participant can verify the correctness of the result by comparing  $R_w$  with  $H(V_w || r_w)$ .

## 6 SECURITY GOALS FOR SEALED BID AUCTIONS

Sealed bid auctions possess unique security goals due to the fact that bids must be sealed during the bid submission. The following outlines the additional security goals for sealed bid auctions:

- Confidentiality**- Once a submitted bid has been sealed, it must remain secret until the winner determination phase of the auction.
- Unforgeability**- Nobody can impersonate a certain bidder.
- Anonymity**- Nobody including an authority can identify the loser bidders even after the opening phase.
- Non-Reputation**- A winner cannot deny that she/he submitted the highest bid after the winner decision procedure. A winner is exactly identified.
- Fairness**- All bids should be fairly dealt with.
- Efficiency of bidding**- The computational and communicational amount in both bidding and verifying a bid is practical.
- Privacy of Losing Bids**- The winner determination phase, the values of the losing bids must be kept secret. Further questions also arise over how much information is disclosed regarding the winner and winning bid. For example, should only the seller and winner know the



amount of the winning bid? What information should subprotocol the Auctioneer learn about winning/losing bids?

- (i) **Unlink ability**-The Auctioneer should not be able to learn information about individual bidders based on previous auctions conducted. This information could be used in future auctions in a manner that disadvantages the bidder.

Our protocol satisfy all the above conditions.

## 7 Conclusion

Electronic auctions are continuing to grow in popularity, but fraud seems to be a major problem. We presented a secure first price sealed auction protocol, where information is shared among part of the auction house (Servers). Our protocol achieving correctness, privacy, non reputation and efficiency.

## References

- [1] Masayuki Abe and Koutarou Suzuki. M+1-st price auction using homomorphic encryption. *Proceedings of public Key Cryptography 2002*, 2002.
- [2] A.Juels and M.Szydlo. A two server, sealed bid auction protocol. *In proceedings of Financial Cryptography 02*, 2002.
- [3] O. Baudron and J. Strn. Non-interactive private auctions. *Proceedings of Financial Cryptography 2001*, 2001.
- [4] C.Cachin. Efficient private bidding and auctions with an oblivious third party. *Proceedings of 6th ACM Conference on Computer and Communications Security*, pages 120-127, 1999.
- [5] Cas Cremers. Scyther- semantics and verification of security protocols. *Ph D thesis, Eindhoven University of Technology*, 2006.
- [6] Michael Harkavy Doug Tygar and Hiroaki Kikuchi. Electronic auctions with private bids. *Proc. 3rd USENIX Workshop on Electronic Commerce*, pp.61-74, 1998.
- [7] M. Franklin and M.Reiter. The design and implementation of a secure auction. *IEEE Trans on Software Engineering*, Vol.22,No.5, 1996.
- [8] Michael Harkavy Hiroaki Kikuchi and J. D. Tygar. Multi-round anonymous auction protocols. *Proceedings of first IEEE Workshop on Dependable and Real-Time E-Commerce Systems*, page 62-69, 1998.
- [9] H. morita M. Hakuta K. Suzuki, K. Kobayashi. Efficient sealed-bid auction by using one way functions. *IEICE Trans Fundamentals*, E84-A(1), 2001.
- [10] K.Kobayashi K.Chida and h. Morita. Efficient sealed bid auctions for massive numbers of bidders with lump comperison. *Proceedings of ISC 2001*, 2001.
- [11] Hiroaki Kikuchi. M+1-st price auction protocol. *In Financial Cryptography 2001*, 2001.
- [12] Kuni Kobayashi Koutarou Suzuki and Hikaru Morita. Efficient sealed-bid auction using hash chain. *Proceedings of International Conference Information Security and Cryptography Conference (FC-02)*, 2002.
- [13] K.Sakurai and S. Miyazaki. A bulletin-board based digital auction scheme with bidding down strategy. *Proceedings of 1999 International Workshop on Cryptographic Techniques and E-Commerce*, pages 180-187, 1999.
- [14] M. Kudo. Secure electronic sealed-bid auction protocol with public key cryptography. *IEICE Trams. Fundamentals*, E81-a(1), 1998.
- [15] Hiroaki Kikuchi Michael Harkavy, J.D TYgar. Electronic auctions with private bids. *Proc. 3rd UNENIX Workshop on Electronic Commerce*, pp.61-74, 1998.

- [16] Benny Pinkas Moni Naor and Reuben Sumner. Privacy preserving auctions and mechanism design. *In Proceedings of the First ACM Conference of Electronic Commerce (EC-99), pages 129-139*, 1999.
- [17] H.Lipmaa N.Asokan and V.Niemi. A two server sealed bid auction procol. *In proceedings of Financial Cryptography 02*, 2002.
- [18] K. Omote and A. Miyaji. An anonymous auctionprotocol with a single nontrusted center using binary trees. *Proceedings of ISW2000, pages 104-120*, 2000. LNCS 1975.
- [19] K. Omote and A. Miyaji. A second-price sealed-bid auction with the p-th root. *In proceedings of Financial Cryptography 02*, 2002.
- [20] S. Etalle R. Corin and A. Saptawijaya. A logic for constant-based security protocol analysis. *In IEEE symposium on security and privacy*, 2006.
- [21] K. Sako. Universally verifiable auction protocol which hides losing bids. *Proceedings of Public Key Cryptography 2000, Pages 35-39*, 2000.
- [22] Luke McAven Shuo Zheng and Yi Mu. First price sealed bid auction without auctioneers. *IWCM'07*, 2007.
- [23] Bakoy's Y. The emerging role of electronic marketplaces on the internet. *Communications of the ACM, Volume 41, Number8,pp35-48*, 1998.

# Short Messaging Service (SMS) Based m-Banking System in Context of Bangladesh

Md. Subrun Jamil, Fouzia Ashraf Mousumi<sup>1</sup>

Chittagong Online Limited, Chittagong-4100, Bangladesh.

<sup>1</sup>Computer Science & Engineering, University of Chittagong-4331, Bangladesh.

**Abstract** — m-banking has become one of the most familiar banking service providing technologies in different western countries. Now-a-days billions of inhabitants of Bangladesh are within a network through mobile network coverage. But in the commercial sectors like banking, m-commerce technology has not been adopted broadly yet. Considering m-commerce perspective in Bangladesh a SMS based m-banking system has been proposed which is able to provide several essential banking services only sending SMS to bank server from any remote location. This proposed system is divided into five major phases: Interfacing Module, SMS Technology Adoption Module, SMS Banking Registration Module, Service Generation Module, and Data Failover Module. This system facilitates bank customers by providing four major services like balance enquiry, balance transfer between authenticated customers, DPS payment and bill payment without going to bank physically and save their precious time. At least, after evaluating each module of this developed system a satisfactory accuracy rate 93.18 % is obtained.

**Keywords** — Balance Enquiry, DPS, Data failover, m-banking, SMS

## I. Introduction

The physical world gradually becomes more and more integrated due to the vast development of information and communication technologies. The ICT departments of different modern organizations e.g. Banks, Commercial or non-commercial firms etc are struggling to keep up with the demand for computing, rapid technology changes and the fundamental transformation of work [1]. As a result, the enhancement of Electronic Commerce technologies (E-Commerce) has increased day by day which impacts have been on our economic, social infrastructure E-Commerce brings the universal access of the Internet to the core business process of buying and selling goods and services [2]. Electronic Banking – the execution of financial services of a Bank via the Internet – is one of most appreciable example of E-Commerce. Because it has changed the business of retail banks significantly, at the same time reducing costs and increasing convenience of customer

[3]. Since cellular networks has become an essential component of the economic and social infrastructure of developed and developing countries, a new subset of electronic banking has created, mobile banking (m-banking). M-banking refers that technology which covers account management via mobile devices [3].

Communication purpose email and messaging have become part & parcel of our daily life. But it is true that most of the people of Bangladesh, basically in the rural areas have lack in Internet skills, access to a personal computer and on top of all these, we lack in sufficient Internet infrastructure. But these people of rural areas have a remarkable contribution for earning foreign exchange. Due to the huge advancement of our telecommunication sector, mobile devices have become prominent to every level of people [4]. There are about 1.5 billion people across more than 210 countries and territories, 15 million mobile users in Bangladesh and it is only 10% of our total population. In our country, not only the price of mobile sets gradually reduced but also the cellular operators are providing an acceptable call rate to the subscribers. Even in rural and Chittagong Hill Tract areas, the mobile network coverage has been made available. Day by day's people accepts mobile not only their communication device but also an information transfer media which is highly necessary for their livelihood [5]. Short Messaging Service (SMS) has become the most popular and cheapest service on our mobile phones. It permits us sending message of alphanumeric characters between mobile phones. In Bangladesh the numbers of GSM mobile subscribers are 42.0 million at the end of June, 2008 while the number was 25.39 million at the end of May, 2007. Considering these issues, SMS technique can be integrated in our banking sector. It is the nature of human being to get services within short time. In our country, several private banks have introduced the facilities of online banking, phone banking and at last SMS banking with too limited services. But our nationalized banks i.e. Sonali Bank, Janata Bank, Agrani Banks and other specialized financial institutions like Krishi Bank do not provide such

facilities. A huge amount of people are the customers of these banks.

In this paper, an approach is proposed to develop a system with the integration of SMS technology to automate our banking system which is convenient and hassle free to use. Through this proposed SMS based banking system, a bank customer can be able to get different services only sending text SMS to SMS server while that person is in remote or not been able to attend the bank physically. It saves not only customers time and energy but also gives a mental satisfaction, by giving all updated information about all of his/her banking transactions. The key challenges of m-banking are security, usability and data failover [3]. In our proposed system these key issues are handled significantly.

## II. Related Works and Motivation

The filed of m-banking is not very old. According to the sweeping enthusiasm that characterized much of the news reporting in the years 1999 and 2000 mobile banking should by now have been firmly established as the most important distribution and communication channel for Retail Banking [3]. Dukic and Katic developed such a m-order payment model via SMS for the Republic of Croatia, which give emphasis for increasing security of the user requested information by Steganography method [6]. In developing countries like Iran has been welcomed M-banking. Shinali et al. has developed a system which is capable to be connected through Bluetooth which is installed in the bank and offered banking services. But their developed system, only works within the bank are not from the remote location [7]. Like other countries, in Bangladesh several research works have been done. Mahfuz et al. described in their paper the prospect of using SMS technology as an effective tool in B2E (Business to Customer Commerce) [5]. A. Azi et al. analyzed existing systems for the aspect of developing

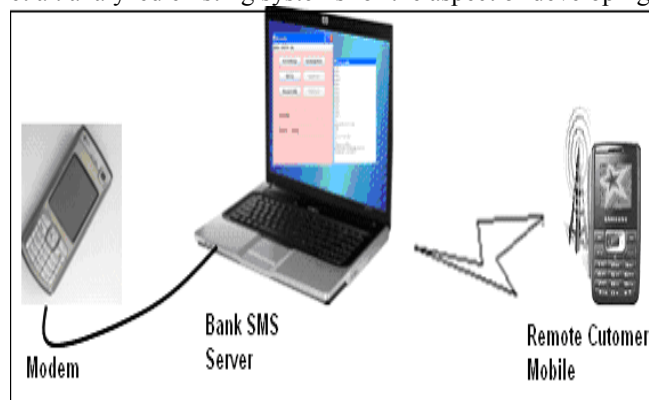


Fig.1. The schematic view of our proposed SMS based m-Banking system

countries, focused on limitations and presented alternative frameworks for e-commerce and electronic payments[2].

From the survey on different research work it is found that in western and developing countries - e- banking, M-banking systems are developed according their country perspective. In Bangladesh, researchers found out SMS as m-banking tool, but yet no remarkable standard SMS banking service model has been determined. In this paper we took an attempt to develop a real time banking service specified SMS based mobile banking system in context of Bangladesh.

## III. The Proposed SMS Based m-Banking System

The basic objective of this paper is to develop a SMS based mobile banking system for Bangladesh which is able to provide several current services that are manually provided by our Nationalized and Private Banks. In our proposed system, when a customer sends a SMS of enquiry or service request to bank, the bank server acquires the SMS through modem. Then the SMS server truncates necessary information from SMS and provides corresponding services to customer by retrieving and updating customer accounts related database. In this system, proper tracks have been kept to face any type of data type of data failure problem due to server failure or unusual behavior of system control software. The proposed system is divided into five major modules are as follows: A) Interfacing Module B) SMS Technology Adoption Module C) SMS Banking Registration Module D) Service Generation Module and E) Data Failover Module. Some well defined sequential methods are applied in these modules. The schematic view of the architecture of our proposed system is depicted Fig.1.

### A. Interfacing Module

In this module, we presented an interfacing technique between bank server PC and server mobile for our proposed system. At first GSM modem has been interfaced with Bank server PC by Bluetooth or RS-232C cable using COM or serial port. The NokiaN70, SimensC-35 and Wavecom mobile devices has been used as server modem of our proposed system.

### B. SMS Technology Adoption Module

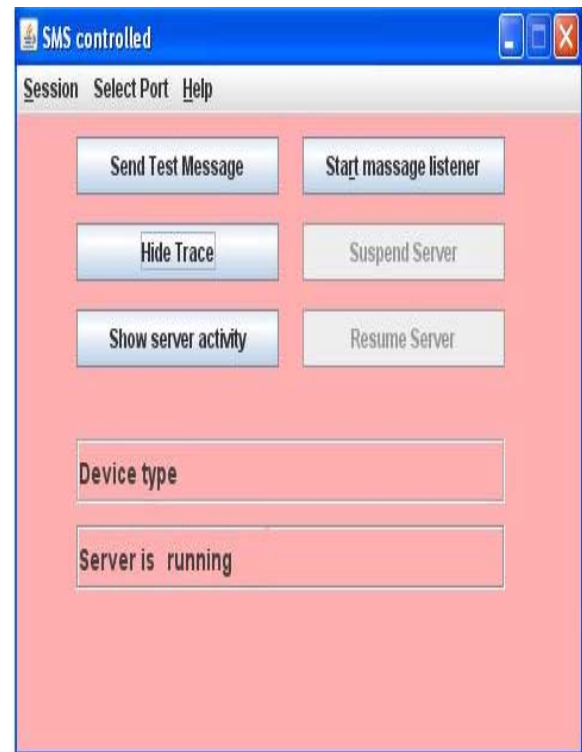
Short Messaging Service (SMS) is a communication protocol allowing the interchange of short text messages between mobile telephone devices. Most SMS messages are mobile-to-mobile text messages, though the standard supports other types of broadcast messaging as well [8]. The SMS is a store and forward service. In other

words, short messages are not sent directly from sender to recipient, but via an SMS center. Each mobile telephone network that supports SMS has one or more messaging centers to handle and manage the short message [9]. This module describes the process of communicating with the GSM modem which is the most important provider of our SMS service. This process of communication has been done through a number of attention commands (AT). This AT commands makes our bank server modem to be able to send and receive SMS from customers mobile and vice-versa. AT commands are used to control a modem which stands for Attention. GSM /GPRS modems and mobile phones support an AT commands set that is specific to the GSM technology, which include SMS. Table I indicates several common AT commands which are used in this proposed system.

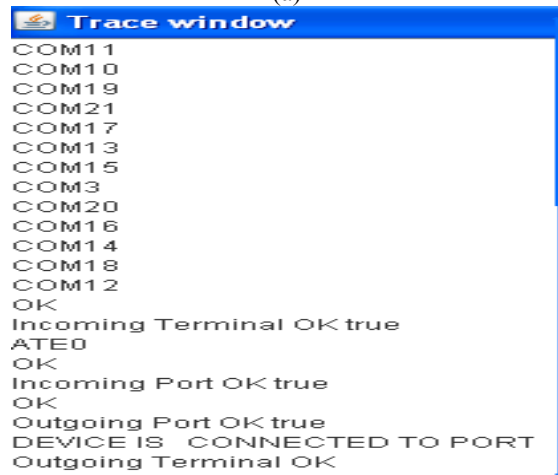
Table I  
List of Several AT Commands

Commands	Description
AT+CMGS	Send SMS Message
AT+CMSS	Send SMS Message from storage
AT+CMGL	List SMS Messages
AT+CMRR	Read SMS Messages

To check the connection between Bank server desktop/laptop computer and GSM mobile/modem, hyper terminals is opened and AT commands are sent. If the modem replies with OK then it can be stated that communication between modem and PC/Laptop are correct. GSM modems can have different AT commands specific of their type but all most all modems supports minimum common commands. The SMS specification has defined two modes in which a GSM/GPRS modem or mobile phone can operate. They are called SMS text mode and SMS PDU (Protocol Data Unit) .Usually string values are defined for text mode while numeric values are defined for PDU mode. We used PDU mode as it supports more features than text mode like it includes detailed Meta information with the raw SMS like actual SMSC number, time stamp, and length of the actual message in hexadecimal. With the help of J2EE and XML for sending and receiving SMS from bank server software to customer mobile an SMS server has been designed. In Fig.2. the interface of the bank server software for the activation of SMS server which is adopted throughout this SMS technology adoption module is represented.



(a)



(b)

Fig.2. The interface of bank server software (a) SMS server activation and (b) tracing window during modem connection establishment

### C. SMS Banking Registration Module

To activate the SMS based m-banking system and getting its facilities, the customers have to be registered according to our proposed model. For this, at first the bank server modem broadcasts SMS giving a notification of their SMS based M-banking service. When a customer wants to register for this service he/she has to send SMS to bank with account number and password. Through the server modem the required information is truncated automatically from the customer SMS and stored in bank's central database. If the account number of the customer is valid, then that customer has been registered for all time and server sends a positive acknowledgement. In Fig.3. a customer registration request process is shown.

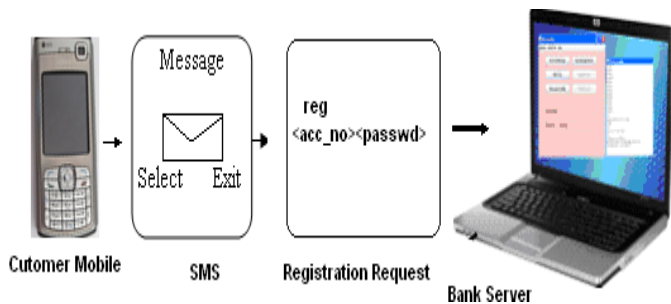


Fig. 3. The SMS request for registration from remote customer mobile

### D. Service Generation Module

Service generation is the most important module of our SMS based banking module. This module composed of four services: Balance Enquiry, Balance transfer between authenticated account holders, DPS utility and Bill Payment. For getting these services a customer of bank must be registered and have to send SMS in pre-specified format. In case of Balance Enquiry service, when a customer sends a SMS for checking current account balance, the bank server modem replies through SMS with balance information. But before replying to customers, the authentication procedure is

performed by matching his/her registered mobile number and password.

The next service of our system is balance transfer only between pre-authenticated account numbers of the same bank that are specified by the account holder. This authentication has been done automatically by extracting related information from customer SMS and matching with bank's central database. In our system, after the transaction, concurrent balance update of both account holder have been done automatically and the bank server sends updated account status to both account holders.

Another effective service of this system is deposit payment. This service gives alert to the deposit account holder before one week of his/her deposit amount payment deadline.

In Bill Payment Service different payment types such as mobile bill payment, electricity bill payment etc are implemented. Before paying bills, the status of current balance of the customer have been checked in parallel with registered mobile number. The whole system working procedure is represented by a flowchart in Fig. 4.

### E. Data Failover Module

The data failover module is the most essential and a new inclusion of our SMS based M-banking system. We have developed an algorithm for overcoming the data failure which can be occurred due to bank server failure during the customer SMS information processing. According to this algorithm if server fails on the halfway of a SMS request processing, at first server stores that message with a not processed flag like NO. Then another flag counts the number of total unprocessed SMS. When server again restores, the first unprocessed SMS request is processed and reply is sent to the customer. Then check for other unprocessed SMS and process one by another. Thus this system overcomes the difficulties in case of SMS data failure due to server unusual failure.

#### IV. Experimental Results and Analysis

For our proposed system, application software has been implemented using J2EE and XML. This application software will be installed in the Bank Server PC which contains the banking database of customers. This central Banking database contains all related information of customers for SMS banking facilities. The accuracy of the system is measured by testing major modules with a set of real time SMS data. The accuracy of the interfacing module is tested in case of both wireless and guided media. The success this module is 90.78% and failure rate is 9.22%. And from analyzing the output it has been found that, the SMS

delivery and receive will not hamper while interfacing media has been changed. We have also tested our interfacing unit using different types of modems as shown in Table II. Similarly SMS Technology Adoption Module, SMS Banking Registration Module, Service Generation Module, and Data Failover Module are evaluated by 300 short text SMS using the following equations:

Success Rate = (Total no of Success/ Total no of sample input)\*100) %. Failure Rate = (100 – Success Rate) %

Accuracy Rate = (100 – Failure Rate) %

Table II  
Experimental results of interfacing module

Server Modem Name	Interfacing Media (Between Server PC & Modem)	No of SMS Delivery & Reception	Successfully Delivered & Received	Success Rate	Failure Rate
Nokia N70	RS232C Bluetooth	50 50	47 47	94%	6%
Siemens C35	RS232C	30	28	93.33%	6.67%
Wavecom	RS232C Bluetooth	40 40	34 34	85%	15%
Average				90.78%	9.22%

Table III  
Result of overall SMS based m-Banking system success and failure rate

Modules	Success Rate	Failure Rate
Interfacing Module	90.78%	9.22%
SMS Technology Adoption	91.58%	8.42%
SMS Banking Registration Module	95.89%	4.11%
Service Generation Module	94.66%	5.34%
Data Failover Module	93%	7%
Average	93.18%	6.82%

## V. Conclusion

In today's busy environment, with so many deadlines to fulfill, appointments to keep and meetings, customers of a bank have little time to visit a branch or to wait for customer services operator to come on phone line. Though SMS Banking in Bangladesh has just been started but this telecom integration with Banking is not yet in full motion. So in this paper, an attempt is taken to develop a SMS based mobile banking system for 24 hours banking convenience which helps customers stay on top of any recent changes made in their current or deposit account through Short Messaging Service (SMS). One of most attractive feature of our proposed banking system is that balance transfer from one account to another of different branches of same bank without attending the bank physically. This system not only warns the customer deposit holders to pay their premiums in due date but also provides the track to pay this deposit premium by sending SMS in a pre-defined format from remote. Through our proposed system, customers would not face problems due to server failure while the server was in processing with customers SMS. Because if bank server become out of order due to some accidental causes, a track of data failure has been kept to handle this situation. In future we have a plan to add more services in our proposed system to enrich flexibility, usability and security of the proposed SMS based system accordance to Bangladesh perspective.

## REFERENCES

- [1] Adel Al -Taiton and Carsten Sorensen, "Supporting mobile professional in global banking : the role of global ICT-support call centers", *IEEE Trans. 37<sup>th</sup> Hawaii International Conference on System Sciences*, 2004.
- [2] A.Azim, A.F. M.S. Kabir and M.F. Zibran, "Alternative frameworks of E-Commerce and electronic payment systems specially suitable for the developing countries like Bangladesh", *ICCIT 2005*.
- [3] K. Pousttchi and M.Schurig "Assesment of today's mobile banking applications from the view of customer requirements", *System Sciences, IEEE Trans, 37<sup>th</sup> Annual Hawaii International Conference*, 2004.
- [4] S. Balasubramanian, R.A. Peterson and S.L.Jarvenpa, "Expolring the implications of M-commerce for markets and marketing", *Journal of Academy of Marketing Science*, PP-348-361, 2002.
- [5] Md. Mahfuz Ashraf and Shusmita Haque, "Short messaging service as a Business to Customer marketing tool: A proposed model in context of Bangladesh ", *ICCIT-2005*
- [6] B. Dukic, Katic,M. "m-order payment via SMS within m-banking" *27<sup>th</sup> International Conference , ITI 2005*, June 20-23, 2005, Cavtat, Croatia.
- [7] Shirali-Shahreza, M. Shirali-shahera, M.H. "Mobile banking Services in the bank area", *SICE* , 2007.
- [8] [http://www.wikipedia.org/wiki/Short\\_message\\_service](http://www.wikipedia.org/wiki/Short_message_service)
- [9] <http://www.mobilecomms-technology.com/projects/sms>

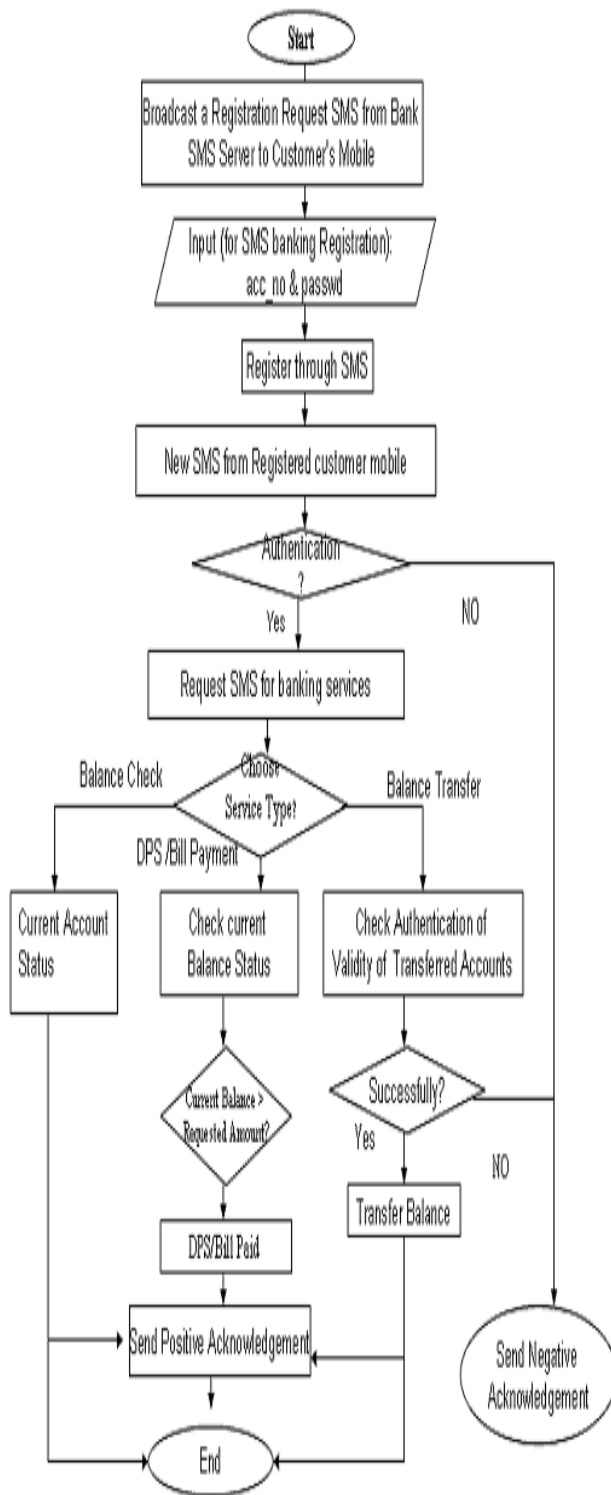


Fig. 4. The flow chart of our proposed SMS based m-Banking system

The whole system performance rate is measured by evaluating each module success or failure rate which is depicted in Table III. The overall system accuracy rate is obtained 93.18%.



# A study of major Mobile payment systems' functionality In Europe

Shahriyar Mohammadi, Hediye Jahanshahi

IT GROUP-Faculty of Industrial Engineering-K.N toosi University of Technology

**Abstract** — Mobile payment is an exciting domain, which will rapidly evolve in the years to come. The key to mobile payment acceptance is in the hands of customers. In this paper first some of the major European mobile payment systems are introduced and evaluated base on security, cost, convenes and functionality requirement criteria, then the results of past researches about the customer and merchant adoption of mobile payment systems (mp) are reviewed and barriers and drivers from both customer and merchant view are compared. Here by, this paper tries to find shortcomings in major mp systems and presents the major barriers and drivers to wide spread adoption of mobile payment from both customer and merchant point of view.

**Index Terms** — mobile Payment system ; Customer adoption ; Mobile commerce.

## I. INTRODUCTION

With the flourishing of electronic commerce and widespread use of mobile devices, a new type of service is emerging that extends e-business using wireless technology by enabling e-commerce services on mobile devices [1]. Mobile payments are payments for goods, services, and bills with a mobile device (such as a mobile phone, smart-phone, or personal digital assistant (PDA)) by taking advantage of wireless and other communication technologies [2]. The key to mobile payment acceptance is in the hands of customers [3]. The most important problem to broad acceptance and usage of an mobile payment (MP) procedure is the obtainment of a critical mass. This tends to be an hen-egg-problem: On the one hand customers will not use the procedure unless a significant number of merchants accepts it, on the other hand merchants will not be willing to accept the procedure unless a significant number of customers uses it [4].

## II. MP PROCIDURES

We A MP transaction can take place in four different general settings, distinguished with regard to the point of sale or service [5]:

- The mobile commerce scenario (MC): New applications and services, e.g. context-sensitive information or video/audio streaming on mobiles

- The electronic commerce scenario (EC): All kinds of B2C and C2C excluding MC, e.g. purchase of goods or content via the Internet

- The stationary merchant scenario: "Bricks and mortar" commerce with transactions between a person (customer) and a person (e.g. cashier, taxi driver) respectively or a machine (e.g. vending of tickets, cigarettes)

- The customer-to-customer scenario (C2C) : Money transfers between individuals, e.g. for settling an E-Bay transaction, as pocket-money for children, for settling debts for small amounts

### A. Customer requirements to MP procedures

MP procedure will be widely accepted by customers if it [6]:

- covers all payment scenarios,
- ensures confidential treatment of data,
- issues payment confirmations via SMS or e-mail,
- incurs none or little costs for customers,
- does not oblige customers to buy a new mobile device,
- easy to handle and responds quickly,
- accepted by a high number of merchants suitable for the execution of bank transfers,
- Suitable for Pico payments (amounts below 20 cents).[6]

At present, there are many examples of successful mobile payment applications such as the mobile content market which has developed into a billion dollar business, PayPal. Wider adoption of mobile payments, however, has not been as rapid or widespread as expected and there is a long list of discontinued mobile payment services such as the Simpay initiative and the Paybox launched in several European countries[7]. There is some debate about types of m-payments. Some parties see m-payments as a new product, whilst others see m-payments as a new service [5]. The European Central Bank categorizes m-payments as m-payments, a new payment initiative that is accessible via mobile phones. The funds, defined as electronic money (or e-money), "an electronic store of monetary value on a technical device that may be widely used for

making payments to undertakings other than the issuer without necessarily involving bank accounts in the transaction, but acting as a prepaid bearer instrument,” are used to pay for products and services [5].

### III MOBILE PAYMENT SYSTEMS IN EUROPE

In this section, we briefly describe some of the MP systems in Europe, Pousttchi and zenker[6] uses a frame work and evaluate 3 main mp procedures in Germany. Factors in their frame work are: “Applicability in all scenarios”, “security”, “cost”, “convenience” and “special functionality requirements”. “security “ includes not only integrity, authorization, authentication, confidentiality and non-repudiation of transactions, but also the issue of subjective security from the customer’s perspective, “costs” includes direct transaction costs and fixed costs of usage plus the cost of the technical infrastructure for the customer (e.g., a new mobile phone), “convenience” includes any issues related to ease and comfort of use[8] . in this paper their frame work was used to analyze mp system in Finland, Germany ,Serbia and compare their functionality, then try to analyze the current deadlock on the mobile payment market in order to develop a set of requirements to an integrative solution for a Mobile Payment System .

#### A. FINEID in Finland

The Finnish electronic identity (FINEID) card along with the supporting infrastructure [1] is a system maintained by the Population Register Centre (PRC) of Finland. The card is a usual microprocessor smart card accepted as a valid ID in all EU countries and several others. The card contains two Citizen Certificates: the authentication certificate of the card holder and the signature certificate of the card holder. Private keys of both certificates are generated by the card and stored exclusively in its protected memory.

**Security:** The Finnish National PKI infrastructure, which is used in FINLAND mobile payment system for authenticating users and for ensuring the non-repudiation of payments. Signatures produced by the FINEID card are legally equivalent to the card holder’s handwritten signature .Additionally, the card contains the PRC’s own Certification Authority (CA) certificate. **Cost:** A mobile subscription customer can request a FINEID compatible PKI-SIM card from their mobile network operator and register it at a police station. Authentication cannot be done without the participation of the mobile network operator, and naturally, operators charge both customers and service providers for authentication. Apparently, this is a major threshold for joining the system

**Convenience:** The current implementation of FINEID on SIM cards is based on a platform developed by SmartTrust The advantages of the SmartTrust platform include its compatibility with almost any GSM mobile phone, and the availability of all the infrastructure needed. Each registered individual gets a unique Finnish Electronic User ID (FINUID). The public keys of each user can be downloaded via a search with the appropriate criteria. Moreover, a revocation list of invalid certificates is available from the FINEID directory [1].

**Functionality Requirements:** FINEID certificates are only available to Finnish citizens and those permanently living in Finland. At the moment, a legal entity such as a merchant or a bank cannot obtain a FINEID certificate[1].

#### B. DINA Card in Serbia

The National credit card, DinaCard, is being issued by almost all banks in `Serbia. That gave a possibility for connecting existing mobile payment system to one point, National center for payment cards (further NCPC), instead of connecting it to each bank separately, which is on the other side connected to all the banks that are issuers of DinaCard. This will give opportunity to many users who didn’t have a chance to use SMS top up, because their banks didn’t have an agreement with mobile operator, to register and start using the service. the system covers Ec ,MC and POS scenario.

**Security:** All the files that are exchanged between NCPC and core of the system are encrypted by file secure program and smart card reader. Encryption and decryption is done by authorized personnel. Personal data about registered users are firstly checked by the bank and NCPC, and then imported automatically into the core database, in order to minimize the possibility of error during manual import. When new user is imported for the first time, mPIN is generated and sent to his registered mobile phone number via SMS. In the database, mPIN is stored in encrypted form, and there is no possibility to view it.

**Cost:** Mobile payment services can be enabled to all

Subscribers of mobile network operator who are owners of DinaCard ,Registration process starts in the bank that issued DinaCard.

**Convenience:** Interface to applications is unified for all types of services. It is based on http and has several methods that can be used in order to realize payment , Single mPIN is used for all defined services and accounts. Mobile payment services are enabled to the user from the moment mPIN is received.

### *C. Mobile Payment systems in Germany*

#### *a. Pay Box*

The most well-known MP procedure in Germany is Paybox. The operator of the procedure was a technology driven payment startup which falls in the category of specialized intermediaries. Having started in May 2000 the service was the first mover in its field. Pay box acted not only as a payment intermediary, but also actively acquired merchants and offered them the technical solution to integrate Pay box in their systems. Within 3 years Paybox enrolled more than 500,000 customers and 6,500 merchants. Payments were authorized using IVR (Interactive Voice Response)[9].

The particular way of authenticating payments imposed restrictions on the use of Paybox within the MC scenario. For example, Paybox could not be used for WAP services since the customer during a WAP session could not be called to conduct the authorization. Paybox was especially targeted to the stationary merchant and EC scenario.

**Security.** The confidential treatment of data was vouched for by Paybox. For the confirmation of payments a number of options were supported depending on the kind of transaction and on the payment scenario, including SMS and e-mail.

**Cost:** The annual basic fee was 10 €, transaction fees did not exist. As it can be assumed that an important part of customers enrolled with a 50-100% discount on the basic fee, the high number of customers does not allow any inference on the acceptance of high basic fees. Paybox could be used with any mobile phone. Thus preliminary costs were low and facilitated customers' participation.

**Convenience:** The procedure was easy to use and understand as the customer only had to know his MS ISDN and could rely on the IVR system to guide him through the transaction or authorization process. Customer experience in the handling of voice-based transactions afforded Paybox the reputation of an easy-to-use and subjectively safe system. Fast processing of transactions was possible, costs were low and facilitated customers' participation[6].

**Functionality Requirements.** Bank transfers through Paybox were possible. Pico payments were technically possible yet not efficient since costs for the used debit procedure and the system's call-back function tended to outweigh transaction amount.

#### *b. I-mode*

The inherent payment procedure of i-mode is an MP procedure that can only be used within the i-mode MC platform. The MP procedure is limited to the MC scenario, e.g. to pay for news or the purchase of tickets.[8]

**Security.** Responsibility for both the confidential treatment of data during payment procedures and the safe handling of customers' data lies with E-Plus. Customers have confidence in the well-known MNO EPlus. Thus, its reputation enhances a feeling of security in customers and the criterion of subjective security is satisfied. Confirmation of payment is handled via a special confirmation page displayed on the mobile device.

**Cost:** The payment does not generate additional transaction fees for customers. I-mode itself can only be used if an E-Plus customer both subscribes to i-mode (monthly basic fee, minimum 5 €) and purchases a special i-mode mobile phone. Subscribers from competing MNO are excluded from the service[6].

**Convenience:** The i-mode MP procedure is easy to use and understand. Payment-related data as a rule do not have to be entered. Thus, the extent of customer experience necessary to handle the system tends to be very low.

**Functionality Requirements.** The requirement of a capacity for the execution of a bank transfer is only partially met. Pico payments can be easily performed since these can be aggregated.

#### *c. Vodafone m-pay*

The latest MP procedure of the three is Vodafone m-pay Which is Devised for the MC scenario but also usable in EC, it can up to now handle payments up to a maximum of 10€ (although this is a deliberate restriction in order to gain experience with the procedure).

**Costs.** The use of m-pay does not entail extra costs for customers, i.e. there are neither basic nor transactions fees. This includes even the SMS being sent by the payment server within the EC scenario. Purchase of a special device other than a WAP-enabled phone is not required.

**Security.** Security requirements like the confidential treatment of data are met. No payment-related data are transmitted from customers to content providers or merchants during the transaction, since these data have been stored by the payment server and used only internally.

**Convenience.** M-pay is basically easy to use and understand. Duration of payment procedures is similar to that of i-mode within the MC scenario and thus comparatively low. In the EC scenario the procedures take slightly more time since the forwarding and entering of a TAN is required. As far as merchant participation is concerned, m-pay is – up to now - at a great disadvantage.

**Functionality Requirements.** As with i-mode the regular MNO billing system is used to process payments. The requirement of a direct bank transfer is met only partially. Picopayments pose no problem since the procedure has been specially designed to enable those[8].

#### d. Rating of current Germany MP procedures

Pousttchi and zenker evaluate mp systems In Germany by a questionnaire based research and base on customer perspective. Base on Pousttchi and zenker research [6] None of the mp procedures in germany reaches the top score since various shortcomings detracted from the procedures' effectiveness. It is apparent, however, that each carries a burden that might deter customers from usage: Paybox' lack of payment in the MC scenario, i-mode's inability to cover scenarios others than MC and the insufficient number of merchants and content providers for m-pay[8] . Because of not accessing to user of mobile payment in Serbia and Finland we couldn't held a questionnaire based survey like the one in [6] and find real evaluation base on customer perspective. but base on our library research We did not find an "ideal type" and do not believe that we will see one in the future.

#### IV. THE PROPOSED CRITERIES OF MP THAT CAN MEASURE CUSTOMER OR MERCHANT SATISFACATION

Roger's Diffusion of innovations [10] is multidisciplinary theory frequently applied in IS adoption research. Empirical research has supported the applicability of the diffusion theory in predicting the adoption of various information systems such as spreadsheets, WWW and operating systems .The diffusion of innovations theory has also been widely used to explain the adoption of a variety of financial and mobile technologies including smart cards, mobile commerce and mobile banking[7] . This paper explores customer and merchant adoption of mobile payment base on a frame work which is based on diffusion of innovation theory:

##### A. Compatibility

Compatibility captures the consistency between an innovation and the values, experiences, and needs of potential adopters . In IS adoption research, the compatibility of the technology has commonly been assessed in relation to the potential adopter's work and tasks .Regarding payment systems, the consumers' ability to integrate them into their daily life is an important aspect of compatibility . Compatibility has been found as a significant determinant of mobile technology and service adoption.

TABLE I  
EVALUATION OFS MP PROCEDURES ON THE GERMAN MARKE  
BASE ON CUSTOMER VIEW [3]

	Pay-box	i-mode	m-pay
<b>Scenarios (0-6)</b>			
MC scenario	1	2	2
Stationary merchant scenario	2	0	2
EC scenario	2	0	2
<b>Security (0-4)</b>			
Confidentiality of data	2	2	2
Confirmation via SMS or e-mail	2	2	2
<b>Costs (0-4)</b>			
None or low costs	1	2	2
No Purchase of a new device	2	0	2
<b>Convenience (0-3)</b>			
Easy handling	1	1	1
Fast processing	1	1	1
Accepting merchants	1	0.5	0
<b>Functionality requirements (0-1)</b>			
Bank transfer possible	0.5	0.25	0.25
Payment < 20ct possible	0	0.5	0.5
<b>Sum (max. 18 points)</b>	<b>15.50</b>	<b>11.25</b>	<b>16.75</b>

##### B. Complexity

In the diffusion of innovations theory, complexity is determined as the "degree to which an innovation is perceived as difficult to understand and use" [10].Complexity and problems with usability have contributed to the low adoption of a variety of payment systems, including smart cards and mobile banking [7]. Mobile payments are commonly expected to increase consumer convenience by reducing the need for coins and cash in small transactions and increasing the availability of payment possibilities [8].Limitations in mobile device features, however, diminish the usability of mobile technologies. Typical limitations include small keypads, limited transmission speeds and memory, and short battery life.

##### C. Cost s

In traditional adoption research, the cost is incorporated in the relative advantage construct .but In recent studies , cost is treated as a separate factor to clearly differentiate it from the relative advantage of time and place independence. Kim[6] found that the perceived fee has a significant effect on the perceived value of mobile Internet. Perceived service cost has also been found as a significant determinant for the intention to use wireless

financial services ,mobile banking ,and mobile commerce transactions .

*D. Network externalities and critical mass* Network externalities are considered as a relevant factor in mobile payment adoption Because payment technologies exhibit indirect network externalities. Failure to create critical mass has impacted the continuance of several previous payment systems .Mobile payments represent a new payment network in the market. A consumer's decision to adopt the network is significantly affected by the amount of merchants using it, since that amount determines the opportunities for consumers to use the new payment service. New consumers adopting the network indirectly increase the value of the network for all consumers because they attract new merchants to join the network[7]

#### *E. Security and trust in mobile payment systems*

The importance of trust is highlighted in electronic and mobile commerce because of the spatial and temporal separation between buyer and seller when buyers are required to give delicate personal information such as telephone number or credit card number to the seller[7] .Consumers' concerns about the privacy and security of mobile payments are commonly related to authentication and confidentiality issues as well as to concerns about secondary use and unauthorized access to payments and user data.

### V. CRITERIA'S IN CONJUNCTUION WITH CUSTOMER AND MERCHANT VIEW

#### *A.. Criteria's in conjunction with customer view*

The findings suggest that the relative advantage of mobile payments is related to the specific benefits provided by the new mobile technology; time and place-independent payments, remote and ubiquitous access to payment services, and the possibility to avoid queuing and to complement cash payments. Furthermore, the findings indicate that the relative advantage becomes more important in certain use situations including the presence of queues, an unexpected need to make the payment, time pressure, and lack of cash or loose change. The most compatible application areas for mobile payments include electronic ticketing, purchases on vending machines, mobile content and services, and payments at POS. Factors found to prevent mobile payment adoption include complex solutions, premium pricing, low adoption rates, perceived risks and perceived incompatibility with large value purchases. These findings suggest that in order to

create volume, mobile payments need to be better integrated with existing financial and telecommunication infrastructures. Proprietary systems with exclusive service providers and infrastructures are not likely to be successful in the long term. Instead, compatibility with consumers' existing services and common standards between different

TABLE II  
DRIVERS FROM CUSTOMER VIEW POINT

Adoption factor	Drivers
Relative advantage	• Time and place independent payments     • Queue avoidance     • Enhanced payment instrument availability     • Complement to cash
Compatibility	• High with digital content and services     • High with small value purchases at POS
Complexity	----
Cost	----
Network externalities	----
Trust and security	• Trust In merchants     • Trust In telecom operators     • Trust In financial institutions

service providers could facilitate adoption[9] . barriers and drivers from customer view point are shown in table 2,3.

#### *B. Criteria's in conjunction with Merchant view*

several barriers are defined to merchant adoption, most clearly in four categories: relative advantage, compatibility, complexity and costs[11].form merchants point of view , Existing mobile payment systems are perceived to be compatible with small value purchases of digital content and services, but lack of technical standards as well as incompatibility with current business practices are considered as significant barriers to more wide-scale adoption of mobile payments. Furthermore, existing solutions are still perceived by most as too complex and difficult to use[9] . premium pricing has prevented the adoption of mobile payments in both merchant and consumer side[9]. ]Payment service providers should price the upcoming solutions competitively so that their provisions would be similar to credit card provisions. The cost of a transaction for\consumers should be independent from the payment method. Costs of the payment solutions

TABLE III  
BARRIERS FROM CUSTOMER VIEW POINT

Adoption factor	Barriers
Relative advantage	----
Compatibility	• Low with large value purchases
Complexity	• Complex data input formats, codes, service Numbers     • Management of separate accounts burdensome     • Complex registration procedures
cost	• Premium pricing &amp; high transaction costs
Network Externalities	• Lack of wide merchant adoption     • Proprietary devices/services
Trust security	• Unauthorized use     • Transaction errors     • Lack of transaction record and documentation     • Vague transactions     • Concerns on device and network reliability     • Concerns on privacy

are still a significant barrier to adoption for the merchants. Security and service provider trustworthiness are essential Requirements in all service relationships and their importance is highlighted when the service involves monetary transactions. The interviewees stated that the reliability of the technical solution and trustworthiness of the payment service provider are essential factors affecting

the payment solution adoption. Merchants and service providers were of the opinion that the role of mobile payment provider would be most suitable to banks and large telephone operators. The security issues were not, however, as important barriers as the issues related to network externalities and costs [9].

#### *C . Comparison of drivers and barriers for mp adoptin from merchant and customer view point*

In this paper barriers and derivers found in [7],[8] and [11] are compared. The main barriers form but customer and merchant view exist in cost, compatibility and complexity criteria's .Also derivers are in Trust and relative advantage domain. both groups trust on payment service providers and mobile network operators and

believe they are trustworthy .it seems that security is yet a barrier for customers . Also both group think that mobile devices are secure and suitable for payment.

#### VII. CONCLUSION

The contribution f this paper can be divided n t three sections: At first This paper analyzed some of the major mobile payment systems in Europe based on five criterion: Scenarios, Security, Cost, Convenience and Functionality Requirements. Seond After survieing the mobile payment systems in finland,serbia and germany (3 cases) the paper contributes to a frame work consists of 6 cirteries. The criteries can be used for evaluating mp system adoption by customers and merchants. Third ,the paper introduces some of the major drivers and barriers that can be effecting on customer and merchant adoption Combination of more detailed survey and using simulations, importance of the mentioned criteria's individually as well as their importance in related to each other can be evaluate

#### REFERENCES

- [1] m. Hassinen,k. Hypponen,E. Trichina":Utilizing national public-key infrastructure in mobile payment systemsElectronic Commerce Research and Applications xxx (2007) xxx-xxx
- [2] T.Dahlberg et al., Past, present and future of zmobilepayments research: literature review , Electronic Commerce Research and application(2007)doi:10.1016/j.elerap.2007.02.001
- [3] K.Pousttchi; Selk, B; Turowski, K.: Acceptance Criteria for Mobile Payment Procedures. In: Hampe, F.; Schwabe, G. (Eds.): Mobile and Collaborative Business 2002. Nürnberg 2002
- [4] K.Pousttchi, .: An Analysis of the Mobile Payment Problem in Europe, in available Online at <http://mpira.ub.uni-muenchen.de/2915/> MPRA Paper No. 2915, posted 07. November 2007 / 02:49
- [5] A.S. Lim, Inter-consortia battles in mobile payments standardisationElectronic Commerce Research and Applications. (2007), doi:10.1016/j.elerap.2007.05.003.
- [6] K. Pousttchi, M. ZenkerCurrent Mobile Payment Procedures on the German Market from the View of Customer RequirementsProceedings of the 14th International Workshop on Database and Expert Systems Applications ,2003, IEEE
- [7] N.Mallat,Exploring consumer adoption of mobile payments – A qualitative study
- [8] N.Mallat, M.Rossi, V.K. Tuunainen,2004. Mobile banking services. Communications of the ACM 47 (5), 42– 46.
- [9] Paybox: Informations on Paybox. (URL: <http://www.paybox.de/3416.html>, browsed 01-17-2003)
- [10] E.M. Rogers,Diffusion of Innovations, fourth ed, 1995, New York:Free Press
- [11] N.Mallat and V.K.Tuunainen, Merchant AdoptionMobile Payment Systems, Proceedings of the International Conference on Mobile Business (ICMB'05) , 2005 IEEE

## The opportunities and challenges of using ICT by State Owned Enterprises in Bangladesh: Case of SOEs under Privatization Commission

Ahsanullah M Dewan<sup>1</sup>, Saifullah M Dewan<sup>2</sup>, Shams Ara Nazmin<sup>1</sup>

<sup>1</sup> East West University, <sup>2</sup> University of New South Wales, Email: amd@ewubd.edu

**Abstract** — The potential contribution of information and communication technology (ICT) to increase the productivity of State Owned Enterprises (SOEs) in developing countries has long been recognized. However, the realization of this potential has been problematic and over recent years there have been a number of initiatives supported by government, non government and foreign agencies which have endeavored to aid and encourage the up take of ICT to enable access to such promised benefits. Despite strong theoretical arguments suggesting that ICT has much to offer to SOEs, the study would seem to suggest that use of ICT by SOEs listed under Privatization Commission is still in its infancy. It also identifies some significant barriers which are impacting upon the level of ICT adoption amongst SOEs, and confidence on ICT of the SOEs in Bangladesh. This paper is based on the study conducted as part of the IDA funded Enterprise Growth and Bank Modernization Project and explores the opportunities and barriers of using ICT to provide a necessary lever to enhance competitiveness and productivity of SOEs in Bangladesh.

**Index Terms** — ICT, e-business, SOE

### I. INTRODUCTION

The prospect of using information and communications technology (ICT) to improve the competitiveness and productivity of state owned enterprises (SOEs) has long been recognized world wide. Despite the recent initiatives by the government, non government and foreign agencies to aid and encourage the up take of ICT to enable access to such promised benefits, the study suggests that the use of ICT by SOEs listed under Privatization Commission is still in its infancy. One of the factors that have been identified as impacting upon the level of ICT adoption amongst SOEs is access to, and confidence of the management on ICT. Climbing the technological ladder does require skills upgrading through explicit learning in the new technologies.

Privatization in Bangladesh has gone through various phases since its Independence. Just after the war of Independence in 16 December 1971, the Government had to step in to fill the vacuum in economy as the then-owners, abandoned their properties and left the country. In addition, the nationalization policy of March, 1972 of major industries like Jute, Textile, Chemical and other basic industries, DFIs, Banks and Insurance etc. brought

into fold nearly 90% of such enterprises under Government control. However these public sector enterprises right from inception could not run profitably and efficiently and many of them ran into severe financial difficulties and serious management problems. By end of 1975, the Government began to reverse its policy of nationalization due to the heavy burdens created by the public sector enterprises on the national exchequer. During the period 1976-1992 about 500 SOEs had been sold or returned to their former owners.

The Privatization Commission which came into being in March, 1993 is entrusted with the overall responsibility of privatizing State-Owned Enterprises identified for privatization, whether small, large, profitable or non-profitable. The Commission is now headed by a Chairman with the rank and status of a State Minister to the Government. The Commission has been placed under the administrative jurisdiction of the Cabinet Division which is headed by the Hon'ble Chief Advisor. Ministries having SOEs under their control have either set up or in the process of setting up Privatization Cells for assisting the Privatization Commission in identifying, preparing, processing, implementing and monitoring SOEs for privatization. Since the creation of Privatization Commission as the special agency in March, 1993 by the Government to implement the program of divestiture of SOEs, till to-date 74 SOEs have been privatized so far. It is worth mentioning that 25 more SOEs have been handed over to Privatization Commission for privatization as well. Moreover, decisions to off-load the shares of some public limited companies have also been taken [1].

Due to the ICT's capability to transfer, collect, manage a great amount of information and to reduce the distance and time barriers new opportunities could well take in place both to create new ventures and to modify the existing businesses. Therefore, SOEs may reduce the transaction costs of information-intensive activities by resorting to ICT. These opportunities may especially favor SOEs that in most cases operate in a dense inter-firm relationship and need to manage a great amount of information.

SOEs performance is highly associated with learning capabilities, level of technology, firm-level knowledge, skills and experience. The successful adoption of ICT tools is likely to enhance individual SOE's productivity

in sectors such as electronics and general machinery. In the more traditional sectors such as jute and textile, there is a propensity for significant rise in product quality and more precise processing. To achieve the goal of better quality products, SOEs should undertake greater training and investment in skills and knowledge upgrading [2].

This paper explores the potential of ICT to provide a necessary lever to enhance competitiveness and productivity of SOEs in Privatization Commission's consideration. The paper in particular addresses issues related to: the opportunities of using ICT applications in the value-creating processes for SOEs, and challenges in adopting these ICT tools by the SOEs in Bangladesh.

## II. CLASSIFICATIONS OF ICT TOOLS

The widespread adoption of ICT worldwide has given rise to an increasing number of studies on this field recently. These studies have dealt with different and complementary aspects, such as: potential effects of ICT on the organizations, the inter-organizational relationships, the market, and the social-economic systems; ICT classification; analysis of the organizational and managerial actions required to support the ICT adoption; and interpretation of the enabled opportunities [3] [4] [5] [6].

ICT tools have been classified and distinguished into several ways where each one is based on different and complementary aspects of ICT. Independent of their classification, thanks to the capabilities of reducing the time and costs of processing and communicating information; storing and elaborating great amount of data and information; organizing and structuring data and information on the basis of the user needs, ICT tools can allow data, information, experiences, and the knowledge owned by individuals and organizations instantaneously available and enable easier sharing. ICT tools in this paper have been classified as follows:

- Integration tools: to help in the integration and coordination of the processes supporting the information transfer (LAN, WAN, database, shared elaboration systems, data modeling support systems, information flows modeling support systems, CASE, group working support systems, EDI, groupware, Internet, ERP, DSS, CAD etc.);
- Knowledge management tools: to support the processes of problem solving and organizational learning, as well as the relationships and integration among individuals and among different organizations (Lotus Notes, software agents, groupware, Internet).
- Process tools: to concur in the transformation of the inputs in output (Dedicated software).

The main capabilities of ICT tools and the related advantages of these tools on the organizational processes are as follows [3] [7].

TABLE 1 - ICT CAPABILITIES AND ADVANTAGES

Capability	Examples of ICT Tools	Advantages
Analytical	Dedicated software, workstations	To bring complex analytical methods to bear on a process
Automational	Artificial intelligence, CAD, EDI, search engine	To replace or reduce human labor in a process
Geographical	EDI, Internet, Intranet, videoconferencing, tele network	To allow an easy and fast information transfer between large distance, making processes independent of geography
Interactivity	CAD/CAM, EDI, Internet, Intranet	To connect parties that would otherwise communicate through an intermediary
Knowledge management	Artificial intelligence, groupware, shared databases, data mining, videoconferencing	To capture and dissemination of knowledge and expertise
Multimedia	Videoconferencing, audio-conferencing, search engine	To deal with and deliver a great amount of information expressed in several ways: graphics, sound, video
Sequential	Groupware technology, shared databases, Internet, Intranet	To enable changes in the sequence of tasks in a process, often allowing multiple tasks to be worked on simultaneously
Tracking	Artificial intelligence, CAM, sensors, Internet-based applications	To allow the detailed tracking of task status, inputs, and outputs
Transactional	Artificial intelligence, dedicated software, groupware technology	To transform unstructured processes into routinized transactions



### III. ICT AND VALUE CREATING PROCESS FOR SOEs

The adoption of ICT by the SOEs can improve the performance of the organizational and inter-organizational processes and, as a consequence, of the value-creating processes. ICT can follow the existing system of inter firm relationships to implement and automate a number of activities involved in the value creating processes carried out across the supply chains, such as vendor seeking, ordering, inventorying, delivering, co-design, etc [8]. The effects of the ICT on three main value-creating processes, namely marketing and customer relations; logistic and networking; and innovation and knowledge management are analyzed in the following.

#### A. *Marketing and customer relations*

The global spread of the Internet and the 24 hour availability of web server have created time and geographical independence and enable customer service to be decoupled from supplier availability [7]. Overcoming the geographical constraints and creating time independence, ICT enable the SOEs to get access, promote and sell goods and services (e-commerce) in a global market. Dealing with and delivering great amount of information in different ways (text, graphics, sound, video), ICT enrich the opportunities for the promotion of goods and services and for the provision of in-depth information.

The new intelligent agent software helps to identify the customer profile (taste and characteristics) and the pattern of purchases. All the information can be easily organized and stored in a database that supports customer service, marketing, and new product development processes. Web technology and the related online trading models (business-to-business and business-to-consumer) provide supplier assets and customer benefits, namely lower transaction costs, better prices, reduction of time to market, affirmation of brand image, market share and access to markets, customer orientation, customer choice increasing, and customer-driven design. Supporting application-to-application as well as person-to-person and person-to-application interactions, ICT allow the direct interaction and continuous links with customer. This in turn provides opportunities for better targeting goods and services. [7] [9]

#### B. *Networking and Logistic*

The impact of ICT on networking processes is due to three main effects: (1) the electronic integration effect, (2) the electronic communication effect and (3) the electronic brokerage effect [9]. The electronic integration effect allows suppliers and procurers to create joint, inter-penetrating processes at the interface between value-adding stages. This effect is due to the combination of different ICT capabilities, such as tracking, sequential, and automational, which enable to

link one activity with the others and make real time data created in one activity available both within the company and with outside suppliers, channels, and customers [11]. ICT allow the supplier and the procuring SOE's inventory management processes to be linked, reducing total inventory costs; the connection between sales activities and order processing as well as manufacturing processes and inventory levels of multiple suppliers, optimizing the supply chain; the integration of the new product development phases carried out in different teams or organizations; and check the status of a payment or a shipping process.

Due to the possibility of communicating more information in the same amount of time decreases the costs of communication, making communication independent of the geographical proximity. This means that ICT enable the SOEs to enrich the content of the information transferred within a client-supplier relationship in the global market. As an example, ICT (CAD/CAM with Web) can support the exchange of complex product models among parties located far away [11].

The electronic brokerage effect is due to the possibility of creating an "electronic broker" that electronically connects a large number of potential buyers and suppliers through a central database, filters the different alternatives, and helps to match one party to the other, thereby reducing the need for buyers and suppliers to contact a large numbers of alternative partners individually. Thus, ICT (Internet, software agents, databases, search engines, etc.) increase the number of alternatives that can be considered, increase the quality of the alternative eventually selected, and decrease the cost of the whole selection process [10]. Low cost and fast expansion of the number of relationships enable the development of new business models based on networking processes, for example, online auction of the machineries of SOEs.

#### C. *Innovation and knowledge management*

The ICT tools supporting knowledge management (i.e. groupware, videoconferencing, Internet, Intranet etc.) are basically aimed at facilitating the inter-connection and integration among individuals through the fast and effective exchange of information and knowledge. SOEs can use the digitization to transmit information, avoiding the contents' modification. The knowledge management technologies allow interactive communication. Besides, thanks to the virtual abolition of the distance, the communication of such kind of knowledge, which has been always limited within a bounded territorial system or a business unit, can now spread in meaningful way. ICT tools can also be used in the new product design and development. It may shorten the development cycle, increase number of alternative designs considered, reduce development costs and produce higher quality products [10].

It is argued that the customers contribute greatly to the development of new products. Listening to the customers and, beyond that, engaging them in a dialogue and listening to their comments on the products are important means of obtaining input for sustained innovation. To that end, SOEs may regularly use, besides traditional means, e-mail response addresses, FAQ lists, discussion forums and online questionnaires. Information for innovation can also be obtained by monitoring online the customer behavior (by tracking visits & downloads) [11] [12].

#### IV. E-BUSINESS MODELS FOR SOES

The adoption and implementation of ICT tools allow businesses to implement new forms of business models. The most relevant e-business models for SOEs along with their implications are listed in Table 2.

TABLE 2: E-BUSINESS MODELS FOR SOES

E-business models	Implications
Collaboration platforms	Electronic business process cooperation, e.g. collaborative design
Electronic customer community	Direct access to a wide community of customers that express their needs and give feedback; targeting and customization
E-mall	Promoting on a global scale a company's products and services
E-procurement	Tendering and procurement on a global scale; automation of the tendering and procurement processes; electronic forms of collaboration and negotiation (electronic implementation of bidding mechanisms) on a global scale
E-shops	Promoting the products and services on a global scale
Professional virtual community	Direct access to a wide community of professional expertise; extended network of complementary competencies

#### V. ROLE OF ICT FOR SOES

The opportunities offered by the adoption and implementation of ICT lie not only in the improvement of the processes carried out at a local level, but also in the possibility of carrying out the same processes on a global scale independent of the firm size. In particular, referring to the three value-creating processes, the implementation and adoption of ICT within SOEs can allow the following positive effects:

- Reinforcing the existing relationships within SOEs, helping the integration of the economic actors

operating along global supply chains (business-to-business, electronic data interchange, Intranet, etc.);

- Providing SOEs with new opportunities of networking with firms located outside the country (business-to-business);
- Expanding the business boundaries of SOEs (web sites, electronic portals, etc.);
- Managing the relationships with the end-markets, offering new services and new ways to create value (electronic commerce, on-line marketing, etc.);
- Supporting both the joint innovation processes developed by SOEs and private industries and by the adoption of exogenous innovations.

#### VI. CHALLENGES TO ADOPT ICT TOOLS BY SOES

There is a wide range of reasons why SOEs are not adopting or using ICT tools in Bangladesh. Reasons do vary widely among sectors but most commonly faced challenged by SOEs in Bangladesh include:

- Unsuitability for business
- Lack of availability of ICT skilled personnel
- Cost of developing and maintaining ICT Tools
- Lack of Network infrastructure issues
- Building security and trust

These barriers are explained below [3, 13]:

##### A. *Unsuitability for business*

The leading reason given by SOEs for not engaging in ICT use is that it is not suited to the nature of their business. Findings of the survey on SOEs showed that among the operational SOEs, around 60% do not use the Internet for their business because they consider that Internet does not suit their type of business and/or products. Other reasons for not e-business models include lack of personnel with appropriate ICT skills, concerns over unbalanced costs and benefits, insufficient customer access to Internet and technology concerns, such as on-line security. During the survey several SOEs recalled that they did receive few inquiries regarding possible e-mailing of orders. Very few of its regular customers had used the Internet, and some preferred receiving confirmation of their order via fax or telephone. These concerns come from the practical judgment of SOEs on the benefits of using ICT.

##### B. *Lack of availability of ICT skilled personnel*

SOEs listed under Privatization Commission generally lack the human resources needed for ICT and e-business, because they focus on day-to-day operations and lack the time to understand the benefits of new technologies. Even when they are aware of the potential benefits of adopting ICT tools, they require know-how or qualified personnel. The SOEs that adopted ICT tools

are likely to have within the SOE someone who has a reasonable amount of knowledge of the specific technology and/or technology in general. Interestingly, the study found that they were not necessarily ICT professionals, but simply people interested in technology. It was also found that ICT are largely adopted by the SOEs which are working in the field of banking and engineering.

#### C. Cost of developing and maintaining ICT Tools

Most SOEs will not adopt ICT tools if the benefits do not outweigh the costs of developing and maintaining the system. The issue is costs relative to benefits expected, not cost itself. Nevertheless, SOEs are generally concerned about the costs of establishing and maintaining ICT tools since they generally suffer from budget constraints and are less sure of the expected returns on the investment. The study found that many SOEs cannot afford to adopt sophisticated ICT tools due to severe fund crisis.

#### D. Lack of Network infrastructure issue

The availability of a wide range of Internet connections and other communication services, preferably at competitive prices, is very important in that it allows SOEs to choose different and appropriate services according to their specific needs and (initial) expectations from ICT tools. Telecommunication networks are likely to continue to serve as the primary means of Internet access for many SOEs because of their relatively lower cost. Availability of broadband connections may affect SOEs' decisions to adopt ICT tools specially Internet usage. Faster Internet connections improve the overall on-line experience for both individuals and businesses, encouraging them to explore more applications and spend more time on line. Slow Internet connections and data transfer have in fact discouraged some SOEs from adopting Internet: over 75 percent of the SOEs indicated excessively slow or unstable data communication as a major hindrance to Internet use.

#### E. Building security and trust

Survey showed that security issues (viruses, hackers) are among the most important perceived barriers to Internet use by SOEs. Barriers to e-business models included payment uncertainties, contract and delivery guarantee uncertainties.

### VII. ICT SPREAD WITHIN SOES: THE CASE OF PRIVATIZATION COMMISSION LISTED SOES

The survey has been carried out by the team on twenty five state owned enterprises listed under Privatization Commission as part of the IDA funded Enterprise Growth and Bank Modernization Project in the financial year 2007-08. It highlights that the process

of adoption of ICT within SOE is still in the very early stage. A large gap exists between Internet connectivity and e-business models adoption and, more in general, between the possession of ICT and the actual use of it.

It could well be pointed out by the survey that the connection to the Internet does not support the unstructured and informal communication processes within SOEs. For example, the utilization of e-mails for the inter-firm relationships or for intra-firm informal communications is very limited. In 59% of SOEs surveyed, there are no significant experiences of informal exchanges of e-mail, 30% of SOEs' management have never used Internet for communicating with buyers. These percentages are almost unthinkable in the case of other forms of ICT-based interactions with customers, such as mailing list and discussion forum.

Further information on the current spread of ICT-based tools within SOEs are shown in Table 3 by using a scale rising from 1 (low) to 5 (high). It clearly shows that SOEs are seriously lagging in the use of Internet. Currently, Internet is used as a simple means of communication and, in particular, it is used more for communication outside Bangladesh rather than inside. However, SOEs that is being run by ICT literate individuals were found very enthusiastic about using ICT tools. However, it was expected, due to the above mentioned challenges that SOEs in Bangladesh would be seriously lagging when it comes to using ICT tools.

TABLE 3: SPREAD OF ICT BASED APPLICATIONS

ICT based applications	Current
Production (ERP) Software usage	1.2
B-to-C activities with on-line sales	0.0
Managerial Software usage	3.1
Promotional sites	1.5
Communications via e-mail, Internet	2.9
Use of Mobile and other traditional methods	5.0

The survey also shows that the degree of diffusion of advance software is still very low. These in fact require relevant investment by the management and, above all, the business processes re-engineering. This aspect, more than the economic one, represents a challenge to the adoption for the greater part of SOEs, which consider traditional way of operating as a distinctive competitive asset [14].

### VIII. CONCLUSION

Shared databases, integrated information systems, and ICT-based applications supporting the marketing and logistic activities can greatly improve the efficiency of the key supply chain business processes such as customer relationship management, customer service management, demand management, order fulfillment, manufacturing flow management, procurement, product

development and commercialization. It is possible to state that the adoption and implementation of ICT by the SOEs in Bangladesh can offer the following main opportunities [15]:

- favoring interactions & information exchanges among SOEs within same business;
- improving communication channels to allow correspondence with potential buyers of the global market;
- optimizing the production processes performance by redefining the relationships with customers and suppliers.

It can be argued that SOEs need institutional support for their survival in the era of globalization. The adoption of ICT by Indian SOEs is a proof of this point [2]. The burden and risk had been shared with the encouragement given by the Government of India (GOI) for private sector participation in providing technological infrastructure. Consequently SOEs have better access to web-enabled and portal based e-business technologies. However, the Government of Bangladesh still has to take initiative for providing uninterrupted utility services so that SOEs can become more competitive in international markets. Finally, the study suggests that SOEs in Bangladesh need much greater infrastructural support in order for them to reap the benefits of ICT and to develop the capabilities to contribute to economic development. Proper policies and programs aimed at providing required infrastructure need to be initiated by the Government of Bangladesh in order to make SOEs more competitive in the domestic and international markets.

#### ACKNOWLEDGEMENT

Ahsanullah M Dewan has served as the MIS Consultant for the year 2007-08 at the Privatization Commission, Chief Advisor's Office under the IDA funded Enterprise Growth and Bank Modernization Project. Authors would take this opportunity to show their gratitude to all his colleagues and management of SOEs for their support and cooperation. Without their support this study was not possible.

#### REFERENCES

1. Annual Report 2007-08, Privatization Commission, 2008
2. Oyeyinka, B. Lal K., *Learning new technologies by Small and Medium Enterprises in developing countries*, Technovation 26, pp 220-231, 2006
3. Davenport, T.H., Short, J.E., *The new industrial engineering: IT and business process redesign*, Sloan Management Review 31, 11-27, 1990.
4. Gurbaxani, V., Whang, S., *The impact of information systems on organizations and markets*, Communications of the ACM 34, 59-73, 1991.
5. Malhotra, Y., *Role of Information Technology in Managing Organizational Change and Organizational Interdependence*, <http://www.brint.com/papers/>, 1993.
6. Laubacher, R.J., Malone, T.W., MIT Scenario Working Group, *Two Scenarios for 21st Century Organisations: Shifting Network of Small Firms or All-Encompassing "Virtual Countries"?* Working Paper, Sloan School of Management, MIT, 1997.
7. Timmers, P., *Electronic Commerce. Strategies and Models for Business-to-Business Trading*, John Wiley & Sons Ltd, Chichester, 1999.
8. Rayport, J.F., Sviokla, J.J., *Exploiting the virtual value chain*. Harvard Business Review 73, 75-85. 1995.
9. Venkatraman, N., Henderson, J.C., *Real strategies for virtual organizing*, Sloan Management Review 40, 33-48, 1998.
10. Malone, T.W., Yates, J., Benjamin, R.I., *Electronic markets and electronic hierarchies*, Communications of the ACM 30, 484-497, 1987.
11. Porter, M., *Strategy and the Internet*, Harvard Business Review (March), 63-78, 2001.
12. Porter, M., Millar, V.E., *How information gives you competitive advantage*. Harvard Business Review (July-August), 1985.
13. OECD, *ICT, e-business & SOE*, France, [www.oecd.org](http://www.oecd.org)
14. Davenport, T.H., *Process innovation. Reengineering Work Through Information Technology*, Harvard Business School Press, Boston, 1993.
15. IDC, *Adoption of the Internet by SOEs in Western Europe*, 2000.

# Parallel Performance of Domain Decomposition Method on Distributed Computing Environment

\*A.M.M.Mukaddes and \*\*Atsushi Uragami

\*Department of Industrial and Production Engineering, Shahjalal University of Science and Technology, Sylhet, Bangladesh

\*\* Department of Intelligent Machinery and Systems, Kyushu University, Japan

**Abstract** - Distributed parallel computing, which uses general-purpose workstations connected by a network as a large parallel computing resource, is one of the most promising trends in parallel computing. Over the last two decades, the growth in the use of such parallel computing environment has ensured the interest on the parallel finite element method. In this work, parallel finite element method using domain decomposition technique has been adapted to the distributed parallel environment of networked workstations and the supercomputer. Using the developed code, several model problems are solved and the parallel performances are analyzed on the network of eight Pentium IV PC cluster and GP7000F workstations. Finally, very large size of problem over 11 million DOFs has been solved using 1024 SR8000 supercomputer.

**Index Terms** – Domain decomposition, finite element, parallel computing, workstation cluster.

## I. INTRODUCTION

Large scale problem, particularly those defined in three dimensional domain, often need substantial computation time and memory to run on ordinary sequential computers. Even if they can be solved, powerful computation capability is required to obtain the accurate and precise results within a reasonable time. Parallel computing can meet requirements of high performance computation [1]. Among the various parallel computing algorithms, domain decomposition method has become much popularity in the recent years [2]. In this method, the whole domain to be solved is first decomposed into a number of subdomains without overlapping. Finite Element Analysis (FEA) is defined and solved in each subdomain in parallel and then partial solutions are assembled together to get the global solution.

This method is exclusively implemented on structural analysis [3], heat conduction problem [4] and fluid analysis [5]. Most of the papers emphasizes only on the implementation of the method without analyzing the parallel performance as well as taking the necessary advantages of the parallel computing environment. In this research, the two parallel approaches (dynamic load distribution [6] and static load distribution) of domain decomposition method

with a suitable preconditioner have been adapted to the distributed parallel environment of networked workstations [7] and the supercomputer.

Using the developed code several models are solved and the performances are measured in different computing environments. The important factors affecting the performance of parallel computing using the domain decomposition method are found and analyzed. This analysis will be useful as a guideline for the user of the domain decomposition based software in the parallel computing environment.

## II. DOMAIN DECOMPOSITION METHOD

*Main method:*

Consider a linear system of algebraic equations,  

$$[K]\{u\} = \{f\} \quad (1)$$

arising from a finite element discretization of a linear elliptic boundary value problem in the domain  $\Omega$ , where  $K$  is the global stiffness matrix,  $u$  is an unknown vector and  $f$  is a known vector. The domain  $\Omega$  to be solved is first decomposed into  $N$  number of subdomains,  $\{\Omega^{(i)}\}_{i=1,\dots,N}$ , for which the union of all subdomains boundaries is

$$\Gamma = \bigcup_{i=1}^N \Gamma^{(i)}. \quad (2)$$

Let  $K^{(i)}$  be the local stiffness matrix corresponding to subdomain  $\Omega^{(i)}$ . Then as usual the global stiffness matrix  $K$  can be generated by subassembling:

$$K = \sum_{i=1}^N R^{(i)} K^{(i)} R^{(i)T} \quad (3)$$

where  $R^{(i)T}$  is the 0-1 matrix which translates the global indices of the nodes into local numbering. With superscript  $T$  we denote the transpose of a matrix (or vector). Let  $u^{(i)}$  be the vector corresponding to the variables in  $\Omega^{(i)}$  and it can be expressed as  $u^{(i)} = R^{(i)T} u$ . Each  $u^{(i)}$  is

split into degrees of freedom  $u_B^{(i)}$ , which correspond to  $\partial\Omega^{(i)} \setminus \Gamma_u$ , called interface degrees of freedom and the remaining interior degrees of freedom  $u_I^{(i)}$ . The subdomain stiffness matrices  $K^{(i)}$ , the vector  $u^{(i)}$  and the 0-1 matrices  $R^{(i)T}$  are then split accordingly:

$$K^{(i)} = \begin{pmatrix} K_{II}^{(i)} & K_{IB}^{(i)} \\ K_{IB}^{(i)T} & K_{BB}^{(i)} \end{pmatrix}, \quad (4)$$

$$u^{(i)} = \begin{pmatrix} u_I^{(i)} \\ u_B^{(i)} \end{pmatrix}, \quad (5)$$

$$\text{and } R^{(i)T} = \begin{pmatrix} R_I^{(i)T} & 0 \\ 0 & R_B^{(i)T} \end{pmatrix}. \quad (6)$$

$$u_B^{(i)} = R_B^{(i)T} u_B, \quad i = 1, \dots, N \quad (7)$$

Let  $V^{(i)}$  be the space of the interface degrees of freedom for the subdomain  $\Omega^{(i)}$  and  $V$  denotes the space of all degrees of freedom on  $\Gamma$ . The unknown in the interior of the subdomain  $\Omega^{(i)}$  is eliminated by Gaussian elimination according to the equation (8) using an initial value of  $u_B^{(i)}$ ,

$$u_I^{(i)} = K_{II}^{(i)-1} (f_I^{(i)} - K_{IB}^{(i)} u_B^{(i)}). \quad (8)$$

Then, after elimination of the interior degrees of freedom, the problem (1) reduces to a problem on the interface space  $V$ ,

$$S u_B = g \quad (9)$$

where  $S$  is the Schur complement matrix:

$$S = \sum_{i=1}^N R_B^{(i)} S^{(i)} R_B^{(i)T}, \quad (10)$$

$$S^{(i)} = K_{BB}^{(i)} - K_{IB}^{(i)T} (K_{II}^{(i)})^{-1} K_{IB}^{(i)}. \quad (11)$$

The problem (9) is solved using the preconditioned conjugate gradient method.

*Preconditioners [8]:*

The present domain decomposition method includes the following preconditioning techniques [8]

**Simplified diagonal scaling:**

$$Q_{DIAG} = \sum_{i=1}^N R_B^{(i)} \left( \text{diag}(K_{BB}^{(i)}) \right)^{-1} R_B^{(i)T} \quad (12)$$

**Balancing domain decomposition (BDD):**

This preconditioning matrix is defined as

$$M_{BDD}^{-1} = Q_c + (I - Q_c S) Q_l (I - S Q_c) \quad (13)$$

where  $Q_l$  is the local level part and  $Q_c$  is the coarse level part of the preconditioner.

**Balancing domain decomposition with diagonal scaling:**

$$M_{BDD-DIAG}^{-1} = (I - Q_c S) Q_{DIAG} \quad (14)$$

**Incomplete balancing domain decomposition (IBDD):**

$$M_{IBDD}^{-1} = \tilde{Q}_c + (I - \tilde{Q}_c S) \tilde{Q}_l (I - S \tilde{Q}_c) \quad (15)$$

where  $\tilde{Q}_c$  is constructed from the incomplete factorized coarse operator

**Incomplete balancing domain decomposition with simplified diagonal scaling (IBDD-DIAG):**

$$M_{IBDD-DIAG}^{-1} = \tilde{Q}_c + (I - \tilde{Q}_c S) \tilde{Q}_{DIAG} (I - S \tilde{Q}_c) \quad (16)$$

### III. PARALLELISM OF THE DOMAIN DECOMPOSITION METHOD

Constructing the DDM algorithms for parallel computers, a good principle is to divide the original domain into parts, which are further decomposed into smaller subdomains. In this research we adopt Hierarchical Domain Decomposition Method (HDDM) [6] which is a well known parallel DDM. This hierarchically structured DDM classifies processors in two different following ways.

*Dynamic load distribution (h-mode):*

The h-mode classifies processors into 3 groups, ‘Grand Parent’, ‘Parent’ and ‘Child’. Fig. 1 illustrates the practical implementation of hierarchical organized processors in the present method. One of the processors is assigned as Grand Parent, a few as Parent, and others as Child. The number of processors assigned as Parent is the same as that of parts. The number of Child processors can be varied; and it affects the parallel performance.

The role of Grand Parent is to organize all processor communications (i.e. message passing) which occur between all processors. Parents prepare mesh data, manage FEA (Finite Element Analysis) results, and coordinate the CG iteration, including convergence decision for the CG iteration. Parents send data to Child processors, where FEA is performed in parallel. After the FEA, Child processors send the results to Parents. This computation will be repeated until the CG iteration is convergent.

*Static load distribution (p-mode):*

However, because almost all computation is performed in Child processors and the most communication time is taken between Parent processors and Child processors, so the communication speed becomes important. In recent years although the communication performance has also been improving by improvement in network technology, the high-speed network is still expensive. On the other hand, for the PC cluster generally used, the network speed becomes a bottleneck to the processing performance of CPU. Moreover, when a parallel processing performance is considered, it is important to reduce the amount of communication time as much as possible. Therefore, the Parent-Only type (static load distribution: p-mode) is useful than the conventional Grand-Parent-Child type (h-mode).

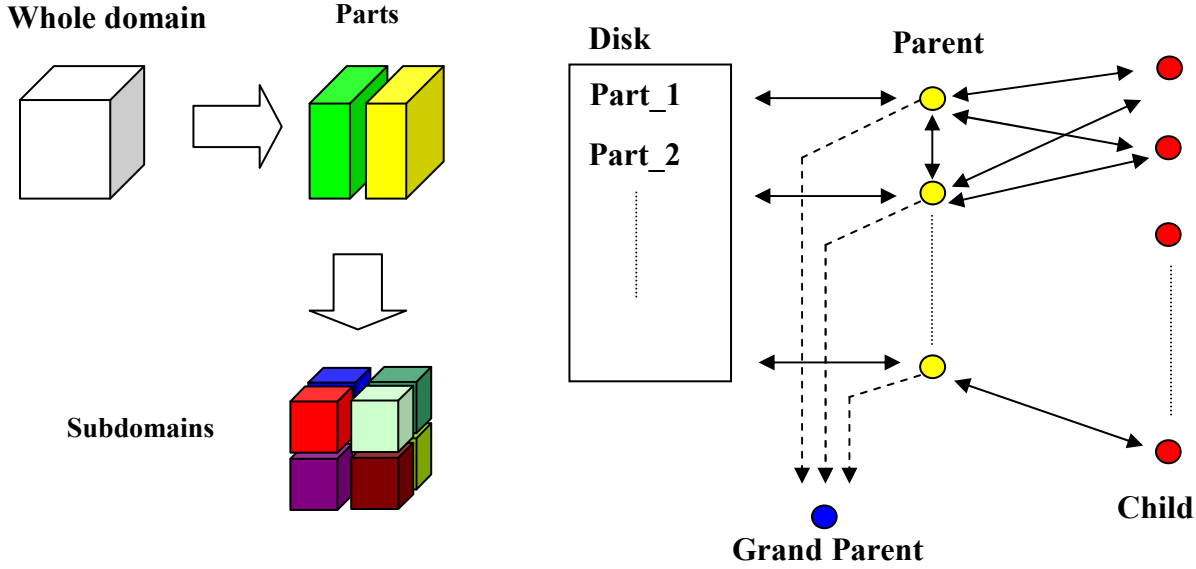


Fig. 1 Hierarchical domain decomposition method (left), dynamic load distribution (right)

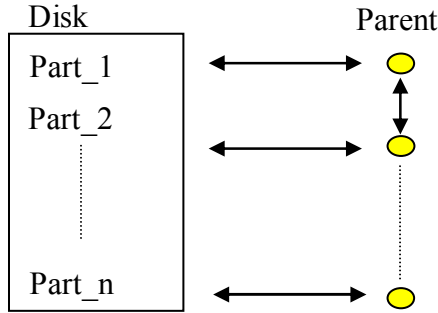


Fig. 2 Parallel processor scheme (p-mode)

Fig. 2 shows the parallel processor scheme of p-mode. In the p-mode, the Parent processors perform the FEA by themselves, which is computed by the Child processors in the h-mode. In the h-mode, although Parent processors store some of the subdomain analysis data and coordinate the CG iteration as the main work, the idling time of CPU increases because of less computation in Parent processors.

On the other hand, in the p-mode, since all processors perform the FEA and CPU can be used without futility, in uniform processor's environment, p-mode is considered to be predominance to h-mode. In the p-mode also, the number of Parent processors should be equal to the number of parts.

#### IV. PARALLEL PERFORMANCE

##### Speed up:

The key issue in the parallel processing in a single application is the speed up achieved, specially its dependence on the number of processors used. For

simplicity speed up with  $n$  processors,  $S_n$  is defined as follows:

$$S_n = \frac{t_1(A)}{t_n(A)} \quad (17)$$

where  $t_1(A)$  and  $t_n(A)$  are total time for solving the problem using one processor and  $n$  processors, respectively. Also scaled speed up  $S'_n$  is defined as follows:

$$S'_n = \frac{n \times t_1(A)}{t_n(nA)} \quad (18)$$

where  $t_n(nA)$  is the total time for solving the problem of size  $nA$ .

##### Sequential processing and parallel processing:

The program might be sequential, parallel or combination of sequential and parallel. The execution time to solve a problem of size  $n$  having the sequential processing is given by the equation (19)

$$T(n) = T_{calc}(n) + T_{i/o}(n) \quad (19)$$

where  $T(n)$ ,  $T_{calc}(n)$ ,  $T_{i/o}(n)$  are the total execution time, calculation time and input/output time. Again the execution time to solve a problem of size  $n$ , having the parallel processing also is given by the equation (20)

$$T(n, p) = T_{calc}(n, p) + T_{i/o}(n, p) + T_{comm}(n, p) \quad (20)$$

where  $T(n, p)$ ,  $T_{calc}(n, p)$ ,  $T_{i/o}(n, p)$  and  $T_{comm}(n, p)$  are the execution time, calculation time input/output time and time for communication respectively.

The ratio of the calculation time, input/output time and communication time over the total execution time is defined as follows

$$R_{calc} = \frac{T_{calc}(n, p)}{T(n, p)}, R_{i/o} = \frac{T_{i/o}(n, p)}{T(n, p)}, R_{comm} = \frac{T_{comm}(n, p)}{T(n, p)} \quad (21)$$

and

$$R_{calc} + R_{i/o} + R_{comm} = 1 \quad (22)$$

## V. NUMERICAL RESULTS AND DISCUSSIONS

### A. Physical model and computational environment:

Using the developed code, two heat transfer models [4] are analyzed here. Models are High Temperature Test Reactor (HTTR) and Advanced Boiling Water Reactor (ABWR) which are shown in fig. 3. The calculations are performed in three different computational environments shown in Table 1 and the model descriptions are shown in Table 2.

### B. Scalability and speed-up:

We evaluate the parallel scalability of the BDD and IBDD-DIAG preconditioner, varying the number of processors employed. Such evaluation is performed using the HITACHI SR8000. The results are shown in fig.4. Fig. 4 reports that BDD and IBDD-DIAG can compute faster solutions of a fixed mesh problem when the number of processors is increased. The speed-up is compared with the number of processors, referring the value for using 8 processors in Fig. 5 while the scaled speed-up with different number of processors refereeing the value of using 8 processors in Fig. 6. The results shown in figures confirm the parallel scalability properties of the BDD and IBDD-DIAG preconditioners. Since the program has both parallelable and nonparallelable portion, results are too far from the ideal one. For the parallelable portion, parallel computing with  $n$  processors can shorten computation time to  $1/n$  in ideal.

### C. Results concerning the number of subdomains:

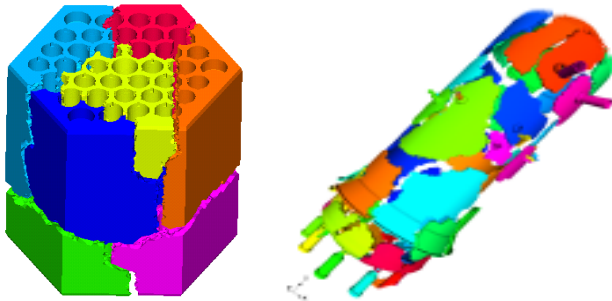


Fig. 3 Part decomposition of HTTR and ABWR model

The number of subdomains is the major concern in the domain decomposition method. We compare the computation time (for IBDD-DIAG) in two different computational environments corresponding to the number of subdomains and results are shown in Fig. 7. CM shows better performance than GP7000F. The ratio of the calculation time, time for input/output data and communication time corresponding to the number of subdomains are shown in the Fig. 8 (p-mode) and Fig. 10 (h-mode). In p-mode with the increase of number of subdomains, ratio of calculation time decreases while ratio of input/output time and communication time increases. Since the network speed of GP7000F is faster than the CM, the communication takes less time compared to the CM clusters. The Fig. 9 shows the computation time for h-mode in two different computation environments, which shows that when the number of subdomain is 3000 the computation time increases in CM cluster compared to the GP7000F. This is because, in the h-mode the more communication takes place between the Parents and Childs as the number of subdomains increases thereby in CM it takes more time for the communication. So in the parallel computing environment having low performance it is better to use the static load distribution mode and having high performance it is better to use the dynamic load distribution mode.

ABWR model is analyzed in the vector type supercomputer SR8000 changing number of processors. Fig. 11 (p-mode) and Fig. 12 (h-mode) shows the variation of the computation time using the different number of processors in SR8000. As the number processor increases, the difference decreases. As a result, the difference between the time required using 512 processors and 1024 processors is too small. So it is not necessary to use more than 1024 processors for this model.

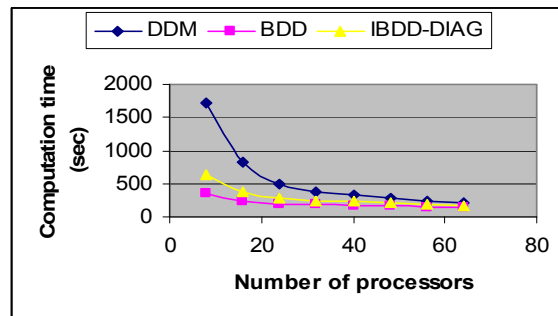


Fig. 4 Computation time in different number of processors

Table 1 Parallel Computational environments

	CPU	Memory	Network
PC Cluster	Pentium4, 2.0GHz	1GB/CPU	Ethernet (100Mbps)
GP7000F	SPARC64 , 333MHz	1GB/CPU	Ethernet (1Gbps)
Hitachi SR8000	250 MHz	1 GB/CPU	



Table 2 Model description

Model	Number of Nodes	Number of Elements
HTTR	1,893,340	1,1,67,26
ABWR	11,794,506	7,486,792

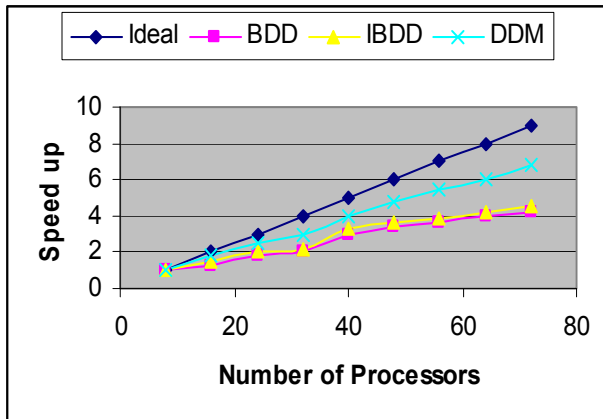


Fig. 5 speed up for different number of processors ( HTGR model)

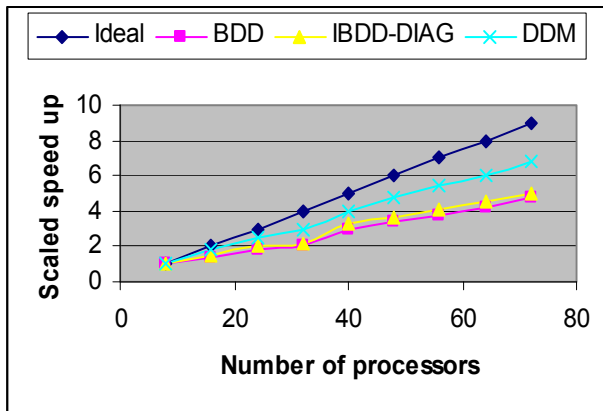


Fig. 6 Scaled speed up for different number of processors ( HTGR model)

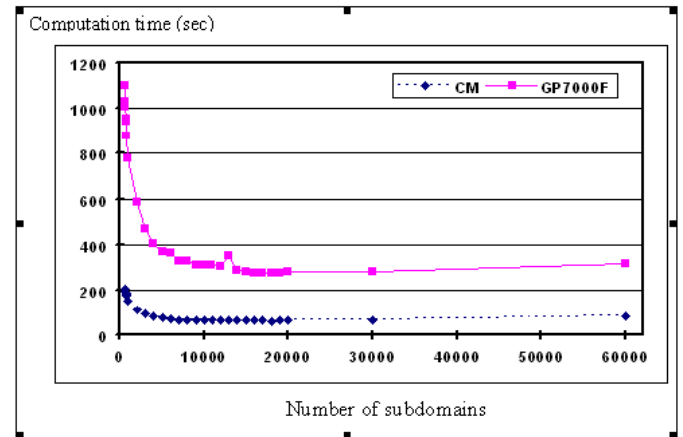


Fig. 7. Computation time in two different computational environments (p-mode)

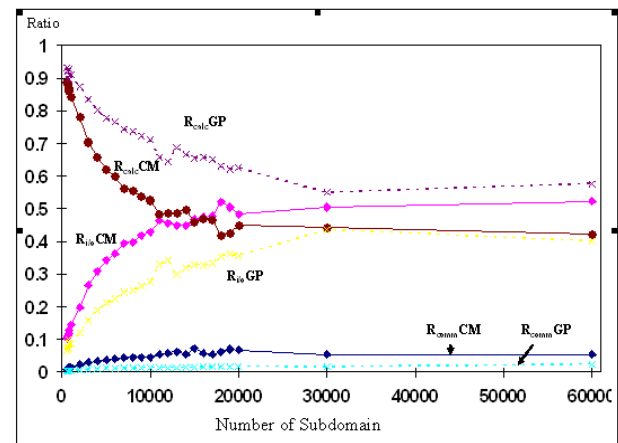


Fig. 8. Ratio of calculation, input/output and communication time (p-mode)

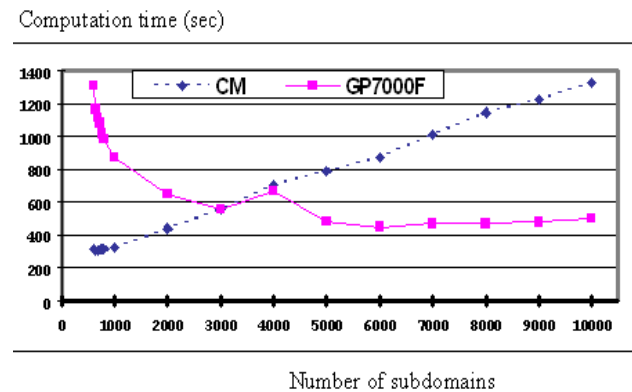


Fig. 9. Execution time in two different computational environments (h-mode)

Ratio

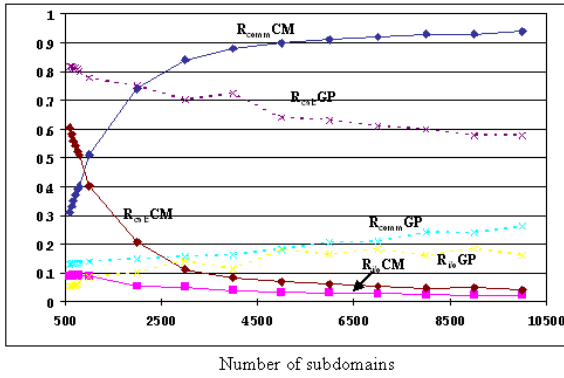


Fig. 10. Ratio of calculation, input/output and communication time (h-mode)

Computation time (sec)

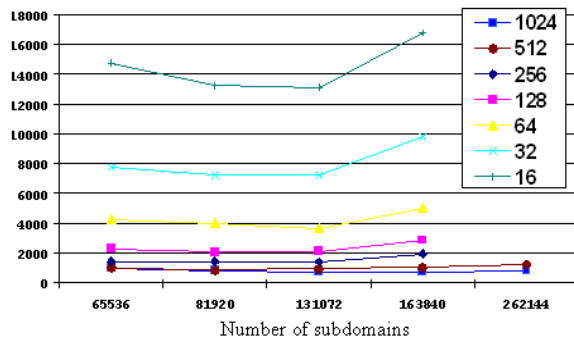


Fig. 11. Computation time vs number of subdomains in different number of processors (p-mode)

Computation time (sec)

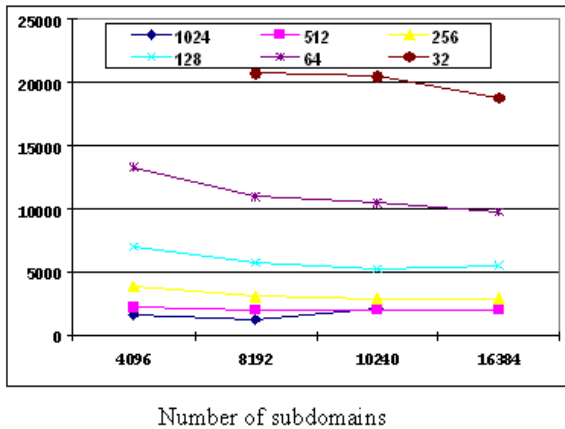


Fig. 12. Computation time vs number of subdomains in different number of processors (h-mode)

## VI. CONCLUSION

Two parallel approaches, static load distribution and dynamic load distribution of the domain decomposition method are adapted to the parallel computing environment and their parallel performances are analyzed.

Static load distribution seems to be more efficient in the parallel computer having low network speed while dynamic load distribution shows good performance in the high speed networked parallel computers. By the performance test, the effectiveness of the domain decomposition method in the parallel computer is verified. Important factors affecting the performance of the distributed parallel computing are found and analyzed. By using workstation cluster, a huge size of problem having over 11 million DOFs is solved successfully.

## REFERENCES

- [1] Almasi G S and Gottlieb A, Highly parallel computing [M]. *The Benjamin/Cummings Publishing Company, Inc.*, Redwood City, Ca, 1990
- [2] Farhat, C., Chen, P.S. and Mandel, J., A scalable lagrange multiplier based domain decomposition method for time-dependent problems, *International Journal for Numerical Methods in Engineering*, Vol. 38 , p. 3831-3853, 1995
- [3] Shioya, R., Ogino, M., Kanayama, H. and Tagami, D., Large scale finite element analysis with a balancing domain decomposition method, *Key Eng. Mate.*, 243-244 , p.21-26,2003.
- [4] A.M.M.Mukaddes, M. Ogino, H. Kanayama and R. Shioya, A scalable balancing domain decomposition based preconditioner for large scale heat transfer problems. *JSME International Journal*, Series B, vol 49, No. 2 pp-533-540, 2006
- [5] Kanayama, H., Ogino, M., Takesue, N and A.M.M.Mukaddes, Finite element analysis for stationary incompressible viscous flows using balancing domain decomposition, *Theoretical and Applied Mechanics*, 54, p.211-219, 2005.
- [6] Yagawa, G. and Shioya, R., Parallel finite elements on a massively parallel computer with domain decomposition, *Computing Systems in Engineering* 4:4-6 , p.495-503, 1993.
- [7] Kim S J, Kim J H. Large-scale structural analysis using domain decomposition method on distributed parallel computing environment. *Proceedings of High-Performance Computing in the Information Superhighway. HPC-Asia*, 573-678,1997.
- [8] Mandel, J., Balancing domain decomposition, *Comm. on Num. Meth. in Eng.*, Vol. 9 , p. 223-241,1993.

# Routing Algorithm for Vertically Stacked Optical Banyan Networks with the Time Complexity $O(N)$ using Pipelined Processors

Md. Mamun-ur-Rashid Khandker

Department of Applied Physics and Electronic Engineering, University of Rajshahi, Bangladesh.

**Abstract** — Centralized control routing algorithm for optical banyan networks on vertical stacking scheme has time complexity  $O(N \log_2 N)$  for rearrangeably nonblocking structure. A distributed algorithm with  $O(\log_2 N)$  has been recently proposed in which authors have considered an  $N$  completely connected processors to take routing decision which practically would be very difficult to implement for large  $N$ . In this paper we have proposed a distributed routing algorithm with processors work in pipelined fashion and take the routing decision in linear time. Also they do not need to be completely connected which makes it more practical for implementation.

**Index Terms** — Vertically Stacked optical Switch networks, Nonblocking Switch Networks, Completely connected graph, Bipartite graph.

## I. INTRODUCTION

The ubiquitous presence of the multimedia applications and the tremendous popularity of the World Wide Web have dramatically increased the amount of traffic over the Internet, which results the urgent requirement for high bandwidth links and switches. Due to the progress in DWDM technology, link bandwidth for carrying such traffic is no longer a problem. For DWDM networks, optical cross-connect (XOC) in nodes are key functional elements that steer whole network traffic by simultaneously switching a huge amount of optical flows. To build a large-scale OXC, generally small (e.g.  $2 \times 2$ ) switching elements (e.g. Directional couplers) are used as the basic building blocks, and are interconnected in a particular fashion to satisfy the required connectivity. Here we refer to the interconnection pattern of the optical links, the basic SEs, and the input/output ports of the switch, as optical switch network. Directional couplers (DC) [1] are considered as a promising candidate for the basic SE since it can handle optical signal of some terabits per second using WDM technology. It is notable that DC suffers from an intrinsic crosstalk problem [1][2], in which a portion of optical power in one waveguide of a DC will be coupled into the other waveguide unintentionally when two optical flows pass through the DC at the same time. This undesirable coupling effect is called first-order crosstalk, which may propagate downstream stage by stage, leading to a higher order crosstalk in each downstream stage with a decreasing magnitude.

According to the blocking properties, switches networks are classified as blocking and nonblocking. If a free input cannot be connected to a free output due to internal link-

blocking then it is considered as blocking. However, if multiple inputs intend to be connected with the same output, it causes output contention resulting in blocking to all requests but one, and mainly regarded as the property of the traffic pattern given to the network. Crosstalk in photonic switching networks adds a new dimension of blocking, called crosstalk blocking, which happens when an SE has two active inputs/outputs. A cost-effective solution to the crosstalk problem is to make sure that only one signal passes through each SE of a network at a time. Although this kind of space dilation of the network can eliminate the blocking and crosstalk completely, it may increase the hardware cost significantly if the interconnection pattern of the switching network is not chosen carefully. Since optical memory is not in a matured state to be used in an efficient queuing system, nonblocking networks have been favored in optical switching system. There are three different types of nonblocking networks: strictly nonblocking (SNB), wide-sense nonblocking (WSNB) and rearrangeably nonblocking (RNB). In both SNB and WSNB networks, a connection between a free input and free output can always be established without interrupting existing connections; however there is a difference. In SNB, any of the conflict-free path between a pair of input/output can be chosen, where as in WSNB, a smart algorithm is used to choose the path so that this connection does not result in blocking to any future connections. Usually a SNB network has got dedicated path for every input/output pair, and WSNB shares portion (links) of a path among different connections. Therefore, SNB networks have the higher hardware cost than WSNB networks but the latter have higher time complexity than the former one. RNB networks can also establish connections among idle inputs and outputs but has to allow rearrangements of the existing connections. It has the lowest hardware cost among the three, and the routing time complexity is often in between WSNB and SNB. Banyan networks [3], although a blocking networks, are attractive for constructing DC-based optical switches for their small depth, absolute loss uniformity (each path goes through exactly the same number of DCs), simple switch setting ability (self-routing, and therefore time complexity  $O(\log N)$ ), and for lower hardware cost (switch complexity  $O(N \log N)$ ). However, optical implementation of banyan networks results in severe blocking and crosstalk. To deal with this situation, it is a novel approach of keeping the whole network nonblocking as well as crosstalk-free by vertical stacking multiple copies (planes) of an optical banyan network [4]. This network is called vertically

stacked optical banyan (VSOB) network which preserves all the property of banyan networks except self-routing characteristics. It has been shown that total number of  $T = 2^{\lfloor (\log N + 1)/2 \rfloor}$  planes are required to make the switch network rearrangeably nonblocking with a routing algorithm that has the time complexity  $O(N \log N)$ . This time complexity may be considered too high when  $N$  is large. For optical interconnections where data rates are very high, and has no viable optical memory system, routing decision should be taken as fast as possible. Zheng *et. al.* [5] have proposed a  $O(\log N)$  parallel routing algorithm where  $N$  completely connected processors have been used, and information is exchanged among  $N$  nodes at every computation steps. They have recursively used balanced 2-coloring algorithm of a bipartite graph  $G(N, k, g)$  to get the  $g$ -edge coloring of the graph.  $G(N, k, g)$  is a bipartite graph with  $N/g$  vertices (in each of  $V_1$  and  $V_2$ ) and  $k$  edges where at most  $g$  edges are incident at any vertex. When  $k = N$ , creating such a graph would take  $O(N)$  time if used single processor, but in this paper it has been considered as  $O(1)$  as all  $N$  processors are completely connected. It is practically very difficult to develop a completely connected network when  $N$  is large. Therefore, we have been encouraged to develop a more pragmatic solution to this problem in which  $\sqrt{N}/2$  processors work in pipeline fashion, and therefore, are connected very loosely. Information is exchanged only between two adjacent processors.

In this paper we propose a routing algorithm in which controls are distributed over groups of inputs to achieve parallelism in computing crosstalk free realizable partial permutation in the  $O(N)$  time complexity. All logarithms used in this article are of base 2.

## II. BACKGROUND

### A. Vertically Stacked Banyan (VSB) Networks

Vertically Stacked Banyan (VSB) networks [4][6] are constructed by stacking  $T$  copies of  $N \times N$  optical Banyan networks. If  $T$  is number of banyan networks required to make the network rearrangeable nonblocking as well as crosstalk-free then,

$$T = 2^{\lfloor (n+1)/2 \rfloor} \quad (1)$$

Here  $n$  is the number of stages in the banyan networks, i.e.  $n = \log N$ . Figure 1 shows an  $N \times N$  VSB network. Circles at inputs represents  $1:T$  input switches and circles at outputs represents  $T:1$  combiner. An input request is sent only one Banyan network (plane) at a time. So, only one of the  $T$  links is busy at any time. On the output side, a combiner receives signal from only one plane since the request is considered to be a permutation. Therefore, switching is not required at the output side. The input request is split into partial permutations (CRPP) in such a way that each of those CRPP can be realized in a Banyan networks without

first-order crosstalk. A central control determines  $T$  partial permutations, and sets the input switch accordingly to route all signals of a CRPP to a Banyan plane. Being a Banyan network, each plane is self-routing

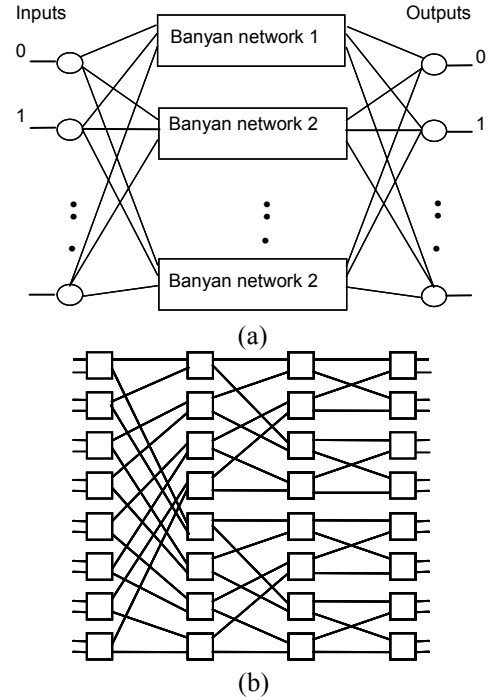


Fig. 1. a)  $N \times N$  VSB network. b) A 16x16 Banyan network.

### i) Crosstalk-free Realizable Partial Permutation (CRPP)

A full permutation is split into  $T$  partial permutation in such way that each of them can be realized in a banyan network without any crosstalk and blocking. Each of these partial permutations is called Crosstalk-free Realizable Partial Permutation (CRPP) [4]. Considering the traffic arrives in the form of permutation, routing of the request is being done by (a) Decomposing a permutation into CRPPs (b) sending  $T$  CRPP to  $T$  planes of Banyan networks.

### ii) CRPP Decomposition Algorithm

As indicated in [4], a CRPP decomposition algorithm can easily be obtained by repetition of the simple bipartite graph coloring with two colors. Due to space limitation here we describe the algorithm in short and simple way.

A bipartite graph  $G = (V, E)$  is constructed where each input switch is considered a node in set  $V_1$  (input set) and each output switch is considered a node in set  $V_2$  (output set). An edge represents that there is a connection request from an input switch to an output switch. Since each switch is a  $2 \times 2$  DC, the node degree is always 2. Using Euler's split  $G$  is split into  $G_1$  (consisting of all forward orientation) and  $G_2$  (consisting of all reverse orientation). Each of these bipartite graphs has  $N/2$  input nodes,  $N/2$  output nodes and

$N/2$  edges, and represents a partial permutation. However, these partial permutations only guarantee that they are crosstalk free in the input and the output stage.

In the next iteration,  $G_1$  is reconstructed where adjacent two nodes (from input set and output set) are merged. Therefore, this new  $G_1$  has  $N/4$  input nodes,  $N/4$  output nodes and  $N/4$  edges. The node degree is again 2. Using the

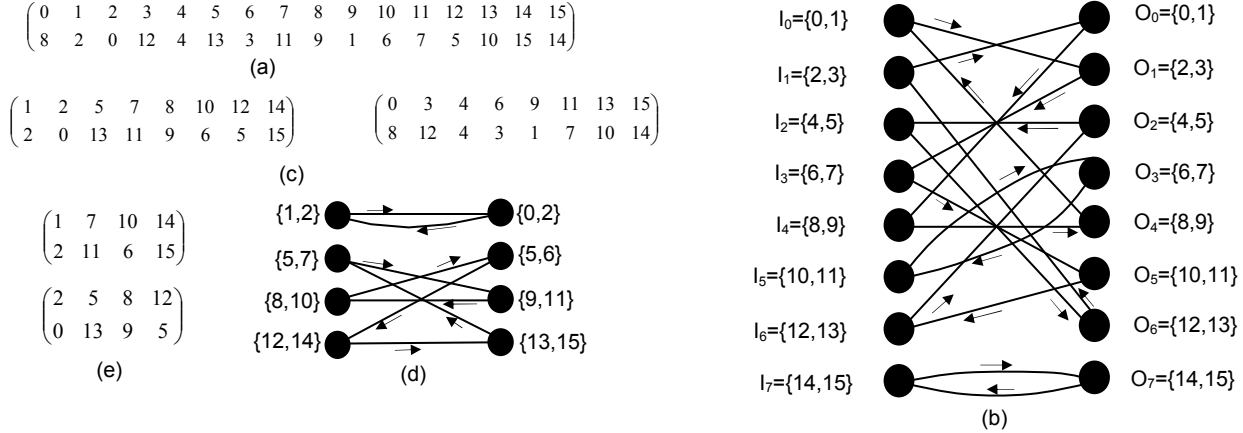


Fig. 2. Illustration of CRPP algorithm. (a) The input request (b) Bipartite graph  $G=(V_1, V_2, E)$  (c) Two partial permutations resulted from G (d) Graph  $G_1'$  for the left partial permutation. (e) Two CRPP resulted from graph  $G_1'$ .

Considering that it takes  $O(\sqrt{N})$  time to send T CRPP to T planes, the time complexity for routing a permutation in VSB networks is  $O(\sqrt{N} + N \log N) = O(N \log N)$ .

### III. DISTRIBUTED CONTROL ALGORITHM

#### A. Problem Formulation

First we show that the input-output request patterns can be represented by a  $2^t$  regular bipartite graph. Then we prove that the graph has such perfect matching which can be routed through a banyan plane without any blocking. We will use few pipelined processors for finding and routing these perfect matching in the  $O(N)$  time complexity.

For a networks having odd number of stages,  $n$ , stage  $(n+1)/2$ -th is the middle stage, and for even  $n$ , stages  $n/2$ -th and  $(n/2+1)$ -th are middle stage switches. Since each SE is connected to two SEs of the previous stage (as well as of the next stage), the number of inputs (outputs) intersects at a middle stage SE is  $2^{\lfloor (n+1)/2 \rfloor}$  i.e.  $T$ . Therefore, we propose input and output groups as per following definitions.

**Definition 1:** Input Groups,  $I_i$ , is defined as the set of inputs that intersect each other in a middle-stage SE. Therefore,

Euler's split algorithm in the same way as mentioned before, the  $G_1$  is split into two graphs.  $G_2$  is also split in the same way. So, we get 4 partial permutations each is represented by a graph. Using this technique recursively,  $T$  partial permutations are generated. Each of these  $T$  partial permutations is a CRPP. Xiang *et. al* [4] has shown that it takes  $O(M \log N)$  time to generate these CRPPs.

$$I_i = \{iT + 0, iT + 1, \dots, iT + (T - 1)\}, \text{ where } 0 \leq i \leq \frac{N}{T} - 1 \quad (2)$$

**Definition 2:** Output groups,  $O_i$ , is defined as the set of outputs that intersect each other at a middle-stage SE. Therefore,

$$O_i = \{iT + 0, iT + 1, \dots, iT + (T - 1)\}, \text{ where } 0 \leq i \leq \frac{N}{T} - 1 \quad (3)$$

**Definition 3:** Two connections,  $xy$  and  $pq$  are said to have group conflict if  $p, x \in I_i$  and/or  $q, y \in O_j$ , where  $0 \leq i, j \leq T - 1$ .

Figure 1 shows the grouping of inputs and outputs for a 16x16 banyan networks.

It is interesting to see that two requests,  $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 4 \\ 7 \end{pmatrix}$

have different input groups and output groups, and they have disjoint paths from input to output. Therefore, these two requests can be routes through a 16x16 banyan network without blocking and crosstalk. We need to know the minimum number of such connections for a Banyan network. We give two essential Lemmas in this regard.

**Lemma 1:** Any middle stage switch can carry connections generated (destined) from (to) only one input (output) group.

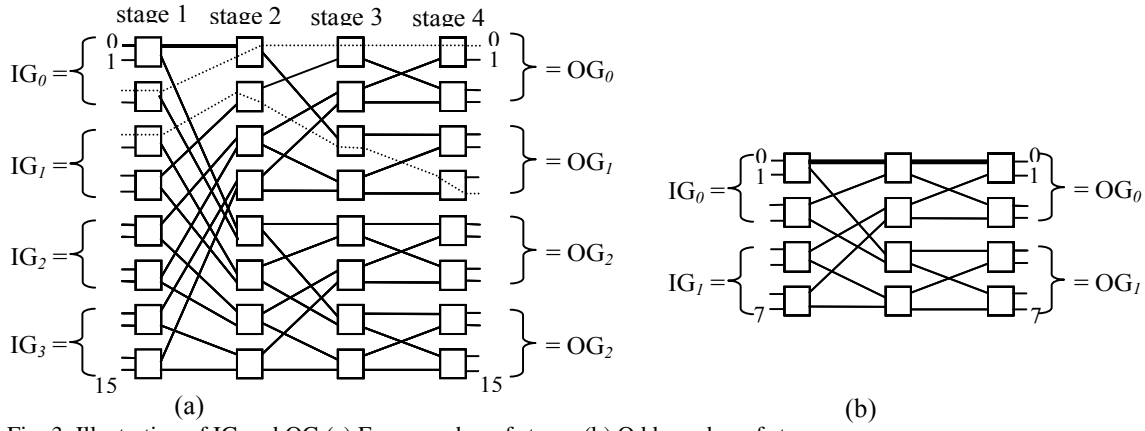


Fig. 3. Illustration of IG and OG (a) Even number of stages (b) Odd number of stages

**Proof:** Let's first prove it for inputs. A switch in stage  $i$  is connected to two adjacent switches of stage  $i-1$  (and  $i+1$ ). A switch at stage  $i$  are connected to only  $2^i$  adjacent inputs.

Therefore, a middle stage switch is connected to only  $2^{\lfloor \frac{n+1}{2} \rfloor}$  (i.e  $T$ ) adjacent inputs. For example, first  $T$  inputs are connected to  $0^{\text{th}}$  and  $k \cdot 2^{\lfloor \frac{n}{2} \rfloor}$ -th SEs of the middle stage, where  $0 < k < 2^{\lfloor \frac{n}{2} \rfloor}$  (for odd  $n$ ) and  $0 < k < 2^{\lfloor \frac{n-1}{2} \rfloor}$  (for even  $n$ ). No other connections can pass through these switches. This group of inputs corresponds to an input group. Therefore, any SE in the middle stage carry connections generated from only one input groups.

An SE in the middle stage is connected to two SEs of  $2^{n-\lfloor \frac{n+1}{2} \rfloor} \times 2^{n-\lfloor \frac{n+1}{2} \rfloor}$  banyan networks. Therefore, an SE of the middle stage carry connections destined to these two  $2^{n-\lfloor \frac{n+1}{2} \rfloor} \times 2^{n-\lfloor \frac{n+1}{2} \rfloor}$  networks. The total number of inputs of these two networks is  $2 \cdot 2^{n-\lfloor \frac{n+1}{2} \rfloor} = 2^{\lfloor \frac{n+1}{2} \rfloor} = T = |O_i|$ . So, we conclude that a middle stage SE can carry connections destined to only one output group. *QED*

**Lemma 2:** Connections without group conflicts can be established in a Banyan networks without crosstalk.

**Proof:** Let us consider any two connections  $xy$  and  $pq$  those do not have any group conflict. Since  $x \in I_i, p \in I_j$  and  $i \neq j$ , they belongs to different middle-stage SEs. Therefore, their routes up to middle stage are switch-disjoint. Similarly,  $y \in O_i, q \in O_j$  and  $i \neq j$ , so their routes from output to the middle stage are also switch-disjoint. Since Banyan network is a unique-path network, two sections of a route of the same connection must coincide at one and only one middle-stage SE. Therefore, the whole routes for  $xy$  and  $pq$  are switch-disjoint. Since they are disjoint, all SEs along their paths contain only one signal. Thereby, both  $pq$  and  $xy$  can be established crosstalk-free.

When  $n$  is even, there are two middle stages; stage  $(n/2)$  and stage  $(n/2+1)$ . Two connections generated from different input groups are switch-disjoint up to stage  $n/2$ , and two connections destined to two different output groups are switch-disjoint up to stage  $n/2+1$ . An SE in the  $n/2+1$ -th stage is connected to two SEs in the  $n/2$ -th stage those belong to different input groups. On the other hand two SEs in the  $n/2$ -th stage are connected to two SEs in the  $n/2+1$ -th stage those belong to different output groups. That means there is only one link between these two middle stages for a connection generated in an input group and destined to a particular output group. Therefore if two connections do not have group conflict they pass through different SEs in stages  $n/2$  and  $n/2+1$ , which ensures that both the connections have switch-disjoint paths from input to output. *QED*

Needless to say that if two connections have disjoint paths from input to output they will certainly not produce any blocking. Therefore, a crosstalk free permutation must be nonblocking in the network. Now we give theorem 1 by which we can determine the minimum number of crosstalk-free connections in an  $N \times N$  optical Banyan switch networks.

**Theorem 1:** An  $N \times N$  Optical Banyan networks can establish at least  $N/T$  connections crosstalk-free.

**Proof:** Let  $S$  be the size of input or output groups. Then  $S = |I_i| = |O_i|$ . By Lemma 2, one pair of input-output groups can have at least one connection crosstalk-free. For a switch network of size  $N \times N$ , there are  $N/S$  pairs of input-output groups. Therefore, at least  $N/S$  connections can be established without any crosstalk. Again,  $S = 2^{\lfloor \frac{n+1}{2} \rfloor} = T$ . Therefore, an  $N \times N$  optical Banyan switch network can establish at least  $N/T$  connections crosstalk-free.

Since an  $N \times N$  VSB network has  $T$  vertically stacked  $N \times N$  optical Banyan networks it can establish  $T \cdot \frac{N}{T} = N$  connections simultaneously without any crosstalk. Considering a group as a node we present a bipartite graph  $G(V_1, V_2, E)$ , in which  $V_1$  is the set of input nodes,  $V_2$  is the set of output nodes and  $E$  is the set of edges connecting them. An edge corresponds to a connection request. Therefore,  $E$  represents a request pattern which is essentially a permutation. Figure 2(b) shows a bipartite graph for the connection pattern in Figure 2(a).

Finding a set of  $N/T$  crosstalk-free connections is a problem of finding a perfect matching in  $G$ . To realize all  $N$  connections  $T$  such perfect matching is required to be found. We prove that  $G$  is  $T (=2^t)$  regular bipartite graph and it has  $T$  perfect matchings (PMs). Then we propose an algorithm for calculating these PMs in  $O(N)$  time. Each PM corresponds to a CRPP. First we discuss PM and its properties.

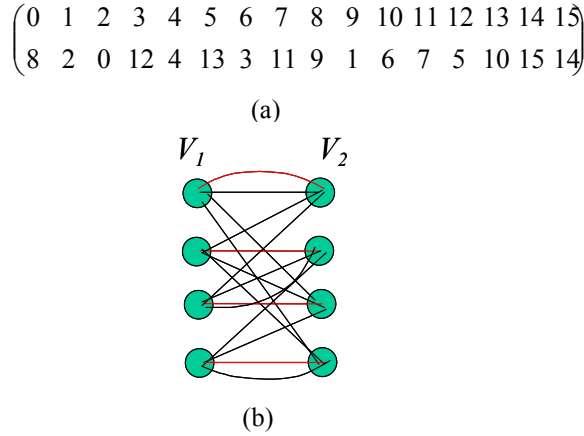


Fig. 4. (a) A connections pattern. (b) Corresponding  $G(V_1, V_2, E)$ .

A matching in  $G$  is a subset of  $E$  such that no two edges share a common vertex. A perfect matching in  $G$  is such that every  $v \in V_1$  is the endpoint of an edge. For example, set of red edges in Figure 2 is a perfect matching in  $G$ .

P. Hall's Theorem [8] states that in a bipartite graph  $G(X, Y, E)$  there exists a perfect matching of  $X$  into  $Y$  if and only if, for any  $S \subseteq X$ ,  $|S| \leq |\Gamma(S)|$ , where  $\Gamma(S)$  is the set of nodes adjacent to nodes in  $S$ .

Our bipartite graph  $G(V_1, V_2, E)$  is a  $T$ -regular graph. For any  $v \in V_1$ ,  $1 \leq |\Gamma(v)| \leq T$ . Therefore,

$$|\Gamma(S)| = \sum_{v \in S} |\Gamma(v)| \geq |S| \quad (3)$$

So,  $G$  has a perfect matching.

## B. Proposed Routing Algorithm

Our algorithm for decomposing a permutation into CFPP is based on Gabow's algorithm for edge coloring of bipartite multigraphs [7] in which author showed that a PM can be found in  $O(m)$  time where  $m$  is the number of edges in the graph. We have introduced parallelization in process of partitioning a  $2^t$  regular graph into  $2^{t-1}$  regular graph. We summarize our algorithm as follows:

Step 1: Construct the graph  $G(V_1, V_2, E)$ .

Step 2: Using Euler orientation, split  $G$  into two  $2^{t-1}$  regular subgraph  $G_1(V_1, V_2, E_1)$  and  $G_2(V_1, V_2, E_2)$ . All forward edges (oriented from  $V_1$  to  $V_2$ ) form  $E_1$  and all backward edges (oriented from  $V_2$  to  $V_1$ ) form  $E_2$ .

Step 3: process each of the subgraphs simultaneously by different processors.

Step 4: Repeat step 2 until all subgraphs are 1-regular bipartite graph.

Step 1 and step 2 require  $O(N)$  time. A processor has to visit all  $N$  edges in the first call of the step 2. In the second call the processor visits  $N/2$ , in the third call  $N/4$  and so on. Therefore the time complexity for one processor is  $O(N + N/2 + N/4 + \dots) = O(2N) = O(N)$ . However,  $2^{\text{nd}}$  processor has to visit  $N/2$  edges,  $3^{\text{rd}}$  processor visits  $N/4$  edges, and so on. Therefore, all the processor ends their operation at the same time. Since all PMs are found at the same time, the overall time complexity is  $O(N)$ .

Each of the  $T/2$  processors takes the responsibility of sending 2 CFPP to 2 planes. First two planes are assigned to first processor,  $2^{\text{nd}}$  two planes to  $2^{\text{nd}}$  processor and so on.

## IV. CONCLUSION

A simple way to parallelizing the routing process of VSOB networks has been proposed. Only  $\sqrt{N}/2$  processor are required to achieve the routing time complexity of the  $O(N)$ . We hope the findings of this paper will help switch designers implement a faster nonblocking optical switch network.

A centralized routing algorithm of the  $O(N)$  is under investigation.

## REFERENCE:

- [1] V. R. Chinni et.al., "Crosstalk in lossy directional coupler switch," *Journal of Lightwave Technology*, vol.13, no.7, pp.1530-1535, July 1995.
- [2] H. S. Hinton, *An Introduction to Photonic Switching Fabrics*, New York: Plenum publisher, 1993.
- [3] G. R. Goke and G. J. Lipovski, "Banyan networks for partitioning multiprocessor systems," *Proc. 1<sup>st</sup> Annu. Symp. Comp. Architecture*, pp.21-28, 1973.
- [4] Xiaohong Jiang, M. M. R. Khandker, Hong Shen and Susumu Horiguchi, "A Nonblocking Optical Switching Network for Crosstalk-free Permutation," *IEICE Trans. Communications*, Vol.E86-B, No.12, Dec.2003.

- [5] E. Lu and S.Q. Zheng, "Parallel routing algorithms for nonblocking electronic and photonic multistage switch networks," <http://www.utdallas.edu/~enyue/research.html>.
- [6] G. Maier and A. Pattavina, "Design of photonic rearrangeable networks with zero first-order switching-element-crosstalk," *IEEE Trans. Commun.*, vol. 49, no. 7, pp. 1268-1279, July.2001.
- [7] H. N. Gabow, "Using Euler partitions to edge color bipartite multigraphs," *Intl. J. Comput. Inform. Sci.* 5 (1976) 345-355.
- [8] [Heckman, Chris](#). *Hall's Condition*, From *MathWorld*--A Wolfram Web Resource, created by [Eric W. Weisstein](#). <http://mathworld.wolfram.com/HallsCondition.html>



# A Crosstalk Free Routing Algorithm of Generalized Recursive Non-blocking Network

Most. Arjuman Sultana, Gita Chowdhury, M. M. Hafizur Rahman

Dept. of Computer Science and Engineering,  
Khulna University of Engineering and Technology (KUET), Khulna-9203, Bangladesh  
E-mail: arjucse03@yahoo.com, gita\_csekh@yahoo.com, hafiz90305@gmail.com

**Abstract**—Crosstalk is an intrinsic drawback of an optical network, and avoiding crosstalk is important for making fruitful application of optical switching network. In this paper, we apply the well-known semi-permutation technique on Generalized Recursive Network (GRN) to make it crosstalk free. GRN has reasonable path independent signal loss and crosstalk. Using the methodology presented in this paper, that reasonable crosstalk is avoided. For the GRN network, we show that any semi-permutation is realizable in one pass and any permutation is realizable in two passes under the constraint of avoiding crosstalk.

**Index Terms**—GRN, Crosstalk, Semi-permutation, and Crosstalk-free routing algorithm.

## I. INTRODUCTION

Two major problems in designing optical switch networks are signal loss and crosstalk. There are two ways in which optical paths can interact in planar switching networks. First, two optical channels on different waveguides may cross each other in order to obtain a particular topology. This is called channel crosstalk. Alternatively, two paths sharing a switching element will experience some undesired coupling from one path to the other. This is called optical crosstalk [1].

Generalized Recursive Non-blocking Network (GRN) is a new approach for building non-blocking optical Multistage Interconnection Networks (MINs) [1]. It can be built from strictly nonblocking to blocking depending on building blocks and thereby it can be traded of cost with performance. It has very easy self routing algorithm [1, 4]. However, using this routing algorithm the network is not crosstalk free. Therefore, the main objective of this paper is to make the GRN crosstalk free using well known semi-permutation [2]. In this paper, a new crosstalk free routing algorithm is proposed for Generalized Recursive Network (GRN). The algorithm is derived based on the idea of the semi-permutation.

The remainder of the paper is organized as follows. In section II, we give an overview idea of preliminaries. In section 4, we discuss details about the crosstalk-free routing algorithm of GRN along with the proposed routing. In section IV, we discuss the performance evaluation of our proposed routing algorithm. Finally, Section V concludes the paper.

## II. PRELIMINARIES

When two optical paths interact to each other, then crosstalk occurs. The number of crosstalk switches along a path of a signal from input to output is a representative measure of crosstalk of the switch architecture. The number of switching elements in a signal path is a measure of signal loss. Clearly signal loss is proportional to the number of switches that has to be crossed by a signal before reaching the output. The total number of switches required, to build a network is a representative measure of the hardware complexity of the network.

Centralized routing control of switches is centralized; the decision about how to send the input signal to its desired output is taken at a central location for any input. Distributed routing, control units may be distributed among several switches. Self-routing, every switching elements in the architecture can decide about the routing without having any information of other switches states [4].

Self-routing reduces the connection complexity for the control lines [1]. As the control function is distributed among different switching modules, the network is less susceptible to faults of a switch. Most of the existing switch networks are not self-routing for all permutations. Banyan and its equivalent networks are self-routing, but they cannot route all  $N!$  permutations. Benes network with more than five inputs are not self-routing. Clos network whose first stage contains more than two switches are not self-routing. Crossbar network is self-routing for all permutations but it requires the header of size  $O(N)$ .

### A. Generalized Recursive Network (GRN)

GRN has better performance than any other indirect interconnection networks. Because it has low signal loss and crosstalk. In this architecture, any  $M \times N$  non-blocking switch network can be built with building blocks of given size  $m \times n$ , where  $M, N, m, n$  are all powers of integers. GRN is self-routing for all  $N!$  permutations and the propagation delay is  $O(\log N)$ . The switches of GRN maintains the connection pattern of shuffle-exchange fashion. If the size of the building blocks is  $m \times n$ , then the number of building blocks required is  $\frac{M}{m}$  and the

number of switching elements required is  $(n(\frac{M}{m} - 1))$ . The number of recursions in GRN is  $\log(N/n)$ .

An  $M \times N$  GRN constructed from  $m \times n$  strictly non-blocking building blocks is also a strictly non-blocking network and the parameters switch count ( $T_N$ ), maximum signal loss ( $S_N$ ) and maximum crosstalk ( $C_N$ ) are given by the following equations, respectively [1]:

$$T_N = \left( \frac{MN}{m} + \frac{MN}{n} - M - N \right) + \frac{MN}{mn}$$

where  $\frac{MN}{mn}$  represents the total number of building blocks of size  $m \times n$

$$S_N = \log\left(\frac{M}{m}\right) + \log\left(\frac{N}{n}\right) + S_m$$

where  $S_m$  is the maximum signal loss of the building block.

$$C_N = C_m$$

where  $C_m$  is the maximum crosstalk of the building block. It means that the crosstalk is only available in the building block.

### B. Limitation of GRN

There are three stages in GRN. The stages are input stage, output stage, and building-block-stage. In  $M \times N$  GRN, with building block size of  $m \times n$ , a signal has to cross  $\log(N/n)$  input stage switches, and  $\log(M/m)$  output stage switches in addition to one building block. Since the input stage switches are  $1 \times 2$  switches, they do not contribute to crosstalk. Likewise, the output stage switches are  $2 \times 1$  and they do not have to bear signals at both of their inputs at the same time. So, they do not contribute to crosstalk as well. The only possible elements that contribute to crosstalk are the building blocks. And our work is concentrate on this issue to avoid the crosstalk.

## III. OUR APPROACHES TO SOLVE THE CROSSTALK OF GRN

One way to solve the crosstalk problem is a space domain approach, where a MIN is duplicated and combined to avoid crosstalk. But this approach uses more than double of the original network hardware to achieve the same permutation capability.

Another way to implement this idea is a time domain approach [2]. A set of connections is partitioned into several subsets such that the connections in each subset can be established simultaneously in a network without crosstalk. In our work, we will consider the idea of time domain approach. Because each switch can only pass one signal at a given time.

There is a nice property for permutation that any permutation can be decomposed into semi-permutations. A

routing algorithm for realizing a semi-permutation in a generalized recursive non-blocking network (GRN) is also presented. Routing is the act of moving information across an inter network from a source to a destination. Semi-permutation, which is a special type of partial permutation and has the maximum potential to be realized in an optical network without crosstalk [2]. This permutation ensures that there is only one active link passing through each input switch and output switch that is it eliminates crosstalk. The semi-permutation concept is accomplished based on the idea of the bipartite graph. Bipartite graph (i.e, a set of graph vertices decomposed into two disjoint sets such that no two graph vertices within the same set are adjacent) such that every pair of graph vertices in the two sets are adjacent.

### A. Semi-permutation

Semi-permutation for a network is a pairing of its inputs and outputs such that each input appears in exactly one pair and each output appears also in exactly one pair. In other words, a permutation is a full one-to-one mapping between the network inputs and outputs. For an  $n \times n$  network, suppose input  $x_i$  is mapped to output  $y_i$ , where  $x_i = i$  and  $y_i \in 0, 1 \dots n-1$  for  $0 \leq i \leq n-1$ , we denote this permutation as,

$$\begin{pmatrix} x_0 & x_1 & \dots & x_{n-1} \\ y_0 & y_1 & \dots & y_{n-1} \end{pmatrix}$$

A crosstalk-free optical network can not realize permutation in a single pass, since at least the two input links on an input switch or the two output links on an output switch cannot be active in the same pass. For an optical network consisting of  $2 \times 2$  switches, it is useful to introduce the following,

A partial permutation,  $\begin{pmatrix} x_0 & x_1 & \dots & x_{n/2-1} \\ y_0 & y_1 & \dots & y_{n/2-1} \end{pmatrix}$  of an  $n$ -element set  $0, 1, \dots, n-1$ , where  $n$  is an even integer,  $x_i, y_i \in 0, 1, \dots, n-1$  and  $x_0 < x_1 < \dots < x_{n/2-1}$ , is referred to as a semi-permutation of the  $n$ -element set, if  $\{\lfloor \frac{x_0}{2} \rfloor, \lfloor \frac{x_1}{2} \rfloor, \dots, \lfloor \frac{x_{n/2-1}}{2} \rfloor\} = \{0, 1, \dots, n/2-1\}$  and  $\{\lfloor \frac{y_0}{2} \rfloor, \lfloor \frac{y_1}{2} \rfloor, \dots, \lfloor \frac{y_{n/2-1}}{2} \rfloor\} = \{0, 1, \dots, n/2-1\}$

### B. Number of Semi-permutation in GRN

Total number of semi-permutation for the GRN is  $\#semi-permutation = 2 \times \log 2^{\lfloor \log_2 buildingblock \rfloor}$ . Our proposed crosstalk free routing algorithm of GRN is properly crosstalk free at building block. For an  $8 \times 8$  GRN network with  $2 \times 2$  building block, maximum number of semi-permutation is  $\log_2^2 = 1$ . For a  $4 \times 4$  building block, maximum number of semi-permutation is  $\log_2^4 = 2$ . Thus, we can find out the total number of semi-permutation in GRN network. But if we replace the building block by  $4 \times 4$  wide-sense non-blocking Benes network then the number of semi-permutation is reduced.

### C. Bipartite Graph

A bipartite graph is a graph whose vertices can be divided into two disjoint sets  $V_1$  and  $V_2$  such that every edge connects a vertex in  $V_1$  and one in  $V_2$ ; that is, there is no edge between two vertices in the same set [1].

Now if we considered an undirected bipartite graph  $G = (V_1, V_2, E)$ . The vertex sets of  $G$  are defined as  $V_1 = \{A_0^{[1]}, A_1^{[1]}, \dots, A_{n/2-1}^{[1]}\}$ ,  $V_2 = \{A_0^{[2]}, A_1^{[2]}, \dots, A_{n/2-1}^{[2]}\}$  and the edge set  $E$  of  $G$  is defined as: for any one-pair mapping  $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$  in the permutation. If  $x_i \in A_{j_1}^{[1]}$  and  $y_i \in A_{j_2}^{[2]}$ , then there is an edge between vertex  $A_{j_1}^{[1]}$  and vertex  $A_{j_2}^{[2]}$  in  $E$ . We also assign each edge in  $E$  a label representing the corresponding one-pair mapping in the permutation. Here  $n$  is an integer value.

This bipartite graph has some properties. This graph may have parallel edges and each vertex of this graph has even degree. This graph may consist of more than one connected component and each edge traversed exactly once. It's vertices  $|V_1| = |V_2| = n/2$  and edges  $|E| = n$ . Each bipartite graph is a perfect graph. In this case, we considered cyclic bipartite graph, that each vertex of this graph has degree two.

### D. Decomposition algorithm of Bipartite Graph

Any permutation can be decomposed into two semi-permutation. Here we considered  $8 \times 8$  network, with  $2 \times 2$  switching building block. The input and output can be decomposed into two semi-permutation as follows,

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 7 & 5 & 3 & 4 & 6 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 3 & 4 & 6 \\ 1 & 5 & 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 5 & 7 \\ 0 & 7 & 4 & 2 \end{pmatrix}$$

- First: Construct a bipartite graph  $G$  for the given permutation.
- Second: We started from a vertex of this component in  $V_1$ , traverse through an unvisited edge to the neighbor vertex in  $V_2$ , back and forth until return to the starting vertex. If the traverse direction is from  $V_1$  to  $V_2$ , is called "forward" edges and marked "backward" if the direction is opposite.
- Third: take all one-pair mappings corresponding to the edges marked with "forward", to form one semi-permutation and "backward" to form another semi-permutation.

The above algorithm is correct because this algorithm maintains the properties of the bipartite graph  $G$ . Now if we apply the above decomposition algorithm in one permutation that decomposed into two semi-permutation,

then the bipartite graph and edge traverses are shown in below,

$$e_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, e_3 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, e_4 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, e_5 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, e_6 = \begin{pmatrix} 6 \\ 6 \end{pmatrix}, e_7 = \begin{pmatrix} 7 \\ 2 \end{pmatrix},$$

Then we find the "forward" pairs  $e_0, e_3, e_4$ , and  $e_6$  to form,

$$\begin{pmatrix} 0 & 3 & 4 & 6 \\ 1 & 5 & 3 & 6 \end{pmatrix} \quad (1)$$

and the "backward" pairs  $e_1, e_2, e_5$ , and  $e_7$  to form,

$$\begin{pmatrix} 1 & 2 & 5 & 7 \\ 0 & 7 & 4 & 2 \end{pmatrix} \quad (2)$$

which completes the decomposition shown in Figure 1,

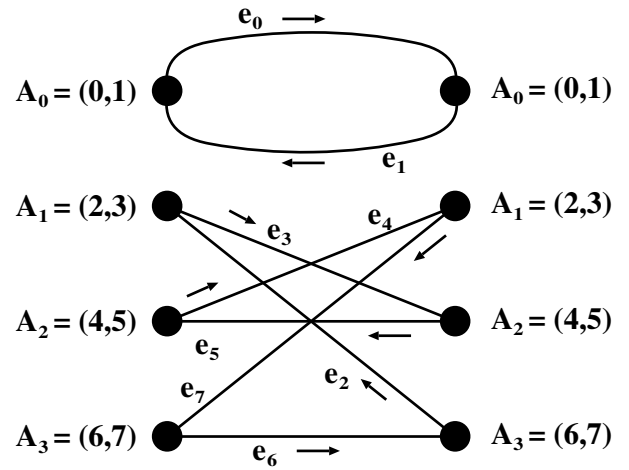


Fig. 1. The bipartite graph and edge traverses of the decomposition

The set of all edges marked with "forward" is a perfect matching of the bipartite graph  $G$ , and so is that marked with "backward". The time to construct the bipartite graph is proportional to the number of pairs in the permutation, i.e.  $O(n)$ , and the time to traverse all edges is  $O(|E|) = O(n)$ . The complexity of the decomposition algorithm is  $O(n)$  [3]. Now we can again decompose the "forward" pair from Eq. 1, according to bipartite graph and we get "forward" pairs  $e_0$  and  $e_3$  and "backward" pairs  $e_1$  and  $e_2$ .

Forward =  $\begin{pmatrix} 0 & 6 \\ 1 & 6 \end{pmatrix}$ , Backward =  $\begin{pmatrix} 3 & 4 \\ 5 & 3 \end{pmatrix}$ , and the bipartite graph for the semi-permutation shown in Eq. 1 is depicts in figure 2.

Similarly, we can also decompose the backward pair of equation 2, according to bipartite graph, we get again the "forward" pairs  $e_0$  and  $e_2$  and "backward" pairs  $e_1$  and  $e_3$ .

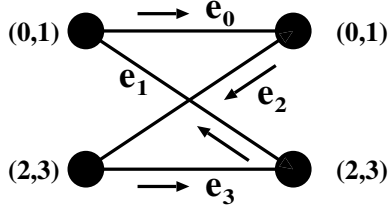


Fig. 2. Forward and backward edges of Eq. 1 traverse along the decomposition

Forward =  $\begin{pmatrix} 1 & 5 \\ 0 & 4 \end{pmatrix}$ , Backward =  $\begin{pmatrix} 2 & 7 \\ 7 & 2 \end{pmatrix}$ , and the bipartite graph for the semi-permutation shown in Eq. 2 is depicts in figure 3.

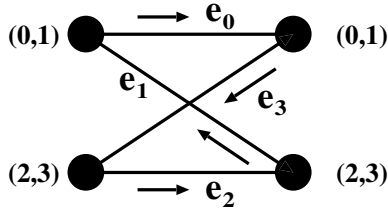


Fig. 3. Forward and backward edges of Eq. 2 traverse along the decomposition

#### E. Routing Mechanism of GRN

GRN networks maintain self-routing principle. The whole routing mechanism of GRN from an input to an output is divided into three steps:

- 1) Routing in the input stage
- 2) Routing in the building-block-stage, and
- 3) Routing in the output stage

Each input and output switches uses 1 bit to decide it's state. Each building block has its own routing strategy. It can route a signal properly with the given source and destination addresses. The input switch uses the first tag bit arriving at the switch, sets the state of the switch accordingly and sends rest of the tag bits to the next switch along the path up to the building block. Starting from the left, the first  $\log(N/n)$  bits are used by the input stage switches. In this stage the signal is routed from the leaf to the root of the output tree. The last  $\log(M/m)$  bit is used by the output stage switches. Building block use the remaining middle bits for it's own routing.

#### F. Proposed Routing Algorithm of GRN

In GRN, crosstalk occurs only at building block. To remove this crosstalk, we propose a new crosstalk free routing algorithm of GRN. The proposed routing algorithm is given below:

**Routing Algorithm** (integer  $N$ , SEMI-PERM semi-perm)

**Step 1:** If  $N$  is 2, compute only the routing mechanism of GRN and exit, no semi-permutation is needed.

**Step 2:** Construct a bipartite graph  $G = (V_1, V_2, E)$  from a given permutation, when  $N > 2$ , where  $N$  is the number of inputs/outputs.

**Step 3:** Decompose the permutation according to the decomposition algorithm of bipartite graph.

**Step 4:** Obtain forward and backward edges from the bipartite graph.

**Step 5:** Then apply, the routing mechanism of GRN upon the obtaining semi-permutations for a particular permutation.

**Step 6:** Recursively call Routing Algorithm (integer  $N$ , SEMI-PERM semi-perm) until reach  $\#Semi-permutation = 2 \times \log_2^{[Building-block]}$

**End**

#### IV. REALIZATION OF A ROUTING

In this section, we have shown the simulation result using an example of routing proposed in this paper. The routing paths for various sets of inputs-outputs are clearly shown in Figure 4 with different colors. The messages to be sent are directed from all input ports to its corresponding output ports.

Let us consider a input and output sequence, where the output sequence is generated randomly. An example is shown below:

$$\begin{pmatrix} \text{Input Sequence} = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \text{Output Sequence} = & 0 & 6 & 1 & 2 & 3 & 4 & 7 & 5 \end{pmatrix},$$

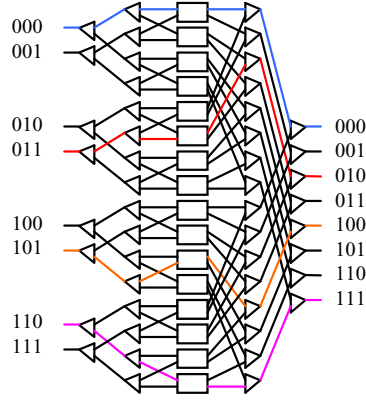
Then we decompose, these sequences according to proposed routing algorithm of Generalized Recursive Non-blocking Network (GRN) and we get "forward" and "backward" pairs as below.

$$\text{Forward} = \begin{pmatrix} 0 & 3 & 5 & 6 \\ 0 & 2 & 4 & 7 \end{pmatrix}$$

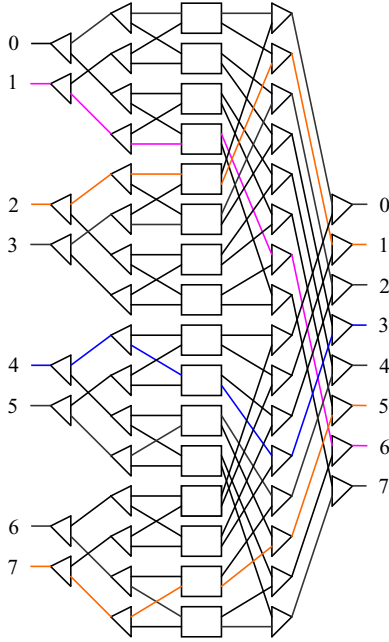
$$\text{Backward} = \begin{pmatrix} 1 & 2 & 4 & 7 \\ 6 & 1 & 3 & 5 \end{pmatrix}$$

In Figure 4, only one signal comes in a building block for a particular time for both forward pair and backward pair. Therefore, there is no possibility to occur crosstalk at its building block. The input and output trees are cross-talk free [1]. Therefore, the whole network is crosstalk free in two passes.

Now, we consider  $8 \times 8$  GRN network with  $(4 \times 4)$  building block, for the same input and output sequences, as mentioned earlier. In Figure 5, at the same time two signals pass in a building block switch. Therefore, it is possible to occur crosstalk in the building block switch. To overcome this problem, we have decomposes this semi-permutation again into forward pair and backward pair.



(a)



(b)

Fig. 4. Routing of various input-output combination with (a) forward pair (b) backward pair

$$\text{Forward} = \begin{pmatrix} 0 & 5 \\ 0 & 4 \end{pmatrix} \quad (3)$$

$$\text{Backward} = \begin{pmatrix} 3 & 6 \\ 2 & 7 \end{pmatrix} \quad (4)$$

The realization of this decomposed permutation is shown in Figure 6. Figure 6(a) is for forward pair shown in Eq. 3 and Fig. 6(b) for backward pair shown in Eq. 4. It is shown that the only one signal passes through the building block. Therefore the network is crosstalk free.

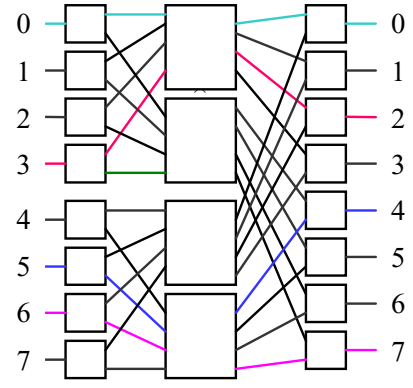


Fig. 5. Routing in  $8 \times 8$  GRN with  $4 \times 4$  building block for forward pair

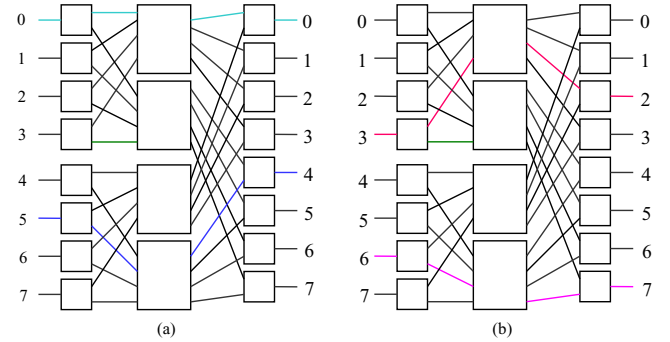


Fig. 6. Routing in  $8 \times 8$  GRN with  $4 \times 4$  after second decomposition. (a) forward pair (b) backward pair

Figure 7 shows the realization of the backward decomposition. Here, it is shown that only one signal is pass through the building block. Thus the building block, and hence the whole network crosstalk-free.

But if we replace the building block by the  $4 \times 4$  wide-sense non-blocking Benes network, then only one signal comes in a building block switch at a time, and there is no possibility to occur crosstalk in the building block switch. Again for backward pair only one signal comes in a building block switch at a time, and it is shown above.

We have shown an example of crosstalk-free routing. In our study study all the input-output permutations, where the output is generated randomly, the GRN is deadlock free with the number of Semi-permutation =  $2 \times \log_2^{[Building-block]}$ .

## V. CONCLUSION

We have proposed a new crosstalk-free routing algorithm of GRN. Each switch can only pass one signal at a given time. Therefore, there is no possibility to occur crosstalk in a building block switch. This algorithm is derived based on the idea of the semi-permutation, and the main advantage of our proposed routing algorithm is

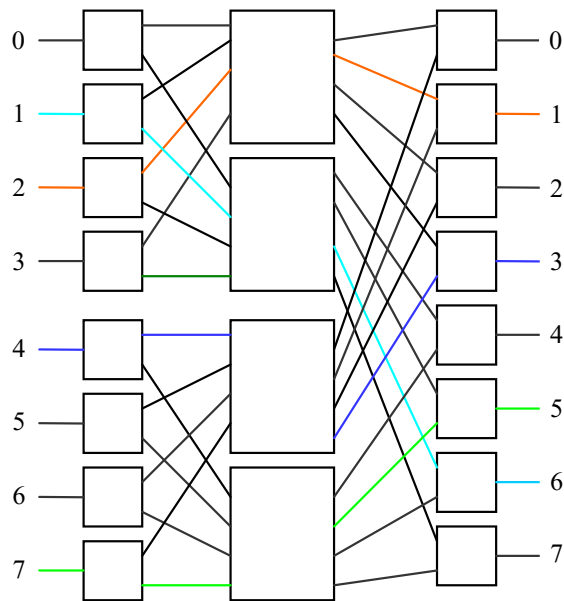


Fig. 7. Routing in  $8 \times 8$  GRN with  $4 \times 4$  building block for backward pair

that it reduces the number of passes of a permutation under the constraint of avoiding crosstalk.

#### REFERENCES

- [1] M.R. Khandker, X. Jiang, S. Horiguchi, and H. Shen, "Generalized Recursive Network: A New Architecture for self routing Non-blocking Optical Switching Networks," *Optics Communication*, Vol.208, No.4-6, pp.299-308, 2002.
- [2] Yuanyuan Yang, Jianchao Wang, and Yi Pan, "Permutation Capability of Optical Multistage Interconnection Networks," *Journal of Parallel and Distributed Computing*, Vol.60, No.1, pp.71-91, 2000.
- [3] X. Jiang, H. Shen, M. R. Khandker and S. Horiguchi, "Crosstalk-free Permutation in Photonic Rearrangeable Networks Built on a Combination of Horizontal Expansion and Vertical Stacking of Banyan Networks," *IEICE Trans. on Information and Systems*, Vol.E86-D, No.9, pp.1525-1533, 2003.
- [4] M. R. Khandker, X. Jiang and S. Horiguchi, "New Nonblocking Optical Switch Network," *Proc. of the First Workshop of International Joint Research: parallel computing, data mining and optical networks*, Tatsunokuchi, Japan, pp. 87-97, 2001.

# Lower Bound on Blocking Probability for Vertically Stacked Optical Banyan Networks with Given Crosstalk Constraint

Basra Sultana and M. R. Khandker

Department of Applied Physics and Electronic Engineering,  
University of Rajshahi, Rajshahi – 6205, Bangladesh.  
E-mail: basra\_apee@walla.com, khandker@ru.ac.bd

**Abstract** — Vertical stacking of multiple copies of an optical banyan network is a novel scheme for building nonblocking optical switching networks. The resulting network, namely vertically stacked optical banyan (VSOB) network, preserves all the properties of the banyan network, but increases the hardware cost significantly under first order crosstalk-free constraint. Therefore, blocking behavior analysis could be an effective approach to studying network performance, and finding a graceful compromise between hardware costs and blocking probability with different degree of crosstalk constraint. However, lower bound on number of planes for such networks with zero crosstalk constraint has been determined previously. In this paper, we present the simulation results for lower bound on blocking probability with given degree of crosstalk. The results show that blocking probability decreases to low value if we allow some CSEs to the network. The simulation results presented in this paper can guide network designer in finding a tradeoff among the blocking probability, and the degree of crosstalk of VSOB networks.

**Index Terms** — Lower bound, blocking probability, crosstalk.

## I. INTRODUCTION

Optical mesh networks are considered more capacity-efficient and survivable for serving as the backbones for next generation internet. A key network element equipped with a switching node of optical mesh networks is the optical switch, which has the capability of switching huge data at an ultra-high speed. The main factors those have to be considered while designing any optical switching networks are hardware cost, blocking probability, crosstalk, switching speed etc. The basic  $2 \times 2$  switching element (SE) in a large optical switching network is usually a directional coupler (DC) [1,7]. DC's can switch multiple wavelengths at the same time, which is important for the future optical cross-connects (OXC's).

Crosstalk is an intrinsic shortcoming of the DC. It is the effect of the undesirable coupling between the signals carried in the two waveguides of the coupler [1,2]. When two optical signals meet at a DC, a small portion of the signal power will be directed to the unintended output channel. Crosstalk suppression becomes particularly important in networks, where a signal propagates through many nodes and accumulates crosstalk from different elements at each node from the system view. In order to obtain an approximate idea of the crosstalk

requirements, suppose that a signal accumulates crosstalk from  $N$  sources, each with crosstalk level  $\alpha$ . This neglects the fact that some interfering channels may have higher powers than the desired channel. Networks are very likely to contain amplifiers and to be limited by signal-spontaneous beat noise. For example, if we have 10 interfering equal-power crosstalk elements, each producing intrachannel crosstalk, then we must have a crosstalk suppression of below 35dB in each element, in order to have an overall penalty of less than 1dB [7]. Thus, Crosstalk reduction is an important issue in designing the systems that are based on DC's. The crosstalk issue can be tackled at either the device level or the system level. The two methods complement each other. The focus of this paper is on the system-level approach. As will be seen, crosstalk adds a new dimension to the theory of building a nonblocking and negligible blocking switching network.

Banyan networks [8,4,3,5] are a class of attractive switching structures for constructing DC-based optical switches, because they have a smaller and exactly the same number of SEs along any path between an input-output pair such that an absolutely loss uniformity and smaller attenuation of optical signals are guaranteed in this class of switching networks. However, with the banyan topology only a unique path can be found from each network input to each network output, in which the network is degraded as a blocking one. The general scheme for building banyan-based nonblocking optical switching networks is to vertically stack the multiple copies of regular optical banyan network [6,10] as illustrated in Fig.1. We use VSOB to denote the optical switching networks built on vertical stacking scheme of banyan network.

Numerous result are available for VSOB networks, such as [6], and their main focus has been on determining the maximum and minimum number of stacked copies (planes) required for a nonblocking VSOB networks if conservative routing algorithm is used for routing a request to a plane. These results indicate that VSOB structure, although is attractive, usually requires either a high hardware cost or a relatively larger depth for building a nonblocking network. In paper [11], the maximum number of planes required for nonblocking a VSOB

networks with link-failures and the blocking probability of VSOB networks for given failure probability & load have been determined. In paper [12], the lower bound on number of planes for VSOB networks with link-failures have been determined. In this paper, we analyze the blocking behaviour of VSOB networks with allow some CSEs to the networks. The simulation results can guide networks designers to initiate a compromise among the hardware cost, and the blocking probability of a VSOB network under different degree of crosstalk.

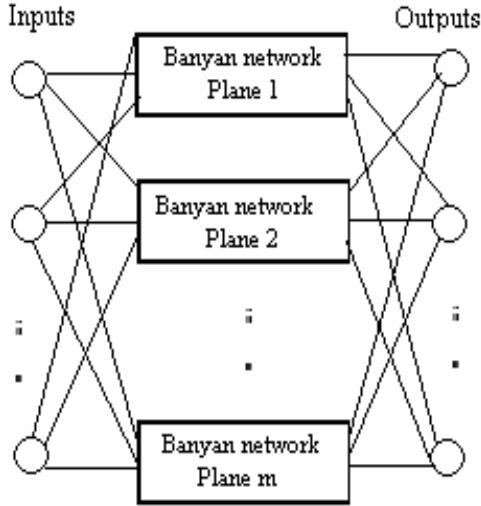


Fig. 1. Vertically stacked optical banyan (VSOB) network.

The rest of the paper is organized as follows: Section 2 provides preliminaries that facilitate our further discussions. Section 3 presents our contribution. Section 4 concludes the paper.

## II. PRELIMINARIES

In VSOB network, blocking happens when two connections intend to use the same link, which is referred to as link-blocking. However, there is another type of connection-related blocking, which occurs when some paths (including the new one) violate the crosstalk-free constraint after adding the new connection. In such a situation, the connection is not allowed to be allocated even if the path is available. We refer to this second type of blocking as crosstalk-blocking throughout the paper. It is the combination of the two types of blocking that makes the design principle different.

For A typical  $N \times N$  banyan network consists of  $\log_2 N$  stages, each containing  $N/2$   $2 \times 2$  SEs. Regular banyan network has a unique path between an input-output pair, and a basic technique for creating multiple paths between an input-output pair is the vertical stacking of multiple copies of the banyan network. In this paper, we focus on the banyan network that has  $m$  multiple copies as shown in fig.2.

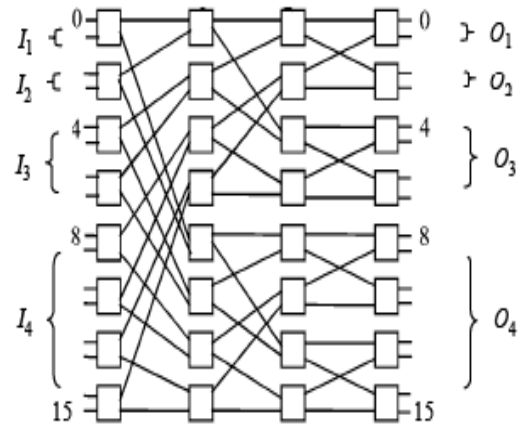


Fig. 2. 16x16 banyan network.

Due to the topological symmetry, all paths in banyan networks have the same property in terms of blocking. We define the blocking probability to be the probability that a feasible connection request is blocked, where a feasible connection request is a connection request between an idle input port and an idle output port of the network. Without loss of generality, we focus on the path between the first input and the first output (which is termed as the tagged path hereafter). All the SEs and links on the tagged path are called tagged SEs and tagged links respectively. The stages of SEs are numbered from left (stage 1) to right (stage  $\log N$ ) and the stages of links are also numbered from left (stage 1) to right (stage  $\log N + 1$ ). For the tagged path, an input intersecting set  $I_i = \{2^{-i}, 2^{-i+1}, \dots, 2^{-i-1}\}$  at stage  $i$  is defined as the set of all inputs that intersect a tagged SE at stage  $i$ . Likewise, an output intersecting set  $O_i = \{2^{-i}, 2^{-i+1}, \dots, 2^{-i-1}\}$  associated with stage  $i$  contains all the outputs that intersect a tagged SE at stage  $\log N - i + 1$ .

When two light signals go through an SE simultaneously, crosstalk is generated at the SE. Such SE is referred to as a crosstalk SE (CSE). The degree of crosstalk of the switching system is defined as the number of CSE's allowed along a path.

A restricted SE (RSE) is a  $2 \times 2$  SE which carries only one light signal at a time. Although crosstalk at an RSE is very small, it may not be entirely zero. For example, when a light signal passes through an RSE, a small portion of the signal will leave at the other unintended output channel. This stray signal can arrive at the input of the next stage SE and generate some crosstalk. Since crosstalk generated by the stray signal is much smaller than the regular crosstalk, we will ignore it in our analysis.

Following the typical assumption as in [6,9] on probabilistic analysis of multistage interconnection networks, we neglect the correlation among signals arriving at input and outputs ports, and consider that the statuses (busy or idle) of individual input and output ports in the network are independent. This assumption is justified by the fact that the correlation among signals at inputs and outputs, though exists for fixed communication



patterns, and becomes negligible for arbitrary communication patterns in large size networks, which is the trend of future optical switching networks that can switch huge data at high speeds.

#### A. Lower Bound on Number of Planes

The lower bound on the number of planes required to make VSOB networks nonblocking is shown in Fig.3.

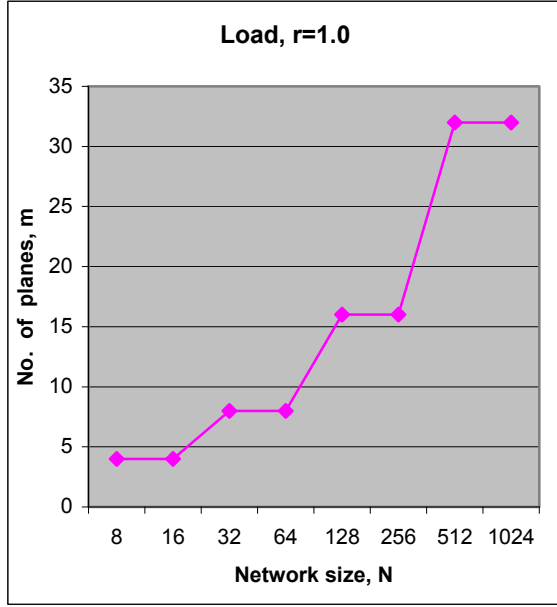


Fig. 3. Nonblocking condition for VSOB network.

### III. VSOB NETWORKS WITH ALLOW CROSSTALK

In this section, we analyzed the blocking behavior of VSOB networks with given crosstalk constraint.

#### A. Network Simulator

The network simulator we developed consists of four major modules as shown in fig. 4.

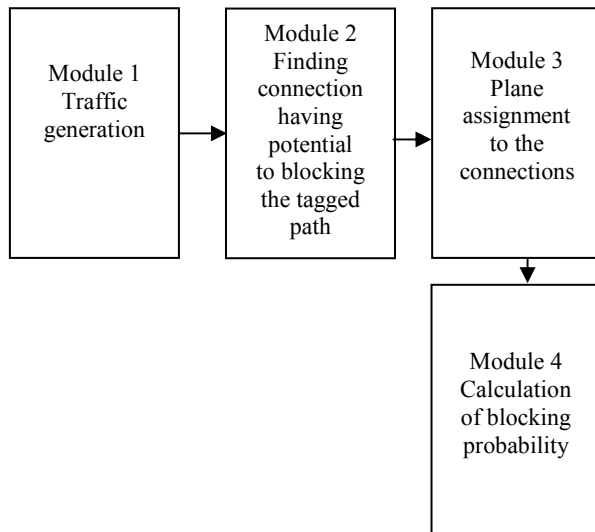


Fig. 4. Block diagram of network simulator.

We consider here the permutation request as the traffic. Due to the symmetric architecture of VSOB  $(N, T)$  network, every connection request has the same probability to be blocked. In our simulation, we fix the connection request of input-output pair 0-0 and investigate the blocking probability of this connection request only that may result by other contentious connections.

**Module 1:** This module randomly generates a permutation request for the VSOB  $(N, T)$  network based on the workload  $r$  (here workload  $r$  is defined as the occupancy probability of a port).

In this section we present some definitions, which are used in the discussion of our algorithm:

**Definition 1 Blocking connections,  $C_{bc}$ :** The connections those are potential to blocking the tagged path.

**Definition 2 Crosstalk-blocking connections,  $C_{cbc}$ :** The connections those are potential to crosstalk-blocking the tagged path.

**Definition 3 Maximum allowed crosstalk,  $C_m$ :**  $C_m$  is defined the maximum allowed crosstalk along the tagged path.

**Definition 4 Allowed crosstalk,  $c$ :**  $c$  is defined the allowed crosstalk along the tagged path at any time.

**Module 2:** This module finds  $C_{bc}$  which is determined by the following relation:

$$I_i + O_j < \log_2 N + 2 \quad (1)$$

Here, input  $i$  has originated from input intersecting group  $I_i$  and destined to output  $j$  that belongs to output intersecting group  $O_j$  (see Fig.2).

We find  $C_{cbc}$  that must satisfy the following relation

$$I_i + O_j = \log_2 N + 1 \quad (2)$$

**Module 3:** This module assigns connection requests to different planes in the following way. The tagged path is assigned first to a plane randomly chosen. Then a connection from the list of  $C_{cbc}$  is assigned to the plane in which tagged path assign and allow one  $c$ . If  $c$  less than  $C_m$  then the next connection from the list of  $C_{cbc}$  is picked and assigned to the plane as mentioned above. When  $C_m$  reached then the other connections from the list of  $C_{bc}$  are assigned to the other free planes. The connections share the planes without conflicting each others. This plane assignment algorithm ensures the use of minimum number of planes for routing a permutation request. At last if there is a connection for which no such plane is available then the connection request pattern is recorded as a blocked connection pattern.

**Module 4:** In this module the blocking probability is estimated by the ratio of number of connection requests in which the 0-0 request is blocked to the total number of connection requests generated.

#### A.1. Lower Bound on Number of Planes Given Crosstalk Constraint

The simulation result for lower bound on blocking probability of VSOB networks under crosstalk constraint is given below:

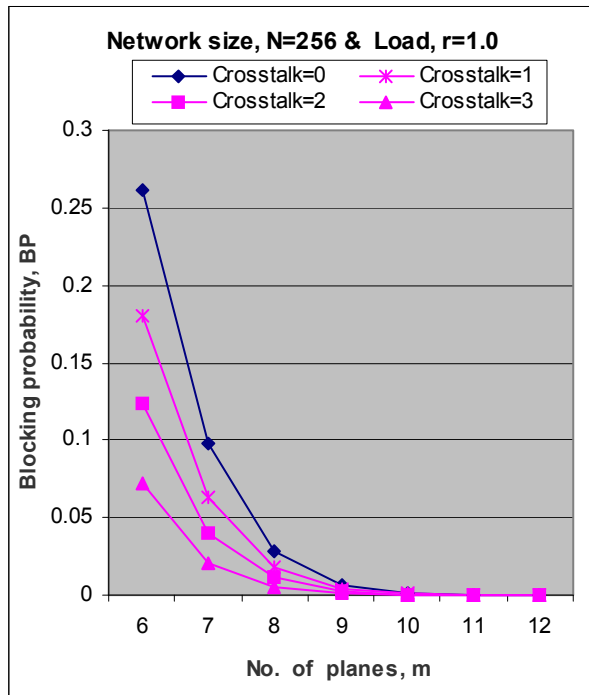


Fig 5. Number of planes vs blocking probability for different crosstalk of VSOB (256,m) networks.

It is shown from Fig.5 that for  $N=256$ ,  $r=1.0$ ; if we allow 2 CSEs to the network then blocking probability decreases 52% for  $m=6$  and 60% for  $m=8$ . It is also shown from Fig.5 that for  $N=256$ ,  $r=1$ ; if we allow 3 CSEs to the network then BP decreases by 72% than  $c=0$  and if we increase the number of planes from 6 to 7 then BP decreases by 62%. This picture derives the fact that if we allow small amount of crosstalk then the blocking probability decreases rate is high compare the number of planes.

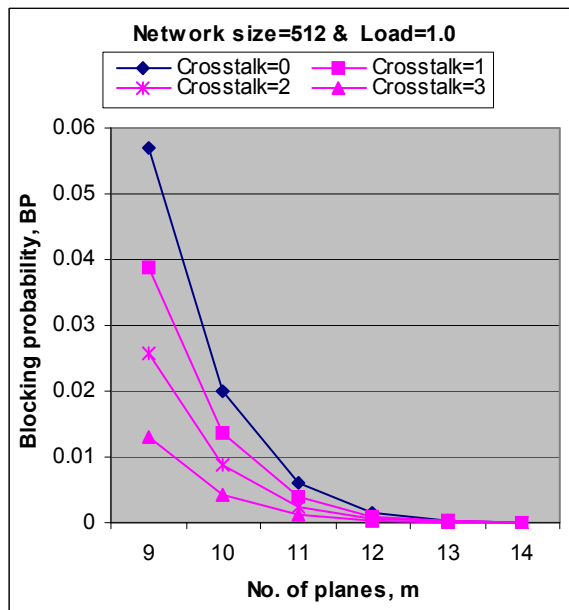


Fig. 6. Number of planes vs blocking probability for different crosstalk of VSOB (512,m) networks.

It is shown from Fig.6 that for  $N=512$ ,  $r=1.0$ ; if we allow 2 CSEs to the network then blocking probability decreases 54.6% for  $m=9$  and 56% for  $m=10$ .

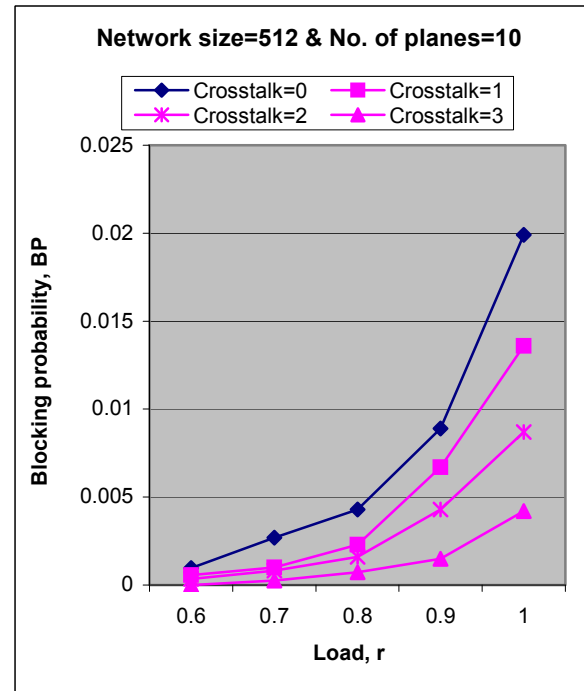


Fig.7. load vs blocking probability for different Crosstalk.

The effect on blocking probability with load for various crosstalks for a fixed network size and number of planes is shown in Fig.7. If we decreases load from 1.0 to 0.8 then BP decreases 78.4% for  $c=0$  and 83% for  $c=1$ .

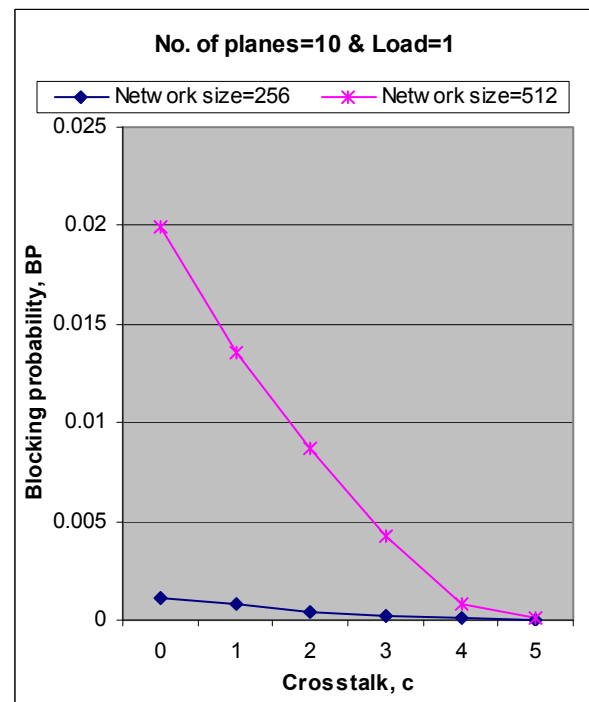


Fig. 8. Crosstalk vs blocking probability for different network size.

It is observed from fig.8 that for  $m=10$  &  $\text{load}=1$ ; if we allow 2 CSEs to the network then blocking probability decreases to 59% for  $N=256$  and 56% for  $N=512$ .

Fig.5 & 6 shows that the trend of variation of the blocking probability remains the same for different network size.

The simulation results show that blocking probability decreases for allowing CSEs because if we allow some CSEs to the network then the network SEs can be shared by more connections.

#### IV. CONCLUSION

We have presented approximate results of lower bound on blocking probability for VSOB networks with given certain degree of crosstalk constraint. The simulated results in section 3 can provide network developers with a guidance of quantitatively determining the effects of various degrees of crosstalk and reduction in number of planes on the overall blocking behaviors of VSOB networks. The results show that if we allow some CSEs which are negligible for higher network size then blocking probability decreases to approximate zero. The results can also show how crosstalk adds a new dimension to the performance analysis of VSOB networks and effects of crosstalk on hardware cost and blocking probability. A mathematical model for blocking probability of this network is under research.

#### References

1. H. S. Hinton, An introduction to photonic switching fabrics, New York: Plenum, 1993.
2. R. Chinni et al., Crosstalk in a lossy directional coupler switch, IEEE/OSA Journal of Lightwave Technology, vol. 13, no. 7, pp. 1530-1535, July 1995.
3. G. R. Goke and G. J. Lipovski, Banyan networks for partitioning multiprocessor systems, Proc. 1<sup>st</sup> Annu. Symp. Comp. Arch, pp.21-28, 1973.
4. F. T. Leighton, Introduction to parallel algorithms and architectures: Arrays, Trees, Hypercubes, Morgan Kaufmann, 1992.
5. J. H. Patel, Performance of processor-memory interconnections for multiprocessors, IEEE Trans. Comput., vol. C-30, pp. 771-780, Oct. 1981.
6. X. Jiang, H. Shen, Md. Mamun-ur-R. Khandker, S. Horiguchi, Blocking Behavior of Crosstalk free Optical Banyan Networks on Vertical Stacking, IEEE/ACM Transactions on Networking, Vol. 11, No. 6, Dec. 2003.
7. Rajib Ramaswami and Kumer N. Sivarajan, Optical networks, Morgan Kaufmann Publishers, ISBN:1-55860-655-6.
8. C.Yu, X. Jiang and S. Horiguchi, Performance Modeling for Vertically Stacked Optical Banyan Networks with Extra Stage, International Journal of Computational Science and Engineering (IJCE), Vol.2, No.1-2, pp.81-87, May 2006.
9. Xiaohong Jiang, H. Shan and S. Horiguchi, Blocking Probability of Vertically Stacked Optical Banyan Networks Under Random Routing, Proc. Of GLOBECOM 2003, Dec.1-5, San Francisco, USA.
10. Xiaohong Jiang, H. Shan, S. Horiguchi and P.H.Ho, Performance Modeling for All-Optical Photonic Switches Based on the Vertical Stacking of Banyan Network Structures, IEEE Journal of Selected Areas on Communications, vol.23/8, pp.1620-1631, Aug.2005.
11. Basra Sultana, M.R.Khandker, X.Jiang and S.Horiguchi: "Blocking Probability of Vertically Stacked Optical Banyan Networks with Link Failures", Proc. The International Workshop on High Performance and Highly Survivable Routers and Networks, Tohoku University, Sendai, Japan, pp. 157-169, March 14, 2007.
12. Basra Sultana, M.R.Khandker: "Lower Bound on Number of Planes for Vertically Stacked Optical Banyan Networks with Link Failures", Proc. 10<sup>th</sup> International Conference on Computer and Information Technology, pp. 178-183, 27-29 December, 2007, Dhaka, Bangladesh.

# Performance of Fast Routing Algorithm for Vertically Stacked Optical Banyan Networks with Link Failures and Given Crosstalk Constraint

Basra Sultana<sup>1</sup> and M. R. Khandker<sup>1</sup>

<sup>1</sup>Department of Applied Physics and Electronic Engineering,  
University of Rajshahi, Rajshahi – 6205, Bangladesh.  
E-mail: basra\_apee@walla.com, khandker@ru.ac.

**Abstract** — For faster connection establishment, previously proposed Pruned Vertically Stacked Optical Banyan (P-VSOB) networks used plane fixed routing (PFR) algorithm, and has  $O(\log_2 N)$  routing complexity. Blocking probability has also been analyzed for these kinds of networks with given amount of crosstalk. In EP-VSOB networks, a few regular banyan planes are added with the P-VSOB networks. Necessary routing algorithm, namely, PFR\_LS and PFR\_RS show that this switching network can reduce the blocking probability to very low value while keeping the hardware cost almost the same as that of P-VSOB networks. Both these algorithm also have the time complexity  $O(\log_2 N)$ . This paper deals with the blocking behavior of EP-VSOB (Extended pruned VSOB) networks having link-failures and allows crosstalk constraint. Our simulation results show that crosstalk adds a new dimension to the performance analysis of practical EP-VSOB networks where some links are failed. The results show that if we allow some CSEs to the network then blocking probability decreases to low value for both PFR\_LS and PFR\_RS algorithm. The results also show that the blocking probability of the EP-VSOB networks does not always increase with the increase of link-failures; blocking probability decreases for certain range of link-failures, and then increases again.

**Index Terms** — EP-VSOB network, crosstalk, link-failure.

## I. INTRODUCTION

Photonic switching networks play an important role within wavelength-division multiplexing (WDM) networks, where optical cross-connects (OXC) are greatly preferable to electronic switches. The main factors those have to be considered while designing any photonic switching networks are hardware cost, blocking probability, crosstalk, switching speed etc. Researchers have proposed many switching architectures having different merits in the literature. A few important works have been listed in the references.

Directional-coupler (DC) [1,2] can handle the optical signals of some terabits per second and with multiple wavelengths, and this make it ideal for serving as the basic  $2 \times 2$  switching element (SE) in high-speed optical switches. However, one of the major shortcomings of a DC is crosstalk. When two light signals go through an SE

simultaneously, crosstalk is generated at the SE. Such SE is referred to as a crosstalk SE (CSE). The degree of crosstalk of the switching system is defined as the number of CSE's allowed along a path.

Banyan networks [3,4,5,6] are attractive for constructing DC-based optical switches for their small depth, absolute loss uniformity, lower hardware cost and a simple switch setting ability (self-routing). With banyan structure, only a unique path exists between input-output pair in which the network is simply degraded as a blocking one. The blocking probability of banyan network is high. The whole network can be made nonblocking by vertically stacking multiple copies of an optical banyan network [7], known as vertically stacked optical banyan (VSOB) network. The VSOB network increases the hardware cost but decreases the blocking probability, and has  $O(N \log_2 N)$  time complexity. Due to high hardware cost and routing complexity of VSOB network, researchers proposed Pruned VSOB (P-VSOB) networks [8]. In P-VSOB network, PFR routing algorithm is used in which a plane of the VSOB network is fixed for a group of inputs and achieved  $O(\log_2 N)$  time complexity. Although P-VSOB networks guarantee considerably low blocking probability, but for some high performance applications, this blocking probability may be considered too high. In paper [12], the blocking probability of P-VSOB networks with given crosstalk have been determined. Khandker *et. al.* have proposed Extended-Pruned-VSOB (EP-VSOB) network [9] that has the potential to reducing the blocking probability while keeping the crosstalk constraint to zero. Necessary routing algorithm, namely PFR\_LS and PFR\_RS modifying the PFR algorithms have also been proposed. The EP-VSOB network has almost the same routing complexity as that of the P-VSOB network with significantly lower blocking probability. This switch network achieves the speed close to P-VSOB networks and blocking probability close to VSOB networks.

Due to the increasing importance and requirement for fault-tolerance in optical switches for large mesh WDM networks, performance analysis of VSOB networks at the

presence of probability of link failures becomes critical for the practical adoption of the VSOB networks in the current internet applications. Paper [10] deals the blocking behavior analysis of VSOB networks with link-failures and the author showed that VSOB network is a fault tolerant network. In paper [11], the blocking probabilities of P-VSOB networks having link failures have been determined. P-VSOB networks can accommodate some link failures safely without degrading the performance. In this paper, we analyze the blocking behavior of EP-VSOB networks with given some crosstalk constraint when some links are failed or broken in the network.

The rest of the paper is organized as follows. We discuss the published work briefly in section 2; especially EP-VSOB networks. Section 3 presents our contribution. We conclude this paper in section 4.

## II. PRELIMINARIES

A connection request may be blocked by link failures in a faulty VSOB network, which is referred to as the *failure-blocking*. We assume that the links in VSOB networks may fail independently and these failures are permanent. Thus, both crosstalk-blocking and failure-blocking should be fully considered in the blocking analysis of a faulty VSOB networks as illustrated in Fig.1 for a  $8 \times 8$  network.

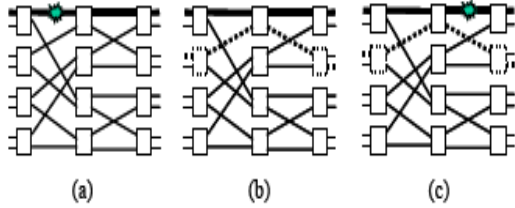


Fig. 1. Blocking in a VSOB network. (a) Failure-blocking. (b) Crosstalk- blocking. (c) Combination of failure-blocking and crosstalk-blocking.

Due to their topological symmetry, all paths in a banyan network have the same property in terms of blocking. Through out this paper, we will select the path between the first input and the first output and try to set up a connection between them. We call the path between this input-output pair the tagged path. In this paper, we consider the banyan networks that support one-to-one communication. We define the blocking probability to be the probability that a feasible connection request is blocked, where a feasible connection request is a connection request between an idle input port and an idle output port of the network

### A. Pruned-VSOB network

P-VSOB  $(N,T)$  denotes an  $N \times N$  Pruned-VSOB network that has  $T$  fixed planes. In P-VSOB network, plane fixed

routing (PFR) algorithm is used for routing input signals to requested outputs. In PFR algorithm, we select one input from each input group  $I_i$  and tie them with a plane of VSOB  $(N, T)$  network permanently. Input groups  $I_i$  is defined as the set of inputs  $\{i, i+1, i+2, \dots, i+T-1\}$  as shown in Fig.2. We first define the following subsets of input taking on from each input sets:

$$G_i = \left\{ i, i+T, i+2T, \dots, i + \left( \frac{N}{T} - 1 \right) T \right\}; 0 \leq i \leq T-1$$

$$T = 2^{\lceil (\log_2 N + 1) / 2 \rceil}$$

In PFR algorithm all  $N$  inputs are always evenly distributed among the planes and each plane has only  $N/T$  connections. Since these  $N/T$  connections are fixed at their corresponding inputs in the plane, only 1 input of a group will be active; all other inputs will remain unused. Therefore, input switches connected to unused inputs and switches in the successive stages corresponding to these unused inputs are redundant. All these redundant switching elements are eliminated. Fig.3 explains the idea.

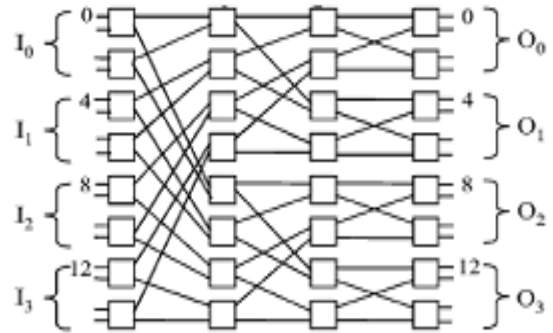


Fig. 2. Illustration of inputs and outputs grouping.

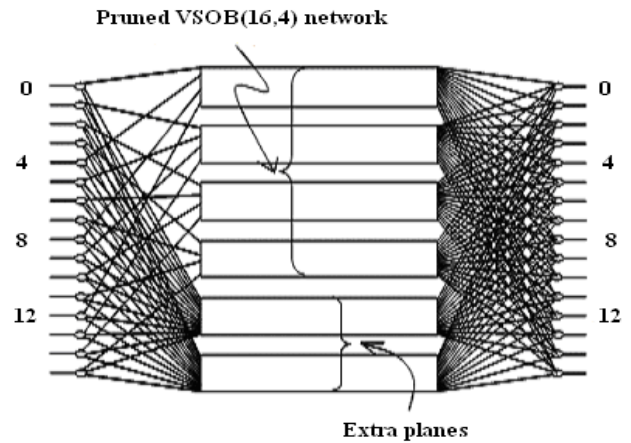


Fig. 3. The structure of a P-VSOB  $(16,4+2)$  network.

### B. Extended Pruned VSOB ( $N, T+K$ ) network

In this section, we describe previously proposed Extended-Pruned-VSOB network. EP-VSOB ( $N, T+K$ ) denotes an  $N \times N$  EP-VSOB network that has  $T$  fixed planes (pruned banyan) and  $K$  extra planes (regular banyan). The main idea of the EP-VSOB is to add a small number of regular banyan planes to the P-VSOB( $N, T$ ) network such that the connections blocked in the fixed  $T$  planes of the P-VSOB( $N, T$ ) network have more chances to be established through these extra  $K$  planes.

Depending on how a plane is chosen from the extra planes, we propose two routing algorithm for fast connection establishment in an EP-VSOB( $N, T+K$ ) network, namely PFR with linear search algorithm and PFR with random selection algorithm.

#### B.1. PFR with Linear Search (PFR\_LS)

In the PFR\_LS algorithm for an EP-VSOB( $N, T+K$ ) network, each connection now has  $(1+K)$  chances to be established through the network. Whenever a request arrives at an input, first it is sent to the fixed plane of that input. If the request cannot be established in the fixed plane, the input searches for a free plane among the extra  $K$  planes. The searching starts with the first plane and continues orderly up to the last one so that each plane is being searched at most one time by an input. If the connection request still fails to find a free plane among all  $K$  extra planes, it is considered as a blocked request. The time complexity of the PFR\_LS algorithm  $O(\log_2 N)$  when  $K$  is constant and much smaller compared to  $T$ .

#### B.2. PFR with Random Selection (PFR\_RS)

In the PFR\_RS algorithm for an EP-VSOB( $N, T+K$ ) network, each connection has two chances to be established through the network. The first chance is to establish the connection through its fixed plane. If the connection request is blocked in its fixed plane, it still has a second chance to be established through another plane selected randomly from  $K$  extra planes. If the connection request still could not be established through the randomly selected plane, it is considered as a blocked request. It is easy to see that the time complexity of the PFR\_RS algorithm remains the optimum  $O(\log_2 N)$ .

The simulation results of EP-VSOB networks without link failures are given below:

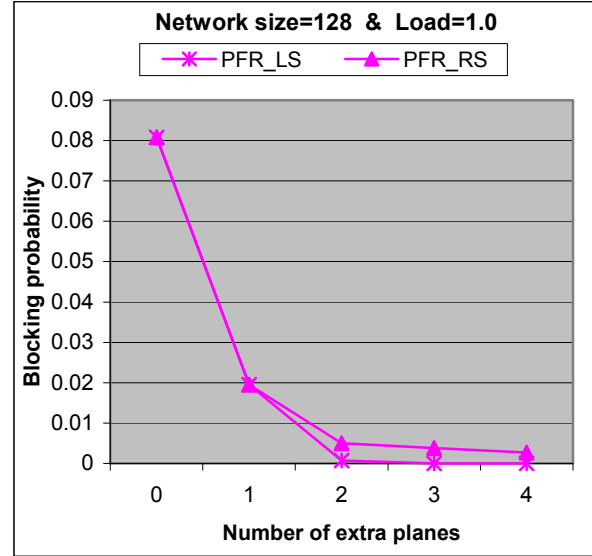


Fig. 4. Blocking probability of EP-VSOB networks without link-failures.

### III. EP-VSOB NETWORKS WITH LINK FAILURES AND GIVEN CROSSTALK

In this section, we analyzed the blocking behaviour of EP-VSOB networks having link-failures with allow some CSEs to the network using PFR\_LS and PFR\_RS algorithm.

#### A. Network simulator

The network simulator we developed consists of six major modules as shown in figure 5.

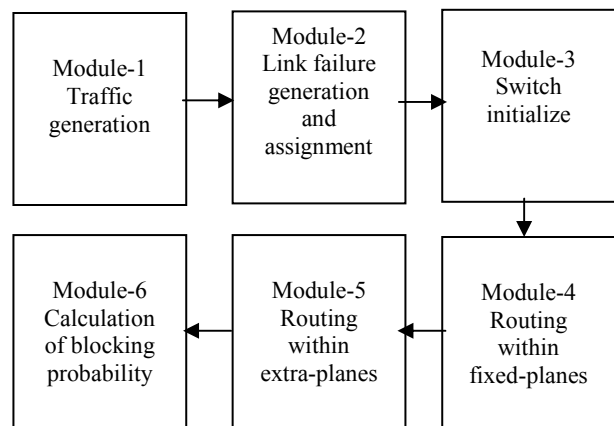


Fig. 5. Block diagram of the network simulator.



We consider here the permutation request as the traffic since a permutation does not have output contention, and therefore, gives real blocking probability of the switch network only. Due to the symmetric architecture of P-VSOB  $(N,T)$  network, every connection request has the same probability to be blocked. In our simulation, we fix the connection request of input-output pair 0-0 and investigate the blocking probability of this connection request only.

**Module 1:** This module randomly generates a permutation request for the VSOB  $(N,T)$  network based on the workload  $r$  (here workload  $r$  is defined as the occupancy probability of a port).

**Module 2:** This module generates link failures based on the given  $pfr$  (here  $pfr$  is defined as the probability that a link is failed or broken) and then assign those failures randomly to different links.

**Module 3:** This module initializes the switches with the permutation request.

**Module 4:** This module attempts to assign connection requests to different planes using PFR algorithm. We only consider the plane that contains the tagged path. We try to establish the tagged path in the selected plane with allow  $C_m$  (maximum allowed crosstalk) CSEs along the tagged path to the network. If the tagged path is not established, then we try to assign other connections in their selected planes.

**Module 5:** If the tagged path is not established in the selected plane then we try to establish the tagged path (along with other connections that are not established in their selected plane with allow  $C_m$  CSEs) in the extra planes with allow  $C_m$  CSEs along all connections path to the network. In case of PFR\_LS algorithm, the searching starts with the first plane and continues orderly up to the last one. If the tagged path still fails to find a free plane among all extra planes, it is considered as a blocked request. In case of PFR\_RS algorithm, we randomly selected a plane from extra planes. If the tagged path still could not be established through the randomly selected plane, it is considered as a blocked request.

**Module 6:** In this module the blocking probability is estimated by the ratio of number of connection requests in which the 0-0 request is blocked to the total number of connection requests generated.

#### A.1. PFR\_LS algorithm

The simulation results of EP-VSOB networks using PFR\_LS algorithm with link failures and allow CSEs are given below:

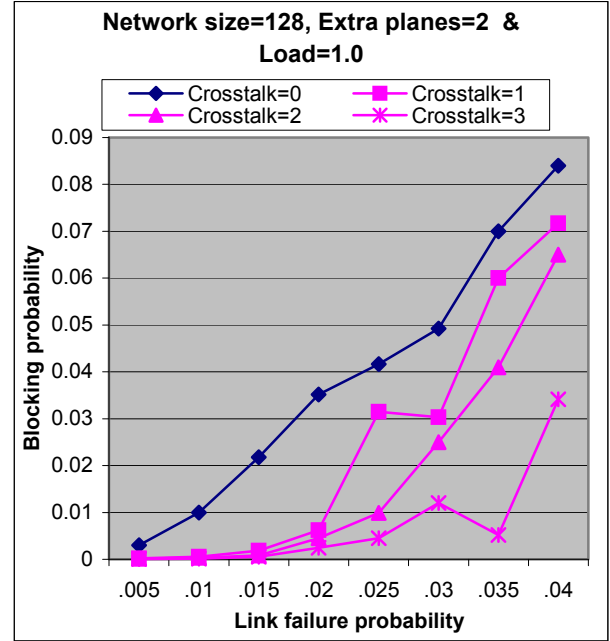


Fig. 6. Blocking probability decreases for certain range of link-failures and then increases again.

It is also interesting to note from Fig.6 that the blocking probability not always increases with the increase of link failures; blocking probability also decreases for certain range of link failures and then increases again. For  $N=128$ ,  $r=1$ , extra planes=2 &  $c=0$ ; if we introduce 0.015 link failure to the networks then BP increases 96.6%. Now if we allow 2 CSEs to the network then BP decreases 96.3%. That means the amount of performance degradation caused by link failures can be compensated by introducing CSEs. This may help switch designer choose a switch networks for different degree of performance with comparably low hardware cost.

The effect on blocking probability with load for different crosstalk is shown in Fig.7. If we reduce load 1 to 0.8 then BP decreases 31.5% for crosstalk=1, BP decreases 45.7% for crosstalk=2.

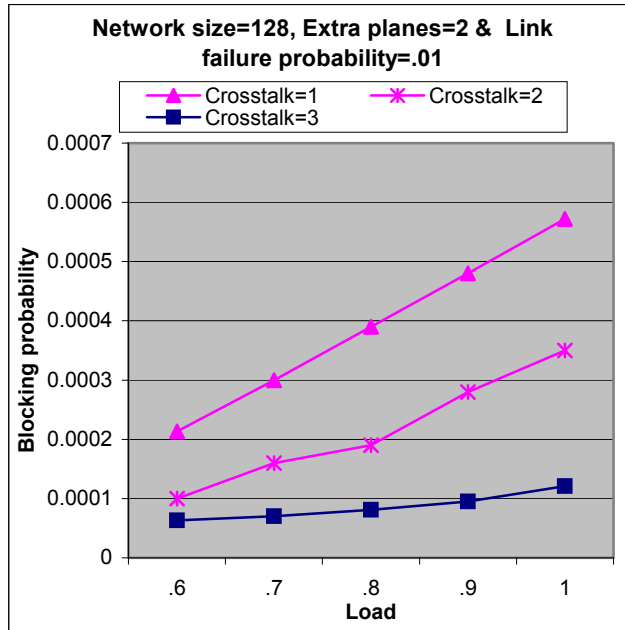


Fig. 7. Load vs blocking probability for different crosstalk.

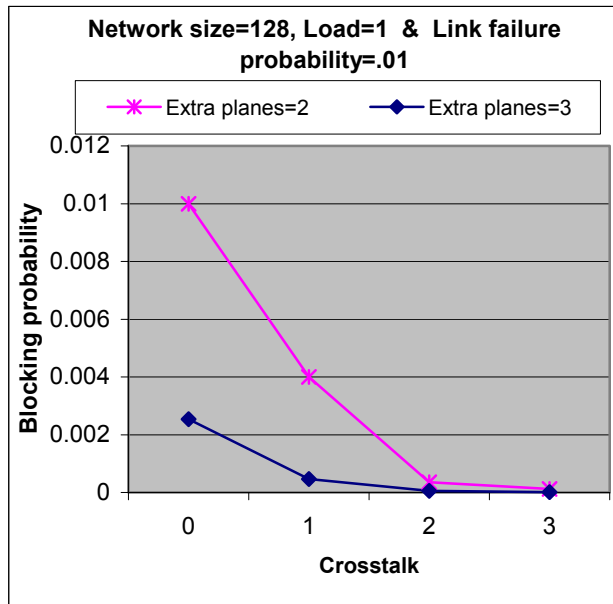


Fig. 8. Blocking probability decreases dramatically with crosstalk.

It is shown from Fig.8 that for  $N=128$ ,  $pfr=0.01$ ,  $r=1$ ; if we allow 1 CSEs to the network then blocking probability decreases to 60% for extra planes=2 and 81% for extra planes=3.

## A.2. PFR\_RS algorithm

The simulation results of EP-VSOB networks using PFR\_RS algorithm with link failures and allow CSEs are given below

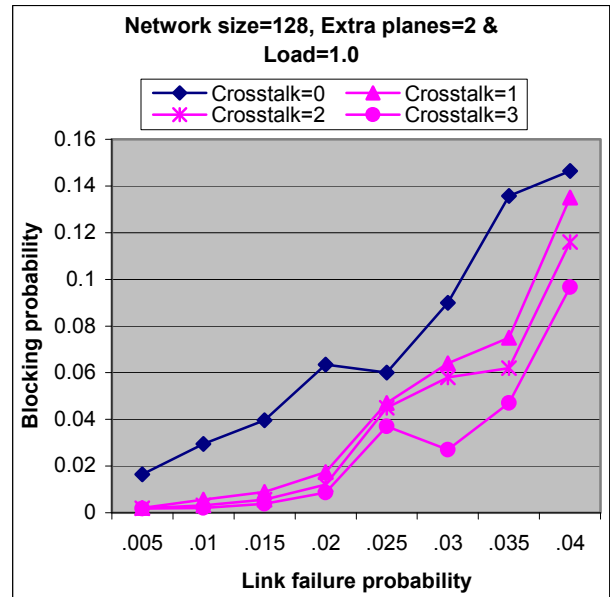


Fig. 9. Link failure vs blocking probability for different crosstalk.

It is shown from Fig.9 that for  $N=128$ ,  $pfr=0.01$ ,  $r=1$ , extra planes=2; if we allow 2 CSEs to the network then BP decreases by 89% than  $c=0$

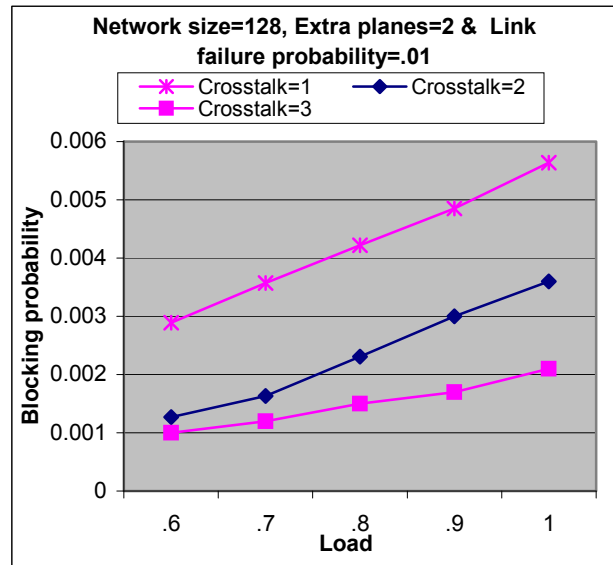


Fig. 10. Blocking probability increases linearly with load.

Blocking probability increases linearly with load as shown in fig.10. If we reduce load 1 to 0.8 then BP



decreases 25% for crosstalk =2, BP decreases 37% for crosstalk=3.

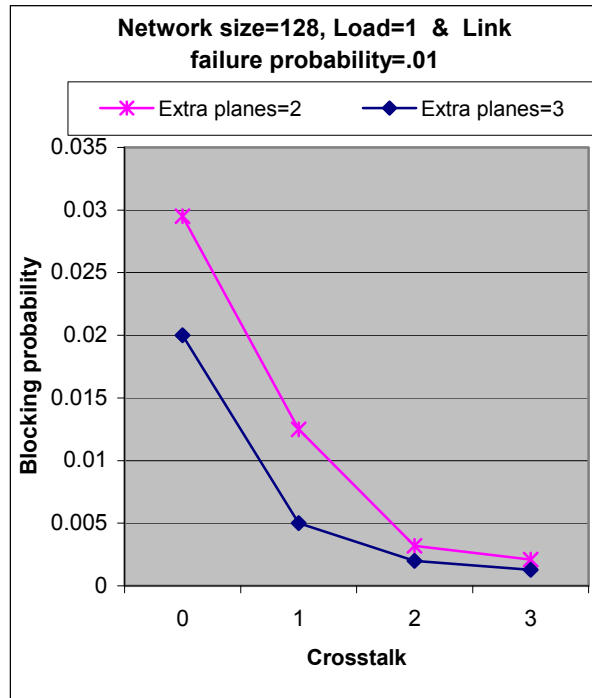


Fig. 11. Crosstalk vs blocking probability for different extra planes.

It is shown from Fig.11 that for  $N=128$ ,  $pfr=0.01$ ,  $r=1$ ; if we allow 1 CSEs to the network then blocking probability decreases to 57.6% for extra planes=2 and 75% for extra planes=3.

From fig. 6 & 9 we can say that for  $N=128$ ,  $r=1$ , extra-planes =2 &  $pfr=0.01$ ; if we allow 2 CSEs to the network then blocking probability decreases to 96.5% for PFR\_LS algorithm and 89% for PFR\_RS algorithm.

From fig. 8 & 11 we observe that for  $N=128$ ,  $r=1$ ,  $c=1$  &  $pfr=0.01$ ; if we add 3 extra-planes to the network then blocking probability decreases to 81% for PFR\_LS algorithm and 75% for PFR\_RS algorithm.

#### IV. CONCLUSION

In this paper, we have presented simulation results of the blocking probability of EP-VSOB networks with link-failures and given crosstalk constraint. To guarantee the fast connection setup in the EP-VSOB networks, we use two PFR based routing algorithm, namely PFR\_LS & PFR\_RS algorithm. The results in section 3 shows that only allow 3 CSEs along a connection can reduce the blocking probability to negligible value. The result also shows the effects of crosstalk on hardware cost and

blocking probability. The performance of PFR\_LS algorithm is better than PFR\_RS algorithm. The simulation results presented in this paper can guide network designer in finding a tradeoff among the blocking probability, the degree of crosstalk, and link failures of EP-VSOB networks. A mathematical model for blocking probability of this network is under research.

#### REFERENCES

- [1] V.R.Chinni et al., "Crosstalk in a lossy directional coupler switch," *J.Lightwave Technol.*, vol.13, no.7, pp.1530-1535, July 1995.
- [2] H.S.Hinton, *An introduction to Photonic Switching Fabrics*, New York: Plenum, 1993.
- [3] G.R.Goke and G.J.Lipovski, "Banyan networks for partitioning multiprocessor systems," *Proc.1<sup>st</sup> Annu. Symp. Comp. Arch.*, pp.21-28, 1973.
- [4] C.P.Kruskal and M.Snir, "The performance of multistage interconnection networks for multiprocessors," *IEEE Trans. Commun.*, vol.COM-32, pp.1091-1098, Dec.1983.
- [5] F. Thomson Leighton, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes*, Morgan Kaufmann, 1992.
- [6] J.H.Patel, "Performance of processor-memory interconnections for multiprocessors," *IEEE Trans. Comput.*, vol.C-30, pp.771-780, Oct.1981.
- [7] C. -T. Lea, "Muti-log<sub>2</sub>N networks and their applications in high speed electronic and photonic switching systems," *IEEE Trans. Commun.*, vol. 38, pp. 1740-1749, Oct. 1990.
- [8] Md. M. Rashid Khandker, X. Jiang, Pin-Han Ho, S. Horiguchi, H. T. Mouftah, "Performance of Fast Routing Algorithm in Large Optical Switches Built on the Vertical Stacking of Banyan Structures," *Cluster Computing: The Journal of Networks, Software Tools and Applications*, Vol 7, No. 3, pp. 219-224, 2004.
- [9] Md. Mamun-ur-Rashid Khandker, Xiaohong Jiang, Masuru Fukushi, Susumu Horiguchi, "Pruned optical banyan networks on vertical stacking scheme for faster connection establishment", Graduate School of Information Sciences, Tohoku University, Sendai, Japan, Received 9 March 2005, accepted 8 September 2005.
- [10] X. Jiang, P.H. Ho, H. Shen and S. Horiguchi: "Fault Tolerance Analysis of Optical Switching Systems Built on the Vertical Stacking of Banyan Network", *Proc. the 2004 IEEE Workshop on High Performance Switching and Routing*.
- [11] Basra Sultana, M.R.Khandker, X.Jiang and S.Horiguchi: "Blocking Probability of Vertically Stacked Optical Banyan Networks with Link Failures", *Proc. The International worksho Workshop on High Performance and Highly Survivable Routers and Networks*, Tohoku University, Sendai, Japan, pp. 157-169, March 14, 2007.
- [12] Md. M. Rashid Khandker, X. Jiang, P.H. Ho and S. Horiguchi, "Blocking Probability of Self-Routing Vertically Stacked Optical Banyan Switching Networks With Given Crosstalk Constraint", *Proc. 3<sup>rd</sup> international conference on electrical and computer Engineering*, 2004, Dhaka, Bangladesh.

# Using Log Likelihood Relation for BER Analysis of QAM in Space Diversity

Md.Zahangir Alam<sup>\*</sup>, Chowdhury Sajadul Islam<sup>\*\*</sup>, Md Zahangir Alom<sup>†</sup>, M. Abdus Sobhan<sup>\*\*</sup>

<sup>\*</sup> Dept. of Electronics and Telecommunication Engineering, Prime University, Bangladesh.

<sup>\*\*</sup> School of Engineering and Computer Science (SECS), Independent University, Bangladesh (IUB)

<sup>†</sup>Dept. of Computer Science & Engineering, University of Rajshahi, Bangladesh

Email: {ripon83\_ru, chowdhury\_iub, zahan\_cse, sobhan30}@yahoo.com.

**Abstract** — The bit error rate (BER) performance for quadrature amplitude modulation (QAM) with different diversity combining scheme is derived for Rayleigh fading channel using Log-likelihood ratio (LLR). Three combining techniques, maximal ratio combiner (MRC), selection combining (SC), and Equal Gain Combiner (EGC) are considered for the BER analysis in this paper. The average BER performance of QAM symbol through a flat fading channel is derived from the individual bits forming the QAM symbol with MRC, SC, and EGC space diversity. The analytical and simulation of BER performance of the QAM symbol through the fading channel with MMC, SC, and EGC combiner is compared. The both analytical and simulation result shows that the probability of bit error decreases with the order of diversity for all the cases of MRC, SC, and EGC. It is also seen from the results that MRC provides better BER performance than the other two combiners (SC and EGC) for the same environment of the fading channel.

**Index Terms** — Rayleigh Fading Channel, QAM, Diversity, Diversity Combiner, LLR.

## I. INTRODUCTION

The radio propagation in Doppler spread environment can be classified by three effects such as multi-path fading or fading, shadowing, and path loss. The signal transmitted from the transmitter to receiver via different path; the path may be line-of-sight (LOS) or Non-line-of-sight (NLOS). In a typical land mobile radio system, the LOS propagation path follows many NLOS reflected propagation paths. This type of propagation path is called multi-path. The channel impulse response of different paths cause individual signals of the corresponding paths arrives at the receiver with different amplitudes, phases and delays [1]. The fading causes the BER degradation of the transmitted symbol. The signal from different paths is added at the receiver. The signals are added constructively if all the incoming signals from different paths are in phase and the signals are added destructively if all the signals are out of phase. The constructive interference increases the signal strength and the destructive interference decreases the signal strength. Equalizer is used to increase the signal strength of the faded signal. The design complexity of the Equalizer is increased with the data rate [2]. The next generation wireless communication (3G and beyond) requires high data

rate, and equalizer is not suitable to minimize the effect of multi-path fading. Diversity is a technique to search for the strongest signal from the multi-path; this property of the diversity technique increases the demand of this technique for digital mobile radio services. The difficulty to send high-speed data is to achieve the limitation of channel capacity for multi-path fading effect. Two types of diversity are used in digital land mobile system, a) transmit diversity and b) receive diversity. Transmit is more efficient than receive diversity to reduce noise and multi-path fading effects [3]. In the proposed diversity technique, modulated complex symbols are transmitted as Space Time Block Code (STBC). In receive diversity one transmitter transmit symbol and two or more antennas received the transmitted symbols that arrive via different multi-path, and in transmit diversity two or more antennas are used to transmit symbol and one antenna is used to receive the multi-path signal. The BER increases when such symbols transmitted through the fading channel and for this reason the transmission channel has bandwidth limitation. The diversity branches carry the same information at the same time period but have uncorrelated multipath fading; the diversity combiner circuit combines those fading signals or selects the path having the highest signal strength. The multi-path fading appearing in the diversity branches becomes uncorrelated if the spacing between the adjacent antennas at receiver site are properly selected.

There are a number of diversity combiner circuits, usually classified into the three categories, such as, Maximal Ratio Combiner (MRC), Selection Combiner (SC), and Equal Gain Combiner (EGC). In MRC, the signal received from different branches are weighted, co-phased and then summed to obtain the highest received signal strength [4]. In SC, the signal having the highest instantaneous signal to noise (SNR) is selected from the multi-path. The EGC receives all the multi-path signals, co-phased and summed but not weighted. The MRC provides highest BER improvement of the multi-path fading channel but SC is more suitable for mobile radio applications because of its simple implementation. The quadrature amplitude modulation (QAM) is one of the effective techniques among all other modulation techniques to achieve highest spectral efficiency for the next broadband wireless communication

system through Rayleigh fading channel under the condition when the receiver is capable to estimate the channel state information (CSI) [2]. Space diversity is used to combat multi-path fading in mobile radio communications, and diversity gain is achieved at the desired level without increasing the corresponding transmission power [5]. The individual bit in QAM symbol is affected independently by the faded channel (frequency selective or frequency non-selective) and noisy channel i.e. Additive White Gaussian Noise (AWGN). The overall BER of QAM symbol is affected by the channel noise because the individual bits are affected independently. The Log-likelihood Ratio (LLR) for individual bit of QAM symbol is derived for flat fading channels [6]. In this paper, we derived the BER expression of each bit in QAM symbol based on LLR approach for a Rayleigh fading channel with MRC, SC and EGC combiner circuits. The MRC weighed each diversity branch in a co-phased manner and for this we obtain the highest SNR at the receiver site. The SNR expression for MRC is applied to find the BER expression based on LLR analysis of each bit. The average BER for QAM in Rayleigh fading channels can also be determined [4]. The SC and EGC use the average SNR for determining the average BER of QAM. A simulation work is done to find the BER performance of M-ary QAM for a Rayleigh fading channel for MRC, SC and EGC. Here, the BER expression of individual bits of QAM is used to find the average BER of the QAM symbol through a fading channel. Finally, we compare the analytical and simulated average BER performance of QAM symbol for Rayleigh fading channel for MRC, SC and EGC.

## II. PROBABILITY OF BER USING LLR SCHEME

Let us consider 16-QAM, where the 4 bits ( $r_1, r_2, r_3, r_4$ ) are mapped on to a complex symbol  $a = a_1 + ja_0$ . If the modulated symbol undergoes multi-path fading, then the received signal  $y$  can be written as

$$y = ha + n \quad (1)$$

where  $h$  is complex channel impulse response, which is assumed to be Rayleigh distributed, and  $n = n_1 + jn_0$  is a complex noise of zero mean and variance  $\sigma^2/2$ .

The LLR of bit  $r_i = 1, 2, 3, 4$  of the received symbol can be given as [6] is:

$$LLR(r_i) = \log \left( \frac{\sum_{\alpha \in S^{(1)}_i} P_r \{a = \alpha | y, h\}}{\sum_{\beta \in S^{(0)}_i} P_r \{a = \beta | y, h\}} \right) \quad (2)$$

For independent symbol using Bayes' rule, we have

$$LLR(r_i) = \log \left( \frac{\sum_{\alpha \in S^{(1)}_i} f_{y|h,a} \{y|h, a = \alpha\}}{\sum_{\beta \in S^{(0)}_i} f_{y|h,a} \{y|h, a = \beta\}} \right) \quad (3)$$

$$f_{y|h,a} \{y|h, a = \alpha\} = \frac{1}{\sigma\sqrt{\pi}} \exp \left( -\frac{1}{\sigma^2} \|y - h\alpha\|^2 \right), \text{ Now,}$$

Eq.(3) can be written as:

$$LLR(r_i) = \log \left( \frac{\sum_{\alpha \in S^{(1)}_i} \exp \left( -\frac{1}{\sigma^2} \|y - h\alpha\|^2 \right)}{\sum_{\beta \in S^{(0)}_i} \exp \left( -\frac{1}{\sigma^2} \|y - h\beta\|^2 \right)} \right) \quad (4)$$

Using the identity  $\log \left( \sum_j \exp(-X_j) \right) \approx -\min_j (X_j)$ ,

we can write Eqn. (4) as:

$$LLR(r_i) = \frac{1}{\sigma^2} \left\{ \min_{\beta \in S^{(0)}_i} \|y - h\beta\|^2 - \min_{\alpha \in S^{(1)}_i} \|y - h\alpha\|^2 \right\} \quad (5)$$

We define a variable,  $z = \frac{y}{h} = a + \frac{n}{h} = a + \hat{n}$ , where  $\hat{n}$

is Gaussian complex random variable and normalizing

$LLR(r_i)$  by  $4/\sigma^2$ , we get:

$$LLR(r_i) = \frac{\|h\|^2}{4} \left\{ \min_{\beta \in S^{(0)}_i} \|z - \beta\|^2 - \min_{\alpha \in S^{(1)}_i} \|z - \alpha\|^2 \right\}$$

$$LLR(r_i) = \frac{\|h\|^2}{4} \left\{ \min_{\beta \in S^{(0)}_i} \left\{ \|\beta\|^2 - 2z_I\beta_I - 2z_Q\beta_Q \right\} - \min_{\alpha \in S^{(1)}_i} \left\{ \|\alpha\|^2 - 2z_I\alpha_I - 2z_Q\alpha_Q \right\} \right\} \quad (6)$$

where  $z = z_I + jz_Q$ ,  $\alpha = \alpha_I + j\alpha_Q$  and  $\beta = \beta_I + j\beta_Q$ . The LLR for bit  $r_1, r_2, r_3$ , and  $r_4$  becomes-

$$LLR(r_1) = \left\{ \begin{array}{l} -\|h\|^2 z_I d, \forall |z_I| \leq 2d \\ 2\|h\|^2 d(d - z_I), \forall z_I > 2d \\ -2\|h\|^2 d(d + z_I), \forall z_I < -2d \end{array} \right\} \quad (7)$$

$$LLR(r_2) = \left\{ \begin{array}{l} -\|h\|^2 z_Q d, \forall |z_Q| \leq 2d \\ 2\|h\|^2 d(d - z_Q), \forall z_Q > 2d \\ -2\|h\|^2 d(d + z_Q), \forall z_Q < -2d \end{array} \right\} \quad (8)$$

$$LLR(r_3) = \|h\|^2 d \{ |z_I| - 2d \} \quad (9)$$

$$LLR(r_4) = \|h\|^2 d \{ |z_Q| - 2d \} \quad (10)$$

Here,  $2d$  is the minimum Euclidean distance, and can be computed according to [7]

$$d = \sqrt{\frac{3m.E_b}{2(M-1)}} \quad (11)$$

where  $m=\log_2 M$ ,  $M$  represents the number of bits per symbol and  $E_b$  represents the average energy per bit.

The probability of error for bit  $r_1$ ,  $P_{b1}$ , is according to [3]:

$$P_{b1} = P_{b1|a_I=-d} P_r\{a_I = -d\} + P_{b1|a_I=-3d} P_r\{a_I = -3d\} + P_{b1|a_I=d} P_r\{a_I = d\} + P_{b1|a_I=3d} P_r\{a_I = 3d\} \quad (12)$$

$$\text{where, } P_{b1|a_I=-d,h} = Q\left(\sqrt{\frac{4E_b\|h\|^2}{5N_0}}\right)$$

It can be written as

$$P_{b1|a_I=-d} = Q\left(\sqrt{\frac{4E_b\|h\|^2}{5N_0}}\right) = \frac{1}{\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right), \quad (13)$$

$$p_1 = \sqrt{\frac{4E_b}{5N_0}}$$

Similarly  $P_{b1|a_I=-3d}$  can be written as:

$$P_{b1|a_I=-3d} = Q\left(\sqrt{\frac{36E_b\|h\|^2}{5N_0}}\right) = \frac{1}{\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right), \quad (14)$$

$$p_2 = \sqrt{\frac{36E_b}{5N_0}}$$

Therefore  $P_{b1}$  is

$$P_{b1} = \frac{1}{2\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right) + \frac{1}{2\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right) \quad (15)$$

For 16-QAM,  $P_{b1}=P_{b2}$ , and the error probability  $P_{b3}$  and  $P_{b4}$  are:

$$P_{b3} = \frac{1}{2} \left( 2P_{b1|a_I=-d} + P_{b1|a_I=-3d} - \frac{1}{\sqrt{2\pi p_3}} \exp\left(-\frac{1}{2} p_3\right) \right) \text{ where,} \quad (16)$$

$$p_3 = \sqrt{\frac{100E_b}{5N_0}}$$

The average bit error rate of 16QAM symbol

$$P_b = \frac{1}{2} (P_{b1} + P_{b3}) \quad (17)$$

### III. SYSTEM MODEL

If a signal  $S(t)$  is transmitted at symbol period  $T$ , the received signal for a postdetection diversity receiver can be written as [9]

$$r_k(t) = h_k S(t) + n_k(t), \quad k=1, 2, \dots, L \quad (18)$$

where  $h_k(t)$  is the channel impulse response for the  $k^{\text{th}}$  diversity branch and  $n_k(t)$  is the AWGN noise for the  $k^{\text{th}}$  fading channel.

#### A. MAXIMAL RATIO COMBINING (MRC)

With MRC, each diversity branch is weighted by their respective fading gain, co-phased, and added. The received signal vectors

$$\bar{r} = (r_1, r_2, \dots, r_L) \quad (19)$$

from each of the  $L$  diversity branch are added to provide maximum SNR. If each branch has fading gain  $h_k$ , then the diversity combiner generates the sum [8]

$$\bar{r} = \sum_{k=1}^L h_k^* r_k \quad (20)$$

The envelope of the composite signal is

$$\alpha_M = \sum_{k=1}^L \alpha_k^2 \quad (21)$$

where  $h_k = \alpha_k \exp(-j\theta_k)$ . The weighted sum of the branch noise power of each branch has the same average noise power  $N_0$  expresses by:

$$N = N_0 \sum_{k=1}^L \alpha_k^2 \quad (22)$$

The SNR at the detector device is,  $\gamma_M = \frac{r^2_M}{2N_T}$

The maximum value of  $\gamma_M$ , according to [9] is:

$$\gamma_M = \frac{1}{2} \sum_{k=1}^L \frac{r_k^2}{N_0} = \sum_{k=1}^L \gamma_k \quad (23)$$

The individual average SNR for each diversity branch is given by

$$SNR = \Gamma = \frac{E_b}{N_0} \alpha^2$$

where  $\alpha^2=1$ .

The average SNR ( $\bar{\gamma}_M$ ) equals to the sum of the individual average SNR [8]:

$$\bar{\gamma}_M = \sum_{k=1}^L \gamma_k = \sum_{k=1}^L \Gamma = L\Gamma \quad (24)$$

The probability of error for distance  $-d$  ( $P_{b1|a_I=-d}$ ) for MRC space diversity is given by applying the average SNR from (24) into (12), and then we get:

$$P_{b1|a_I=-d} = Q\left(\sqrt{L \frac{4E_b}{5N_0}}\right)$$

$$= \frac{1}{\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right), p_1 = \sqrt{L \frac{4E_b}{5N_0}} \quad (25)$$

Similarly, the error for distance  $-3d$  ( $P_{b1|a_I=-3d}$ ) from (13), we have:

$$P_{b1|a_I=-3d} = Q\left(\sqrt{L \frac{36E_b}{5N_0}}\right) = \frac{1}{\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right), p_2 = \sqrt{L \frac{36E_b}{5N_0}} \quad (26)$$

Hence, the probability of error for 1<sup>st</sup> bit in the QAM symbol is  $P_{b1} = \frac{1}{2}(P_{b1|a_I=-d} + P_{b1|a_I=-3d})$ . Using Eqs. (25) and (26) the bit error is expressed as:

$$P_{b1} = \frac{1}{2\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right) + \frac{1}{2\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right) \quad (27)$$

Now, we know  $P_{b3}=P_{b4}$ . Similarly, by the similar argument as above, the probability of error for 3<sup>rd</sup> and 4<sup>th</sup> bit can be given as

$$P_{b3} = \frac{1}{2} \left( 2P_{b1|a_I=-d} + P_{b1|a_I=-3d} - \frac{1}{\sqrt{2\pi p_3}} \exp\left(-\frac{1}{2} p_3\right) \right),$$

where  $p_3 = \sqrt{L \frac{100E_b}{5N_0}}$  (28)

From (27) and (28), the average bit error for MRC space diversity is given as:

$$P_b = \frac{1}{2}(P_{b1} + P_{b3}) = \frac{3}{4\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right) + \frac{1}{2\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right) - \frac{1}{4\sqrt{2\pi p_3}} \exp\left(-\frac{1}{2} p_3\right) \quad (29)$$

## B. SELECTION COMBINING

In SC the detector select the highest SNR path, in this case the diversity combiner performs the operation

$$\bar{r} = \max_{|h_k|} |r_k| \quad (30)$$

The instantaneous SNR ( $\gamma_k$ ) of  $k^{\text{th}}$  branch has the probability density function [8] is expressed as:

$$p(\gamma_k) = \frac{1}{\Gamma} \exp\left(-\frac{\gamma_k}{\Gamma}\right) \quad (31)$$

where  $\Gamma$  is the mean SNR of each branch [9] is:

$$\text{SNR} = \Gamma = (E_0/N_0) \bar{\alpha}^2 \quad (32)$$

We assume  $\bar{\alpha}^2 = 1$ . The average SNR with SC [8] is:

$$\bar{\gamma} = \Gamma \sum_{k=1}^L \frac{1}{k} \quad (33)$$

Let us define the average SNR by  $\beta = \bar{\gamma} = \frac{E_b}{N_0} \sum_{k=1}^L \frac{1}{k}$ ,

then from Eq.(27), we have the bit error for 1<sup>st</sup> bit with SC expressed as:

$$P_{b1} = \frac{1}{2\sqrt{2\pi q_1}} \exp\left(-\frac{1}{2} q_1\right) + \frac{1}{2\sqrt{2\pi q_2}} \exp\left(-\frac{1}{2} q_2\right),$$

where  $q_1 = \sqrt{\beta \frac{4E_b}{5N_0}}$  and  $q_2 = \sqrt{\beta \frac{36E_b}{5N_0}}$

Now from (28) the error for 3<sup>rd</sup> and 4<sup>th</sup> bit of QAM symbol with the assumption  $P_{b3} = P_{b4}$  can be given as:

$$P_{b3} = \frac{1}{2} \left( 2P_{b1|a_I=-d} + P_{b1|a_I=-3d} - \frac{1}{\sqrt{2\pi q_3}} \exp\left(-\frac{1}{2} q_3\right) \right), \text{ where}$$

$$q_3 = \sqrt{\beta \frac{100E_b}{5N_0}}$$

The average probability of bit error for SC space diversity through Rayleigh fading channel can be given as

$$P_b = \frac{1}{2}(P_{b1} + P_{b3}) = \frac{3}{4\sqrt{2\pi q_1}} \exp\left(-\frac{1}{2} q_1\right) + \frac{1}{2\sqrt{2\pi q_2}} \exp\left(-\frac{1}{2} q_2\right) - \frac{1}{4\sqrt{2\pi q_3}} \exp\left(-\frac{1}{2} q_3\right) \quad (34)$$

## C. EQUAL GAIN COMBINER (EGC)

The EGC is same as MRC but different from it that the diversity branches are not weighted. The combiner generates the sum [8]

$$\bar{r} = \sum_{k=1}^L \exp(-\theta_k) \bar{r}_k \quad (35)$$

The envelope of the composite signal is

$$\alpha_E = \sum_{k=1}^L \alpha_k \quad (36)$$

The resulting SNR is

$$\bar{\gamma}_M = \frac{\alpha^2 E_b}{L N_0} = \frac{1}{L} \sum_{k=1}^L \gamma_k = \frac{1}{L} \times L \Gamma = \Gamma \quad (37)$$

The probability of error due to distance  $-d$  ( $P_{b1|a_i=-d}$ ) for EGC space diversity is given by applying the average SNR from Eq.(37) into Eq.(12), we have

$$P_{b1|a_i=-d} = Q\left(\sqrt{\frac{4E_b}{5N_0}}\right) = \frac{1}{\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right), p_1 = \frac{4E_b}{5N_0} \quad (38)$$

Similarly, the error probability for distance  $-3d$  ( $P_{b1|a_i=-3d}$ ) from Eq.(13), we have:

$$P_{b1|a_i=-3d} = Q\left(\sqrt{\frac{36E_b}{5N_0}}\right) = \frac{1}{\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right), p_2 = \frac{36E_b}{5N_0} \quad (39)$$

Hence, the probability of error for 1<sup>st</sup> bit in the QAM

symbol is  $P_{b1} = \frac{1}{2} (P_{b1|a_i=-d} + P_{b1|a_i=-3d})$ . Using

Eqs.(38) and (39) the bit error is expressed as:

$$P_{b1} = \frac{1}{2\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right) + \frac{1}{2\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right) \quad (40)$$

The probability of error for 3<sup>rd</sup> and 4<sup>th</sup> bit can be given with the assumption  $P_{b3}=P_{b4}$  as:

$$P_{b3} = \frac{1}{2} \left( 2P_{b1|a_i=-d} + P_{b1|a_i=-3d} - \frac{1}{\sqrt{2\pi p_3}} \exp\left(-\frac{1}{2} p_3\right) \right),$$

$$\text{where } p_3 = \frac{100E_b}{5N_0} \quad (41)$$

Using Eqs. (40) and (41), the average bit error for EGC is as:

$$P_b = \frac{1}{2} (P_{b1} + P_{b3}) = \frac{3}{4\sqrt{2\pi p_1}} \exp\left(-\frac{1}{2} p_1\right) + \frac{1}{2\sqrt{2\pi p_2}} \exp\left(-\frac{1}{2} p_2\right) - \frac{1}{4\sqrt{2\pi p_3}} \exp\left(-\frac{1}{2} p_3\right) \quad (42)$$

#### IV. RESULTS

The simulation results for the BER performance of MRC, SC and EGC space diversity with 16-QAM for  $2 \times 2$  diversity system for Rayleigh fading channel is shown in

Fig.1. The simulation results show that MRC provides better BER performance than SC and EGC. The analytical BER performance of 16-QAM for Rayleigh fading channel with MRC space diversity is obtained from Eq.(29), and is shown in Fig. 2 for  $L = 1$  to 8. It is seen from Figs. 1 and 2 that at  $E_b/N_0 = 10$  dB with 2 transmit and 2 receive antenna system, the simulated and analytical BER is 0.0015 and 0.0028 respectively. The analytical BER performance of 16-QAM symbol for Rayleigh fading channel with SC space diversity is obtained from Eq.(34), and is shown in Fig. 3 for  $L = 1$  to 8.

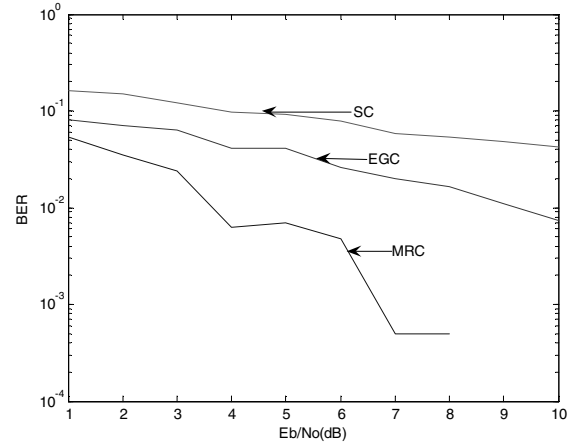


Fig.1. Simulated BER performance of MRC, SC, and EGC with 2 transmits and 2 receive antennas.

It is seen from Figs. 1 and 3 that for  $E_b/N_0 = 10$  dB with 2 transmit and 2 receive antenna system for Rayleigh fading channel, the simulated and analytical values of BER are 0.13 and 0.2 respectively. Similarly, the analytical BER performance of 16-QAM symbol for  $L = 1$  to 8 for Rayleigh fading channel with EGC space diversity is obtained from Eq.(42), and is shown in Fig. 4.

The simulated and analytic BER values of EGC are 0.193 and 0.212 respectively at  $E_b/N_0 = 10$  dB with  $2 \times 2$  space diversity. The theoretical BER performance results of MRC, SC, and EGC from Figs.2, 3, and 4 show that the BER for MRC space diversity are higher than those of SC and EGC values, because MRC weights and adds all branch signals while SC selects one branch having the highest SNR and EGC does not weight each branch but the only co-phased signal.

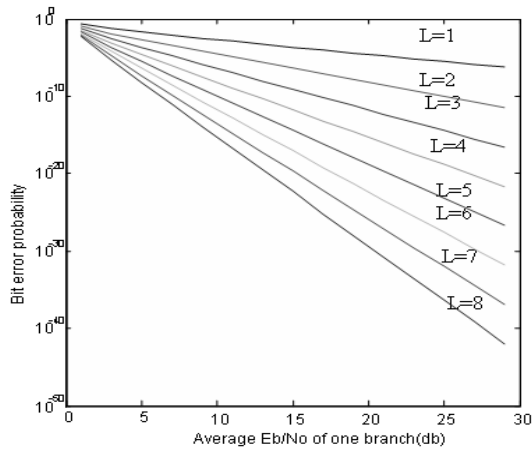


Fig. 2: The analytical BER Performance of 16-QAM with MRC space diversity in Rayleigh fading channels evaluated using (29).

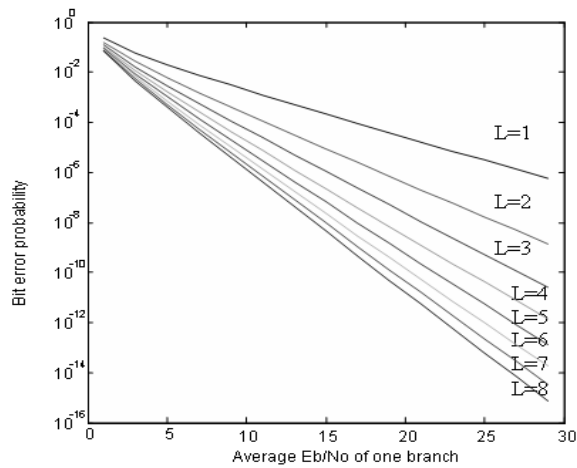


Fig. 3: The BER Performance of 16-QAM with SC space diversity in Rayleigh fading channels according to (34).

## V. CONCLUSION

In this paper, we compare the BER performance of different diversity combiners specially MRC, SC and EGC. The theoretical and simulated BER performance values are shown in this work. The theoretical values are calculated using LLR approach and the simulation values are found by sending a 16-QAM symbol through Rayleigh fading channel.

Both theoretical and analytical BER performance values show that the MRC provides the highest BER performance than those of the other two combiners. The BER performances are also calculated for MRC and SC space diversity for Rayleigh fading channel in work [5]. The simulation results in Fig.1 show that the BER values at  $E_b/N_0 = 10$  dB for MRC and SC are 0.0015 and 0.13 respectively for  $2 \times 2$  diversity. In work [5], the BERs for MRC and SC at  $E_b/N_0 = 10$  dB are 0.03 and 0.24 respectively. We get the BER values at  $E_b/N_0 = 10$  dB to be 0.0028 and 0.2 for MRC and SC from the LLR based BER

derivation as in Eqs.(29) and (34). The key contribution of our paper is that our LLR based BER calculation and simulated BER performance are more accurate than the BER derivation in work [5].

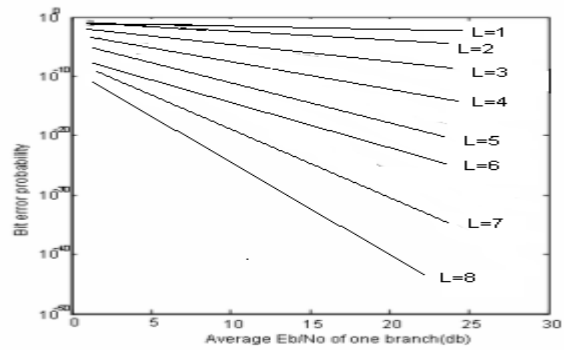


Fig. 4: The BER Performance of 16-QAM with EGC space diversity in Rayleigh fading channels based on (42).

## REFERENCES

- [1] Kamilo feher, "Wireless Digital Communications-Modulation and Spread Spectrum Applications", Prentice-Hall of India, 2005.
- [2] W.Webb and L.Hanzo," Modern Quadrature Amplitude Modulation-Principles and Applications for Fixed and Wireless Communications", IEEE press and pentech press, 1994.
- [3] S.M.Alamouti "A simple transmit diversity Technique for Wireless Communication"pp.1451-1458 IEEE-JSAC,Oct-1998.
- [4] Bernard Sklar,"Digital Communications- Fundamental and Applications", Pearson Education Asia, 2001.
- [5] Chang-Joo kim, Young-Su Kim, Goo-Young Jeong, Jae-kyung mun, and Hyuck-Jac Lee," SER Analysis of QAM With Space Diversity in Rayleigh Fading Channels", ETRI Journal, Volume 17, number 4, January 1996.
- [6] M.Suredra Raju, A. Ramesh and A. Chakalingam, " BER Analysis of QAM with Transmit Diversity in Rayleigh Fading Channels", Proceedings of the IEEE 2003 Global Communications Conference, 1-5 Dec 2003, San Fransisco , USA, PP.641-645 .
- [7] M. K. Simon, S. M. Hinedi, andW. C. Lindsey, "Digital Communication Techniques—Signal Design and Detection. Englewood Cliffs, NJ: PRT", Prentice Hall, 1995.
- [8] Gordon L. Stuber, " Principles of Mobile Communication", Kluwer Academic Publishers, 2001.
- [9] Theodore S. Rappaport,"Wireless Communications-Principles and practice" , Prentice Hall of India private Limited", 2004.

# Learning Intrusion Detection Based on Adaptive Bayesian Algorithm

Dewan Md. Farid, Mohammad Zahidur Rahman

Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh. Emails: [dmfarid@uiu.ac.bd](mailto:dmfarid@uiu.ac.bd), [rmzahid@juniv.edu](mailto:rmzahid@juniv.edu)

**Abstract** — Recent intrusion detection have emerged an important technique for information security systems. It's very important that the security mechanisms for an information system should be designed to prevent unauthorized access of system resources and data. Last few years, many intelligent learning techniques of machine learning applied to the large volumes of complex and dynamic audit data for the construction of efficient intrusion detection systems (IDS). This paper presents, theoretical overview of intrusion detection and a new approach for intrusion detection based on adaptive Bayesian algorithm. This algorithm correctly classify different types of attack of KDD99 benchmark intrusion detection dataset with high detection accuracy in short response time. The experimental result also shows that, this algorithm maximize the detection rate (DR) and minimized the false positive rate (FPR) for intrusion detection.

**Index Terms** — Intrusion detection, Bayesian algorithm, classification, detection rate, false positive rate.

## I. INTRODUCTION

Development of interconnection of computer systems have increased the number of threats to information security systems. To defend various cyber attacks and computer viruses, lots of computer security technique have been introduced in the last few years, which include cryptography, firewalls, and intrusion detection systems (IDS) etc. Among all IDS become more promising for defending complex and dynamic intrusion behaviors. IDS was introduced by Anderson in 1980 [1]. He coined an intrusion as an attempt or a threat to be the potential possibility of a deliberate unauthorized attempt to access information, manipulate information, and render a system unreliable. The main goal of an IDS is to detect computer security incidents accurately in real-time and process the appropriate actions. The most popular IDS uses audit data to detect both pattern-based and anomaly-based detection [2]. An audit data is a record of activities generated by the operating system that are logged to a file in chronologically sorted order. Thus, one of the important problems for IDS is to built patterns to distinguish normal behaviors from abnormal behaviors by observing collected audit data. Because the behavior of an IDS is highly dependent on its environment, it is hard to program an IDS using ordinary procedural programming languages in a constantly changing environment. On the other hand,

large sizes of audit data generated everyday by the computer system makes manual analysis impossible. To solve this problem, earlier IDS usually rely on security experts to analyze the audit data and construct intrusion detection rules manually [3].

Due to the above deficiencies of IDS, intrusion detection techniques are using machine learning algorithms, such as Neural Networks, SVMs, Genetic algorithm, and Naïve Bayesian Classifier, and Decision Tree etc. [4], [5] to detecting known and unknown intrusions. IDS based on machine learning aims to solve the problems of analyzing huge volumes of audit data and realizing performance optimization of detection rules. By making use of machine learning algorithms, intrusion detection models can be automatically constructed based on labeled or unlabeled audit data. Despite of many advances that have been archived, existing IDS still have some difficulties, such as low detection accuracy, unbalanced performance on different attack types and long response time. It is also difficult to improve detection abilities for complex or new attacks without increasing false alarms. Another difficulty of current IDS is to detect high speed network traffics in real time. Because in high speed networks, IDS have to deal with large volumes of data in a short time. So, it is another challenging problem for IDS research to maintain high detection accuracy for high speed networks. In host-base IDS, there still exist other difficult problems for sequential behavior prediction.

The Bayesian algorithm is an efficient method for performing supervised learning [6], [7]. It provides an optimal way to predict the class of an unknown example. It is widely used in many fields of data mining, image processing, bio-informatics, and information retrieval etc. In the basic Bayesian algorithm, conditional probabilities for each attribute value are calculated from a given training set. These conditional probabilities are then used to find out the probabilities of belongingness to different classes. The unseen example is then categorized to that class, which assumes the maximum value. In this paper, based on a comprehensive analysis for the current research challenges in intrusion detection, a new method for intrusion detection using adaptive Bayesian algorithm is presented. Experiment on benchmark data such as KDD99 dataset show that this algorithm correctly classify different



types of attack with high detection accuracy in short response time. Also this algorithm maximize the detection rate (DR) and minimize false positive rate (FPR) for intrusion detection.

In this paper a brief introduction of intrusion detection is presented in section II. In section III our proposed algorithm is presented. Finally experimental result and conclusion are presented in section IV and V.

## II. OVERVIEW OF THE INTRUSION DETECTION

Intrusion detection is the process of identifying unauthorized usage of a computer system. It is an important skill for computer system administrators. It is difficult to learn on the job because it is needed only occasionally but can be critical. The most popular IDS uses audit data to detect intrusions. IDS collect information from a variety of systems and analyze the information for signs of intrusions. In general, there are seven types of IDS that are discussing bellow:

- 1) Network IDS detects intrusions by reading all the incoming packets and trying to find suspicious patterns. It monitors multiple hosts and investigates network traffic by connecting with network devices, like: hub, switch, and router. It also inspects the outgoing or local network traffic as well to identify the internal intruders (who abuse their privileges).
- 2) Protocol-based IDS is installed on a server and analysis the protocol that is use by the server. It sit at the front end of a server, monitoring and analyzing the dynamic behavior and state of the communication protocol between a connected device (a user/PC or system) and the server.
- 3) Application protocol-based IDS normally sits between a group of services/ process, monitor and analysis the behavior and state of the application protocol in use by the system between two connected devices.
- 4) Anomaly-based IDS or misuse intrusion detection detects computer intrusions and misuse by monitoring system activity and classifying it as either normal or anomalous. It attempts to characterize normal operation, and try to detect any deviation from normal behavior. The advantage of anomaly detection is a relatively high detection rate for new types of intrusions.
- 5) Signature-based IDS performs simple pattern matching for match a pattern corresponding to a known attack type. It is also known as pattern-based IDS. The main advantage of

signature-based IDS is that it has a relatively low rate of false alarms, which means is has a relatively high precision. And the main disadvantage of this IDS is that the detection rate of attacks is relatively low, because attacker will try to modify the basic attack signature in such a way that it will not match the known signatures of that attack and it cannot detect a new attack for which a signature is not yet installed in the database.

- 6) Host-based IDS identifies intrusions by analyzing system calls, application logs, file-system modifications, and other host activities. It installed on individual systems to examine the internal interfaces. It monitor the operating system and write data to log files and/or trigger alarms.
- 7) Hybrid IDS combines one or more approaches. Host agent data is combined with network information to form a comprehensive view of the network.

### A. Intrusion Detection Data

Intrusion detection data set was built a predictive model or classifier capable of distinguishing between “bad” connections, called intrusion or attacks, and “good” normal connections. In 1998, to compare the performance of various intrusion detection methods, a simulated environment was set up by the MIT Lincoln Lab. The objective was to survey and evaluate research in intrusion detection. A standard set of data to be audited, which includes a wide variety of intrusions simulated in military network environment. The KDD99 benchmark data set for intrusion detection uses a version of this data set. In KDD99 data set, each data record corresponds to the features of a connection in the network data flow. A connection is a sequence of TCP packets starting and ending at some well defined times, between which data flows to and from a source IP address to a target IP address under some well defined protocol. Each connection is labeled either as normal or as an attack, with exactly one specific attack type. The data records are all labeled with one of the following five classes:

- 1) Normal is a class that generated by simulated daily user normal behavior.
- 2) Probing is a class of attacks where an intruder scans a network to gather information or find known vulnerabilities. An intruder with a map of machines and services that are available on a network can use the information to look for exploits.

- 3) Denial of service is a class of attacks where an intruder makes some computing or memory resource too busy or too full to handle legitimate requests, thus denying legitimate users access to a machine.
- 4) User to root exploits are a class of attacks where an intruder starts out with access to a normal user account on the system and is able to exploit vulnerability to gain root access to system.
- 5) Remote to user attack is a class of attacks where an intruder sends packets to a machine over a network, then exploits machine's vulnerability to illegally gain local access as a user.

### III. PROPOSED METHOD

Correct classification for different types of attack is one of the central problems for intrusion detection. Supervised or unsupervised learning algorithms are using to classify different types of attack in intrusion detection, but their performance are not very satisfactory due to the challenging problem of detecting novel attacks with low false alarms. In this section, we proposed a new algorithm to classify the different types of attack for intrusion detection, using adaptive Bayesian algorithm.

#### A. Naïve Bayesian Algorithm

The naïve Bayesian (NB) algorithm is used for learning task, where a training set with target class is provided. Training set is described by attributed  $A_1$  through  $A_n$  and each attribute is described by attribute values  $a_1, a_2, \dots, a_n$  associated with class  $C$ . The objective is to classify an unseen example, whose class value is unknown but attribute values are known. The Bayesian approach to classifying the unseen example is to assign the most probable target class,  $C_{MAP}$  given the attribute values ( $a_1, a_2, \dots, a_n$ ) that describe the example.

$$C_{MAP} = \underset{C_j}{\operatorname{argmax}} \sum C^P(C_j | a_1, a_2, \dots, a_n)$$

The expression can be rewrite using Bayesian theorem as  $C_{MAP} = \underset{C_j}{\operatorname{argmax}} \sum C^P(a_1, a_2, \dots, a_n | C_j) P(C_j)$  (1) It is easy to estimate each of the  $P(C_j)$  simply by counting the frequency with which each target class  $C_j$  occurs in the training set. The naïve Bayesian algorithm is based on the simplifying assumption that given the target class of the example, the probability of observing the conjunction  $a_1, a_2, \dots, a_n$  is just the product of the probabilities for the individual attributes:  $P(a_1, a_2, \dots, a_n | C_j) = \prod_i P(a_i | C_j)$ . Substituting this into equation 1, we get

$$C_{NB} = \underset{C_j}{\operatorname{argmax}} \sum C^P(C_j \cdot \prod_i P(a_i | C_j)) \quad (2)$$

Where  $C_{NB}$ , denote the target class predicate by the naïve Bayesian classifier. In naïve Bayesian algorithm, the

probability values of equation 2, are estimated from the given training data. These estimated values are then used to classify unknown examples.

#### B. Improved Adaptive Bayesian Algorithm

In our improved Bayesian algorithm, we estimate the class conditional probabilities for each attribute value based on their frequencies in the training dataset. We calculate the sum of weights for each class from the training dataset. Initially all the training dataset have unit weight. Then we calculated the sum of weights for each attribute value with respect to same class. Finally, we calculate the class conditional probabilities for each attribute value from the training dataset by dividing the sum of weights for each attribute value by the sum of weights for each class. After calculating the class conditional probabilities for each attribute value from the training dataset, we classify the test dataset. If any test example is misclassified, we update the weights of training dataset. We compare each of test example with every of training examples and compute the similarity between them (out of four attributes two attribute values are same then similarity becomes 0.5 or one attribute value is same then similarity is 0.25 and so on), and then weights of training dataset are increased by a fixed small value multiplied by the corresponding similarity measure. If the test example is correctly classified, then the weights of training set will remain unchanged. After weights adjustment, the class conditional probabilities for attribute values are recalculated from the modified training dataset. If the new set of probabilities correctly classifies all the test examples, the algorithm terminates. Otherwise, the iteration continues until all the test examples are correctly classified or the target accuracy is achieved. At this stage the algorithm terminates, the class conditional probabilities are preserved for future classification of seen or unseen intrusions.

Algorithm I: Improved Adaptive Bayesian Algorithm (IABA)

Input: training set, test set, weight

Output: intrusion detection model

Procedure:

1. Initialize the weights  $W_1, W_2, \dots, W_n$  of training set:  
 $\text{for } (i = 1 \cdot N)$   
 $W_i = 1.0;$
2. Calculate the sum of weights for each class is training set:  
 $\text{for } (i = C_i \cdot C_N)$   
 $\text{for } (j = 1 \cdot N)$   
 $\text{if } (W_j \cdot C_i) \text{ then } (W_{Ci} = W_{Ci} + W_j);$
3. Calculate the sum of weights for each attribute value in training set:  
 $\text{for } (i = A_i \cdot A_N)$

$for(j = a_i \bullet a_n)$   
 $for(k = C_i \bullet C_n)$   
 $W_{a_{ij}} = W_{a_{ij}} + W_k;$

4. Calculate the class conditional probabilities for each attribute value for training set:

$for(i = A_i \bullet A_n)$   
 $for(j = a_i \bullet a_n)$   
 $P_{a_{ij}} = W_{a_{ij}} / W_{C_i}$

5. Classify the each testing examples:  
 $TargetClass, C_{TC} = \underset{C_j}{\operatorname{argmax}} \sum C_{C_i}^W \bullet_i P_{a_{ij}};$
6. If any test example is misclassified then update the weights of training set: Similarity,  $S = 0.5$  (out of four attributes two attributes values are same)  
 $W_i = W_i + (S * 0.01);$
7. If weight update, go to step (2);
8. If all test examples are correctly classified then the algorithm terminates;

#### IV. EXPERIMENTAL RESULTS

In our experiments, we perform 5-class classification for intrusion detection. The training and test examples are taken from the benchmark KDD99 dataset. In this experiments, 10 percent of the whole KDD99 dataset used for the training examples and 10 percent whole KDD99 dataset used for the testing examples. The training and test examples are taken randomly. This is done with different ratios of positive versus negative instance. In the KDD99 dataset, the normal data belongs to class1, probe belongs to class2, denial of service (Dos) belongs to class3, user to super user (U2R) belongs to class4, and remote to local (R2L) belongs to class5. Table I shows the distribution of connection types for 10% training and 10% testing dataset in the KDD99 dataset, which has a total number of 494021 and 311028 examples respectively.

TABLE I

NUMBER OF CONNECTIONS FOR TRAINING AND TESTING DATASET IN KDD99 DATASET

Class in KDD99	10% Training Examples	10% Testing Examples
Normal	97277	60592
Probe	4107	4166
DoS	391458	237594
U2R	52	70
R2L	1126	8606
Total	494021	311029

In our experiments, firstly we use standard naïve Bayesian algorithm and then our improved adaptive Bayesian algorithm to perform classification and compare their results. The training examples are used to train the

algorithms, and the test examples are used to evaluate the performance of the algorithms. Table II shows the comparisons of training and testing speed of algorithms.

TABLE II

PROCESSING SPEED COMPARISON

Classifiers	Training Time (m)	Testing Time (m)
Naïve Bayesian Algorithm	106.7	26.4
Adaptive Bayesian Algorithm	52.8	13.2

It is clear that that the proposed adaptive Bayesian algorithm is approximately 2 times faster in training and testing than conventional naïve Bayesian algorithm. To evaluate the performance of our improved adaptive Bayesian algorithm for intrusion detection, we are interested in two major indicators of performance: detection rate (DR) and false positive rate (FPR). Detection rate (DR) is defined as the number of intrusion examples detected by the system divided by the total number of intrusion examples present in the dataset. And false positive rate (FPR) is defined as the total number of normal examples. The performance comparison between naïve Bayesian algorithm and our improved adaptive Bayesian algorithm are shown in Table III and Table IV.

TABLE III

PERFORMANCE OF NAÏVE BAYESIAN ALGORITHM

Attack type in KDD99	Training Time (m)	Testing Time (m)
DoS	99.68	0.04
Probe	98.47	0.42
R2L	98.34	7.03
U2R	98.91	0.13

TABLE IV

PERFORMANCE OF ADAPTIVE BAYESIAN ALGORITHM

Attack type in KDD99	Training Time (m)	Testing Time (m)
DoS	99.81	0.03
Probe	99.23	0.39
R2L	99.45	5.96
U2R	99.07	0.11

We also compare the intrusion detection performance among Neural Network (NN), Support Vector Machines (SVM), Naïve Bayesian Classifier, and our improved adaptive Bayesian algorithm (IABA) on KDD99 dataset. The comparative results are summarized in Table V.

TABLE V: COMPARISON OF SEVERAL LEARNING ALGORITHM

Attack type in KDD99	NN(%)	SVM(%)	NB(%)	IABA(%)
DoS	91.82	99.43	99.67	99.78
Probe	97.27	98.53	99.39	99.49
R2L	94.43	96.44	99.01	99.47
U2R	49.05	64.07	64.45	99.35

## V. CONCLUSION

Recently research on machine learning for intrusion detection has received much attention in the computational intelligence community. In intrusion detection algorithm, huge volumes of audit data must be analyzed in order to construct new detection rules for increasing number of novel attacks in high speed network. Intrusion detection algorithm should consider the complex properties of attack behaviors to improved the detection speed and detection accuracy. Analyze the large volume of network dataset and improve the performance of detection accuracy, intrusion detection become an important research field for machine learning. This paper presents a new algorithm for intrusion detection based on machine learning. The future works focus on applying the domain knowledge of security to improve the detection accuracy, visualizing the procedure and result of intrusion detection to promote the understandability of intrusion detection. Also apply machine learning techniques in real commercial intrusion detection systems.

## ACKNOWLEDGEMENT

This work was supported by the Department of Computer Science and Engineering at Jahangirnagar University.

## REFERENCES

- [1] J.P Anderson, "Computer Security Threat Monitoring and Surveillance Technical Report", James P Anderson Co., Fort Washington Pennsylvania, April 1980.
- [2] Neil C. Rowe, and S. Schiavo, "An Intelligent Tutor for Intrusion Detection on Computer Systems", Code CS/Rp, Naval Postgraduate School, Monterey, CA USA 93943, 1998, (Computers and Education) volume 31, pp. 395-404.
- [3] W.k. Lee, S.J.Stolfo. "A data mining framework for building intrusion detection model", In: Gong L., Reiter M.K. (eds.): Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA: IEEE Computer Society Press, pp. 120-132, 1999.
- [4] Cannady J. "Artificial neural networks for misuse detection" In: Proceedings of the 1998 National Information System Security Conference. Arlington: Virginia Press, 1998, pp. 443-456.
- [5] Shon T, Seo J, Moon J, "SVM approach with a genetic algorithm for network intrusion detection". In: Proceedings of the 20<sup>th</sup> International Symposium on Computer and Information Science. Berlin: Springer Verlag, 2005, pp. 224-233.
- [6] Kononenko, I, "Comparison of inductive and naïve Bayesian learning approaches to automatic knowledge acquisition", In: Wieling, B. (Ed), Current trend in knowledge acquisition, Amsterdam, IOS press, 1990.
- [7] Langely, P., Iba, W. and Thomas, K. "An analysis of Bayesian classifier" In: Proceedings of the 10<sup>th</sup> National Conference on Artificial Intelligence (San Mateo, CA: AAAI) 1992, pp. 223-228
- [8] Sebastiaan Tesink, "Improving Intrusion Detection System through Machine learning", Technical report (sl. 07-02) ILK Research Group, Tilburg University, March 2007.
- [9] Cestnik, B., "Estimating probabilities: a crucial task in machine learning". In: proceedings of European Conference on Artificial Intelligence, 1990, pp. 147-149.
- [10] Kohavi, R., "Scaling up the accuracy of Naïve Bayesian Classifier: A decision tree Hybrid", In: Proceedings of 2<sup>nd</sup> International Conference on Knowledge Discovery and Data Mining (Mento Park, CA: The AAAI Press) 1996, pp.202-207.
- [11] Fisher, D.H., "Knowledge acquisition via incremental conceptual clustering", Machine Learning, Kluwer Academic Publishers, Boston, 1987, vol. 2, pp. 139-187.
- [12] Xin Xu, "Adaptive Intrusion Detection Based on Machine Learning: Feature Extraction, Classifier Construction and Sequential Pattern Prediction", In: International Journal of Web Services Practices, 2006, Vol. 2, No. 1-2, pp. 49-58.
- [13] Paverl LasKov, Patrick Dussel, Christin Schafer, and Konrad Rieck, "Learning Intrusion Detection: supervised and unsupervised", Fraunhofer-FIRAT.IDA, Kelulestr. 7, 12489 Berlin, Gernany, 2006.
- [14] M. M. Sebring, E. Shellhouse, M.E. Hanna, and R. Alan Whitehurst, "Expert Systems in Intrusion Detection: A case study", In: Proceedings of the 11<sup>th</sup> Natioanl Computer Security Conference, Baltimore, Maryland, October 1988, pp. 74-81
- [15] Biswanath Mukherjee, L Todd Feberlein, and Kari N Levitt, "Network Intrusion Detection IEEE Network", May/June 1994, pp. 26-41.

# Analytical Performance Evaluation of Space Time Coded MIMO OFDM Systems Impaired by Timing Jitter

M. R. H. Mondal<sup>1</sup> and S. P. Majumder<sup>2</sup>

<sup>1</sup>IICT, BUET, rubaiyat97@iict.buet.ac.bd, Bangladesh

<sup>2</sup>Department of EEE, BUET, spmajumder@eee.buet.ac.bd, Bangladesh

**Abstract** — An analytical approach is presented to determine the impact of time selective fading, timing jitter and AWGN on the bit error rate performance of an OFDM system with DQPSK and QPSK modulation. The expression for the conditional BER conditioned on a given timing error and fading, is derived and the average BER is evaluated. The BER performance results are evaluated for different values of fading and jitter variances. The performance of the OFDM system in Rayleigh and Rician fading channels is also compared. Analytical approach is also developed to evaluate the BER performance of a quasi-orthogonal space-time block coded (STBC) OFDM system having multiple transmitting and single receiving antennas. The analysis is extended for a MIMO-OFDM system using the “selection method” for combining multiple receiving antennas, which offers significant improvement in the system performance. The effects of increase in number of OFDM subcarriers and increase in Doppler frequency are also investigated.

**Index Terms** — Convolutional coding, intersymbol interference, MIMO, OFDM, Rayleigh fading, STBC, timing jitter.

## I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is effective in avoiding intersymbol interference caused by multipath delay. However, it is sensitive to time-selective fading which destroys the orthogonality among different subcarriers in one OFDM symbol leading to intercarrier interference. OFDM signal is also degraded by timing jitter caused by time synchronization errors at the receiver. Several research works have been carried out during the last few years on the performance evaluation of an OFDM system over fading channels in [1]-[5]. The effect of timing jitter on the performance of an OFDM system without considering multipath fading is also reported [6], [7]. The performance of quasi-orthogonal-space-time coded (QOSTC) system over fast fading channel is investigated for single-carrier systems in [8]. The performance of QOSTC-OFDM system over fading channel is reported in terms of carrier to interference (C/I) and signal to interference and noise ratios [9]. The BER is also evaluated through simulations. However in [9] the jitter effect is ignored and only a single receiving antenna is used.

In this paper, we develop an analytical approach in order to find the BER of an OFDM system and a STBC-OFDM system considering all three channel impairments like AWGN, fading and jitter in fading environment. We extend the analysis for a MIMO-OFDM system with switching/selection method for combining multiple receiving antennas. We also apply convolution coding to evaluate the effectiveness of coding in minimizing the effect of system impairments.

## II. SYSTEM MODEL

The STBC-OFDM system model considered for analysis is shown in Fig. 1.

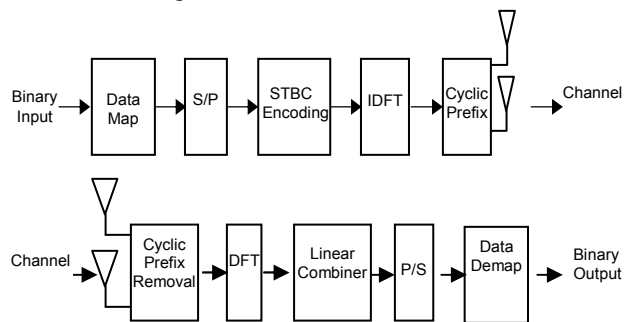


Fig. 1. Block diagram of the system

Binary input data is mapped to modulation symbols. These symbols are serial-to-parallel converted and then mapped to a matrix by using a quasi-orthogonal STBC with constellation rotation scheme. An inverse Fourier transform is used to find the corresponding time waveform. The guard period is then added to reduce the intersymbol interference (ISI). This guard period is a cyclic extension of the symbol to be transmitted. After the guard has been added, the symbols are then converted back to a serial time waveform. This is then the base band signal for the OFDM. The symbols are transmitted from the multiple transmitting antennas simultaneously during every OFDM symbol period.

The receiver basically does the reverse operation to the transmitter. The signal is received in two receiving

antennas. Diversity combiner at the receiver selects the best instantaneous signal of the two antennas. The received signal can be detected through differential or coherent scheme.

### III. THEORETICAL ANALYSIS OF BER

For a stand alone OFDM system the STBC Encoding block and Linear Combiner block in Fig. 1 will not exist. Considering such an OFDM system, the expression of signal-to-noise plus interference power ratio (SNIR) without timing error can be expressed as [5]:

$$\text{SNIR} = \frac{1}{\frac{(k+1)(1+\delta_c)}{2E_b} + \frac{(\pi F_d T_s)^2}{2} + \frac{(\pi F_d T_s)^2}{6(1+\delta_c)^2}} \quad (1)$$

where  $F_d$  is the maximum Doppler frequency,  $N_0$  is the Gaussian noise (AWGN),  $E_b$  is the bit energy, and  $k$  is the Rician parameter/factor. In (1)  $T_s$  is the symbol duration and  $\delta_c$  is the guard interval ratio of an OFDM symbol. Here intra-symbol interference ( $I_s$ ) and intercarrier interference ( $I_c$ ) are considered as:

$$I_s = \frac{(\pi F_d T_s)^2}{2}$$

$$I_c = \frac{(\pi F_d T_s)^2}{6(1+\delta_c)^2}$$

By expressing the intercarrier interference component in terms of number of OFDM subcarriers ( $N_s$ ), the expression of SNIR in (1) is modified as follows:

$$\text{SNIR} = \frac{1}{\frac{(k+1)(1+\delta_c)}{2E_b} + \frac{(\pi F_d T_s)^2}{2} + \sum_{n=1}^{N_s} \frac{1}{n^2} (F_d T_f)^2} \quad (2)$$

where  $T_f$  is the effective symbol period deducting guard period. To incorporate the jitter effect along with fading and AWGN, the jitter term is added in the equation of SNIR. In any communication system timing jitter causes the signal power to degrade by a factor of  $(1-\varepsilon)$  over a time slot. The jitter also causes interference which is added as  $E_b \varepsilon$ . Here  $\varepsilon$  is the timing error normalized by  $T_s$  i.e. ( $\varepsilon = \Delta / T_s$ ) and  $E_b$  is the bit energy. However in fading environment, the received signal energy is random depending on random fade but the noise energy remains unchanged. The received bit energy  $E_b$  is the product of input bit energy  $E_m$  and the multiplicative fade  $\alpha^2$ , i.e.  $E_b = E_m \alpha^2$ . So with this modification the expression of SNIR for a stand alone OFDM system is as follows:

$$\text{SNIR}(\varepsilon, \alpha) = \frac{1}{\frac{(k+1)(1+\delta_c)}{2(E_m \alpha^2)(1-\varepsilon)} + \frac{(\pi F_d T_s)^2}{2} + \sum_{n=1}^{N_s} \frac{1}{n^2} (F_d T_f)^2 + (E_m \alpha^2) \varepsilon} \quad (3)$$

In the presence of time-varying fading, SNIR for quasi-orthogonal STBC-OFDM system has an expression as [9]:

$$\text{SNIR} = \frac{4 \sum_{l=0}^{L-1} [N_s + 2 \sum_{i=1}^{N_s-1} (N_s - i) J_0(2\pi F_d T_s)] e^{-\frac{\tau_l}{\tau_{rms}}}}{4 \sum_{k'=1}^{N_s-1} \sum_{l=0}^{L-1} [N_s + 2 \sum_{i=1}^{N_s-1} (N_s - i) J_0(2\pi F_d T_s) \cos(\frac{2\pi}{N_s} k' i)] e^{-\frac{\tau_l}{\tau_{rms}}} + \sigma^2} \quad (4)$$

The above expression is considered for unit symbol energy and four number of transmitting antennas. In (4)  $\tau_l$  is the delay of the  $l$ th path normalized with respect to the OFDM symbol duration  $T_s$  whereas  $\tau_{rms}$  represents the root-mean square (rms) delay spread. In the expression  $\sigma^2$  represents the variance of AWGN. Now we consider  $E_s$  symbol energy with guard interval ratio  $\delta_c$  and Rician parameter  $k$ . For DQPSK and QPSK symbol energy is twice the bit energy i.e.  $E_s = 2E_b$ . So for DQPSK and QPSK modulation the equation of SNIR for a quasi-orthogonal STBC-OFDM system in (4) can be modified as follows:

$$\text{SNIR} = \frac{N_T \sum_{l=0}^{L-1} \left\{ \frac{2E_b}{(k+1)(1+\delta_c)} \right\} [N_s + 2 \sum_{i=1}^{N_s-1} (N_s - i) J_0(2\pi F_d T_s)] e^{-\frac{\tau_l}{\tau_{rms}}}}{N_T \left\{ \frac{2E_b}{(k+1)(1+\delta_c)} \right\} \left[ \sum_{k'=1}^{N_s-1} \sum_{l=0}^{L-1} \left\{ N_s + 2 \sum_{i=1}^{N_s-1} (N_s - i) J_0(2\pi F_d T_s) \cos(\frac{2\pi}{N_s} k' i) \right\} e^{-\frac{\tau_l}{\tau_{rms}}} \right] + \sigma^2}$$

where  $N_T$  is the number of transmitting antennas. For STBC-OFDM, jitter has the same effect as in OFDM systems. So by simultaneously considering AWGN, jitter and fading the expression of SNIR for a quasi-orthogonal STBC-OFDM system can be expressed as follows:

$$\text{SNIR}(\varepsilon, \alpha) = \frac{N_T \sum_{l=0}^{L-1} \left\{ \frac{(2E_m \alpha^2)(1-\varepsilon)}{(k+1)(1+\delta_c)} \right\} [N_s + 2 \sum_{i=1}^{N_s-1} (N_s - i) J_0(2\pi F_d T_s)] e^{-\frac{\tau_l}{\tau_{rms}}}}{N_T \left\{ \frac{(2E_m \alpha^2)(1-\varepsilon)}{(k+1)(1+\delta_c)} \right\} \left[ \sum_{k'=1}^{N_s-1} \sum_{l=0}^{L-1} \left\{ N_s + 2 \sum_{i=1}^{N_s-1} (N_s - i) J_0(2\pi F_d T_s) \cos(\frac{2\pi}{N_s} k' i) \right\} e^{-\frac{\tau_l}{\tau_{rms}}} \right] + (\sigma^2 + E_m \alpha^2 \varepsilon)}$$

Following [5], the bit error rate (BER) performance for DQPSK/OFDM system over multipath fading channels can be evaluated in presence of jitter as:

$$P_e(\varepsilon, \alpha) = \frac{\{1 - J_0(2\pi F_d T_s)\} \left\{ \frac{SNIR(\varepsilon, \alpha)}{2} \right\} + 1}{2 \left\{ \frac{SNIR(\varepsilon, \alpha)}{2} + 1 \right\}} \exp \left\{ -\frac{k \frac{SNIR(\varepsilon, \alpha)}{2}}{\frac{SNIR(\varepsilon, \alpha)}{2} + 1} \right\}$$

Here  $J_0(2\pi F_d T_s)$  is the time correlation function. Following [10], The BER performance for QPSK/OFDM system can be expressed as follows:

$$P_e(\varepsilon, \alpha) = 0.5 \operatorname{erfc} \sqrt{SNIR(\varepsilon, \alpha)}$$

For a fixed value of  $\alpha$  the probability of bit error is  $P_e(\alpha)$ . This  $P_e(\alpha)$  is considered as a conditional probability of error for a given value of channel attenuation  $\alpha$ . For Rayleigh fading channel,  $\alpha$  is Rayleigh distributed. Then the pdf of  $\alpha$  is:

$$p(\alpha) = \frac{\alpha}{\sigma^2} \exp\left(-\frac{\alpha^2}{2\sigma^2}\right)$$

Here  $\sigma$  is Rayleigh parameter. The mean and variance value of the Rayleigh density function is a function of  $\sigma$ . Mean =  $1.25 \sigma$  and variance =  $0.655 \sigma$ . For Rician fading channel,  $\alpha$  is Rician distributed. Then the pdf of  $\alpha$  is:

$$p(\alpha) = \frac{\alpha}{\sigma^2} \exp\left(-\frac{(\alpha^2 + m^2)}{2\sigma^2}\right) I_0\left(\frac{m\alpha}{\sigma^2}\right)$$

Here  $I_0$  is the zeroth order Bessel function and  $m^2 = 2k\sigma^2$ . The pdf of zero mean Gaussian distributed jitter is

$$p(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{\varepsilon^2}{2\sigma_\varepsilon^2}\right)$$

where  $\sigma_\varepsilon^2$  is the variance of jitter. The unconditional/average BER ( $P_e$ ) can be calculated by averaging the conditional BER,  $P_e(\varepsilon, \alpha)$ , over all possible values of  $\alpha$  and  $\varepsilon$

$$P_e = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_e(\varepsilon, \alpha) p(\alpha) p(\varepsilon) d\alpha d\varepsilon$$

STBC-OFDM with transmitting diversity is transformed into MIMO-OFDM by addition of receiving diversity. For receiving diversity scheme, "selective" diversity combining method is utilized in wireless systems. Considering the average bit error probability without diversity as  $P_{e1}(A)$ , the average BER with diversity can be derived as [10]:

$$P_{e2}(A) = 2P_{e1}(A) - P_{e1}(A/2)$$

where  $A$  denotes the average SNIR.

#### IV. RESULTS AND DISCUSSION

Following the analytical approach, the bit error rate performance results are evaluated at a data rate of 1 Mbps per OFDM subcarrier with maximum Doppler frequency of

60 Hz in fading channel. Keeping number of subcarriers to sixteen and fading variance to 0.1 we compare the performance of the system with and without jitter. Fig. 2 shows the plots of BER vs.  $P_{in}$  (dBm) for DQPSK-OFDM system. From Fig. 2 it is noticed that the BER slightly degrades with increase in jitter. Fig. 3 shows the plots of BER vs.  $P_{in}$  (dBm) for QPSK system. The plots show that OFDM is robust against lower values of jitter variance but the BER is highly degraded when jitter is substantially high and results in BER floor. At a BER of  $10^{-9}$ , the jitter effect is negligible for jitter variance  $\sigma_\varepsilon^2 < 0.125$  with QPSK modulation. It is found that the jitter effect is more pronounced in QPSK than in DQPSK system. The amounts of penalty suffered by the OFDM systems due to timing jitter are shown in Table I. Fig. 4 shows the plots of BER vs.  $P_{in}$  (dBm) for DQPSK system in Rayleigh and Rician channels. It is noticed that BER performance is better in Rician fading channel. This is because in Rician channel there exists one dominant line-of-sight path that is absent in Rayleigh channel. From Fig. 4 it is also revealed that BER performance improves with increase in Rician factor. For example, at 0 dBm input power and DQPSK modulation, the BER of a Rayleigh fading channel is  $10^{-4}$  while Rician channels have a BER in the order of  $10^{-5}$ ,  $10^{-6}$  and  $10^{-7}$  for  $K=0$  dB, 3 dB and 6 dB respectively. This result can be explained from the fact that higher Rician factor indicates more dominant line-of-sight path.

Fig. 5 shows the effect of timing jitter in a DQPSK STBC-OFDM system with one receiving antenna and fading variance of 0.1. From Fig 5, it is revealed that with the increase in jitter the BER performance degrades and results in BER floor for higher values of jitter variance. At a jitter variance  $\sigma_\varepsilon^2$  of 0.2, the BER floor occurs at about  $10^{-3}$  compared to  $10^{-4}$  for the case of without timing error. We also evaluate the similar performance of the system employing QPSK modulation as shown in Fig. 6 and noticed that the jitter causes BER floor. Fig. 7 illustrates the performance comparison of STBC-OFDM with and without receiver diversity. The plot shows that the BER reduces from  $10^{-3}$  to  $10^{-6}$  in presence of jitter  $\sigma_\varepsilon^2 = 0.2$  by deploying diversity combining in receiving side with two receiving antennas. It shows two pair of curves, one without jitter and another with jitter variance of 0.2.

Fig. 8 shows the plots of Space time block coded MIMO-OFDM with DQPSK modulation system with and without convolutional coding in presence of jitter considering four transmitting and two receiving antennas. For convolution code of rate  $1/2$ , the coding gain is 26 dB for constraint length  $K=6$  and 28 dB for  $K=7$  at an uncoded BER of  $10^{-7}$ . It is also noticed that for higher amounts of input power, the coding gain is substantially higher in  $K=7$  than in  $K=6$ . Fig. 9 shows the plots of BER vs. number of subcarriers ( $N_s$ )

for different values of Doppler frequency normalized by symbol period. It is noticed that with the increase in Doppler frequency the BER increases. From Fig. 9 it is also revealed that with increase in  $N_s$ , the BER decreases and becomes minimal and then increases for a particular value of Doppler frequency. It is observed that BER is minimal when number of subcarriers ( $N_s$ ) is 12.

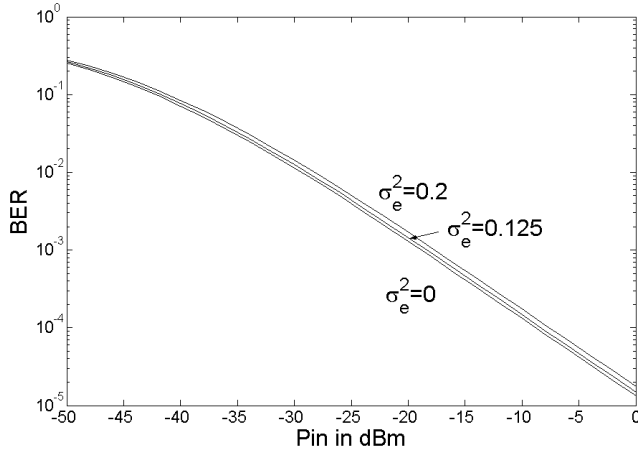


Fig. 2. BER vs.  $P_{in}$  (dBm) in presence of timing jitter for DQPSK-OFDM ( $N_s=16$ ,  $F_d=60$  Hz,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s)

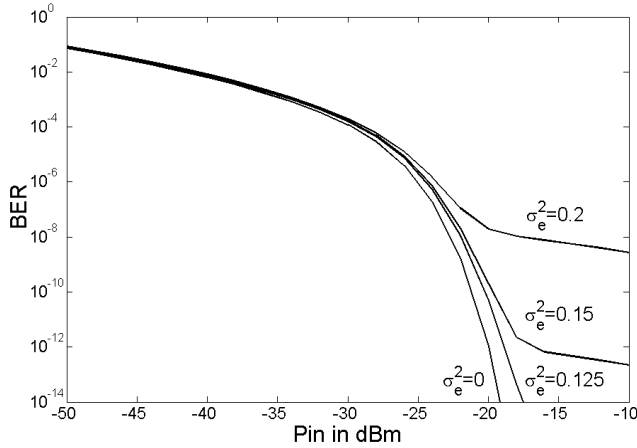


Fig. 3. BER vs.  $P_{in}$  (dBm) in presence of timing jitter for QPSK-OFDM ( $N_s=16$ ,  $F_d=60$  Hz,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s)

TABLE I

Power penalty (in dB) due to jitter for OFDM system at  $BER=10^{-6}$

Modulation	$\sigma_e^2 = 0.2$	$\sigma_e^2 = 0.15$	$\sigma_e^2 = 0.125$
DQPSK	0.85	0.65	0.5
QPSK	1.5	0.8	0.6

Fig. 10 shows the plots of BER vs.  $P_{in}$  (dBm) for the analysis presented by Zhang in [9] and the proposed analysis of this paper. The plots are considered for Rayleigh channel with  $F_d=60$  Hz,  $T_s=10^{-6}$  s, four number of transmitting and single receiving antennas. The proposed analysis shows more BER than the previous analysis presented by Zhang [9], because the proposed analysis considers additional impairments of timing jitter ( $\sigma_e^2=0.1$ ) and channel attenuation ( $\sigma_f^2=0.1$ ). From the findings of the evaluation we can propose an optimum system with following design parameters:

1. DQPSK as modulation scheme as it is more robust against jitter.
2. STBC-OFDM with  $T_x=4$ ,  $R_x=2$
3. Number of OFDM subcarriers 12
4. Convolutional coding of rate  $1/2$  and  $K=7$

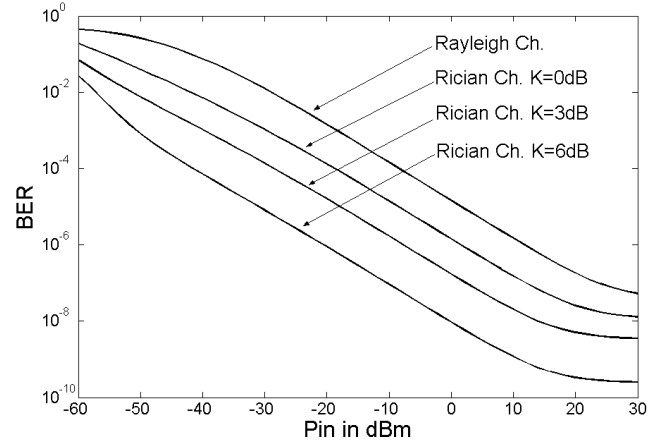


Fig. 4. BER vs.  $P_{in}$  (dBm) for Rayleigh and Rician channels for DQPSK-OFDM ( $N_s=16$ ,  $F_d=60$  Hz,  $\sigma_e^2=0.2$ ,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s)

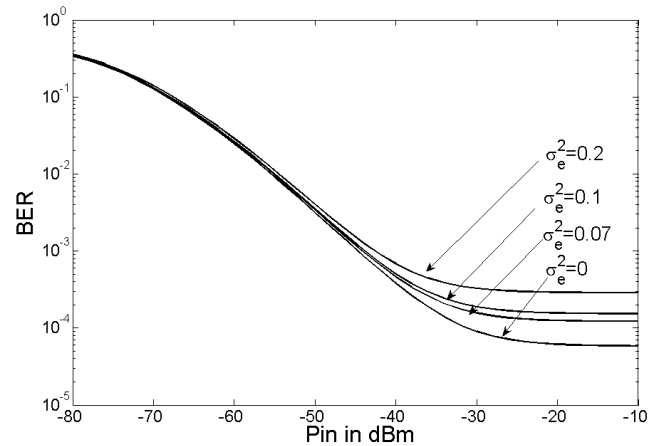


Fig. 5. BER vs.  $P_{in}$  (dBm) in presence of timing jitter for STBC-OFDM (DQPSK) ( $N_s=16$ ,  $F_d=60$  Hz,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s,  $T_x=4$ ,  $R_x=1$ )



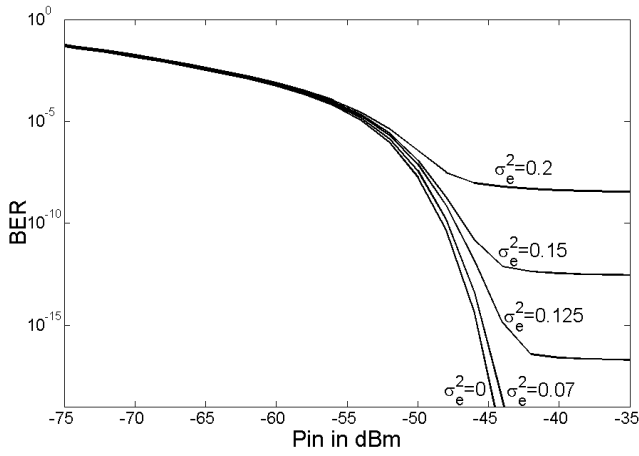


Fig. 6. BER vs.  $P_{in}$  (dBm) in presence of timing jitter for STBC-OFDM (QPSK) ( $N_s=16$ ,  $F_d=60$  Hz,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s,  $T_x=4$ ,  $R_x=1$ )

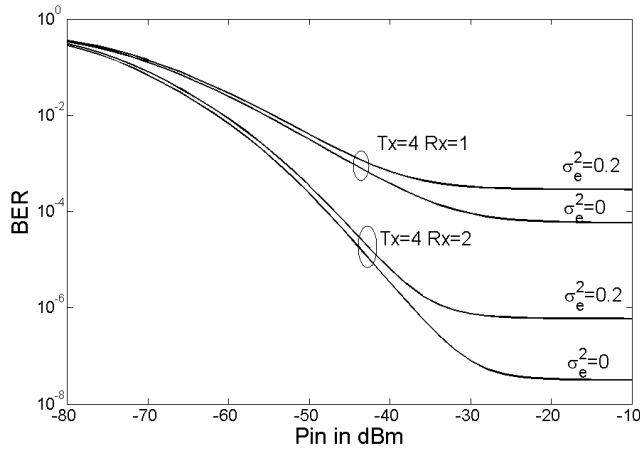


Fig. 7. Plots of BER vs.  $P_{in}$  (dBm) with & without receiving diversity for STBC-OFDM (DQPSK)

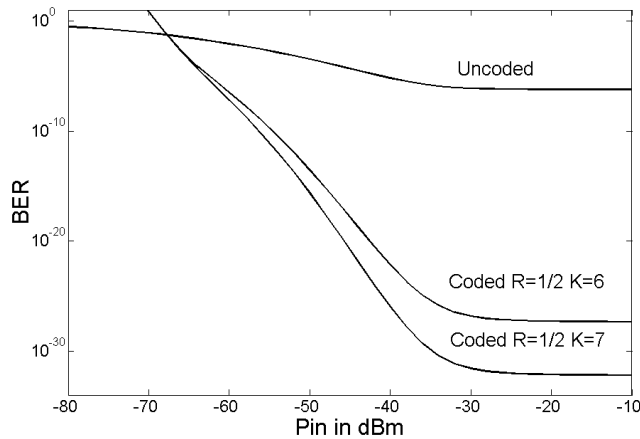


Fig. 8. BER vs.  $P_{in}$  (dBm) with and without coding for MIMO-OFDM (DQPSK) ( $N_s=16$ ,  $F_d=60$  Hz,  $\sigma_e^2=0.2$ ,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s,  $T_x=4$ ,  $R_x=2$ )

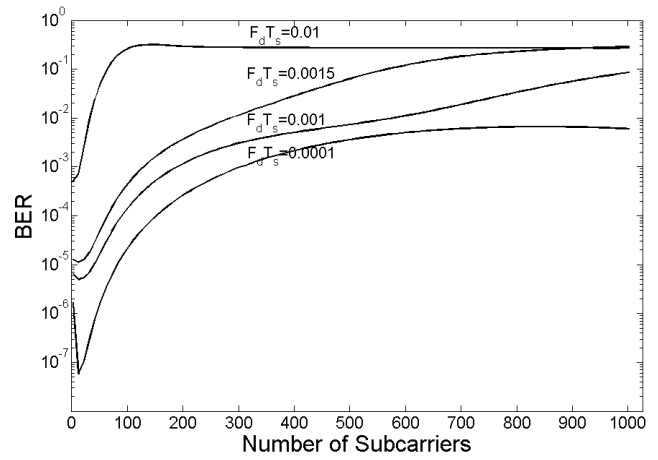


Fig. 9. BER vs.  $N_s$  (number of subcarriers) for MIMO-OFDM (DQPSK) ( $F_d=60$  Hz,  $\sigma_e^2=0.01$ ,  $\sigma_f^2=0.1$ ,  $T_s=10^{-6}$  s,  $T_x=4$ ,  $R_x=2$ )

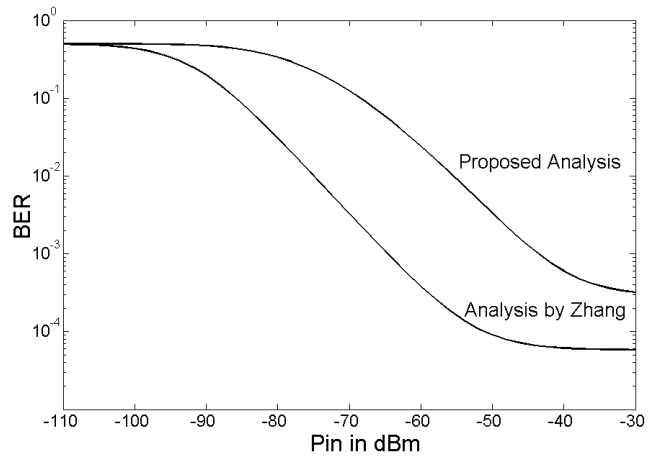


Fig. 10. BER vs.  $P_{in}$  (dBm) for previous and proposed analysis ( $F_d=60$  Hz,  $T_s=10^{-6}$  s,  $T_x=4$ ,  $R_x=1$ )

## V. CONCLUSION

We have analyzed the combined effects of timing jitter, time selective fading and AWGN on the BER performance of an OFDM system. For DQPSK the jitter effect is less pronounced than coherently detected QPSK system. For example, at a jitter variance of 0.2, the penalty at a BER of  $10^{-6}$  is 1.5 dB for QPSK and 0.85 dB for DQPSK. We also observe that the BER of an OFDM system is much higher in Rayleigh fading channel than Rician channel. We compute the BER of OFDM in Rician fading channels with different Rician parameters and numerical results show that the performance is improved with higher values of Rician parameters. It is noticed that STBC-OFDM shows better BER performance than OFDM system. Numerical computations indicate that BER becomes higher with increase in Doppler frequency and increase in number of subcarriers. We extend the analysis for a MIMO-OFDM system with selection method for combining multiple

receiving antennas and find substantial improvement in the system performance. As coded MIMO-OFDM successfully achieves outstanding performance, it can be used in practical applications where fading, jitter and AWGN strongly exist.

#### REFERENCES

- [1] A. Glavieux, P. Y. Cochet, and A. Picart, "Orthogonal Frequency Division Multiplexing with BFSK Modulation in Frequency Selective Rayleigh and Rician Fading Channels", *IEEE Transaction on Communication*, vol. 42, no. 2/3/4, February/March/April 1994.
- [2] H. Steendam and M. Moeneclaey, "Analysis and Optimization of the Performance of OFDM on Frequency-Selective Time-Selective Fading Channels", *IEEE Transaction on Communication*, vol. 47, no. 12, December 1999.
- [3] S. B. Bulumulla, S. A. Kassam and S. S. Venkatesh, "A Systematic Approach to Detecting OFDM Signals in a Fading Channel", *IEEE Transaction on Communication*, vol. 48, no. 5, May 2000.
- [4] F. Sasamori, H. Umeda, S. Handa, F. Maehara, F. Takahata, and S. Oshita, "Approximate Equation of Bit Error Rate in DQPSK/OFDM Systems Over Fading Channels", *European Wireless 2002*, pp 602-607, February 2002.
- [5] F. Sasamori, S. Handa, and S. Oshita, "A Simple Method of BER Calculation in DPSK/OFDM Systems over Fading Channels", *IEICE Trans. Fundamentals*, vol. E88-A, no. 1, pp. 366-373, January 2005
- [6] L. Tomba and W. A. Krzymien, W. A., "A Model for the Analysis of Timing Jitter in OFDM Systems", *IEEE International Conference on Communications (ICC '98)*, vol. 3, pp. 1227-1231, 7-11 June 1998.
- [7] K. N. Manoj and G. Thiagarajan, "The Effect of Sampling Jitter in OFDM Systems", *IEEE International Conference on Communications (ICC '03)*, vol. 3, pp. 2061-2065, 11-15 May 2003.
- [8] H. Jafarkhani, "A Quasi-Orthogonal Space-Time Block Code", *IEEE Transaction on Communication*, vol. 49, no.1, January 2001.
- [9] Y. Zhang and H. Liu, "Impact of Time-Selective Fading on the Performance of Quasi-Orthogonal Space-Time-Coded OFDM Systems", *IEEE Transaction on Communication*, vol. 54, no.2, February 2006
- [10] K. Feher, *Wireless Digital Communication (Modulation and Spread Spectrum Application)*. pp. 209-214, Prentice-Hall. 2005.

# Performance Comparison of IPv4 and IPv6 on Various Windows Operating Systems

Shaneel Narayan, Samad S. Kolahi, Yonathan Sunarto, Du D. T. Nguyen, Paul Mani

School of Computing and Information Technology, UNITEC New Zealand  
snarayan@unitec.ac.nz

**Abstract** — Internet is a ubiquitous part of businesses and individuals worldwide. With its popularity on an incline, operating system vendors are developing end-systems that support the new version of Internet Protocol (IPv6) that eventually will replace IPv4. The new version resolves issues that IPv4 has and takes the Internet into the 21st century. Performance of the IP stack and how it associates with operating systems is critical to the efficiency of all network related activities on any computing infrastructure. Hence it is essential to evaluate performance of IP stack with different operating systems. In this paper, two Microsoft operating systems namely, Windows XP and Windows Server 2003 are configured with the two versions of IP and empirically evaluated for performance difference. The experimental results demonstrate that theoretical and practical performance values are different. It also shows that network performance depends not only on IP version and traffic type, but also on the choice of the operating system.

**Index Terms** — Internet Protocol, IPv4, IPv6, network performance, Windows operating system.

## I. INTRODUCTION

Internet has become the default communication channel for businesses and it continues to grow exponentially beyond technical capacity anticipated by the original engineers and designers. There is now an expectation among experts that the currently used version of Internet Protocol, IPv4, will run out of previously unallocated address space in 2010 or 2011, as only 16% of the total address space remains unallocated in early 2008 [1]. Eventual depletion of Internet addresses in the current schema was foreseen in the late 1990s and Internet Engineering Task Force (IETF) began work on the development of a new version of Internet Protocol, IPv6. This new version not only substantially expands the address space, but also has a raft of additional benefits that were lacking in the predecessor. Today, new versions of popular end-user operating systems come fully integrated with IPv6. Hardware vendors, software developers and Internet Service/Exchange Providers are moving towards offering support for IPv6 [2]. However the overall uptake of this new technology is slower than anticipated and there are considerable challenges that need to be addressed. Performance degradation of network with IPv6 is one such challenge. IPv4 has a 20-Byte header and IPv6 header is

40 Bytes, therefore theoretically the latter has higher overhead. This overhead varies with data payload size -for example, with no fragmentation of datagrams a payload of 128 Bytes has performance difference of 10.3%, and for payload of 1408 Bytes the difference is 1.3% approximately. In this paper, we compare TCP and UDP traffic performance theoretical values with data from test bed experiments for Windows operating systems with IPv4 and IPv6. Our results show that theoretical and practical values are dissimilar. Further, it demonstrates that performance of TCP and UDP varies for different operating systems.

The rest of the paper is organized as follows: Section II discusses some of similar work undertaken by other researchers in evaluating performance of protocols on operating systems. Section III describes Internet Protocols mainly focusing on IPv6 and Section IV outlines the experimental setup used in this research to evaluate the performance of IPv4 and IPv6 on different operating systems. We present the results and discuss the findings in Section V. Finally, conclusions from the research are drawn in Section VI.

## II. RELATED RESEARCH

Research related to performance evaluation of protocols with different operating systems has been undertaken by a number of researchers [Table 1]. Early research focused mostly on one operating system implemented with IPv6 beta (or prototype) and compared performance with IPv4. Results indicated that IPv6 experiences slight performance degradation when implemented on a network (for example, Windows NT with IPv6 downgraded by 2% in [3] and FreeBSD with IPv6 performed slightly inferior in [4], [5] and [6]). In [7], it is shown that both protocols performed the same on FreeBSD. Recent researchers have evaluated the performance of protocols on two or more operating systems comparatively. Windows 2000 and Solaris 8 have been compared in [8] and this work was extended to include Redhat 7.3 in [9]. And Windows 2000 with Linux has been evaluated in [10]. Windows 2003 and Redhat 9.0 have been evaluated with FreeBSD 4.9 in [11], and with FreeBSD 5.3 in [12]. All these researchers have

empirically shown that there are performance differences of various degrees between operating systems with same Internet Protocols. Both TCP and UDP traffic also show interesting patterns when tested on various operating systems.

RESEARCHER(S)	OPERATING SYSTEM(S)
Draves, Mankin & Zill (1998)	Windows NT (MSR IPv6)
Xie (1999)	Windows NT (MSR IPv6 Beta)
Ariga, et al. (2000)	FreeBSD (KAME IPv6)
Ettikan (2000)	FreeBSD (KAME IPv6)
Anand (2000)	Linux kernel 2.4.17 (IPv6)
Zeaddally & Raicu (2003)	Windows 2000, and Solaris 8 (IPv4, IPv6)
Zeaddally, Wasseem & Raicu (2004)	Windows 2000, Solaris 8 & Redhat 7.3 (IPv4, IPv6)
Ahuja & Dendukuri (2006)	Windows 2000 and Linux (IPv4, IPv6)
Mohamed, Buhari & Saleem (2006)	Windows 2003, Redhat 9.0, FreeBSD 4.9 (IPv6)
Visoottiviset & Bureenok (2008)	Windows 2003, Red Hat 9.0, FreeBSD 5.3 (IPv4, IPv6)

Table 1: Research on Performance Evaluation of Operating Systems

The related work indicates that comparative performance analysis of Internet Protocols with different operating systems is imperative. In this undertaking, we extend the list of operating systems examined by other researchers to include Windows XP Professional and Windows Server 2003. At the time of this research, Windows operating systems had approximately 90% market share [13] and the two operating systems evaluated were the current desktop and server versions. We examine their performance with both versions of Internet Protocols. These protocols are discussed next.

### III. INTERNET PROTOCOL OVERVIEW

IPv6 is the communication protocol which is the successor of IPv4. It has been designed to offer improvements over the short comings of the IPv4. The biggest problem with IPv4 is it is running out of the IP addresses, and with the ever increasing number of computers worldwide, this creates a need for a much longer address space. IPv6 offers an answer by increasing the addresses from 32 to 128 bits (an increase from approximately  $4.3 \times 10^9$  binary to  $3.4 \times 10^{38}$  hexadecimal addresses).

IPv6 addresses other inadequacies of IPv4. Being plug-and-play, IPv6 allows novice users to connect nodes to network and be able to use devices instantly without any additional configuration. This is possible due to stateful or stateless auto-configuration features that enable devices to generate link-local address for network interface -this uses Neighbor Discovery (ND) protocol. Secondly, security features that were optional in IPv4 are mandatory in IPv6. Authentication Header (for authentication and integrity) and Encapsulated Security Payload (for integrity and confidentiality) implement security below application layer

in protocol stack and are used to provide deferring levels of security to users and applications. Thirdly, IPv6 has an added scope field to multicast addresses which gives improved multicast routing. Also included is a new type of address called anycast, which is used to send packets to any one of a group of nodes. Fourthly, end-to-end transparency is achieved, which eliminates the need for address translation technologies. And finally, Quality of Service (QoS) guarantees in accomplished in IPv6 by introduction of a flow field. This field identifies IP packets from a particular traffic stream and assigns different priorities to the packets prior to network transmission. QoS when introduced with Internet will provide support for real time services such as IP video and voice.

The challenge of converting from IPv4 to IPv6 is enormous. By design, IPv4 to IPv6 transition is evolutionary and there are a number of mechanisms that can be employed in current interim period [14]. These mechanisms allow for co-existence of both IPv4 and IPv6 networks (maintaining complete compatibility), and the techniques can be broadly grouped into three categories: dual stack, protocol translation, and tunneling:

A. Dual Stack: This is a straightforward transition mechanism where network nodes are stacked with both IPv6 and IPv4 protocols. These IPv6/IPv4 nodes use IPv6 stack to communicate with IPv6 nodes, and IPv4 for communicating with IPv4 networks. Although this solution is practical, all network devices (such as routers, switches, bridges, gateways, etc) end up being dual stacked as well. The additional burden makes network devices inefficient and does not reduce the demand for global routable IPv4 addresses.

B. Protocol Translation: This mechanism works for pure IPv6 network when communicating with nodes on an IPv4 network. Protocol translation techniques solves the problems that dual stack mechanisms exhibit, and depends on some intermediary device or service that convert packet headers at network border. Three protocol translation techniques are prevalent:

i. Dual Stack Application Level Gateway (DS-ALG): A dual stack IP device is used which can access both IPv4 and IPv6 services natively. This is a simple and robust solution, however requires all clients to be configured to use ALG and will work for specific ALG-supported applications only.

ii. Network Address Translator – Protocol Translator (NAT-PT): This is similar to traditional IPv4 NAT; however it also has protocol translation. This uses a pool of IPv4 addresses for assignment to IPv6 nodes on a dynamic basis as sessions are initiated across IPv4-IPv6 boundaries [15]. The IPv6 address binds to IPv4 address to give transparent routing for datagrams traversing

between address realms. Assignment of IPv4 addresses is done by the border router.

iii. Dual Stack Transition Mechanism (DSTM). A DSTM server provides a temporary IPv4 address to dual stack nodes on needs to have basis when IPv6 node wishes to correspond with a node on IPv4 network. The IPv6 node configures its IPv4 stack with the allocated address and tunnels the packets to the DSTM gateway at the edge of the network. The gateway has the responsibility to encapsulate or decapsulate packets as traffic transits through network border.

C. Tunneling: This provides a method to carry IPv6 traffic to other networks using current IPv4 routing infrastructure. IPv6 datagram are encapsulated within IPv4 packets and routed through IPv4 regions of the internetwork. The begin and the end points of the tunnel has to be IPv6/IPv4 dual stacked nodes. Router-to-router, host-to-router, router-to-host and host-to-host are the four possible configurations and there are several types of tunnels:

i. Configured Tunnel: IPv6 data is encapsulated within IPv4 packet at the originating end point and then transported through a tunnel. At the exit point of the tunnel, the packet is decapsulated. The tunnel endpoint addresses are determined from the configuration information that is stored at the encapsulating endpoint (hence named configured tunnel). Configured tunnels can be defined to go between router-to-router, host-to-router, host-to-host or router-to-host. Normally a tunnel broker (a server that manages tunnel requests from users) is employed with this technique.

ii. Automatic Tunnel: IPv6 packet is tunneled all the way to the final destination. There are no pre-configured tunnels and the nodes that perform automatic tunneling are assigned an IPv4-compatible IPv6 address. Since IPv6 packet's destination is the tunnel endpoint, IPv6 packet's destination address determines the tunnels endpoint. This technique is named automatic tunneling because the endpoint address is automatically determined from the embedded IPv4 address of the data's IPv6 address.

iii. 6 to 4: This allows communication from one IPv6 site to another over an IPv4 network without the need for an explicitly configured tunnel or requiring IPv4-compatible IPv6 addresses. There is no end node configuration and edge router configuration is minimal. When transporting data, node sends IPv6 packets to the destination and these packets are tunneled in IPv4 between the edge 6 to 4 routers of the two networks.

iv. 6 over 4: This is an automatic mechanism for establishment of a tunnel that is used to provide unicast and muticast connectivity between IPv6 nodes via an IPv4 network. A virtual link is created between the two

networks and IPv6 multicast addresses are mapped to IPv4 addresses to enable network discovery.

The above discussed techniques are all good, but only in transition period where there is coexistence of both Internet Protocols. Performance evaluation of above mentioned transition mechanisms have been undertaken by numerous researchers, including [16], [17], [18], and [19]. They have comparatively analyzed different traffic types to gauge performance with various mechanisms. In this research, we undertake to compare data transfer rate differences between pure IPv4 and IPv6 networks with Windows operating systems –experimental setup used for analysis is described next.

#### IV. EXPERIMENTAL SETUP

To measure performance of Internet Protocols on various Windows operating systems, two computers with identical hardware (CPU: Intel P4 3GHz, RAM: 1GB, Network card: PCI Intel Pro 100, Hard drive: Seagate Barracuda 7200 [20GB]) were connected point-to-point using a 100Mbps cross-over cable. Connecting test computers such a way eliminate effects that intermediary devices (switches, routers, etc) have on network performance. This testbed setup remained constant for all experiments conducted.

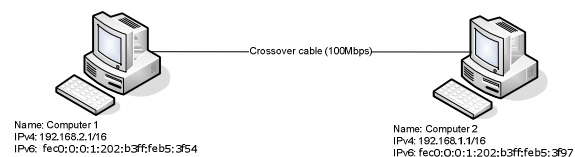


Figure 1: Network test bed

Initially, both computers were loaded with Microsoft Windows Server 2003 SP2. They were configured as application servers (i.e. not as part of a Domain) and no additional services such as DHCP or DNS were installed. Since Windows Server 2003 has by default both IPv4 and IPv6 protocol stacks, only one version was enabled and configured on both computers simultaneously. In the second set of experiments, operating systems were changed to Microsoft Windows XP Professional SP2. Windows XP natively has IPv4 only, therefore IPv6 had to be explicitly installed and enabled when required. For both operating systems, default configurations as per vendor instructions, were configured and no performance optimizations were tweaked. No additional third party software or vendor updates/patches were installed except service packs. The IPv4 and IPv6 parameters were statically configured.

IP Traffic 2.4.2 [20] was used as the primary tool to evaluate performance of protocols on operating systems. IP Traffic generates, receives, captures and replays TCP,

UDP and ICMP traffic end-to-end and measures performance over IP network. These three traffic types are used by mail exchanges, file transfer, ping program, and World Wide Web transmissions. Since IP Traffic uses Windows TCP/IP stack (Winsock2 interface) of the end computers, IP Traffic works independent of transmission link between the nodes. Test bed Computer 1 was configured as IP Generator (node that generates IP traffic) and Computer 2 was IP Answering (node that receives IP traffic). The Generator sends packets through the cable to the Answering, and the Answering echoes the packet back to the Generator. Payload sizes ranging from 128 Bytes to 1408 Bytes were used in the experiment. This range caters for most packet sizes on networks and the Internet [21].

To ensure high data accuracy, all tests were executed with multiple runs for sufficient duration. One million packets were sent 20 times for each payload size. The time and amount of transferred data for each run depends on the payload. For small payload (e.g. 128 Bytes), the transferred time was few seconds and approximate 25MB of data was sent. For larger payload size, such as 1028 Bytes, two minutes was taken to transfer the approximate 2GB of data from Generator and Answering. Standard deviation for data generated for each payload size was calculated and if the results fell outside 95% confidence interval, the experiment was rerun to rule out inconsistencies.

The metric used in the experiment is throughput (measured in Mbps). This metric gives a valuable insight into performance of a system since it is the rate at which data gets transmitted from one node to another over a network. Throughput values were measured for TCP and UDP traffic that was generated and sent between the two nodes. The results are presented and discussed next.

## V. EXPERIMENTAL RESULTS

We present the findings of this research in this section. Data packets were gradually increased in size for TCP and UDP traffic and resulting throughput values were plotted. This was done for both Windows XP and Windows Server 2003 operating systems.

Figure 2 shows TCP throughput results for operating systems with IPv4 and IPv6 protocols for different packet sizes. Also shown are the results for Windows 2000 experiments conducted by Zeadally [8] (indicated by \* in Figure 2). From TCP throughput values, it is observed that for all operating systems there are performance differences between IPv4 and IPv6 throughput. However, the variations in throughput are different for various operating systems.

For Windows Server 2003, IPv4 outperforms IPv6 for all packet sizes. IPv4 throughput range from 61.02 to 76.09, and IPv6 values are from 57.84 to 70.24 Bytes. On an average this indicates a performance difference of 6.2% between the two protocols (lowest difference 3.5% for

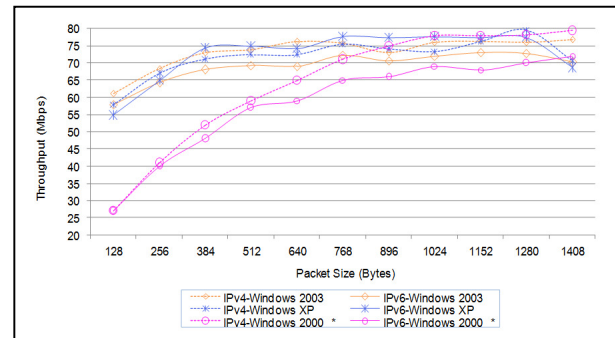


Figure 2: TCP Throughput of Windows Operating Systems

packet size 896 Bytes, and highest difference 10.4% for packet size 640 Bytes). For smaller packet sizes (up to 384 Bytes) there is a significant incline in throughput as the pack size increase (a gain averaging 18.5% for both the protocols). However for larger packet sizes, the graphs flatten off, indicating only slight variations in throughput for each protocol. For packet sizes larger than 512 Bytes, throughput means are 75.5 and 71.1 Mbps for IPv4 and IPv6 respectively.

Windows XP throughput exhibits slightly different behavior to Windows Server 2003. For packet sizes up to 256 Bytes, IPv4 outperforms IPv6 by 4.6%. Also, there is a significant incline in throughput (average gain: 17.2%) for both protocols for these packet sizes. For packets sizes between 256 and 1152 Bytes, IPv6 throughput values are higher than IPv4. Average throughput difference for packet sizes in this range is 3.6% (lowest difference 1.7% for packet size 1152 Bytes, and highest difference 5.7% for packet size 1024 Bytes). Also, the graphs flatten off in this range with mean throughput values of 73.6 and 76.2 Mbps for IPv4 and IPv6 respectively. For packet sizes larger than 1152 Bytes, IPv4 outperforms IPv6 by 2.8%. For these packet sizes, throughput performance declines by exactly 11.4% for both the protocols.

Although both Windows Server 2003 and Windows XP evidently show different throughput values for different protocols, there are some similarities in their performance. For small packet sizes, they both incline in performance with similar gains (18.5% and 17.2%); IPv4 outperforms IPv6 for both operating systems; and all throughput values are within +/-7 Mbps of each other. For packet sizes between 384 and 1152 Bytes, all throughput values are



within  $\pm 5$  Mbps, and throughput values for all increase on an average 4.3% ( $\pm 1\%$ ) with increased packet size.

On comparing these results with Zeadally's Windows 2000 throughput values, the following is evident; For packet sizes less than 640 Bytes, Windows Server 2003 and XP throughput values for both protocols are significantly higher than Windows 2000 (more than double for 128 Bytes packet size); IPv6 throughput values are comparatively higher for Windows Server 2003 and XP than Windows 2000; and IPv4 when implemented on Windows 2000 averages the highest throughput (78.4 Mbps).

UDP results obtained from the test bed for Windows operating systems with the two protocols are presented in Figure 3. Also plotted are Windows 2000 results from work of Zeadally [4] (indicated by \*). Like TCP results, UDP throughput values for all operating systems are different for IPv4 and IPv6. And the variations in throughput are also different for various operating systems.

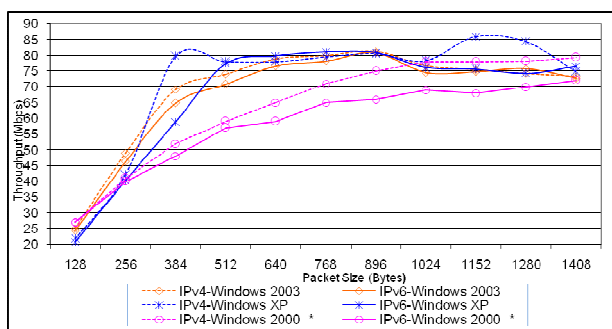


Figure 3: UDP Throughput of Windows Operating Systems

Windows Server 2003 throughput values are higher for IPv4 protocol stack than IPv6. For packet sizes smaller than 384 Bytes, IPv4 outperforms IPv6 by approximately 5%. In this range, there is a throughput gain of almost 170% for both protocols as the packet size increase. For packet sizes between 384 and 640 Bytes, IPv4 throughput values supersede by approximately 5% on IPv6. Also, throughput values for both protocols increase by approximately 16% as packet size increase. The difference in throughput narrows to approximately 1.5% for packet sizes in the range 640 to 896 Bytes. In this range there is throughput gain averaging 4% for both IPv4 and IPv6, peaking to maximum throughput of approximately 80 Mbps. As packet size increases from 896 to 1020 Bytes the throughput values for both protocols decrease; however IPv4 still outperforms IPv6 by 1.7%. For packet sizes larger than 1024 Bytes, throughput values for IPv4 and IPv6 stays constant averaging 75 Mbps. In this packet size range, IPv4 throughput values on average are 3%

Windows XP UDP throughput values show an interesting pattern. For small payload sizes (up to 256 Bytes), IPv4 outperforms IPv6 by 5%, however as the packet size increases to 512 Bytes, IPv4 performance becomes significantly higher than IPv6 – extending to a 35% difference). Throughput values are also substantially small for smaller packet sizes and increases four fold with increased payload. For mid-sized packets (512 to 896 Bytes), IPv6 shows better throughput than IPv4 by approximately 2%. This is the only instance in all experiments where the newer protocol is a better performer than the counterpart. For large packet sizes, IPv4 outperforms IPv6 by almost 12% - with throughput averages of 85 and 75 Mbps respectively for IPv4 and IPv6.

Both Windows Server 2003 and XP show some similarities for UDP throughput. For smaller payload, throughput values are comparatively low and increase in value with increase payload size. For average packet sizes (512 to 896 Bytes) all values band together with an average of 78.5 Mbps with 3% standard deviation. And for large payloads, all values except for Windows XP with IPv4, all throughput values are tightly grouped around 75 Mbps. For packet sizes less than 1000 Bytes, Windows 2000 throughput values are significantly lower than both Windows Server 2003 and XP. For larger payloads, all three operating systems with both the protocols band together in throughput values except Window 2000 with IPv6 – this shows the lowest throughput for most sizes of payloads.

The operating systems evaluated in this research all have at the core the same Window NT kernel architecture and yet portray different throughput performance. In the absence of availability of the operating system source codes, we can only speculate reasons for differences in performance of the protocols. Comparing Windows 2000 IPv4/IPv6 performance for both TCP and UDP protocols with the newer operating systems (Windows XP and 2003), it is observed that Windows 2000 throughput values are significantly lower than both the counterparts for smaller packet sizes. We believe this difference can be attributed to how efficiently an operating system handles TCP/IP packet buffer allocated by the kernel. Packet buffer is a shared memory space on a computer that holds data as it transits between devices. In [8] the authors compared Windows 2000 with UNIX platforms (Solaris and Linux) and concluded that kernel buffer allocation strategies are less efficient in Windows 2000 platform when compared to UNIX counterparts. We see a similar performance difference with Windows XP and 2003 banding together at a higher average throughput value than Windows 2000 for smaller packet sizes. Therefore we believe packet buffer handling strategies have improved

significantly in the newer Windows operating systems, however there may still be differences between Windows XP and Windows 2003. Microsoft has also incorporated changes in newer kernels to improve its overall kernel scalability [22]. This includes improvements in scheduling, memory management, kernel spinlocks, and heap. While discussing how each of these mentioned changes affects performance is beyond the scope of this paper, we believe they have contributed to overall performance enhancement of newer Windows operating systems.

## VI. CONCLUSIONS AND FUTURE WORK

In this paper, we conducted unbiased empirical performance comparison of IPv4 and IPv6 protocol stack on Windows operating systems. We evaluated Windows XP Professional & Server 2003 and compared results with previously studied Windows 2000. Using TCP and UDP traffic between two nodes, and throughput as the metric, we found that:

- For small packet sizes, performance difference between IPv4 and IPv6 is lower than theoretical value of approximately 10.3%. Both Windows XP and Server 2003 have throughput difference of approximately 5% evident for both TCP and UDP traffic.
- For large packet sizes, IPv4 and IPv6 performance difference is higher than the theoretical value of 1.3%. TCP traffic on Windows Server 2003 shows a difference of 10.4% and UDP traffic on Windows XP shows 12%.
- Although two operating systems may have the same kernel, network performance can be different. Comparing throughput values obtained in this research, we conclude that network throughput depends not only on traffic type and Internet Protocol version, but is also governed by the choice of the operating system.

We plan to extend this study by incorporating more operating systems and the range of metrics. Other areas that will be investigated are Internet Protocols with wireless media.

## REFERENCES

- [1] OECD report, Internet Address Space: Economic Considerations in the Management of IPv4 and in the Deployment of IPv6, OECD Ministerial Meeting, Seoul, Korea, 17-18 June 2008.
- [2] P. Bieringer, 'Status of IPv6 (Information and Workshop)', in the Proceedings of the Symposium on Applications and the Internet Workshop, 31 Dec-4 Jan 2005, pp 45-57.
- [3] R. P. Draves, A. Mankin, and B.D. Zill, 'Implementing IPv6 for Windows NT'. In the Proceedings of the 2nd USENIX Windows NT Symposium, Seattle, WA, USA, August 1998.
- [4] K. K. Ettikan, 'IPv6 Dual stack Transition Technique Performance Analysis: KAME on FreeBSD as the Case'. Faculty of Information Technology, Multimedia University, Jalan Multimedia, October 2000.
- [5] P. P. Xie, Network Protocol Performance Evaluation of IPv6 for Windows NT. Masters thesis, California Polytechnic State University, San Luis Obispo, June 1999.
- [6] M. Anand, Netperf3 TCP Network Performance on IPv6 using 2.4.17 Kernel. IBM Linux Technology Center, [www-124.ibm.com/developworks/opensource/linuxperf/netperf/results/may\\_02/netperf3\\_ipv6\\_2.4.17\\_results.htm](http://www-124.ibm.com/developworks/opensource/linuxperf/netperf/results/may_02/netperf3_ipv6_2.4.17_results.htm).
- [7] S. Ariga, K. Nagahashi, M. Minami, H. Esaki, J. Murai, 'Performance Evaluation of Data Transmission Using IPsec over IPv6 Networks', INET 2000 Proceedings, Japan, July 18th, 2000.
- [8] S. Zeadally, and I. Raicu, 'Evaluating IPv6 on Windows and Solaris', IEEE Internet Computing, Volume 7, Issue 3, May-June 2003, pp 51-57.
- [9] S. Zeadally, R. Wasseem, and I. Raicu, 'Comparison of End-system IPv6 Protocol Stack', in Proceedings of IEEE Internet Protocols, Technology and Applications (VolIP), Volume 151, Issue 3, 25 June 2004, pp 238 - 242.
- [10] S. P. Ahuja, and K. Dendukuri, 'An Empirical Evaluation of IPv6 in Linux and Windows Environments', Annual Review of Communications, Volume 59, November 2006, pp 361-370.
- [11] S. S. Mohamed, A. Y. M. Abusin, D. Chieng, 'Evaluation of IPv6 and Comparison Study with Different Operating Systems', in Proceedings of the Information Technology and Applications 2005 (ICITA 2005), Volume 2, July 2005, pp. 665-670.
- [12] V. Visootviseth, and N. Bureenok, 'Performance Comparison of ISATAP Implementations on FreeBSD, RedHat, and Windows 2003', in Proceedings of IEEE 22nd International Conference on Advanced Information Networking and Applications - Workshops (AINAW), Okinawa, 25-28 March 2008, pp. 547-542.
- [13] NetApplications: Top Operating System Market Share Trend for September, 2006 to August, 2007. Available: <http://marketshare.hitslink.com/report.aspx?qprid=5&qprdt=1&qpct=4&qptimeframe=M&qpsp=92ppn=12>, March 2008.
- [14] J. Govil, J. Govil, N. Kaur, and H. Kaur, 'An Examination of IPv4 and IPv6 Networks: Constraints and Various Transition Mechanisms', Proceedings of IEEE Southeastcon, 3-6 April 2008, pp 178-185.
- [15] G. Tsirtsis, P. Srisuresh, 'Network Address Translation - Protocol Translation (NAT-PT)', IETF RFC 2766, 2000.
- [16] I. Raicu, and S. Zeadally, 'Evaluating IPv4 to IPv6 Transition Mechanisms', in Proceedings of IEEE 10th International Conference on Telecommunications, Volume 2, 23 Feb-1 March 2003, pp 1091-1098.
- [17] J. Chen, Y. Chang, C. Lin, 'Performance investigation of IPv4/IPv6 transition mechanisms', in Proceedings of the 6th International Conference on Advanced Communication Technology, 2004, Volume. 2, pp. 545-550.
- [18] T. Sanguankotchakorn, and M. Somrobru, 'Performance Evaluation of IPv6/IPv4 Deployment over Dedicated Data Links', in Proceedings of the 5th International Conference on Information Communications and Signal Processing, 6-9 December 2005, pp 244-248.
- [19] M. Shin, H. Kim, D. Santay, and D. Montgomery, 'An empirical analysis of IPv6 transition mechanisms', in Proceedings of the 8th International Conference on Advanced Communication Technology (ICACT), Volume 3, 20-23 Feb. 2006, pp 1991-1996.
- [20] ZTI Telecom, IP Traffic - test & measure. Available: <http://www.zti-telecom.com>.
- [21] K. Thompson, G. J. Miller, and R. Wilder, 'Wide-area Internet traffic patterns and characteristics', IEEE Network, Nov - Dec 1997, Volume 11, Issue 6, pp 10-23.
- [22] Microsoft Corporation, Windows Server 2003 Kernel Scaling Improvements, October 2003. Available: <http://www.microsoft.com/windowserver2003/evaluation/performance/scaling.mspx>.



# A New Proposal for Choosing a Deflection Link in a Deflection Network

M. Atiqur Rahman, M. Tanvir Hossain, and Shahed Anwar

Department of Computer Science & Engineering, University of Dhaka, Dhaka - 1000, Bangladesh

**Abstract** — In high-speed networks with greater connectivity, packets can be deflected through a sub-optimal path to reach the destination if the desired path is congested, failed or the buffer is full. However, in a combined store-and-forward and deflection network, packets are deflected when there is congestion in the network. For choosing a link for deflection in such a combined network, currently existing strategy considers only local information to the node in need to deflect a packet. In this paper, a new proposal has been introduced to select an outgoing link for deflection. Under the proposed strategy, packets are guaranteed to traverse through the best cost path available to the destination when they are deflected. Simulation shows better results than that of the existing strategy in terms of packet delivery, discard and average cost per delivered packet.

**Index Terms** — Computer Networks, Routing, Deflection Routing.

## I. INTRODUCTION

Currently computer networks supporting multi-gigabit transmission rates are feasible by using high-speed optical networks. Optical buffering, wave length conversion and deflection routing are technologies that can be used in optical networks for switching purposes to resolve contention [1]. However, as current optical buffers are typically limited to providing a few tens of micro seconds delay; it is not easy to store longer optical bursts. Compared to optical buffering and wavelength conversion, it is relatively easy to optically switching burst data [2]. Accordingly, another technique to control traffic flow and resolve contention in burst optical networks becomes attractive – “deflection routing”.

Because of high connectivity of the routers in current computer networks, there is generally more than one path that can be followed to reach a destination from the same source. With the high transmission rate, the degradation of performance can be kept within a tolerable range, if a packet is deviated from its optimal or shortest path even for real-time applications. In the case of store-and-forward routing, it is natural that, more packets can be stored using larger buffers. However, larger buffers can force packets to wait for longer period of time at each node. These packets could reach the destination through a suboptimal path instead of waiting in the queue for the preferred link. In the worst case, packets could be discarded when the queue or the buffers for the preferred link is full. These

packets could be given a second chance to follow an alternate route to the destination to improve the overall packet delivery or throughput [3]. This type of routing where a packet is intentionally deviated from its optimal path to the destination to a different path that can also lead the packet to its desired destination is known as Deflection routing. Work has been done in [3] to combine the store-and-forward approach of routing with the Deflection approach where nodes are allowed to have limited buffers to work just like store-and-forward routing when the network is in normal condition. But in case of congestion or failure of the preferred link, the combined routing strategy selects another link to deflect the packet exactly in the way the Deflection routing protocols work. But the choice for a deflection link proposed in [3] is not that much efficient in terms of the number of packets delivered, average cost per delivered packet etc. This paper has concentrated on this area of choosing a better link for deflection.

The rest of the paper is organized as follows. Section II explores related works on deflection routing and store-and-forward routing in detail. This section also includes a mixed approach for these two routing styles. In Section III, a new proposal to select a link for deflection in a combined store-and-forward and deflection network is given that results in better performance than that of what proposed in [3]. Section IV is dedicated to simulation and comparative study of the proposed strategy with the existing one. And finally, section V concludes the paper.

## II. EMPIRICAL STUDY

### A. Store-and-Forward Network

In a store-and-forward network, input and output queues (buffers) are attached to each link  $L$ . These buffers are used to temporarily store the incoming packets through link  $L$  and outgoing packets that will be forwarded on the link  $L$ , respectively. Since these buffers are of finite size, the output buffer can become full if the packet incoming rate for the link  $L$  is greater than the packet outgoing rate. Under pure store-and-forward routing, new packets are generally discarded when the output buffer becomes full.

### B. Deflection Network

In a deflection network, if two or more packets choose the same outgoing link  $L$  simultaneously, one packet is forwarded on link  $L$  and the other packet(s) are forwarded on alternate links, if possible. It can be assumed that, if some of the load is distributed among the idle links, the packets will not follow the optimal route, but the overall performance and number of delivered packets will be improved. Deflection is suitable for optical networks with no or limited buffering because, it is difficult to store optical messages [5].

The idea of buffer-less Deflection routing was first suggested by Baran [7]. The focus of Baran's investigation concerned the reliability of a communications network. He suggested that keeping packets "moving" regardless of the optimality of the path was a desirable property of a survivable network.

Lawrie and Padua [8] analyzed the performance of routing with limited buffers in the folded shuffle-exchange network. In this analysis, packets were not routed along shortest paths; a slightly longer path was sometimes chosen to simplify implementation.

### C. Mixed Approach

Deflection routing can be combined with store-and-forward routing in order to achieve certain performance gains from the following perspectives.

- 1) Retransmission and More Loads: Networks having too much buffer space is more harmful than having too little, since the packets or their retransmissions have to be dropped after they have consumed precious network resources [6].
- 2) Real-time application: For real-time applications long buffers should be avoided to satisfy the stringent delay requirements. For example, voice, video-conferencing, movie etc. These types of applications must satisfy some maximum delay [4].
- 3) No or Limited Buffer: Limited buffers can speed up the switching significantly in case of optical networks, so the process can follow the speed of the links as closely as possible.
- 4) Link Failure: In deflection routing, the packets for the failed link can be forwarded through other active idle links to balance the load of the failed link.
- 5) A hop-count bounded adaptive hybrid routing strategy that combines the store-and-forward approach with the deflection approach is proposed in [3].

When the outgoing queue of the link  $L$  is full and a new packet arrives that chooses link  $L$  then in pure store-and-forward routing, the packet is discarded generally. But, in [3] the author gives a proposal which, instead of discarding the packet, gives it a second chance to go to the network again and append it to the outgoing queue of a link chosen for deflection of the packet, if possible. The

author proposes that the datagram is allowed to be deflected no more than  $N$ -times, where  $N$  is the max. hop count value.

In [3], for choosing the deflection link, a scanning is performed among all the idle links and the first underloaded link found was used as the deflection link. When the output queue of the preferred link overflows, the output queue lengths of the available links are checked to find a deflection link. The first such link  $L$  is used as the deflection link. On the other hand, if there is no underloaded link, the packet is discarded.

## III. A NEW PROPOSAL FOR SELECTING A DEFLECTION LINK

Selection of a deflection link in a combined store-and-forward and deflection network is crucial, since it affects the performance of the overall routing strategy. In [3], the load of a link was used to reflect link utilization. As the estimation of the load, the current queue length of the link was used to keep the proposed strategy simple. For choosing a deflection link, [3] considered only information local to a node. The first found link with an underloaded outgoing queue was chosen as a deflection link regardless of the cost of the path the selected deflection link would direct the packet to. It had a potential problem of not forwarding the packet through the best cost path among all the available ones. Thus the probability of following the best cost path via the selected deflection link was,  $1/(k-1)$ , where  $k$  is the connectivity degree of the network.

To raise this probability to exactly 1, while choosing the deflection link, we have considered all the available links having an underloaded outgoing queue. It then selects the link as the deflection link that forwards the packet to the neighbor node with the minimum cost path to the destination of the packet. This definitely improves the performance of the network in terms of the number of packets delivered as well as the average cost per delivered packet. Here is how the proposed strategy selects the best cost path for deflection.

This strategy assumes that the routing tables of the nodes are updated by exchanging information with neighbors or adjacent nodes using Distance Vector Routing (DVR). Usually the routing table of each node contains an entry for each of the destinations in an one dimensional array recording the next hop to the destination with the associated cost to reach that destination as a set of triples of the form as follows -

(Destination, Cost, NextHop)

For example, Table I shows the complete routing table maintained at node B for the network in Fig. 1.

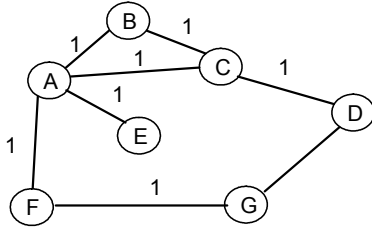


Fig. 1. An example network topology

TABLE I  
ROUTING TABLE MAINTAINED AT NODE B

Destination	Cost	NextHop
A	1	A
C	1	C
D	2	C
E	2	A
F	2	A
G	3	A

Instead of storing information only about the destination nodes, in our proposed strategy, the routing table of a node also stores the routing tables advertised by its directly connected nodes; i.e. its neighbors in a two-dimensional array as follows.

TABLE II  
ROUTING TABLE MAINTAINED AT NODE B ACCORDING TO PROPOSED APPROACH

Destination	Cost	NextHop	Routing Table
A	1	A	Routing table of A
C	1	C	Routing table of C
D	2	C	NULL
E	2	A	NULL
F	2	A	NULL
G	3	A	NULL

For a node, the routing tables of its neighboring nodes can be inserted into its routing table using the same DVR algorithm without increasing the time complexity of the algorithm but with an increase in space complexity. While nodes exchange their routing tables with their neighbors, each node, in addition to computing the costs to the destinations, also stores the advertised routing tables for future reference.

Whenever a node needs to deflect a packet, it scans its entire outgoing queues to find underloaded ones. Whenever such a queue is found, it finds out the link for which this outgoing queue is maintained. It then accesses the routing table of the node to which this node is connected through the link by accessing the appropriate

column of the routing table of the current node. From the accessed routing table, the cost of the best path to the destination of the packet to be deflected is found and stored in a variable called *min\_cost*. At the same time, another variable called *def\_link* is updated to point to the link corresponding to the cost stored in *min\_cost*. Initially *min\_cost* is assigned a very large value and *def\_link* is assigned to NULL. When all the outgoing queues are finished with searching, the current node decrements the deflection count of the packet by 1 and then deflects the packet through the link pointed to by the variable *def\_link*, if it is not NULL. Otherwise, it discards the packet.

In this strategy, the algorithm for routing a packet currently staying at node *i* is as follows –

1. Consult routing table of node *i*
2. Let  $L_{i1}$  be the preferred link to the next hop on the best route to destination.
3. If the queue for  $L_{i1}$  is full then
4. Set *def\_link* = NULL
5. Set *min\_cost* = MAX
6. For each outgoing queue  $q_i$  of node *i*
7. if  $q_i$  is underloaded
8. Consult routing table of *i* and find the cost of the path from the neighbor node for  $q_i$  to the destination node of the packet
9. if the received cost  $C_i < \text{min\_cost}$  then
10. Set *def\_link* = the outgoing link for queue  $q_i$  and *min\_cost* =  $C_i$
11. End of If structure
12. End of If structure
13. End of For
14. If *def\_link*  $\neq$  NULL then
15. Put the packet in the incoming queue of the node connected to node *i* through *def\_link*
16. Else
17. Discard the packet.
18. End of if structure
19. Else
20. Put the packet in the queue of link  $L_{i1}$
21. End of if structure

**Complexity Study of the Proposed Strategy:** A node in need of finding a deflection link, explores all its *k* outgoing links which takes  $O(k)$  time, where *k* is average degree of connectivity. Since  $k \ll n$ , where *n* is the total number of nodes, we can say that the proposed strategy calculates the deflection link in constant time. Again, since it requires for each routing table of a node to store the routing tables of its neighbors, the space complexity of this approach is for storing the routing table is  $O(n^2)$  instead of  $O(n)$ .

#### IV. SIMULATION

In this chapter simulation technique, comparisons and results are discussed. Simulations have been performed for datagram service network based on the approach proposed in [3] and the newly proposed strategy using a program written in Java. Simulation results are observed in terms of number of packets successfully delivered and the total cost.

##### A. Major Parameters Observed

Each Router has queues on its entire links. Queue Capacity specifies the capacity of these queues. In a particular test case all the queues have equal capacities. Maximum Bandwidth specifies the maximum bandwidth of the links in a network. During simulation all links are given a bandwidth value within the range 2 to max. bandwidth. Packets/Timeslot means how many packets are generated on each of the routers in every timeslot. Maximum Deflections Allowed indicates the maximum number of times a packet is allowed to be deflected. Total Packets Generated indicates the total number of packets generated in a particular test case. Average Packet Delivery indicates how many packets on an average could be successfully delivered in a given test case. Average Packet Discards indicates how many packets on an average were discarded in a given test case. All Queues Full (AQF) indicates how many packets were on an average discarded because of the case that all of the outgoing queues in a Router were full. Maximum Deflection Reached (MDR) indicates how many packets were, on an average, discarded due to exceeding the maximum allowed deflections. Each link on the network is assigned a cost randomly. When a packet travels through a link, it adds up the cost of that link to its own cost overhead metric. In this paper, Average Cost specifies the total cost overhead of all of the packets incurred during a particular test case divided by the number of delivered packets.

In all the test cases the following parameters were kept constant:

Total Number of Routers in the Simulated Network: 20;  
Maximum Cost of a Link: 20; Number of Timeslots Simulated in a Test Case: 2000

##### B. Simulation Technique

The simulation technique assumes time as discrete, non-overlapping and consecutive time slots of same duration. Each time slot is considered as unit time. The simulation is performed for the time period T between two consecutive routing updates. It starts after collecting all the required routing information from the other nodes of the network and assumes that, during simulation routing tables do not change. It first computes the routing tables that contain the cost of paths, next hop on that path. Then, it is executed for certain number of timeslots. In each time slot, constant load (fixed number of packets) is generated at each router and the packets waiting in the queues of the nodes advance one hop to their destination.

Result tables are presented below, comparing the outcomes obtained after applying the previous as well as the proposed approach respectively under different test conditions. First of all we have used different Queue Capacities and measured the Average Packet Delivery and Costs, both for the old and proposed approach in Table III. In Table IV you will see the results found by varying the Maximum Bandwidth value. Finally, Table V shows the results found by altering the network load. Each table is also followed by two charts depicting the difference between the two approaches. The first chart shows how better the proposed approach performed in terms of Average Packet Delivery and the second chart shows how much cost savvy the proposed approach is compared to the old approach. In almost all cases the proposed approach has resulted in far better results.

##### Case 1: Varying the queue capacity

In this test case the following parameters were kept constant and only the Queue Capacity was varied:

Maximum Bandwidth = 10; Packets/Timeslot = 20;  
Maximum Deflections Allowed = 5; Total Packets Generated = 8,00,000

TABLE III  
VARYING THE QUEUE CAPACITY

	Queue Capacity	Avg. Packets Delivered	Delivery in Percentage	Avg. Packets Discarded	Discards in Percentage	AQF	MDR	Avg. Cost
Old Approach	1	51157	6.39	748661	93.58	748661	0	139.29
	5	216341	27.04	582801	72.85	555601	27200	137.86
	10	296573	37.07	501633	62.7	412311	89322	123.54
	20	295143	36.89	500993	62.62	393147	107846	139.43
Proposed Approach	1	85633	10.7	714167	89.27	714167	0	85.33
	5	375518	46.94	423554	52.94	383992	39562	75.89
	10	411822	51.48	386400	48.3	303249	83151	80.6
	20	462911	57.86	333351	41.67	225102	108249	74.94

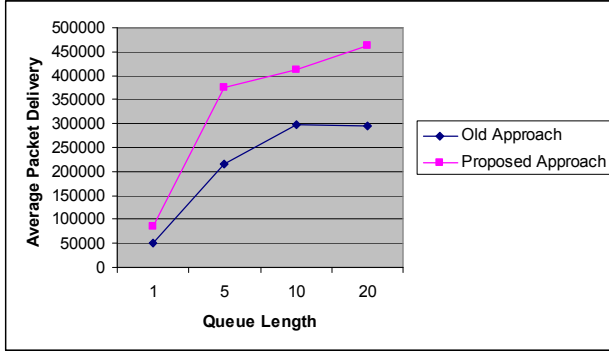


Fig. 2. Impact of varying queue capacity on average packet delivery

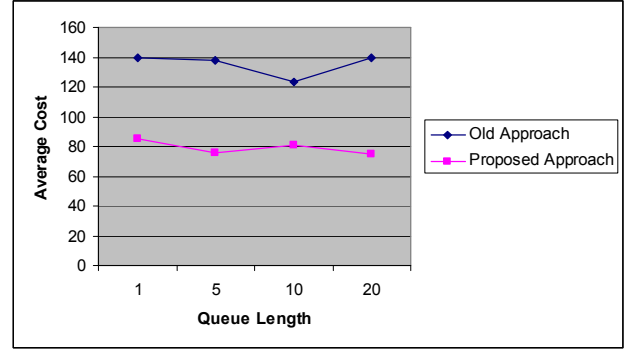


Fig. 3. Impact of varying queue capacity on average cost

#### Case 2: Varying the bandwidth

In this test case the following parameters were kept constant while the Maximum Bandwidth of the links was varied:

Queue Length = 5; Maximum Bandwidth = 10;  
Maximum Deflections Allowed = 5

For datagram service, the proposed strategy performs better in terms of number of packets delivered and total

TABLE IV  
VARYING THE MAXIMUM BANDWIDTH

	Max. Bandwidth	Avg. Packets Delivered	Delivery in Percentage	Avg. Packets Discarded	Discards in Percentage	AQF	MDR	Avg. Cost
Old Approach	2	118913	14.86	680077	85.01	680013	64	134.49
	10	216341	27.04	582801	72.85	555601	27200	137.86
	16	267263	33.41	531713	66.46	458298	73415	128.74
Proposed Approach	2	176567	22.07	622476	77.81	622476	0	78.57
	10	375518	46.94	423554	52.94	383992	39562	75.89
	16	429500	53.69	369399	46.17	299670	69729	74.29

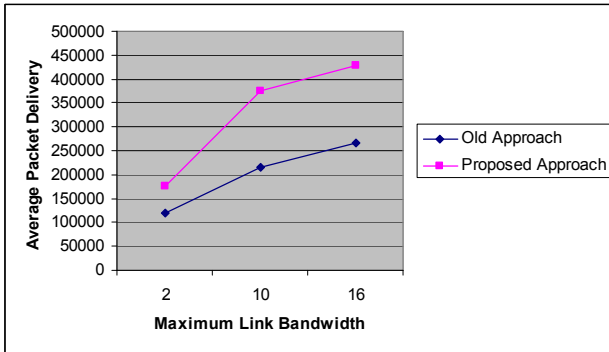


Fig. 4. Impact of varying bandwidth on average packet delivery

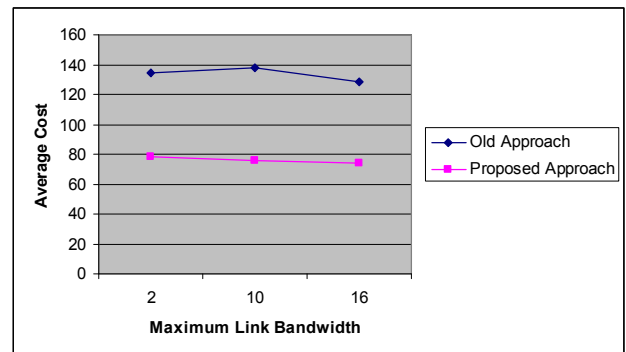


Fig. 5. Impact of varying bandwidth on average cost

Queue Length = 5; Packets/Timeslot = 20; Maximum Deflections Allowed = 5; Total Packets Generated = 8,00,000

#### Case 3: Varying the load

In this test case the following parameters were kept constant while the Number of Packets Generated per timeslot was varied:

path cost traversed by packets. From the simulations performed above it is understood that the proposed approach gives better results if there is more traffic in the network. Also as in the proposed approach, packets get the opportunity to go through the next possible sub-optimal path, in case an optimal (or sub-optimal) path is engaged, the average cost incurred by all the packets are always less than the previous approach.

TABLE V  
VARYING THE MAXIMUM BANDWIDTH

	Packets/ Timeslot	Total Packets Generated	Avg. Packets Delivered	Delivery in Percentage	Avg. Packets Discarded	Discards in Percentage	AQF	MDR	Avg. Cost
<b>Old Approach</b>	5	200000	188191	94.1	11458	5.73	0	11458	30.15
	10	400000	227763	56.94	171378	42.85	51872	119506	115.23
	20	800000	216341	27.04	582801	72.85	555601	27200	137.86
	30	1200000	256365	21.36	942621	78.55	940141	2479	125.22
<b>Proposed Approach</b>	5	200000	183184	91.6	16468	8.23	0	16468	26.61
	10	400000	296995	74.25	102234	25.56	12431	89803	63.21
	20	800000	375518	46.94	423554	52.94	383992	39562	75.89
	30	1200000	385293	32.11	813733	67.81	809117	4616	78.2

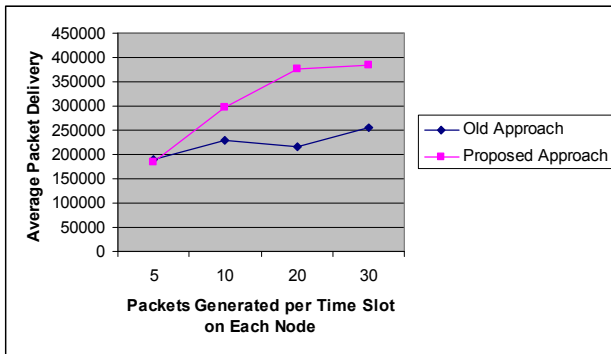


Fig. 6. Impact of varying network load on average packet delivery

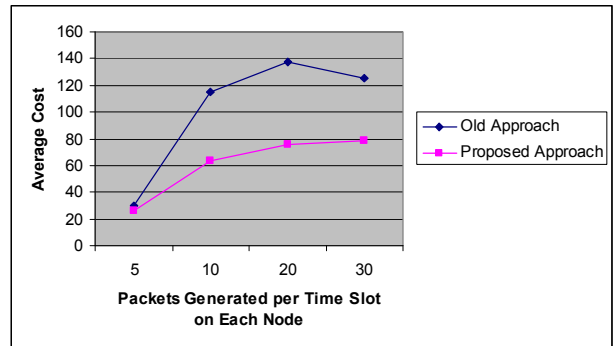


Fig. 7. Impact of varying network load on average cost

## V. CONCLUDING REMARKS

In order to select the deflection link, further trick can be applied. Even after choosing the deflection link according to the strategy proposed in this paper it may appear that, the next link has all its outgoing links congested. If it is possible to be aware of this situation well ahead of time, the performance of the routing can be improved even better.

## REFERENCES

- [1] X. Wang, H. Morikawa and T. Aoyama, "Deflection routing protocol for burst switching WDM mesh networks," School of Engineering, The University of Tokyo, JAPAN, 2003.
- [2] T. Kato, J. Sasaki, T. Shimoda, H. Hatakeyama, T. Tamanuki, M. Yamaguchi, M. Kitamura and M. Itoh, "10Gb/s Photonic Cell Switching with Hybrid 4x4 Optical Matrix Switch Module on Silica based Planar Waveguide Platform," in Proc. OFC'98, post deadline paper, San Jose, 1998.
- [3] M. S. Hasan, "A New Hop-count Bounded Adaptive Hybrid Routing Strategy for Datagram and Virtual Circuit Service," Department of Computer Science, University of Dhaka, Thesis for the Degree of Masters in Science, March 2001.
- [4] C. Baransel, W. Doboseiwicz and P. Gburzynski, "Routing in Multihop Packet switching Network: Gb/s Challenge," IEEE Network, Vol. 9, No. 7, May-June, 1995.
- [5] C. Busch, M. Herlihy and R. Wattenhofer, "Randomized Greedy Hot-Potato Routing," Computer Science Department, Brown University, Providence, RI 02912, 2000.
- [6] R. Jain, "Congestion Control in Computer Networks: Issues and Trends," IEEE Networks Magazine, 4, 3, 24-30, May 1990.
- [7] P. Baran., "On distributed Communications Networks," IEEE Transactions on Communications, 12:19, 1964.
- [8] D. H. Lawrie and D. A. Padua, "Analysis of Message Switching with Shuffle-Exchanges in Multiprocessors," Proceedings of Workshop on Interconnection Networks for Parallel and Distributed Processing, pp. 116-123, ACM, April, 1980.

# Design and Construction of Direct Sequence Spread Spectrum CDMA Transmitter and Receiver

M. Habib Ullah, Akhmad Unggul Priantoro, M. Jasim Uddin,  
Department of Electrical and Computer Engineering  
International Islamic University Malaysia  
habib\_ctg@yahoo.com

**Abstract** - Spread Spectrum Communication techniques have been widely accepted in mobile and wireless communications. They have very beneficial and tempting features, like Antijam, security, and multiple accesses. It is the purpose of this paper to describe the features of Spread Spectrum systems. The emphasis will be on the Direct Sequence Spread Spectrum (DS-SS) scheme, Pseudo Noise signals (PN), Modulators and Demodulators, Hardware implementation and illustrate some of the DS-SS system features.

**Index Terms** - CDMA, PN Code, Modulation, Demodulation, DS-SS.

## I. INTRODUCTION

Spread Spectrum Communications is one of the widely used data communication schemes nowadays. It has many features that make it suitable for secure, multiple accesses, and many other properties that are needed in a communication system. Spread Spectrum is defined as: Spread Spectrum is a means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information. The band spread is accomplished by means of a code, which is independent of the data, and synchronized reception with the code at the receiver is used for despreading, and subsequently data recovery. Basically spread spectrum is a coding technique for digital transmission. It was originally developed for the military under a veil of secrecy. The purpose of coding is to transform an information signal so that it looks more like noise. In the receiver, the incoming signal is decoded, and the decoding operation provides resistance to interference and multi path fading. The spreading or dilution of energy in spread spectrum systems over a wide bandwidth results in several possible advantages, short-range interferences-free overlays on their emissions, and resistance to interference, from other emissions, and low detestability. The low spectral density needed for spread spectrum communication systems, as well as ability of some of these systems

to process signals that are carried far into the noise, offer a potential for shared spectrum use with existing Systems on a non-interference basis. Finally, Spread spectrum systems could be useful in applications to control multi path interference [1].

## II. CDMA OVERVIEW

Spread Spectrum communication techniques provide a new method of multiple access, known as CDMA (Code Division Multiple Access) that is proposed as an interesting alternative with respect to the traditional methods, i.e.: the well-known TDMA (Time Division Multiple Access) and FDMA (Frequency Division Multiple Access) [2]. The main parameter in spread spectrum systems is the processing gain the ratio of transmission and bandwidth has shown in equation 1.

$$G_p = \frac{BW_t}{BW_i} \quad (1)$$

The processing gain determines the number of users that can be allowed in a system, the amount of multi-path effect reduction, the difficulty to jam or detect a signal etc. For spread spectrum systems it is advantageous to have a processing gain as high as possible. A PNcode is a sequence of chips valued -1 and 1 (polar) or 0 and 1 (non-polar) and has noise like properties. This results in low cross-correlation values among the codes and the difficulty to jam or detect a data message. Several families of binary PNcode exist, they are addressed in another. A usual way to create a PNcode is by means of at least one shift-register. When the length of such a shift-register is n, the following equation 2 can be said about the period of the above mentioned code-families:

$$N_{DS} = 2^n - 1 \quad (2)$$

In direct-sequence systems the length of the code is the same as the spreading-factor with the consequence has shown in equation 3:

$$G_p(DS) = N_{DS} \quad (3)$$

This can also be seen from figure 1, where it shows how the PN code is combined with the data signal ( $N_{DS}=7$ ). The bandwidth of the data signal is now multiplied by a factor  $N_{DS}$ . The power contents however stay the same, with the result that the power spectral density lowers.

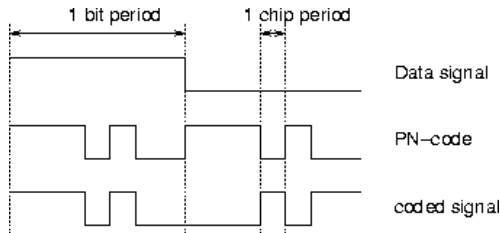
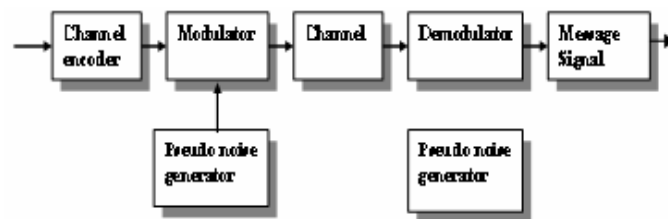


Figure 1: direct-sequence spreading

The generation of PN codes is a number of shift-registers are all that is required. For this reason it is easy to introduce a large processing-gain in Direct-Sequence systems. The main problem with applying Direct Sequence spreading is the so-called Near-Far effect. This effect is present when an interfering transmitter is much closer to the receiver than the intended transmitter. Although the cross-correlation between codes A and B is low, the correlation between the received signal from the interfering transmitter and code A can be higher than the correlation between the received signal from the intended transmitter and code.

#### A. General model of a spread spectrum digital communication

The input is fed into a channel encoder that produces an analog signal with a relatively narrow bandwidth around some center frequency. This signal is further modulated using the sequence of digits known as spreading code or spreading sequence. Typically, but not always, the spreading code is generated by pseudo noise or pseudo random number, generator. Figure 2 shows General model of a spread spectrum digital communication [1].



I/O  
DATA

Figure 2: General model of a spread spectrum digital communication

The effect of this modulation is to significantly increase the bandwidth (spread the spectrum) of the signal to be transmitted. On the receiving end, the same digit sequence is used to demodulate the spread spectrum signal. Finally, the signal is fed in to a channel decoder to recover the data. Spread spectrum uses wide band, noise-like signals. Because spread spectrum signals are noisy they are hard to detect. Spread spectrum signals are also hard to intercept or demodulate. Further spread spectrum signals are harder to jam (interfere with) than narrow band signals due to the increased bandwidth. To qualify as a spread spectrum signal, two criteria should be met:

1. The transmitted signal bandwidth is much greater than the information bandwidth.
2. Some function other than information being transmitted is employed to determine the resultant transmitted bandwidth.

#### B. Pseudo-Random Noise Codes:

A PNcode used for DS-spreading exists of  $N_{DS}$  units called chips, these chips can have 2 values: -1/1 (polar) or 0/1. As we combine every data symbol with a complete PNcode, the DS processing gain is equal to the code-length.

The sequences must be building from 2-leveled numbers. The codes must have a sharp (1-chip wide) autocorrelation peak to enable code synchronization. The codes must have a low cross-correlation value, the lower this cross correlation, the more users we can allow in the system. This holds for both full-code correlation and partial-code correlation. The latter because in most situations there will not be a full-period correlation of two codes, it is more likely that codes will only correlate partially (due to random-access nature). The codes should be "balanced": the difference between ones and zeros in the code may only be 1. This last requirement stands for good spectral density properties (equally spreading the energy over the whole frequency-band).

#### C. Direct Sequence-spread Spectrum Transmitter and Receiver

In Direct Sequence-Spread Spectrum the baseband waveform is multiplied by the PN sequence. The PN is produced using a PN generator. Frequency of the PN is higher than the Data signal. This generator consists of a shift register, and a logic



circuit that determines the PN signal. After spreading, the signal is modulated and transmitted. The most widely modulation scheme is BPSK (Binary Phase Shift Keying) [10]. Figure 3 shows the PN Generator Block Diagram.

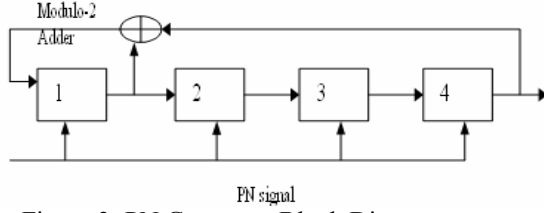


Figure 3: PN Generator Block Diagram

The equation 4 represents this DS-SS signal.

$$S_n = \sqrt{\frac{2E_s}{T_s}} m(t) p(t) \cos(2\pi f_c t + \theta) \quad (4)$$

where  $m(t)$  is the data sequence,  $p(t)$  is the PN spreading sequence,  $f_c$  is the carrier frequency, and  $\theta$  is the carrier phase angle at  $t=0$ . Each symbol in  $m(t)$  represents a data symbol and has a duration of  $T_s$ . Each pulse in  $p(t)$  represents a chip, and has a duration of  $T_c$ . The transitions of the data symbols and chips coincide such that the ratio  $T_s$  to  $T_c$  is an integer. The waveforms  $m(t)$  and  $p(t)$  are shown in fig. 5. Here we notice the higher frequency of the spreading signal  $p(t)$ . The resulting spread signal is then modulated using the BPSK scheme. The carrier frequency  $f_c$  should have a frequency at least 5 times the chip frequency  $p(t)$ . In the demodulator section, there has been simply done reverse the process. We Demodulate the BPSK signal first, Low Pass Filter the signal, and then despread the filtered signal, to obtain the original message. The process has shown in equation 2.1. In the demodulator section, we simply reverse the process. Demodulate the BPSK signal first, Low Pass Filter the signal, and then despread the filtered signal, to obtain the original message. The process has been described in equation 5 [3].

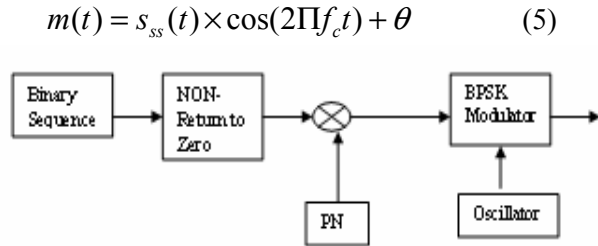


Figure 4: DS-SS Transmitter Block Diagram

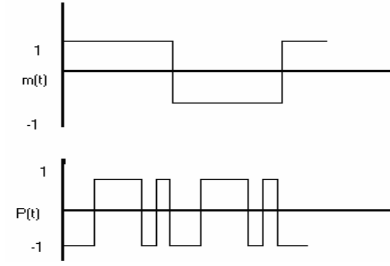


Figure 5:  $m(t)$  and  $p(t)$  for a Maximum Length PN

$$m(t) = \sqrt{\frac{2E_i}{T_i}} m(t) p(t) \frac{1}{2} [1 + \cos(4\pi f_c t + 2\theta)] \quad (6)$$

As shown in equation 5 and 6 when we multiply two cosine signals together, we will obtain two expressions, one of which has twice the frequency of the original message. And this part can be removed by a LPF. The output is  $m_{ss}(t)$  as shown in figure 5. This design is based on Coherent Detection BPSK, so we don't have to worry about carrier synchronization issues. Basically we have tried to construct the modulated part in our project. But for receiving data we have designed the Demodulator circuit by the software of 'Electronic Workbench'. As for the PN sequence in the receiver, we mentioned earlier that it should be an exact replica of the one used in the transmitter, with no delays, because this might cause severe errors in the incoming messages. Again, my design is based on the idea that PN sequences are matched, and actually we are going to use the same generator for both to ease the design. There are various techniques that deals with PN delay problems and mismatches, but we are not going to encounter any in this design [4]. After the signal gets multiplied with the PN sequence, the signal despreads, and we obtain the original bit signal  $m(t)$ , that was transmitted. The block diagram of the receiver is shown in figure 6 [5].

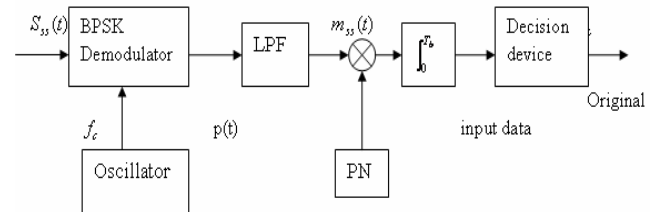


Figure 6: DS-SS receiver

This simple straightforward description of DS-SS systems, will allow us to design the Modulator

circuits with some ease. We are going to take advantage of the block diagrams for each one of them.

#### D. The Direct sequence spread spectrum Modulator Design:

In this chapter, we are going to illustrate the circuit design and show the simulation of the DS-SS Modulator (Transmitter). Simulations were conducted using the EWB (Electronics workbench) software and also Hardware. Ideal components were used for the ease of the design, and to try to avoid other factors, that might directly affect the performance of the system. And since we are going to design a simple circuit for educational purposes, we decided to use the ideal components offered inside the software although I constructed it in hardware[9].

#### E. The PN Generator Design

The design of the PN generator is based on the block diagram shown in figure 7. Here we are going to need 4 D-Flip-Flops (FFs), a XNOR gate, and a Clock source. Since we are using 4 FFs,  $m=4$  and the PN signal will repeat every 15 clock cycles. We have chosen the CLK frequency,  $f_{CLK}=100$  KHz. The period of the PN chip is  $TPN_{chip} = 1/100k = 10^{-5}$  seconds. The total period of the PN sequence  $TPN = 15 * TPN_{chip} = 15 * 10^{-5}$  seconds. The period of the Binary Input signal (per bit) is going to be  $T_b = TPN = 15 * 10^{-5}$  seconds, that is a frequency  $f_b = 6.6667$  KHz. The Clock signal was taken from an external source (a Function Generator) [6, 7]. We can design a clock using an oscillator. But we are not going to go into that. The Circuit that resembles this PN is now easily understood. It is shown in figure 7. In our project we used only PN without data. We can let PN is the data. From this design, the output level from the PN is not 1,-1 as was indicated in the signal diagram of  $p(t)$ . The output PN sequence is shown in figure 8.

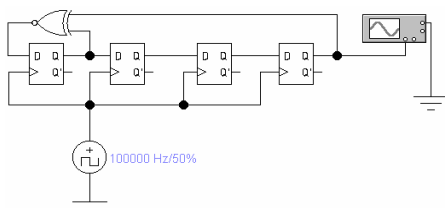


Figure 7: The PN Generator Circuit Diagram.

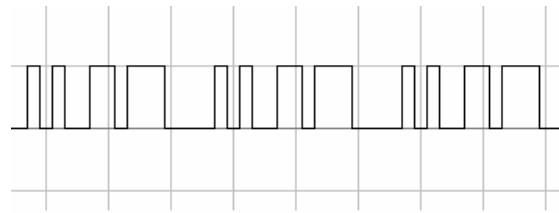


Figure 8: The PN Signal

#### F. The Input Word Generator Design:

The Word Generator Circuit design is based on the one designed and illustrated in equation 6. The Word Generator consists of two ICs, an Inverter, a Clock signal with frequency of 6.666KHz, and a DC supply. The 74163 IC is a TTL 4-bit binary counter, and the 74166 is a TTL 8-bit Parallel-In/Serial-Out Shift Register. The Clock signal feeds both ICs as shown in 9. The shift register loads the levels on its inputs when the Parallel Load signal is activated. Then it starts to shift the data, and sends them serially. The Parallel Load signal is taken from a 4-bit binary counter, so that after the counter finishes its count cycle, it will issue the RCO 1 signal, which is used to trigger the shift counter to activate the Parallel Load function. There would be obtaining a controllable 8-bit periodic Word Generator signal. The Signal obtained is shown in figure 10. The 8-bit Word is then multiplied by the PN signal. In the block diagram there has been used a NRZE 2 to adjust the Word level voltages. Now, assuming that we are using the NRZE, the output of the multiplier given the word signal shown in figure 10.

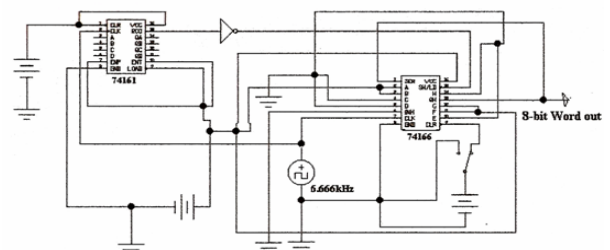


Figure 9: The 8-bit Generator Circuit diagram with a data rate (period of one bit) of 6.666 KHz



Figure 10: The 8-bit word signal.

There has been seen the PN signal affects the incoming word signal that is 1/15 of its frequency.

When a  $\bullet 5$  is multiplied with a  $+5$  the result is  $\bullet 25$ . And when a  $\bullet 5$  is multiplied by a  $\bullet 5$  the result is a  $+25$ . I mentioned in the previous section that the output of the PN is a 0–5 volts, not a  $\bullet 5$ –5 level. Also, the Word levels are 0–5 not  $\bullet 5$ –5. This means that i can't use the multiplication procedure here, because a  $0 \cdot 0$  is not equal to 1. To overcome this problem, and to eliminate the use of a NRZE for the word signal, i substituted the Multiplier with a XNOR gate.

#### G. The BPSK Modulator Design:

This section illustrates the design of the BPSK modulator. The design is directly based on the block diagram shown in figure 9. Figure 11 shown in BPSK Modulator circuit. The design and simulations are based on the use of ideal components. The BPSK Modulator is based on the idea of changing the phase of the carrier signal whenever the incoming Bit changes its state, for example, if the incoming message changes its state from 0–1, the carrier changes its phase by  $+180^\circ$ , and if it changes its state from 1–0 the carrier changes its phase by  $-180^\circ$ .

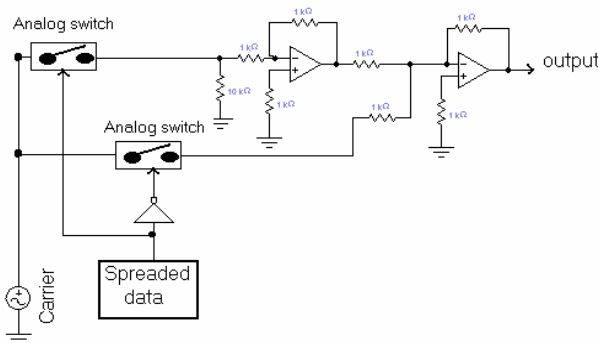


Figure 11: The BPSK Modulator circuit

For the design of BPSK modulator we used two inverting OP-AMP with the gain 1 and also two analogue switch and one inverter. At operation of the BPSK, the carrier signals are entered into the OP-AMPs through the analog switches which are active at the positive pulse. So switches are controlled by the data. When PN is +ve, the switch one is on and the signal pass and enter into the 1st OP-AMP it inverts this signal [8]. At the same time 2nd switch remains off because it is connected with an inverter. When this signal cross 2nd OP-AMP the signal is inverted again and comes to the original position. On the other hand the 2nd switch will be on when the PN is –ve and it will enter into only the 2nd OP-AMP and invert. In this way it transmits

data. The input carrier frequency used in this design is 5 times f<sub>pn</sub> chip, which is 500 kHz. But in our circuit implementation we used carrier of 100 kHz if we give more as like simulation, it is difficult to find out the phase shift in the oscillator. Though one can rise the frequency of the signal in wireless applications to end up with a shorter antenna for the devices. Anyway, this isn't our concern in this report. The modulation process, we finally get the  $S_{ss}(t)$  which is the signal obtained from the output of the Modulator circuit. The output of the DS-SS Modulator is shown in figure 12 and precise mode figure 13. Here, a couple of the bits from  $S_{ss}(t)$  has illustrate the output signal.

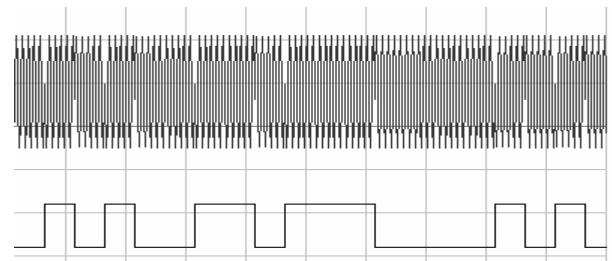


Figure 12:  $S_{ss}(t)$  The DS-SS signal coming out

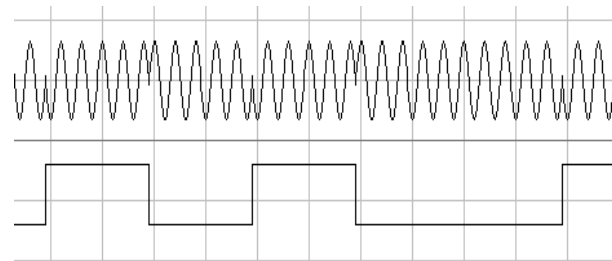


Figure 13:  $s_{ss}(t)$  The DS-SS signal coming out of of the Modulator (Precise mode).

#### H. Bandwidth consideration of BPSK:

A balance modulator is a product modulator, the output signal is the product of the two input signal. In a BPSK modulator, the circuit output signal is multiplied by the binary data. If +1V is assigned to a logic 1 and -1V is assigned to a logic 0, the input carrier ( $\sin \omega_c t$ ) is multiplied by either a +1 or -1. Consequently, the output signal is either  $+1\sin \omega_c t$  or  $-1\sin \omega_c t$  [9, 10]. This paper described the design of the modulation technique of direct sequence-spread spectrum with its hardware implementation. It also describes the characteristics of BPSK modulator. There has been found the circuit's output characteristics and the BPSK modulated signal's characteristic is same

although a fault was found in our circuit's OP-AMP (at BPSK modulator circuit). As in obvious form the observed out put and simulated out put there is same anomaly. The amplitude of the observed output (fig 14) is not uniform all along. There are sudden jumps in the output wave shape. The cause of these jumps may be manufacturing non uniformly of the integrated circuits.

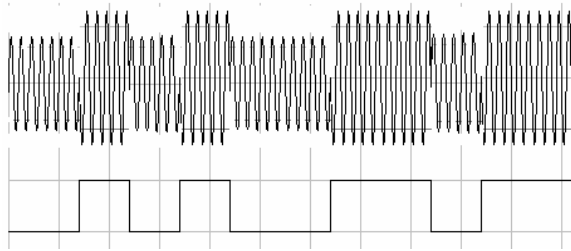


Figure 14: Circuit's output

In the physical structure, it has been used available components. Firstly there has been first designed it by the simulation software named EWB (Electronics Workbench). There also has been designed the Demodulator circuit with the software and it works successfully.

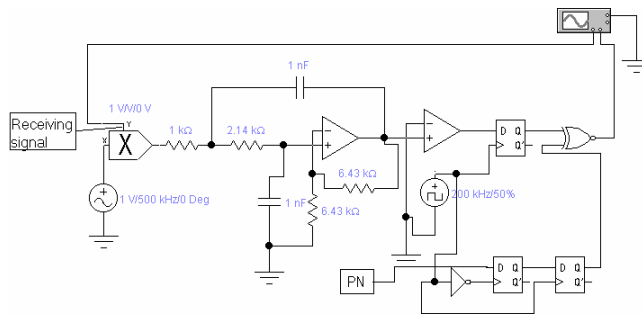


Figure 15: DS-SS receiver circuit diagram

Figure 14 shows the diagram of the filter (the Sallen-Key second order active filter) and figure 15 shows the full diagram of the demodulator circuit. One who is interested as for Multiplier (Balance Modulator) he can use MC1496 IC. Here all equipments are common and available. Finally, hope that this report has supplied you with some new information, techniques, and ideas. Again, it is intended to be as an educational supplement for any advanced communications course, where the student wants to see a practical application on what is studying.

### III. Conclusion

This paper demonstrated the design of a simplified Direct Sequence Spread Spectrum CDMA system. Design of 8-bit word generator is given. Spreading technique is clearly implemented. Finally, original data are recovered by despreading in receiver.

### REFERENCES

- [1] "An Introduction to Direct-Sequence Spread-Spectrum Communications". <http://www.cs.clemson.edu/~westall/851/spread-spectrum.pdf>
- [2] Simon Haykin, "Communication Systems", fourth edition, published by JOHAN WILEY & SONS, INC.
- [3] "Digital downconverter/despreader for direct sequence spread spectrum CDMA communications system". <http://www.patentstorm.us/patents/6141372-description.html>
- [4] "Acquisition of Direct Sequence Spread Spectrum Acoustic Communication Signals". <http://www.mit.edu/~millitsa/resources/pdfs/acq.pdf>
- [5] Samuel C., Yany, "CDMA RF System Engineering", First edition
- [6] "Supplemental Text Book of CDMA-One Basic Technologies" Training text book-FUJITSU, July 2000.
- [7] "CDMA Technology Resources: Welcome to the world of CDMA: Common air interface" 2003 CDMA Development group <http://www.cdg.org/technology/index.asp>
- [8] "COFDM as a modulation technique of wireless telecommunication with a CDMA comparison" Thesis-Erc Lawrey-1997
- [9] J.K. Thomacy "Digital communication system".
- [10] S.L. Gupta & Dr. V. Kumar "Handbook of electronics", 31<sup>st</sup> revised and enlarged edition.

# Comparative performance analysis of DSR and AODV protocol based on Mobility factor

A. S. M. Ashique Mahmood, Mohammad Shahid Hossain, Faisal Aman Aziz

Dept. of Computer Science and Information Technology(CIT)  
Islamic University of Technology (IUT), Boardbazar, Gazipur, Bangladesh  
ashique.mahmood@gmail.com, geteamon@yahoo.com, faisal\_aman\_23@yahoo.com

**Abstract** — Ad hoc wireless networks are decentralized in nature and hence routing is a central challenge in this type of network. In this paper we describe a simulation model to simulate the routing protocols for wireless ad hoc networks and analyze the results from the simulation for performance evaluation. The simulation model incorporates the number of parameter values with changing environmental factors. Specially we have chosen the Dynamic Source Routing (DSR) and Ad-hoc On Demand Distance Vector Routing (AODV) protocol to be implemented in the simulation software with that model. We carried out the extensive simulations and we conclude with the comparative statistical analysis of those two protocols based on mobility factor.

**Index Terms** — MANET, Wireless ad-hoc, routing, DSR, AODV.

## I. INTRODUCTION

Both Dynamic Source Routing (DSR) [1] and Ad-hoc On Demand Distance Vector Routing (AODV) [2] are reactive and on-demand routing protocols [3] where source initiates the route discovery process to find the route to the destination. DSR takes the advantage of route cache, in which it caches source routes that it has learned, maintained by each host in the ad hoc network [4]. In contrast, in AODV the packet header contains only the next hop address for a desired destination.

A packet-level simulation model that was used allowed us to observe and measure the protocols' performance under a variety of conditions. The basic simulation parameters were chosen to model an ad hoc network consisting of a collection of mobile hosts moving around in a certain space. In addition to a number of parameter choices in the protocol, the simulator allows us to vary certain environmental factors such as –

- The number of mobile hosts
- Mobility Pattern of the hosts
- Traffic pattern
- Simulation time and the distribution of the hosts in space.

Each host is initially placed at a random position in the simulation area. As the simulation progresses, each host pauses at its current location for a period (pause time), and then randomly chooses a new location to move with a

velocity in a predefined range. Thus the uses of constantly changing factors gave us an opportunity to analyze the protocol in a wide range of situations.

Some performance metrics were measured from the simulation results and those metrics were analyzed for evaluating the protocols' performance. These performance metrics include -

- Packet delivery ratio – The ratio of data packets delivered to the destinations and data packets originated by the sources.
- Average end-to-end delay of data packets – This includes all possible delays of propagating the packets.

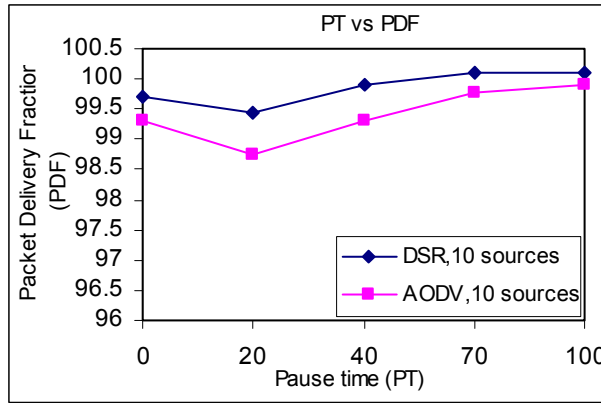
## II. SIMULATION MODEL AND METHODOLOGY

The modeling under the simulating software is divided into 3 levels. The upper level is the network level, and consists in the building of the network topology. The second level is the description of each node that composed this network topology, namely node model. The third and last level is the description of each of these processes that composed a node model. In next section we present the node model for the DSR protocol.

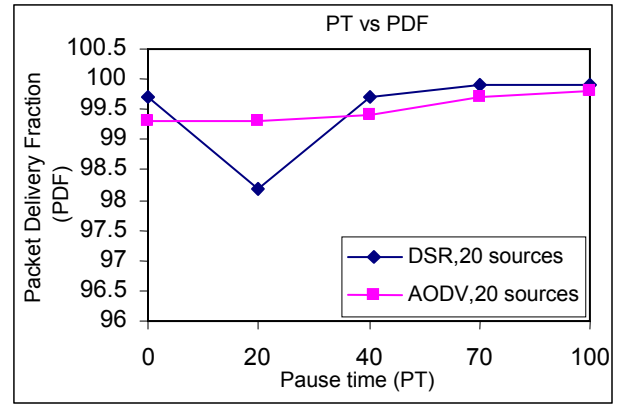
### A. The Node Model

Here is a description of each layer of the DSR node model:

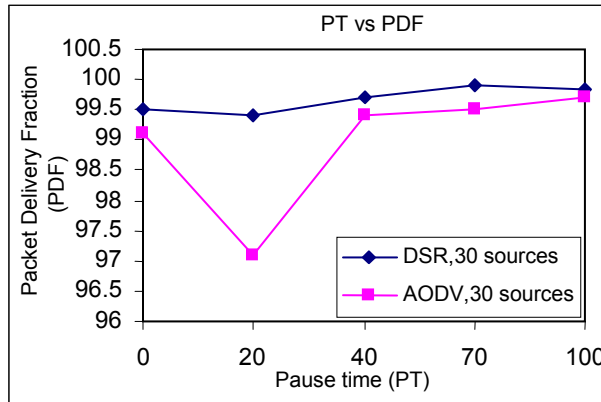
- The physical layer is composed of a transmitter and a receiver. Actually, both the physical and the link layer of this DSR node model come from the Wireless LAN model. Both the transmitter and receiver communicate with upper layers through the 802.11 MAC protocol.
- The link layer used in the model is the 802.11 model. Some small modifications are included in order to link this MAC layer with our DSR routing layer (sending of acknowledgement and error messages, addition of the promiscuous mode). We choose to use 802.11 since this protocol is presently in fashion in the MANET community, and seems to be the protocol that will be used in this kind of network.



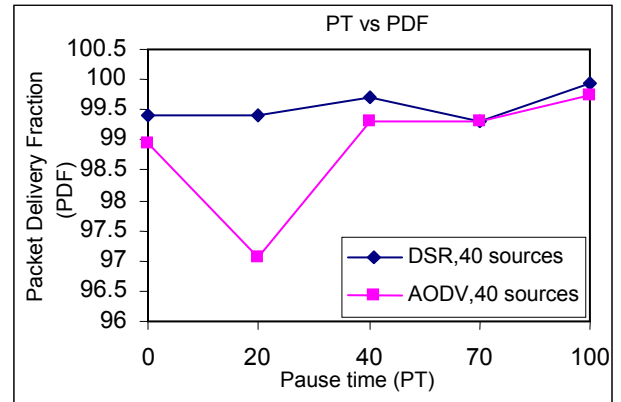
1(a)



1(b)



1(c)



1(d)

Figure 1: Pause Time vs Packet Delivery Fraction (PDF)

- The network layer is the core of our node model since it contains our DSR routing process model. As the 802.11 layer, it is divided in two processes. The first one is the DSR routing module, and the second one an interface with the upper layer that chooses for instance a random destination address for each data packet that must be transmitted on the network.
- The upper layers are mainly composed of two processes. The source is a process that generates data packet traffic. The receiver is a sink that just destroys the packet after the reception and processing completion.
- We placed another process model in the upper layers, called mobility. This one is in charge of the node mobility. This process defines the mobility pattern of the node, selects the destination point of that node. Upon reaching this point, the node pauses for a wait timer and then computes a new destination.

We use a detailed simulation model based on ns-2 in our evaluation. In a paper, the Monarch research group

at Carnegie-Mellon University[5] developed support for simulating multi hop wireless networks complete with physical, data link, and medium access control (MAC) layer models on ns-2.

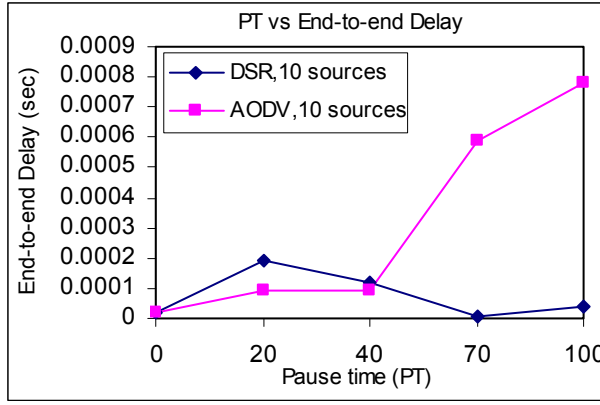
#### B. The Traffic and Mobility Model

Traffic sources are continuous bit rate (CBR). The source-destination pairs are spread randomly over the network. Only 512-byte data packets are used. The number of source-destination pairs and the packet sending rate in each pair is varied to change the offered load in the network.

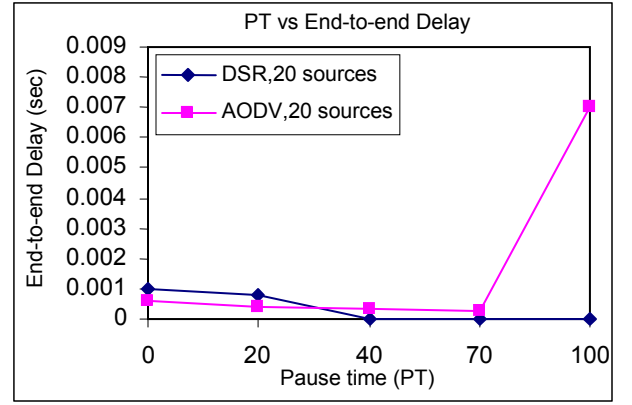
For the simulations carried out, traffic models were generated for 50 nodes with cbr traffic sources, with maximum connections of 10,20,30,40 at a rate of 4kbps. Now we varied the the no. of sources within the network. For example our simulation ran for each instance with 10, 20, 30, 40 sources.

Mobility models were created for the simulations using 50 nodes, with pause times of 0,20,40,70,100 seconds, maximum speed of 20m/s, topology boundary of 500x500 and simulation time of 100secs.

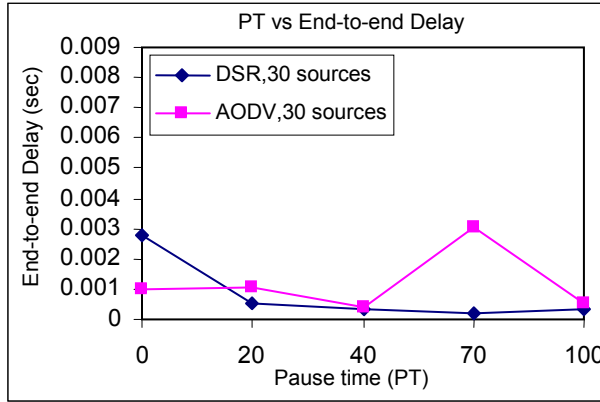




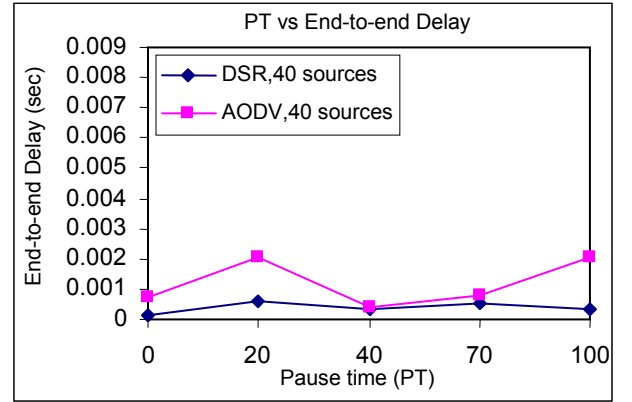
2(a)



2(b)



2(c)



2(d)

Figure 2: Pause Time vs Average end-to-end Delay

Simulations could be carried out with more complex scenarios using higher traffic rates. However, the simulations take much longer to complete and the trace file generated by each run ends up to more than 50 Mb of user space.

### III. OBSERVATIONS

To analysis any routing protocol we have to measure some performance metrics which are discussed in the previous part. Among those performance metrics we have chosen two important metrics to analyze our concerned protocols. Those are –

- Packet Delivery Fraction
- Average end-to-end delay of data packets

We have measured these four performance metrics based on the *Pause Time* environmental factor that actually represents the mobility pattern.

#### A. Pause Time

Pause time indicates how much time a mobile host keeps stagnant at a particular position. In a mobile adhoc network simulation a host randomly selects a particular direction and goes to that direction. After reaching that position the node waits for a certain amount of time. This time is called pause time. In our simulation we vary this pause time and find the value of different performance metrics for those pause time.

Now we will see the result of the simulation of both DSR and AODV protocol where we vary the pause time as 0,20,40,70,100 and plot the performance metrics against them.

The first set of experiments uses differing numbers of sources with a moderate packet rate and varying pause times. For the 50 node experiments we used 10, 20, 30, and 40 traffic sources and a packet rate of 4 packets/s, except for 40 sources, which use 3 packets/s. We used a slower rate with 40 sources because the network congestion was too high otherwise for a meaningful comparison. The packet delivery fractions for DSR and AODV are significantly different with 10 and 20 sources

(Fig. 1a and b). DSR is likely to perform well in these situations. With 30 and 40 sources, AODV tends to perform almost similarly with DSR (Fig. 1c and d) at lower pause times (higher mobility). For higher pause times (low mobility), however, DSR has a better delivery fraction than AODV. The relative performance of both protocols with respect to delays is similar to that with delivery fractions. DSR and AODV have almost identical delays with 10 and 20 sources (Fig. 2a and b) and with higher mobility. With 30 and 40 sources, AODV has higher end-to-end delay than DSR (Fig. 2c and d) for lower pause times. But for higher pause times, DSR has better (10–20 percent lower) delay than AODV.

One interesting observation is that the delays for both protocols increase with 40 sources with very low mobility (Fig. 2d). This is due to a high level of network congestion and multiple access interferences at certain regions of the ad hoc network. Neither protocol has any mechanism for load balancing, that is, choosing routes in such a way that the data traffic can be more evenly distributed in the network. This phenomenon is less visible with higher mobility where traffic automatically gets more evenly distributed due to source movements.

#### IV. DISCUSSION

##### A. Effect of Mobility

Our simulation results show that mobility affects the performance of AODV and DSR differently. In the presence of high mobility, link failures can happen very frequently. Link failures trigger new route discoveries in AODV since it has at most one route per destination in its routing table. Thus, the frequency of route discoveries in AODV is directly proportional to the number of route breaks. The reaction of DSR to link failures in comparison is mild and causes route discovery less often. The reason is the abundance of cached routes at each node. Thus, the route discovery is delayed in DSR until all cached routes fail. But with high mobility, the chances of the caches being stale are quite high in DSR. Eventually when a route discovery is initiated, the large numbers of replies received in response are associated with high MAC overhead and cause increased interference to data traffic. Hence, the cache staleness and high MAC overhead together result in significant degradation in performance for DSR in high mobility scenarios. Our simulation results show that this effect is more severe with large numbers of sources and for larger networks.

With low mobility, the possibility of link failures is low. However, nodes usually get clustered with low mobility, an artifact of our node movement (random waypoint) model. This leads to network congestion in

certain regions in the presence of high traffic. Congestion in turn causes link layer feedback to report link failures even when the nodes are relatively static and the physical link exists between nodes. Such spurious link failures lead to new route discoveries in AODV. DSR, in contrast, is largely unaffected by this problem at low mobility. DSR caches are nearly up to date in low-mobility cases. Thus, even when a spurious link failure is reported, DSR benefits from caching considerably by salvaging at intermediate nodes and using alternate routes at the sources. Also, the AODV timer-based route expiry mechanism could result in unnecessary route invalidations since the spacing between data packets using a route is critical to refreshing timers associated with that route at different nodes. The above effects of mobility are particularly visible for high traffic scenarios. Also, in reality, a combination of nodes with different mobility (different speeds and different pause times) can form an ad hoc network. In that case, it is hard to predict the relative performance of AODV and DSR.

##### B. Packet Delivery and Choice of Routes

DSR fared comparatively well in our application-oriented metrics (delivery fraction and delay) in more “stressful” situations (i.e., larger numbers of nodes, sources, and/or higher mobility). However, AODV also performed better in less stressful situations. The reason for both of these phenomena is the aggressive use of route caching in DSR. In our observation, such caching provides a significant benefit up to a certain extent. But with extreme loads the extent of caching is deemed too large to benefit performance. The simulation time span was at best 200 sec but in reality more stressful situation can be created with larger time frame. It’s an assumption that AODV will outperform DSR in more stressful conditions. Often, stale routes are chosen since route length (and not any freshness criterion) is the only metric used to pick routes from the cache when faced with multiple choices. Picking stale routes causes two problems:

- Consumption of additional network bandwidth and interface queue slots even though the packet is eventually dropped or delayed
- Possible pollution of caches in other nodes

We have also independently observed that cache expiry using suitable timeouts and wider propagation of route errors can improve the performance of DSR significantly. When compared to AODV, a much smaller number of packets was dropped in DSR for lack of route (e.g., indicating a high cache hit ratio); however, significantly more packets were dropped due to the interface queue being full. An efficient mechanism to “age” packets and



drop aged packets from the network will improve delays in both protocols, particularly DSR.

## V. FUTURE DIRECTIONS

The simulations were not done for larger time span. Actually the ad-hoc network does not exist for only 200 sec which was our highest implementation. An ad-hoc network operates in a larger time span. So if we can increase the simulation time then we can further trace the activities of the nodes and hence analyze the performance in more details and with higher accuracy.

## VI. CONCLUSION

The simulation was done with the specified environmental conditions and with the changing parameters. In cases of high mobility (low pause time), it was found that the DSR protocol produced a higher PDF (packet delivery fraction) than AODV. One important thing to note is that PDF values converge to same values with decreasing mobility i.e. both the protocols tend to perform similarly in stable, non-mobile conditions. Now, for end-to-end delay, it was seen that in highly mobile conditions (when nodes have less pause time) both AODV and DSR performs almost similarly. But, with the

increase of pause time, DSR exhibits notable improvement over AODV, because AODV produces higher end-to-end delay with decreasing mobility. So, in short, DSR is likely to perform better than AODV in more stressful situation.

## REFERENCES

- [1] J. Broch, D. Johnson, and D. Maltz, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks", *IETF Internet draft*, draft-ietf-manet-dsr-03.txt, Oct. 1999.
- [2] C. E. Perkins and E. M. Royer, "Ad-hoc OnDemand Distance Vector Routing," *Proc. 2nd IEEE Wksp. Mobile Comp. Sys. And Apps.*, pp. 90-100, Feb. 1999.
- [3] Elizabeth M. Royer, C-K Toh, "A Review of Current Routing Protocols for Ad-Hoc Mobile wireless Networks", *IEEE Personal Communications*, Vol.6, No.2, pp.46-55, April, 1999.
- [4] D. Johnson and D. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks", *In T. Imielinski and H. Korth (Eds.), Mobile Computing*, Kluwer Academic Publishers, Chapter 5, pp. 152-181, 1996.
- [5] J. Broch et al., "A performance comparison of multihop wireless ad hoc network routing protocols", *Proc. 4<sup>th</sup> annual ICM/IEEE International Conference on Mobile Computing and Networking(MobiCom 98)*, ACM Dallas, pp 85-97, Oct, 1998.

# A Least Square Approach for TDOA/AOA Wireless Location in WCDMA System

Mohammad Nur-A-Alam\*, Mohammed Mezbaul Haque\*\*

Department of Communication Engineering, Tampere University of Technology, Finland\*

Telecom Malaysia International Bangladesh, Bangladesh\*\*

Emails: mohammad.nur-a-alam@tut.fi, mezbaul.haque@aktel.com

**Abstract**—Cellular mobile networks use time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA) or received signal strength (RSS) based measurement techniques for mobile position estimation. Each of these measurement techniques has its degree of accuracy and level of complexity. Individual methods are usually not able to give adequate correct position of a mobile user. Therefore, a common trend is to combine these methods with other suitable techniques to attain enhanced performance. There are several algorithms to compensate the position errors by combining different methods. In this paper, a two step least square algorithm for wireless position estimation technique is implemented in WCDMA system where an angle of arrival (AOA) assisted and correlation slope-based time difference of arrival (TDOA) measurement method is considered and their performance is presented in signal to noise ratio conditions

**Index Terms**—Least squares, WCDMA, TDOA, AOA, Performance analysis.

## I. INTRODUCTION

Cellular network based wireless location systems have been researched and successfully deployed for a few decades, but with limitations in accuracy due to multipath and other complications. Code-division multiple access (CDMA) systems overcome [4] this moderately, as their wide bandwidth is able to resolve some of the multipath components. CDMA systems assign each user a unique pseudonoise (PN) spreading code, which makes the identification of multipath components with delay over one PN code chip epoch efficient by signal correlation processing. The correlation peak of a multipath with a delay over one chip does not affect the correlation peak position of the direct signal. One chip epoch in CDMA system is about 800 ns which equals to 240 metres in distance and for WCDMA one chip epoch is about 260 ns [1]. The reflected rays arriving within a chip period of the first ray cannot be resolved by a correlation technique, which is a significant factor limiting the location accuracy. In order to meet the accuracy requirements in handset-based solutions, the mitigation of the effects of intra-chip multipath must be considered.

There are two major methods to implement a cellular network-based location system. In the first approach, the mobile station (MS) receives signals from the base stations

(BS) to calculate its own position, as in GPS. In the second approach, the BS's process the signals from the MS to locate the user. The second approach has the advantage of using the existing receiver. There are several systems using cellular network based or GPS-assisted cellular network-based solution in which a cell phone user receives GPS signals and also connected to a network-based location server which give the actual user location.

The primary objective of this work is to find out the best i.e. the most accurate radio location measurement technique. The target is also to develop a practical signal processing method in mobile based location systems to estimate the AOA and the TDOA for the LOS signal under multipath environments.

This paper is divided into several sections. First mathematical model of AOA/TDOA measurement techniques and a least square algorithm are described. Furthermore a detailed description of the implemented AOA/TDOA algorithms presented with simulation results. Also the results are analyzed and compared the accuracy performance of the explained method.

## II. MEASUREMENT MODELS

As static positioning is considered, a unique single location is estimated from several concurrently taken measurements at each time, which are not dependent with its past or future estimations. And efficient position estimation needs proper utilization of all measured data's in a well defined mathematical model. This mathematical model is pursued by different estimation technique for better position estimation. In this section, the models and assumptions for the TDOA/AOA and hybrid measurements are described.

The system equation of a positioning problem can be written as

$$A(\alpha) + e = b \quad (1)$$

where,  $A(\alpha)$  is the known vector function,  $e$  is the measurement error vector and  $b$  is the collection of all measured values.

For the two-dimensional (2-D) case, if the MS position to be determined is  $\alpha = [x, y]^T$  and the known coordinates

of the  $i$  th BS be  $\alpha_i = [x_i, y_i]^T, i = 1, 2, \dots, N$ , where the superscript  $T$  denotes the transpose operation and  $N$  is the total number of receiving BSs.

The distance between the MS and the  $i$  th BS, denoted by  $d_i$ , is given by

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad (2)$$

#### A. Time Difference of Arrival (TDOA) measurements

TDOA system uses time difference measurements rather than absolute time measurement as TOA does. It is often referred to as the hyperbolic system [7] because the time difference is converted to a constant distance difference between two base stations to define a hyperbolic curve. The intersection of two hyperbolas determines the position. In this approach at least three base stations are needed.

TDOA method does not require knowledge of the transmit time from the source therefore it doesn't need any strict clock synchronization between the source and receiver. As a result, TDOA techniques do not require additional hardware or software implementation within the mobile unit. Also, this method is able to reduce common errors experienced at all receivers due to the channel.

TDOA are based on range measurements and can directly obtained from the TOA data. In TDOA it is needed to take the difference in TOAs of the MS signal at a pair of BSs. Any BS can be assigned as the reference though in most of the cases nearest BS is fixed as the reference. Now, the range measurements based on the TDOAs without any disturbances is

$$ct_{i,j} = \frac{\sqrt{(x_i - x)^2 + (y_i - y)^2} - \sqrt{(x_j - x)^2 + (y_j - y)^2}}{c}, i, j \in \mathbf{S}; i \neq j \quad (3)$$

where  $i, j = 1, 2, 3, \dots, N$  respectively and  $\mathbf{S}$  is the set of all BS.

Now, in case of disturbances, range measurements of TDOA can be expressed simply as

$$\begin{aligned} r_{TDOA,i,j} &= r_{TOA,i} - r_{TOA,j} \\ &= d_i + e_{TOA,i} - d_j - e_{TOA,j} \\ &= (d_i - d_j) + (e_{TOA,i} - e_{TOA,j}) \\ &= \sqrt{(x_i - x)^2 + (y_i - y)^2} - \sqrt{(x_j - x)^2 + (y_j - y)^2} + e_{TDOA,i,j} \end{aligned} \quad (4)$$

where  $e_{TDOA,i,j}$  is the range measurement error in  $r_{TDOA,i,j}$ .

This equation (4) can be written in vector form as,

$$\mathbf{r}_{TDOA} = \mathbf{A}_{TDOA}(\alpha) + \mathbf{e}_{TDOA} \quad (5)$$

where,  $\mathbf{r}_{TDOA}$  is the collection of range measurement vector,  $\mathbf{e}_{TDOA}$  is the range error vector and  $\mathbf{A}_{TDOA}(\alpha)$  is the actual range measurement vector function from MS to BS, which contains the original MS position. For simplicity lets just assume that BS 1 is the reference then these vectors can be shown in below

$$\begin{aligned} \mathbf{r}_{TDOA} &= [r_{TDOA,2}, r_{TDOA,3}, \dots, r_{TDOA,N}]^T, \\ \mathbf{e}_{TDOA} &= [e_{TDOA,2}, e_{TDOA,3}, \dots, e_{TDOA,N}]^T, \\ \mathbf{A}_{TDOA}(\alpha) &= \begin{bmatrix} \sqrt{(x_2 - x)^2 + (y_2 - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \\ \sqrt{(x_3 - x)^2 + (y_3 - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \\ \vdots \\ \sqrt{(x_N - x)^2 + (y_N - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \end{bmatrix} \end{aligned} \quad (6)$$

#### B. Angle of Arrival (AOA) measurements

The Angle of Arrival measurement technique determines the location based on the direction of arrival of the mobile station signal in base station, by using directive antennas or antennas array. The minimum number of BS's needed for the location process is two, which makes AOA a very attractive method in situations where detect ability of multiple BS signal is a problem. Another advantage is that the AOA method does not need synchronization with the BS clock. The main disadvantages of this method are multipath effects, and that an antenna array is needed. The accuracy of the AOA method is highly dependent on the propagation environment since multipath around the MS will affect the measured AOA. Multipath components affect the angle measurement in a way that the arriving signal may appear to arrive from a different direction.

In adaptive antenna array BS's signals are received with AOA information from MS and these signals are needed for position estimation purpose. MS is at the intersection of several lines corresponding to AOA measurements. The AOA of the transmitted signal from the MS at the  $i$ th BS, denoted by  $\theta_i$  is related to [6]

$$\tan(\theta_i) = \frac{y_i - y}{x_i - x} \quad (7)$$

where,  $\theta_i$  is the angle between the linked object from the  $i$ th BS to the MS and the  $x$ -axis. In presense of angle errors the AOA measurements can be denoted as  $r_{AOA,i}$  and are expressed as

$$\begin{aligned} r_{AOA,i} &= \theta_i + e_{AOA,i} \\ &= \tan^{-1} \frac{y_i - y}{x_i - x} + e_{AOA,i} \end{aligned} \quad (8)$$

where  $e_{AOA,i}$  is the angle measurement error in  $r_{AOA,i}$ . Equation (8) can also be expressed in vector form as

$$\mathbf{r}_{AOA} = \mathbf{A}_{AOA}(\alpha) + \mathbf{e}_{AOA} \quad (9)$$

where,  $\mathbf{r}_{AOA}$  is the collection of angle measurement vector,  $\mathbf{e}_{AOA}$  is the angle error vector and  $\mathbf{A}_{AOA}(\alpha)$  is the actual angle measurement vector function from MS to BS, which contains the original MS position. These vectors can be shown in below

$$\begin{aligned} r_{AOA} &= [r_{AOA,1} r_{AOA,2} \dots r_{AOA,N}]^T, \\ e_{AOA} &= [e_{AOA,1} e_{AOA,2} \dots e_{AOA,N}]^T, \\ A_{AOA}(\alpha) &= \begin{bmatrix} \tan^{-1} \frac{y_1 - y}{x_1 - x} \\ \tan^{-1} \frac{y_2 - y}{x_2 - x} \\ \vdots \\ \tan^{-1} \frac{y_N - y}{x_N - x} \end{bmatrix} \end{aligned} \quad (10)$$

### C. AOA/TDOA Hybrid measurements

In this hybrid measurement technique, several base stations (at least three) receive the target mobile signal, and the target position is determined by combining AOA estimates from each base station and TDOA estimates between multiple base stations. AOA estimates are made using an adaptive array or other adjustable arrangements. Times of signal arrival are measured at each base station and are time stamped with a MS time reference to determine time difference of arrival estimates.

Though multipath fading will occur in this case, still this type of hybrid system is probably the most accurate and reliable type of cellular positioning system that can be used. However, it is an important issue in such systems to make sure that both methods combine synergistically and that the inaccuracies from one method do not adversely affect the overall position estimate.

Now equation (8) can be re-written as

$$\begin{aligned} x \sin(\theta_i + e_{AOA,i}) - y \cos(\theta_i + e_{AOA,i}) &= \\ x_i \sin(\theta_i + e_{AOA,i}) - y_i \cos(\theta_i + e_{AOA,i}) \end{aligned} \quad (11)$$

By simply combining equation (4) and discarding the angle error from (11) we get the combined AOA/TDOA

measurements as follows

$$\begin{bmatrix} x_2 - x_1 & y_2 - y_1 & d_{2,1} \\ \vdots & \vdots & \vdots \\ x_N - x_1 & y_N - y_1 & d_{N,1} \\ \sin(\theta_1) & -\cos(\theta_1) & 0 \\ \vdots & \vdots & \vdots \\ \sin(\theta_N) & -\cos(\theta_N) & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ d \end{bmatrix} = \begin{bmatrix} \frac{B_2 - B_1 - d_{2,1}^2}{2} \\ \vdots \\ \frac{B_N - B_1 - d_{N,1}^2}{2} \\ x_1 \sin(\theta_1) - y_1 \cos(\theta_1) \\ \vdots \\ x_N \sin(\theta_N) - y_N \cos(\theta_N) \end{bmatrix} \quad (12)$$

where  $B_i$  is the collecting BS's and written as

$$B_i = (x_i + y_i)^2 - 2x_i y_i \quad (13)$$

Equation (12) can also be expressed in vector form as

$$\mathbf{H}_{hybrid} \mathbf{I} = \mathbf{K}_{hybrid} \quad (14)$$

where,  $\mathbf{I}$  contains the original MS position.

### III. TDOA AND AOA BASED POSITIONING ALGORITHM

A two step LS algorithm is applied for individual TDOA/AOA and their combined hybrid estimation measurements. Parameters of TDOA scheme provide a non-linear system of equations. So, it determines the MS position by simply solving (5) via standard LS and the location estimate is found from the following optimization problem [5]

$$\hat{\alpha} = \arg \min_{i,j \in S; i \neq j} \sum (r_{i,j} - ||\alpha - X_i|| - ||\alpha - X_j||)^2 \quad (15)$$

And AOA system also individually determine the MS position by solving (9) LS and the location estimate is found from the following optimization problem

$$\hat{\alpha} = \arg \min_{\alpha} \sum_{i \in S} |- \sin \theta_i (x - x_i) + \cos \theta_i (y - y_i)|^2 \quad (16)$$

To improve positioning accuracy, it is better to use as much information as possible. A hybrid solution is obtained by simply combining TDOA and AOA measurements as stated above and then by solving iteratively an estimation of mobile position can be obtained.

#### IV. RESULTS

In order to evaluate the accuracy of the hybrid mobile position estimation technique, we consider root-mean square (RMS) position estimation error. RMS error is calculated as the square root of mean squared error (MSE). For comparing the performance of the hybrid algorithm, the RMS position estimation errors of TDOA and AOA are also calculated. This section presents the simulation result that is obtained from the hybrid mobile position estimation algorithm for single and multipath path channel.

The performance of TDOA, AOA and hybrid position estimations for single path channel is presented in figure 1. In this figure RMSE values are plotted against SNR values. The red line indicates the performance of TDOA algorithm while the cyan line represents the AOA algorithm and blue line represents the Hybrid algorithm against SNR.

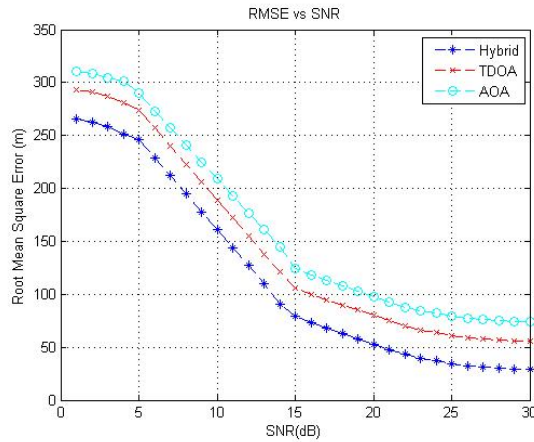


Fig. 1. RMSE performance comparison among TDOA, AOA and hybrid method in single path channel

From the above figure we can find that TDOA is giving lower RMS error than AOA. As the SNR value increases, the error reduced significantly for both cases. At 30 dB SNR, AOA's RMS error is 75 meter whereas TDOA'S RMS error is 52 meters. Error difference in this case is 23 meters, so TDOA is giving much better location accuracy compare to AOA.

Now, we examine the RMSE performance of hybrid algorithm against the SNR values. For low SNR values, the performance is not very impressive as the result is showing quite large RMS error. But as the SNR value increases the RMS error is reduced. Also it can be seen that when SNR value goes above 15 dB then the system is providing quite accurate result. At SNR of 30 dB the RMS error remains 35 meter.

It is clear from the figure that hybrid algorithm provides better result, when TDOA and AOA algorithms are implemented independently. At optimum level (SNR 30 dB) the hybrid method's RMS error is 35 meter which is 17 meter

less than TDOA's and 40 meter less than AOA's RMS error estimation.

For single path channel path delay and power vector of [0] are assumed. But for simulation on 4 path multipath channel the path delay vector of [0 0.3 0.5 0.8] and the power vector of [0 0.7 0.4 0.2] are taken as a channel parameter.

The performance of TDOA, AOA and hybrid position estimation algorithms for multipath path channel is presented in figure 2. It is clear from the figure that, hybrid algorithm gives better performance compare to individual performance of TDOA and AOA, also in case of multipath channel.

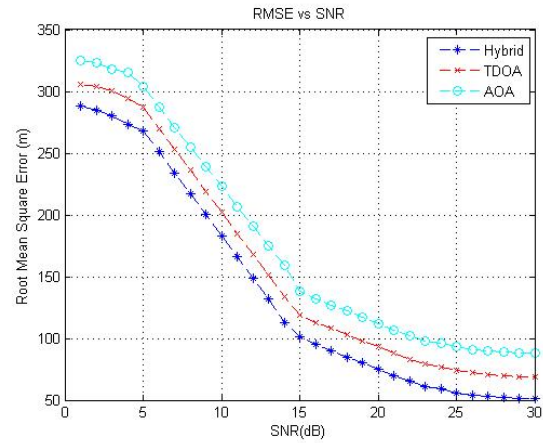


Fig. 2. RMSE performance comparison among TDOA, AOA and hybrid method in Multipath channel

#### V. CONCLUSION

In this paper a TDOA/AOA and their hybrid measurement approach is implemented by a two step LS algorithm in WCDMA system. Their implementation shows that this algorithm works quite accurately in WCDMA environment. And also found that this system can combine with GPS system to improve the location accuracy when GPS signal gives poor coverage specially in indoor. Multipath channel mitigation is also tried in the WCDMA channel model.

#### ACKNOWLEDGMENT

This work was the individual interest of the author's. But, we wanted to thank Department of Communication Engineering, Tampere university of Technology, Finland for the supports in simulation period.

#### VI. REFERENCE

- [1] Lu. Dingchen, "Multipath Mitigation in TOA Estimation Based on AOA," Phd thesis, University of Calgary, 2007.

- [2] George A. Mizusawa, "Performance of hyperbolic position location techniques for code division multiple access," Master's thesis, Virginia Polytechnic Institute and State University, 1996.
- [3] I. Reza Rahman, "Data fusion for improved TOA/TDOA position determination in wireless Systems," Master's thesis, Virginia Polytechnic Institute and State University, 2000.
- [4] J. Caffery and G. L. "Subscriber location in CDMA cellular networks," in IEEE International Transaction on Vehicular Technology, May, 1998.
- [5] M. Changlin, R. Klukas, and G. Lachapelle, "An enhanced two-step least squared approach for TDOA/AOA wireless location," in Proceedings of the IEEE International Conference on Communications, May 2003, pp. 987-991.
- [6] K.W.Cheung, H.C.So, W.K.Ma and Y.T.Chan, "A constrained least squares Approach to mobile positioning: Algorithm and optimality" in EURASIP journal on Applied Signal Processing, 2006.
- [7] Y.T.Chan and K.C.Ho, "A simple and efficient estimator for hyperbolic location," in E IEEE Transaction on Signal processing, August, 1994.
- [8] L. Cong and W. Zhuang, "Hybrid TDOA/AOA mobile user location for wide band CDMA cellular systems," IEEE Transactions on Wireless Communications, 2002.

# Prediction of the Density of Active Wireless Device Using Markov Model

MD.Osman Gani<sup>1</sup>, Iftekhar Jahan Mehedi<sup>1</sup>, Mehnaz Seraj<sup>1</sup>,  
Hasan Sarwar<sup>2</sup> and Chowdhury Mofizur Rahman<sup>2</sup>

<sup>1</sup>Department of CSE, Military Institute of Science and Technology, <sup>2</sup>Department of CSE, United International University, Dhaka, Bangladesh

**Abstract** — Location management is one of the key issues in wireless networks to provide an efficient and low-cost service. Realistic modeling of user mobility is a critical research area in wireless network. Mobility data based on real human behaviors may give us the opportunity to improve wireless and Mobile services for users in many ways. At present, several mobility models are proposed based on the analysis of real traces [2], [3]. In this paper, we investigate the feasibility of next state prediction using sequences of previously observed state and analyze the efficiency of MARKOV MODEL [1], [4]. The scenario concerns servicing wireless devices by wireless access point in the Dartmouth college [6] campus over some period of time. The performance of the method has been verified for prediction accuracy. It is found that, on average, the choice of training data leads to prediction accuracy of 78.45%, in some cases the accuracy achieves about 95.55%.

**Index Terms** — Location management, state prediction, Markov model, Access Point.

## I. INTRODUCTION

We investigate and predict the number of wireless devices in an Access Point (AP) within the campus of Dartmouth College [6] based on number of devices previously serviced. Our hypothesis is based on the First Order Markov Process. We analyze the Markov O(1) predictor's efficiency based on the training data set size.

The number of devices serviced by an AP is time dependent. Our aim is to see how far the Markov model dynamically predicts the number of devices which will be serviced by the AP. We have done the prediction over the data collected from CRAWAD [6] which includes data of 6202 wireless devices and 624 AP's over 300 days.

## II. MARKOV MODEL

A Markov chain is a stochastic process with Markov property [4]. Markov property means that, given a present state, future states are independent of the past states. The description of the present state fully captures all the information that could influence the future evolution of the process. Future states will be reached through a probabilistic process instead of a deterministic one. At each step the system may change its state from the current state to another state, or remain in the same state, according to a certain probability distribution.

The changes of state are called transitions, and the probabilities associated with various state-changes are called transition probabilities.

A Markov chain is a sequence of random variables  $X_1, X_2, X_3, \dots$  with the Markov property, namely that, given the present state, the future and past states are independent. Formally,

$$\Pr(X_{n+1} = x | X_n = x_n, \dots, X_1 = x_1) = \Pr(X_{n+1} = x | X_n = x_n).$$

The possible values of  $X_i$  form a countable set  $S$  called the state space of the chain. Markov chains are often described by a directed graph, where the edges are labeled by the probabilities of going from one state to the other states. An  $N$ -state Markov (processes) model for a discrete system is defined by a number of parameters. The set of states  $S$  is denoted by

$$S = \{1, 2, 3, \dots, N\}$$

We have interest in initial probability,  $PI_{i,t}$ , probability that the Markov model will be in state  $i$  at time  $t$ . We denote this as

$$PI_{i,t} = \Pr[S_t = i], i = 1, 2, 3, \dots, N$$

$S_t$  denotes the state at time  $t$ . Thus,  $S_t$  ranges over the set  $S$ . The transition probabilities,  $A_{ij}$  [7], denote the probability of going from state  $i$  at time  $t$  ( $S_t = i$ ) to state  $j$  at time  $t + 1$  ( $S_{t+1} = j$ ). In equation form

$$A_{ij} = \Pr[S_{t+1} = j | S_t = i], i, j = 1, 2, 3, \dots, N$$

Here  $t+1$  do not denote a delay of 1 second from time  $t$ .

## III. DATA SETS

CRAWAD (Community Resource for Archiving Wireless Data)[6] places 6202 wireless devices and 624 wireless access points in the Dartmouth campus to observe the behavior of network for 300 days. Total number of Devices: 6202. Total number of Access Points: 624 (include "OFF"). Total time: 300 days. The observed data from the network are placed into several files and saved as follows.

TABLE I  
ACCESS POINT TRACE FILE - NUMBER OF DEVICES ATTACHED TO AN AP AT DIFFERENT TIME INTERVAL

AP	t <sub>0</sub>	t <sub>1</sub>	t <sub>2</sub>	t <sub>3</sub>	t <sub>4</sub>	t <sub>5</sub>	t <sub>6</sub>	...	t <sub>48</sub>
AcadBldg10AP10	0	0	4	24	0	0	0	...	0
AcadBldg10AP11	0	2	0	0	0	0	0	...	0
AcadBldg10AP12	0	0	0	0	0	0	0	...	0
AcadBldg10AP13	0	0	5	0	0	0	23	...	0
AcadBldg10AP14	0	0	0	0	0	0	0	...	0
AcadBldg10AP15	9	0	0	0	8	0	0	...	17
AcadBldg10AP16	0	0	2	0	0	0	0	...	0

**File Name:** day [1...300].csv

These files describe the number of wireless devices associated with one Access Point per day in a 30-minute interval. This file has the tabular form showed in TABLE I. In the above table, the 1<sup>st</sup> column AP showing the location of the wireless access points. The location format for the wireless access points is as follows:

AP location format:

[Building Name][Building Number][AP][AP Number]

The building names are:

Academic Building (AcadBldg), Administration Building (AdmBldg), Athletic Building (AthlBldg), Library Building (LibBldg), Residential Building (ResBldg)

Example: AcadBldg [10] AP [10]

From the 2<sup>nd</sup> to 49<sup>th</sup> column shows the number of wireless devices that accessed a specific AP during a particular period of 30 minutes time interval. Time intervals are taken as t<sub>1</sub> = 1200AM to 1230AM, t<sub>2</sub> = 1230AM to 0100AM ... .. t<sub>47</sub> = 1100PM to 1130PM.

There is a special symbol 'OFF' in the mobility data which means the card is off the wireless network. We determined the time of OFF state by the syslog message "Disauthentication" from the last associated AP with reason of "Inactivity." The Cisco AP sends this message when the card has no activity for 30 minutes. Therefore the time that the card goes to OFF is 30 minutes earlier than the time of the message. For each time interval, count is the number of connections made by wireless devices to each and every access point. A unique wireless device that is connected to the same access point in separate instances during the same time interval will increment the counter in multiple occasions.

#### IV. MARKOV MODEL IMPLEMENTATION

To set the stage for Markov models, we consider a three state model based on the number of wireless devices serviced by the AP. We assume the states low, mid and high for the

number of wireless devices ranges 0 to 5, 6 to 10 and greater than 10 respectively.

$$S = \{\text{low, mid, high}\}$$

As time progresses the AP makes transition between the low, mid and high states. Prior to the transmission of each new interval of time, the environment may change state or remain in the current state. This transition between states takes place with a set of transition probabilities,  $A_{ij}$ . These are conditional probabilities, and for our three-state model there are nine probabilities of interest. Let  $S_t$  and  $S_{t+1}$  denote the state at time  $t$  and at time  $t+1$ , respectively. With this notation, we define the four transition probabilities to be

$$\begin{aligned} A_{\text{low, low}}(t) &= \Pr \{S_{t+1} = \text{low} \mid S_t = \text{low}\} \\ A_{\text{low, mid}}(t) &= \Pr \{S_{t+1} = \text{mid} \mid S_t = \text{low}\} \\ A_{\text{low, high}}(t) &= \Pr \{S_{t+1} = \text{high} \mid S_t = \text{low}\} \\ A_{\text{mid, low}}(t) &= \Pr \{S_{t+1} = \text{low} \mid S_t = \text{mid}\} \\ A_{\text{mid, mid}}(t) &= \Pr \{S_{t+1} = \text{mid} \mid S_t = \text{mid}\} \\ A_{\text{mid, high}}(t) &= \Pr \{S_{t+1} = \text{high} \mid S_t = \text{mid}\} \\ A_{\text{high, low}}(t) &= \Pr \{S_{t+1} = \text{low} \mid S_t = \text{high}\} \\ A_{\text{high, mid}}(t) &= \Pr \{S_{t+1} = \text{mid} \mid S_t = \text{high}\} \\ A_{\text{high, high}}(t) &= \Pr \{S_{t+1} = \text{high} \mid S_t = \text{high}\} \end{aligned}$$

This can be represented by the state transition matrix

$$A(t) = \begin{bmatrix} A_{\text{low, low}} & A_{\text{low, mid}} & A_{\text{low, high}} \\ A_{\text{mid, low}} & A_{\text{mid, mid}} & A_{\text{mid, high}} \\ A_{\text{high, low}} & A_{\text{high, mid}} & A_{\text{high, high}} \end{bmatrix}$$

In the work to follow we will assume that the model is stationary and therefore the state transition matrix  $A(t)$  is fixed so that  $A(t) = A$ . However, if a simulation is performed in which the initial state probability distribution  $PI_0$  is different from the steady-state distribution, some time is required for the state distribution to evolve to the steady-state value  $PI_{ss}$ . The probability of finding the model in a given state is of interest. We define  $PI_t$  as the state probability distribution at time  $t$ . specifically,

$$PI_t = [PI_{t, \text{low}} \quad PI_{t, \text{mid}} \quad PI_{t, \text{high}}]$$



where  $PI_{t,low}$  and  $PI_{t,mid}$  and  $PI_{t,high}$  represent the probabilities of finding the AP in the low state or in the mid state or in the high state at time  $t$ , respectively. By definition of the state transition matrix, the state distribution at time  $t + 1$  is given by

$$PI_{t+1} = PI_t A$$

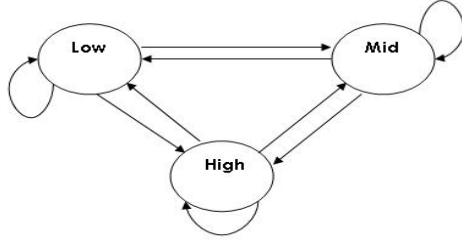


Fig.1. The transition between the three states

In a similar manner, the state distribution at time  $t + 2$  is

$$PI_{t+2} = PI_{t+1} A = (PI_t A) A = PI_t A^2$$

Thus, in general  $PI_{t+k} = PI_t A^k$

Where  $A^k$  represents the  $k$  step transition matrix. Most, but not all, Markov processes settle to a steady-state probability distribution as time evolves. Assuming that the Markov process converges to a steady-state value distribution  $PI_{t+k} = PI_t$  for sufficiently large  $t$  and arbitrary  $k$ , it follows that

$$PI_{ss} = PI_{ss} A^k = [PI_{ss,low} \ PI_{ss,mid} \ PI_{ss,high}]$$

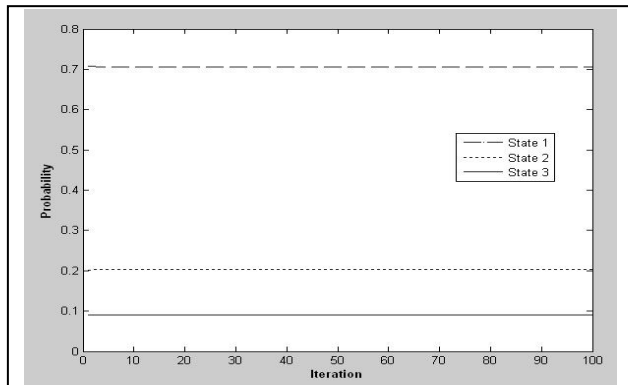


Fig.2. Steady-state probability of the three state models

The Markov process of order  $m$  (or memory  $m$ ), defined as the probability of the state at time  $t + 1$ ,  $S_{t+1}$ , is given by

$$Pr[S_{t+1} | S_t, S_{t-1}, \dots] = Pr[S_{t+1} | S_t, S_{t-1}, \dots, S_{t-m+1}]$$

Which states that the probability is dependent upon the previous  $m$  states. For a first-order Markov process, the probability of the state at time  $t + 1$  is a function only of the previous state. In other words:

$$Pr[S_{t+1} | S_t, S_{t-1}, \dots] = Pr[S_{t+1} | S_t]$$

For our three state models we will use a first-order Markov model.

Assume that the model is in state  $k$  at time  $t$  and that the next possible states are to be simulated. The first step is to generate uniformly distributed random variable,  $U1$ . The random variable,  $U1$ , is used to determine the new state. Given that the model transition matrix  $A$ . The  $k^{th}$  row of  $A$  is designated by the vector  $A_k$  where

$$A_k = [a_{k1} \ a_{k2} \ \dots \ a_{ki} \ a_{k,i+1} \ \dots \ a_{kN}]$$

The cumulative probabilities associated with  $A_k$  are denoted as  $BI_k$  and are given by

$$BI_k = \sum_{j=1}^k a_{kj}$$

The probability of making a transition from state  $k$  to state  $i$  is, by definition,  $a_{ki}$ , and is given by

$$a_{ki} = BI_i - BI_{i-1}$$

This probability is represented by the shaded area in the figure below:

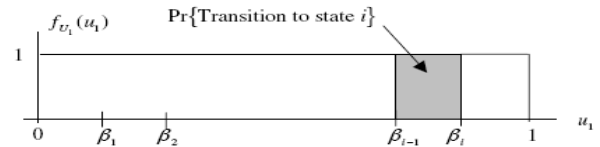


Fig.3. Determination of the transition probability

The implementation of the simulation procedure is realized using the following pseudo code:

```

A [3] [3] is the transition matrix
N is the total state
STATE ← i
FOR T ← 1 to N
 Begin
 State_Seq [T] ← STATE
 U1 ← Random (1)
 Cum_Sum ← CUMSUM (A (State, :))
 FOR I ← 1 TO Total_States
 Begin
 If U1 >= Cum_Sum(I) AND U1 < Cum_Sum(I+1)
 STATE = I
 END IF
 END FOR
 State_Seq [T] ← STATE
 END FOR

```

Fig.4. Pseudo Code for the simulation procedure

We have run the Markov O(1) predictor on the data set. We deal with the AP name ResBldg101AP1. Accuracy is measured as percentage of correct predictions of the next state of the AP. The result of the predictor's accuracy to predict the next state of AP is tabulated below based on the training data size. We vary the data size and duration to train the predictor to see how well it can predict the state of the AP. We run the predictor for first 100 days, 200 days and 300 days for different training sets. The average accuracy to predict the first 100 days data is 86%, first 200 days data is 76.44% and first 300 days data is 70.38%.

The result for the 100 days, 200 days and 300 days are plotted based on the training data size. We observe that the prediction accuracy increases, if we increase the training data set size for the first 100 days and 200 days. But for the 300 days the predictor shows a different characteristic for the increasing of training data size. For the increasing of data size (shown in TABLE II), the accuracy decreases (No-09 and No-10) and same size of data taken from different time of interval (No-10 and No-11) shows different accuracy.

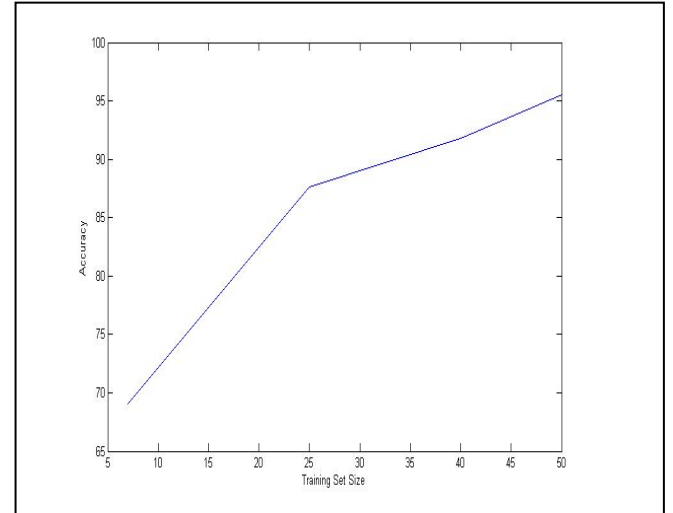


Fig.6. Markov O(1) predictor's accuracy for 200 days  
Data shown in TABLE II.

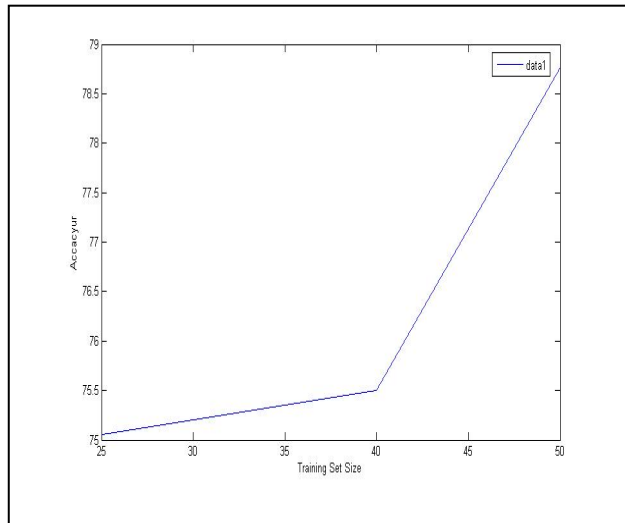


Fig.5. Markov O(1) predictor's accuracy for 100 days  
Data shown in TABLE II.

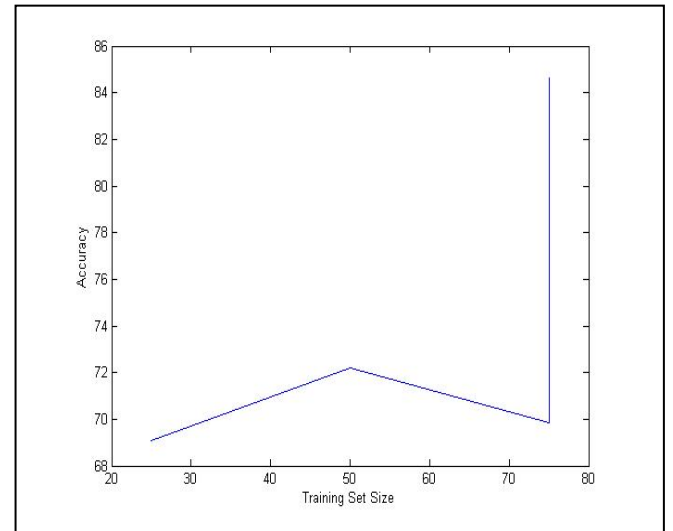


Fig.7. Markov O(1) predictor's accuracy for 300 days  
Data shown in TABLE II.

TABLE II  
TRAINING DATA, PREDICT DATA, AND ACCURACY OF MARKOV O(1) PREDICTOR

No	Training Set	Predicted Set	Accuracy
1	Day 1 to Day 7	Day 8 to Day 100	69.05%
2	Day 1 to Day 25	Day 26 to Day 100	87.61%
3	Day 1 to Day 40	Day 41 to Day 100	91.80%
4	Day 1 to Day 50	Day 51 to Day 100	95.55%
5	Day 1 to Day 25	Day 26 to Day 200	75.05%
6	Day 1 to 40	Day 41 to 200	75.50%
7	Day 1 to Day 25 and Day 100 to Day 125	Day 26 to Day 200	78.77%
8	Day 1 to Day 25	Day 26 to Day 300	69.06%
9	Day 1 to Day 25 and Day 100 to Day 125	Day 26 to Day 300	72.22%
10	Day 1 to Day 25 and Day 100 to Day 125 and Day 200 to Day 225	Day 26 to Day 300	69.86%
11	Day 1 to Day 25 and Day 100 to Day 125 and Day 226 to Day 250	Day 26 to Day 300	84.62%

#### V. CONCLUSION AND FUTURE WORK

We conclude by saying that the choice of training data creates an impact on prediction accuracy, since same size data sets taken in different interval produced an accuracy of even 95.55%. Our findings open the door for improved prediction and modeling of highly mobile users.

#### REFERENCES

- [1] R. A. Howard, Dynamic Probabilistic Systems, Vol. I: Markov Models, New York: Wiley, 1971.
- [2] W. Hsu, T. Spyropoulos, K. Psounis and A. Helmy, "Modeling Time-variant User Mobility in Wireless Mobile Networks", Proceedings of IEEE INFOCOM 2007.
- [3] C. Tudece, T. Gross, "A mobility model based on WLAN traces and its validation", Proceedings of INFOCOM 2005: Miami, FL, USA 664-674
- [4] Ehrhard Behrends. Introduction to Markov Chains. Vieweg, 1999.
- [5] H. S. Wang and P.-C. Chang, "On Verifying the First Order Markovian Assumption for a Rayleigh Fading Channel Model," IEEE Transactions on Vehicular Technology, Vol. 45, May 1996, pp. 353-357.
- [6] <http://crawdad.cs.dartmouth.edu/>
- [7] A Tutorial On Hidden Markov Model and Selected Applications in Speech Recognition, Lawrence R. Rabiner, Fellow, IEEE

# End-to-End Mobility Management Solutions for TCP: An Analysis

<sup>1</sup>Peer Azmat Shah, <sup>2</sup>Muhammad Yousaf

<sup>1</sup>Department of Computer Science Mohammad Ali Jinnah University Islamabad, Pakistan

peer\_shah@hotmail.com

<sup>2</sup>Center for Advanced Studies in Engineering (CASE), Islamabad Pakistan

myousaf@case.edu.pk

**Abstract** - Today's mobile devices are equipped with multiple network interfaces belonging to different access networks, these interfaces may be used in different requirements. A network connection should be established using the best possible interface. During communication, characteristics of a network may change or a new access network is reached by Mobile Node thus an already established connection should be moved from one interface to another or parallel transmission is done over two or more connections. In the last several years' variety of mobility management solutions have been proposed that operate at different layers of the communication protocol stack (e.g link layer, network layer, transport layer and application layer), but the network and transport layer mobility management solutions are more mature. In this paper we have discussed various end-to-end mobility management solutions for TCP. We have also identified a number of evaluation parameters and on the basis of these parameters different mobility management techniques are analyzed.

**Index Terms** - Mobility Management, vertical handover, transport layer, TCP, end-to-end

## I. INTRODUCTION

Next generation mobile communication system (4G) will provide all time access and seamless mobility across different heterogeneous networks like WLAN, WiMAX, UMTS and WWAN etc. In wireless networks, mobility of a Mobile Node (MN) can be defined as "changing the point of attachment without losing its ability to communicate in the network" with correspondent node (CN). Whenever a MN changes its point of attachment, handover occurs. Handover is divided into two types: *horizontal* and *vertical*. A horizontal handover is made between different base stations within the same link layer technology; it is usually due to the geographical movement of a MN in a network. A vertical handover means that MN hand-off from one underlying access network to another. A common goal [1] among all the approaches for mobility management is that mobility should not break the existing connections between two hosts and also the MN should not become unreachable for future connections. Mobility can be handled at different layers [1] of the traditional TCP/IP protocol stack but in any case link layer support is necessary for providing meaningful information to initiate the handover. It can be handled at Network layer (MobileIP [2] etc), Transport layer (SCTP, variants of TCP etc.) and at session layer (SIP, SLM etc.). At each of these layers mobility management schemes have strengths and weaknesses. At Network layer mobility management has several shortcomings [3], e.g. an

intermediate entity is introduced between MN and CN. These approaches (Mobile IP) do not require any changes to be made on the fixed side but require changes in the existing network infrastructure. This is the main problem as service providers are not willing to change the infrastructure due to additional deployment cost, thus end user suffers. Session layer mobility management approaches are quite favorable but they are only for real time data, so cannot be considered as the best technique. Transport layer mobility management techniques are considered strongest despite the fact that they require changes to the existing protocols like TCP. Mobility management scheme operating at any layer should have following characteristics that attempt to keep the Internet infrastructure unchanged by allowing the end hosts to take care of mobility.

- There should be no change in the network infrastructure.
- No need to maintain states in the network [4].
- Low handover latency should be involved [5].
- Direct communication should be done between MN and CN, avoiding the packet redirection [5].
- Application should remain unaware of the handover.
- Authentication should be simple and highly secure and no third parties should be involved.
- Mobility management scheme should not affect the TCP's reliability and in-order delivery features.

In the last several years variety of end-to-end mobility management schemes have been proposed that include variants of TCP for mobility, SCTP [7] based, HIP based mobility management schemes and Session Layer mobility Management (SLM). SCTP based schemes does not use the existing TCP but they are based on SCTP protocol. Host identity protocol (HIP) [8] is a new secure mobility management protocol that handles mobility by separating the dual role [9] of IP address as end point identifier and locator.

In this survey paper we have discussed various end-to-end mobility management schemes for TCP and compared them under common evaluation parameters. We have done an analysis of some selected approaches on the basis of handover delay. Rest of the paper is organized as follows:

In section II related works that has been done so far in this domain are discussed. In section III various evaluation criterions have been identified for mobility schemes. Section IV covers a detailed survey of existing solutions for end-to-end mobility management. In section V a comparison is made among different techniques and in section VI conclusions is made.

## II. RELATED WORK

Mobility management schemes have been discussed and analyzed in the literature at various layers and in varying contexts. For instance in [3] an empirical analysis of handover performance for SIP, Mobile IP, and SCTP Protocols is made but SIP and MobileIP are not based on end-to-end philosophy. In [5] a comprehensive performance analysis of Mobile IP, TCP-Migrate, and SIP is done with the derivation of mathematical model, but the analysis is not made for TCP based connections. [6] has discussed various transport layer mobility management solutions and classified them into different classes on the basis of their functionality. In [10] handover performance is compared for HIP and Mobile IPv6. Different protocols for improving TCP performance during mobility are discussed in [11]. A comparison of host mobility for IP networks using TCP-Migrate, HIP and MobileIP is done in [12] but it does not consider end-to-end mobility. All these works consider some mobility management solutions that are based on different protocols. In this paper we have discussed solutions for end-to-end mobility management for TCP based communication. A comprehensive analysis is made for each solution under some common parameters and strengths and weaknesses are identified on the basis of these parameters.

## III. EVALUATION PARAMETERS

When a MN moves from one location to another then the mobility management scheme should be able to manage communication without disruption. It should handle handover and thereafter sequenced delivery of data. We have selected a number of evaluation parameters and divided them into three classes.

- Handover related issues
- Infrastructure and Design of TCP/IP model related issues
- Performance related issues

### A. Handover related issues

In this class those parameters are included that are related to the type of handover and handover cost etc.

*Soft handover:* A soft handover allows the MN to continue communication without any disruption. Thus soft handover can be defined as make-before-break.

*Handover overhead:* During the handover some control messages have to be exchanged between the two communicating nodes. These messages are transmitted and processed by the communicating parties. The number of these control messages varies from mobility scheme to scheme, and may create overhead.

*Low layers awareness:* A mobility scheme that takes information about signal strengths etc. from the lower layers, for handover decision, is said to be Low layer aware. Usually this information is provided by the link layer through Media Independent Handover (MIH) triggers (802.21) [30].

*Willful handover:* User centric approach [13] can be said as the willful handover. The handover must be according to the requirements of the user. It is possible that there may be an

overlapping area of the two or more networks i.e. strong signals of different access technologies may be received by the MN, but user may not want to do handover.

*Handover latency:* It is the time elapsed after the MN receives the last packet in the old network until it receives the next packet in the new network.

### B. Infrastructure and Design of TCP/IP model related issues

In this class, parameters regarding the network infrastructure and the current implementation of TCP/IP model are included.

*Changes in the current implementation of TCP:* Some mobility management schemes require the changes in the current implementation of TCP means that they do not use the conventional TCP control and data signaling. These schemes may be termed as the TCP-variants for mobility management.

*Requires changes at both ends:* A mobility scheme may require changes in the communication protocol stack at either end or at the both ends.

*Support from Network:* A mobility management scheme may require support from the network for control and data signaling, e.g. use of gateways, proxies, mobility agent etc.

*Cross layer:* Cross layer optimization may be used by some approaches to handle mobility.

### C. Performance related issues

Here the parameters that regard the extra functionality and performance of handover are discussed.

*Location updates:* Whenever a MN moves into a foreign network and gets a new IP address, then the new nodes wanting to communicate with this MN would not be able to locate it. For this purpose location updates are required.

*Multi-homing:* Today's mobile devices come up with multiple network interfaces. Thus it is possible for a mobile device to have multiple IP addresses. These IP addresses may be obtained from different subnets. Using multi-homing [14] an association between MN and CN can be spanned across multiple IP addresses.

*Bandwidth Aggregation:* When a multi-homed MN has access to the coverage areas of different wireless networks, it can use multiple interfaces simultaneously for sending and receiving data [14]. Some mobility management schemes use multiple TCP connections between MN and the CN, in this situation a scheduler is required for efficient bandwidth aggregation on parallel connections.

*Simultaneous movement:* It may be possible that two communicating nodes may move simultaneously. A mobility management scheme is said to be simultaneous movement aware if it supports simultaneous movement of MN and CN.

*Application transparency:* A mobility management scheme is application transparent when applications are unaware of the handover process.

*Security issues:* Different mobility management schemes

have security flaws and/or in some schemes existing security solutions are implemented which do not meet the requirements of that scheme.

#### IV. MOBILITY MANAGEMENT SOLUTIONS FOR TCP

In this section we have discussed various mobility management solutions for TCP. Strengths and weaknesses of each scheme are highlighted. Table.1 shows the comparison among different schemes.

##### A. Freeze TCP

In Freeze TCP [15] an ongoing TCP connection is *frozen* or stopped when the signal strength is fading by advertising a zero window size to CN, thus stopping all transmission before handover. After handover, when MN has moved to some other location and reconnected, the existing TCP *frozen* connection is *unfrozen* by sending a message. As soon as the connection is re-established, MN sends Triplicate ACKs for the last data segment received before handover.

This scheme reduces the chances of packet loss during handover at the cost of higher delays. It does not require any intermediaries like proxy, gateway etc. No changes are required in TCP protocol at CN but needed at MN. It does not support multi-homing but requires support from lower layers for handover. No option for bandwidth aggregation as there are no options for multi-homing and parallel connections.

##### B. TCP-Migrate

In TCP-Migrate [16] [17] for connection migration author has proposed a *Migrate* option for TCP that allows an existing TCP connection to be migrated by either host from an old IP address to a new IP address. This TCP connection migration is done by exchanging newly introduced TCP SYN with Migrate option. In *migrate* option a *token* (identifier) is included that identifies the previously established connection. This *token* is negotiated during initial connection establishment using *migrate-permitted* option.

This scheme is a complete end-to-end mobility management solution but the handover involved is hard handover. Just like *Freeze-TCP* the connection is stopped. It does not require any changes in the network infrastructure but TCP at both ends needs to be changed. No option for multi-homing and bandwidth aggregation. TCP-Migrate handles encrypted traffic and can be implemented with the existing security solutions. This scheme lacks simultaneous movement support and does not preserve location privacy. An enhanced version of TCP-Migrate with simultaneous movement support is proposed in [18].

##### C. TCR-Redirection

For continuous operation on Mobile Node, TCP-R [19] revises the pair of IP addresses and port numbers in the existing connection, whenever there is a change in the MN's IP address. For new connection establishment TCP-R performs two operations:

1. check whether CN is TCP-R aware

2. exchange authentication keys

After getting a new IP address, MN sends a *redirect* message to CN. On receiving the *redirect* message CN sends authentication request to MN. MN replies with *connection-authentication* message. CN checks the connection authenticator, if found correct then revises the pair of addresses of existing TCP connection. In the meanwhile MN also revises its pair of addresses.

The scheme can be implemented as a complete mobility solution and is application transparent. It does not require any support from lower layers for connection migration. DNS is proposed for location management. TCP-R does not support multi-homing and bandwidth aggregation.

##### D. VA-TCP

Vertical handoff Aware-TCP (VA-TCP) [20] is an end-to-end mobility management scheme in which a new abstraction layer is inserted between Network layer and Transport layer for connection continuity. This new layer adds some new messages to support vertical handover and also maintains a Local Connection Translation (LCT) table for maintaining a relationship between original connection information and the new connection. When a MN roams in a new network, then VA-TCP dynamically estimates the connection parameters from the information provided by the lower layers, so that it can fit into the new environment.

VA-TCP is designed to achieve TCP's best performance for end-to-end mobility management that requires changes in the protocol stack by introducing an abstraction layer. Low layer triggers are required for handover initiation. No option for multi-homing and thus no aggregate bandwidth management is required. Handover is not done according to user requirements but is done according to the new network conditions.

##### E. MSOCKS

Maltz et.al proposed MSOCKS [21], Transport layer mobility management architecture based on TCP Splice. MSOCKS is built around a proxy inserted in the path of MN-CN. Connection from MN to CN is split at proxy and two interfaces are made at proxy. Each interface has its own IP address. Scheme introduces a RECONNECT message that is sent to the Proxy upon getting a new IP address in a new network. MSOCKS enabled MN creates a new socket and sends a RECONNECT message to Proxy with connection ID of old session. Proxy establishes new connection and sends a reply message to MN. Upon receiving reply MN resumes transmission.

Mobility in MSOCKS is limited to the coverage area of Proxy as location management is done by Proxy. It also needs low layer information for handover initiation. A node may have multiple IP addresses but communication is done only through one interface so no option for bandwidth aggregation. MSOCKS requires changes in the current implementation of TCP but only at the MN and Proxy. This protocol cannot work with IP layer cryptographic techniques like IP-SEC.

### F. Tsukamoto's scheme with Multiple TCP Connections

Tsukamoto proposed a new mobility management scheme that modifies the transport protocol for handling multiple connections [22] during handover process. A cross layer approach is used that detects any changes in the status of wireless networks and selects the best optimal interface. A Handover Manager (HM) is implemented at transport layer that detects the changes in the status of wireless networks. Link layer information is sent to the HM at transport layer bypassing the intermediate layers. HM using this information can detect a link-up or link-down state of a wireless network. If during communication HM detects a link-up state for a new wireless network on a new interface, then it starts a new TCP connection via that interface and starts parallel transmission using two TCP connections. For each connection end-to-end performance measures are calculated and compared. On the basis of this comparison handover takes place.

In this scheme no details of the control and data flow are given. It is not mature for implementation and a lot of work needs to be done. No provision of willful handover is there. In this scheme bandwidth aggregation can be possible but author has not given any details about this.

### G. pTCP

pTCP [24] [25] is a transport layer solution for multi-homed MN. It creates and maintains one TCP state for each network interface. *TCP-virtual* (TCP-v) pipes are created for each active interface in a pTCP connection. Whenever a new pTCP socket is opened by application for a connection then a TCP-v pipe for that connection using active interface is created. Application in the MN opens a pTCP socket with CN and pTCP creates the first TCP-v pipe. This pTCP with only one pipe behaves as the same as TCP. When MN moves to a new network and gets a new IP address then pTCP creates a new pipe but does not closes the previously

established pipe.

pTCP provides the reliable and sequenced delivery just like the simple TCP. Both ends need to be pTCP aware. No support is required from the network like proxy etc. It does not support willful handovers but bandwidth aggregation is performed [26] for effective throughput.

### H. I-TCP

Indirect TCP (I-TCP) [23] proposes to split the TCP connection at the mobility support station that is typically a gateway/MSR. Two TCP connections are used, one is an I-TCP connection between MN and gateway and the other is regular TCP between gateway and CN. When a MN moves and gets a new IP address then only the TCP connection between MN and gateway is changed and replaced by the new TCP connection. Hard handover is done and protocol requires changes at the MN. No option for multi-homing is available. Network support is required with gateway/MSR.

### I. SLM

Session Layer Mobility management (SLM) [28] is a framework that supports mobility management at session layer. It extends the TCP/IP model by adding session layer. This layer operates above TCP and switches data streams between multiple connections. The end-to-end network path is divided into three separate paths. Two paths are between the applications on the two hosts and socket connectors and the third path between the socket connections of two hosts. Location management is done through User Location server (ULS) in the new network.

In this scheme both ends need to be SLM aware and support is required from the network in the form of proxy / User Location Server (ULS). Multi-homing is available and bandwidth aggregation can be done. This scheme also supports simultaneous movement of two nodes. No option is available for willful handover.

TABLE. 1 Comparison of different mobility management techniques  
Y: Yes N: No

Evaluation Parameters	Freeze-TCP	TCP-Migrate	TCP-R	VA-TCP	MSOCKS	Tsukamoto TCP	pTCP	I-TCP	SLM	HIP based	EMF
Low layers aware	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Willful handover	N	N	N	N	N	N	N	N	N	N	Y
Requires changes in TCP	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N
Changes required on both sides	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y
Multi-homing	N	N	N	Y	Y	Y	Y	N	Y	Y	Y
Bandwidth aggregation	N	N	N	N	N	N	Y	N	Y	N	Y
Requires support from network	N	N	N	N	Y	N	N	Y	Y	N	N
Cross layer approach	N	N	N	N	N	Y	N	N	N	N	Y
Location updates	N	Y	Y	Y	Y	Y	Y	N	Y	N	Y
Application transparency	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Handover	Hard	Hard	Hard	Soft	Hard	Soft	Soft	Hard	Soft	Soft	Soft
Security issues	N	N	N	N	Y	N	N	Y	N	N	N
Handover overhead	N	N	N	N	N	N	N	N	N	N	N
Handover latency	Y	Y	Y	N	Y	N	N	Y	N	N	N
Simultaneous movement aware	N	N	Y	Y	N	N	Y	N	Y	Y	Y

### J. HIP base mobility management techniques

In current Internet IP address is used as both topological locator and an end point identifier. Host Identity Protocol (HIP) [8] separates the dual role of IP address as end point identifier and locator. It introduces a Host Identifier (HI) used as end point name at transport layer and IP address is used as locator at network layer. This is termed as a new layer - Host Identity Layer, which is seen as a 3.5 layer [9].

In this approach multi-homing is allowed but parallel transmission on two links is limited to the fact that when there is data flowing on one link, control packets for handover flow on other link. Therefore no explicit mechanism for bandwidth aggregation has been proposed so far. Willful handover is not considered and both end need to be HIP aware.

### K. EMF

End-to-end Mobility management Framework [29] is a new idea proposed for continuous transmission of data on multiple links between MN and CN. In this scheme cross layer architecture is proposed in which Host Agent gets information from lower layers through MIH triggers [30] and passes to the User Agent on application layer and to Control Agent on transport layer. Host Agent takes handover decision on the basis of information provided by User agent and lower layers. Data handler module is present that is responsible for soft handover and in-order delivery of packets from different TCP connections to the application. In each TCP connection traditional TCP provides reliable data delivery, congestion control and flow control mechanisms. Multiple TCP connections are established under a single EMF association for transmission of single stream of data between MN and CN.

In this approach no support is required from the network but changes are to be made on the both sides in transport layer. Willful handover is also possible that makes this approach unique from others. This scheme inherently provides bandwidth aggregation but no details are given.

## V. ANALYSIS OF DIFFERENT APPROACHES

In any mobility scheme handover delay/latency and throughput degradation are the major factors to be determined. Handover latency depends upon number of handover control messages, their number of retransmissions, and RTT. Throughput degradation depends upon handover latency and the time for TCP slow start.

### TCP slow start time:

Assuming no packet loss and TCP doubles its congestion window on each RTT, [5] has defined the time for TCP to reach steady state as:

$$T_s = [\log_2(CW_s + 1)] RTT \quad (1)$$

Here

$CW_s$  = Congestion window at steady state

$T_s$  = Time to reach steady state

### Handover latency:

Assume that three control messages are being transmitted for handover. Message1 (MSG1) with end-to-end packet loss probability  $p_1$  is sent  $(i - 1) \geq 0$  times unsuccessfully and  $i^{\text{th}}$  time successfully. MSG2 with end-to-end packet loss probability  $p_2$  is sent  $(j - 1) \geq 0$  times unsuccessfully and  $j^{\text{th}}$  time successfully and MSG3 with end-to-end packet loss probability  $p_3$  is sent  $(k - 1) \geq 0$  times unsuccessfully and  $k^{\text{th}}$  time successfully and on each retransmission doubling the RTO, handover latency for control messages as explained by [5] is

$$Latency_{handover} = 1.5RTT + \sum_{a=0}^{i-1} 2^a RTO + \sum_{b=0}^{j-1} 2^b RTO + \sum_{c=0}^{k-1} 2^c RTO$$

Or it may be written as

$$Latency_{handover} = 1.5RTT + RTO(2^i + 2^j + 2^k - 3) \quad (2)$$

Total handover delay will be:

$$T_{handover} = Latency_{handover} + T_{L_2} + T_\alpha \quad (3)$$

Where  $T_{L_2}$  is the time required for layer2 handover,  $T_\alpha$  is the time required for the acquisition of new IP address in the new network.

### Throughput degradation time:

It is the sum of total handover delay and time to reach steady state for TCP. From eq (1) and (3) as defined in [5]

$$T_{Thput Deg} = T_s + T_{handover} \quad (4)$$

To compare the performance of different approaches, we consider the following values for different parameters as defined in [6]:  $T_{L_2} = 10$  ms  $T_\alpha = 20$  ms.  $T_{L_2}$  and  $T_\alpha$  are defined in (3). End-to-end packet transportation delay = 100 ms, then  $RTT = 2 \times 100 = 200$  ms. RTO can be taken as  $RTO = 1.5 \times RTT$  i.e.  $RTO = 300$  ms. We plotted the graph for handover latency by using different number of retransmissions of handover control messages. In Fig.1 a comparison is made for schemes having three control messages (pTCP, TCP-Migrate, HIP) and two control messages (VA-TCP).

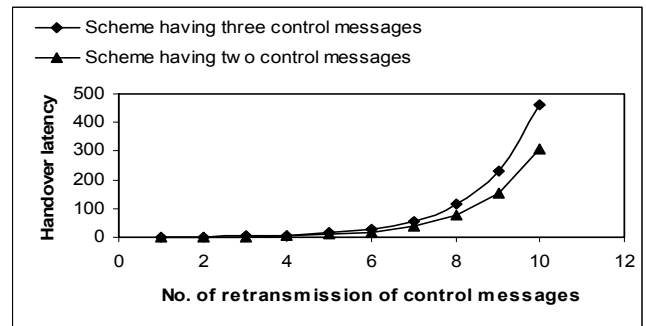


Fig.1. Handover Latency comparison for schemes having two and three control messages

The results show that greater the number of control messages a scheme has, greater will be the handover latency and vice versa. Similarly in Fig.2 a comparison is made for a single technique (pTCP) by doubling the RTT and RTO values and



found that greater values of RTT and RTO leads to longer handover delay.

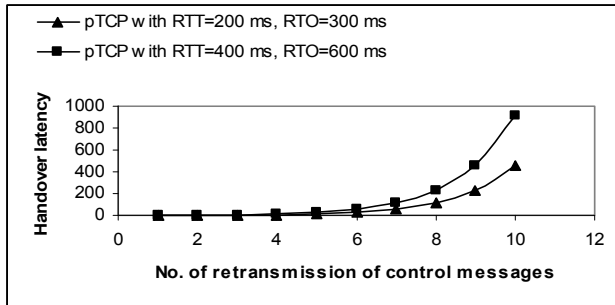


Fig.2. Handover Latency comparison with respect to RTT and RTO

Throughput degradation occurs when there is no provision for parallel TCP connections i.e. for handover initiation connection is stopped or freezed and whenever data is sent on new interface TCP will start from slow start and will take time  $T_s$  as defined in (1) to reach at steady state thus resulting in throughput degradation. Whereas in schemes like pTCP, EMF data can be transmitted over multiple interfaces enjoying the capability of bandwidth aggregation, thus there is no throughput degradation.

## VI. CONCLUSION

In this paper we have surveyed various mobility management solutions for TCP and evaluated them under common parameters. Each technique has strengths and weaknesses. A tabular comparison is made highlighting these strengths and weaknesses. Finally an analysis of some approaches is done on the basis of handover delay and throughput degradation and it is found that handover latency depends upon the number of control messages an approach uses for handover initiation, the RTT and RTO values and the link conditions. Similarly throughput degradation occurs when there are no options for bandwidth aggregation. Solutions that can make use of parallel TCP connections between communicating nodes, like pTCP, EMF, Tsukamoto's TCP, should have lower handover latency and there should be minimum throughput degradation time for such schemes. In our opinion a solution for mobility management will be comprehensive if it supports multi-homing with bandwidth aggregation option and is application transparent. It may take low layer support for handover and should provide willful handover with simultaneous movement of MN and CN having low handover latency and no security issues. Also solution must not depend upon network infrastructure (proxies, gateways, MSR).

## REFERENCES

- [1] Wesley M. Eddy "At What Layer Does Mobility Belong?" Communications Magazine, IEEE, Oct. 2004, Vol: 42 pp: 155- 159
- [2] Charles E. Perkins "Mobile Networking Through Mobile IP" Internet Computing January/February 1998 (Vol. 2, No. 1) pp. 58-69
- [3] Sherali Zeadally, Farhan Siddiqui "An Empirical Analysis of Handoff Performance for SIP, Mobile IP, and SCTP Protocols" Wireless Personal Communications (2007) 43:589-603
- [4] Christopher Ross "Mobility in the Transport Layer Mobility Without Touching the Network" Siemens Mobile Networks magazine
- [5] Shantidev Mohanty, Ian Akyildiz "Performance Analysis of Handoff Techniques Based on Mobile IP, TCP-Migrate, and STP" IEEE Transaction on computing, July 2007 (Vol. 6,) pp. 731-747
- [6] Atiquzzaman, M. Reaz, A.S "Survey and classification of transport layer mobility management schemes" PIMRC 2005. page(s): 2109-2115 Vol. 4
- [7] R. Stewart et. al. RFC 2960 "Stream Control Transmission Protocol" October 2000
- [8] R. Moskowitz RFC 4423 "Host Identity Protocol (HIP) Architecture"
- [9] Joseph, Jidong Wang "HIP Based Mobility Management for UMTS/WLAN Integrated Networks"
- [10] Jokela, P. Rinta-aho, T. "Handover performance with HIP and MIPv6".Wireless Communication Systems, 2004 page(s): 324- 328
- [11] H. Elaearg "Improving TCP performance over mobile networks" ACM Computing Surveys Volume 34 , Issue 3 Pages: 357- 374, 2002
- [12] T. R. Henderson "Host Mobility for IP Networks: A Comparison" Network, IEEE Nov.-Dec. 2003 Volume: 17, Issue: 6 pp: 18- 26
- [13] A. Calvagna, G. D. Modica "A user-centric analysis of vertical handovers" Proceedings of the 2nd ACM international workshop on Wireless mobile applications and services on WLAN hotspots Philadelphia, PA, USA Pages: 137 - 146
- [14] K. H. Kim, Y. Zhu "A receiver-centric transport protocol for mobile hosts with heterogeneous wireless interfaces" Wireless Networks Volume 11, Issue 4 (July 2005) Pages: 363 - 382
- [15] T. Goff "Freeze-TCP: A true end-to-end TCP enhancement mechanism for mobile environments" INFOCOM 2000. page(s): 1537-1545 vol.3
- [16] Alex C. Snoeren and Hari Balakrishnan "An End-to-End Approach to Host Mobility" 6th ACM/IEEE MobiCom 2000
- [17] Snoeren, A.C. Balakrishnan, H. Kaashoek, M.F. "Reconsidering Internet mobility" This paper appears in: Hot Topics in Operating Systems, 2001. May 2001 page(s): 41- 46
- [18] Yi Wu; Yanqun Le; Dongmei Zhang "An Improved TCP Migrate Scheme with DNS Handover Assistant for End-to-End Mobility" ICC apos:07. 24-28 June 2007 Page(s):1923 - 1928
- [19] Funato, D. et al. "TCP-R: TCP mobility support for continuous operation" Network Protocols, 1997. Page: 229-236
- [20] Yu-Chieh Lin "VA-TCP: a vertical handoff-aware TCP" Symposium on Applied Computing 2007 Seoul, Korea, Pages: 237 - 238
- [21] Maltz, D.A. Bhagwat, P. "MSOCKS: an architecture for transport layer mobility" INFOCOM '98, page(s): 1037-1045 vol.3
- [22] Kazuya Tsukamoto et al. "Vertical Handover Management Scheme Using Multiple TCP Connections for Heterogeneous Networks" WiNTE CH'07, September 10, 2007, Montr éal, Q uébec, Canada.
- [23] A. Bakre "I-TCP: indirect TCP for mobile hosts" Proceedings of the 15th International Conference on Distributed Computing Systems
- [24] Hung-Yun Hsieh "An end-to-end approach for transparent mobility across heterogeneous wireless networks" Mobile Networks and Applications Volume 9 , Issue 4 (August 2004) Pages: 363 - 378
- [25] H. Y. Hsieh and R. Sivakumar "pTCP: An End-to-End Transport Layer Protocol for Striped Connections" Proceedings of 10th IEEE International Conference Network protocols Nov. 2002 pages: 24- 33
- [26] Hung-Yun H. and R. Sivakumar "A transport layer approach for achieving aggregate bandwidths on multi-homed mobile hosts" Wireless Networks Volume 11 , 2005 Pages: 99 - 114
- [27] H. Sivakumar, S. Bailey, and R. L. Grossman, "Psockets: The case for application-level network striping for data intensive applications using high speed wide area networks" in Proceedings of Supercomputing 2000. Dallas, TX: November 2000, pp. 240-246.
- [28] Landfeldt, B.; Larsson, T.; Ismailov, Y.; Seneviratne, A. "SLM, a framework for session layer mobility management" Computer Communications and Networks, 1999. Proceedings. Eight International Conference Volume , Page(s):452 - 456
- [29] M.Yousaf and A. Qayyum "On end-to-end- Mobility Management in 4G Heterogeneous Wireless Networks" Proceedings of IEEE 2<sup>nd</sup> International Networking and Communication Conference (INCC-2008) pp- 118-123 May 2008.
- [30] George Lampropoulos et.al "Media-Independent Handover for Seamless Service provision in Heterogeneous Networks" IEEE Communication Magazine January 2008.

# Peer-to-Peer Mobile Applications Using JXTA/JXME

Tahmina Tahsin, Lazeeb Faiz Choudhury, Md. Lutfur Rahman

Department of Computer Science and Engineering, University of Dhaka, Dhaka – 1000, Bangladesh

**Abstract** — With the advancement in mobile wireless communication technology and the increasing number of mobile users, Peer-to-Peer (P2P) computing, in both academic research and industrial development, has found a new dimension of communication, collaboration and resource sharing. In the recent years P2P have evolved rapidly in internet especially for file-sharing. The major issue of this popularity is that no longer the networks are constrained on unreliable central servers; sharing and connectivity is the main focus. The popularity of Peer-to-Peer file sharing applications such as Gnutella, KaZaA, or Napster has created an outbreak in recent internet history. But P2P File Sharing applications have not yet been widely adopted in mobile devices. In this paper, possibilities of P2P applications in mobile devices are presented. To illustrate the possibilities two demo applications are presented here. One is a Message Passing application and another is a File Sharing application. The open source JXTA/JXME framework is used here.

**Index Terms** — P2P, mobile applications, JXTA, JXME, file sharing, JSR75 API.

## I. INTRODUCTION

P2P networks are typically used for connecting nodes via largely ad-hoc connections. Such networks are useful for many purposes. Sharing content files which contains audio, video, data or anything in digital format is very common, and real time data, such as telephony traffic, is also passed using P2P technology [1]. In the same manner that the Internet provides domain name lookup (DNS), World Wide Web, email, and other services by spreading responsibility among millions of servers, P2P has the capacity to power a whole new set of robust applications by leveraging resources spread across all corners of the Internet [2].

Mobile phones are now-a-days evolved into personal digital assistants. People have been very dependent on the information available on the Internet and with the rapid evolution of mobile devices during the last years. Everywhere people go today, they also bring their mobile phones. Because of this the demand for new applications and services are growing. To meet the demand, Sun Microsystems Inc. introduced J2ME which is a highly optimized Java runtime environment intended for small wireless handheld devices with little CPU, memory and screen resolution.

On the other hand, Project JXTA is an open source project that defines a set of open, generalized, P2P protocols, defined as XML messages, which allows any connected device on the network – from mobile phones and

wireless PDAs, to PCs and servers – to communicate and collaborate in a P2P manner. Because of the explosion of the mobile devices, Project JXTA designed JXTA for J2ME, also known as JXME [3].

Almost all of the mobile phones today use GPRS which provides data rates from 56 up to 114 kbit/s [4]. The most recent mobile phones also support EDGE and 3G data. With EDGE/GPRS the mobile phone is always connected to a network, and ready to receive and send data. For the nature of mobility, these mobile devices can be connected or disconnected at any time from any network and may join in another network. P2P is well suited in these environments. Information is stored on several peers, so it does not matter if some peers get disconnected. With P2P we are able to develop instant messaging applications for mobile phones which might cost the price of using a P2P IM application instead of sending SMS (it depends on the charge applied by carriers for transferring data and SMS). The same idea goes with MMS, which is relatively expensive today, and it might be cheaper to have a P2P application that does the same thing [3].

With the advancement of technology mobile devices are in a rapid growth and these mobile devices are capable of communicating with other mobile devices. So in our research we wanted to show some methods to make these kinds of communications more efficient and more cost effective. And we were motivated by the limitation that a user has to depend on only to the services provided by mobile phone's service providers. In this research, the implementation part is focused mainly on the GPRS/EDGE communication channel.

## II. JXTA/JXME PLATFORM

### A. Background

The current applications of P2P are incompatible with one another. They use their own protocols and form their own community while operating and are incapable of communicating with other applications. So these applications cannot leverage their services together. To evolve P2P into a mature solution platform, developers need to refocus their efforts from programming P2P network fundamentals to creating P2P applications on a solid, well-defined base. To do this, P2P developers need a common language to allow peers to communicate and perform the fundamentals of P2P networking [5].

Realizing this need for a common P2P language, Sun Microsystems Inc. formed Project JXTA, a small development team under the guidance of Bill Joy and Mike Clary, to design a solution to serve all P2P applications. The name “JXTA” is not an acronym. It is short hand for juxtapose, as in side by side. It is a recognition that P2P is juxtaposed to client-server or Web-based computing, which is today’s traditional distributed computing model [6]. At its core, JXTA is simply a set of protocol specifications. Anyone who wants to produce a new P2P application is spared the difficulty of properly designing protocols to handle the core functions of P2P communication. The JXTA Protocol Specification states that JXTA does not make any assumption about communicating devices. In addition to operating systems, programming language and computing system’s independency, JXTA is not dependent on network transport mechanism as well. JXTA assumes that peers and their resources may appear and disappear from the network spontaneously and peer’s location can be changed radically or be masked by Network Address Translation (NAT) or firewall equipment [7]. Apart from the specifications, JXTA makes some recommendations also. JXTA recommends that peers cache some information to reduce network traffic and provide message routing to peers that are not directly connected to the network.

### *B. Justification of Using JXTA*

The vision of Project JXTA is a world where each peer, independent of software and hardware platform, has the benefit of being connected to millions of other peers. Project JXTA was initiated to standardize a common set of protocols for building P2P applications and was designed to allow a wide range of devices to communicate in a decentralized manner.

Project JXTA has a set of objectives to address the shortcomings in the already existing, or under development, P2P systems [8]:

- Interoperability
- Platform independence
- Ubiquity
- Security and monitoring

**Interoperability:** P2P systems of today are designed to deliver a particular service, such as Napster provides music file sharing, Gnutella provides generic file sharing and AIM provides instant messaging. But as they use their own protocols to develop applications, they are incompatible with one another. Project JXTA aims to bring to the P2P world what HTTP and the browser brought to the Internet [7, 9].

**Platform Independence:** P2P applications today are meant for a particular operating system. Software vendors are compelled to generate different version of the same application for different operating systems or make some kind of bridge between the platforms. JXTA technology is designed to be embraced by all developers, independent of

preferred programming languages, development environment, or deployment platform [3].

**Ubiquity:** The Project JXTA protocols are designed to be implemented on any device with a digital heartbeat, which include sensors, consumer electronics, PDAs, appliances, network routers, desktop computers, data-centre servers and storage systems [3, 8].

**Security and monitoring:** For commercial and enterprise deployment, JXTA provides proper security and monitoring services. JXTA peers operate in a role-based trust model, in which an individual peer acts under the authority granted to it by another trusted peer to perform a particular task [8].

### *C. JXTA Components*

Several JXTA components are mentioned below:

- Peers
- Peer Groups
- Virtual Network Overlay
- Uniform Peer Addressing
- Message Relaying
- Messages
- Advertisements
- Universal Resource Binding and Rendezvous
- Pipes

### *D. JXTA Protocols*

JXTA uses several protocols for communication purpose. All of these are asynchronous and are based on query/response model. Using these protocols peers can discover one another or can discover network resources.

Peers do not require to implement all the protocols, they just implement what they need [6].

Six protocols of JXTA are mentioned below:

- The Peer Endpoint Protocol
- The Rendezvous Protocol
- The Peer Resolver Protocol
- The Peer Discovery Protocol
- The Peer Information Protocol
- The Pipe Binding Protocol

### *E. JXTA for J2ME*

Because of the rapid explosion of wireless devices, Project JXTA designed its stripped down version named JXME which lets MIDP compliant devices participate in Peer-to-peer activities with JXTA peers running on larger platforms. Two important ideas in understanding JXTA/JXME are:

- The JXTA Relay
- The JXME API

**The JXTA Relay:** As mobile devices are limited in processing power, most of these devices cannot handle most of the requirements of a regular JXTA peer. As for example these devices cannot parse large XML messages, so some extra component called JXTA Relay is needed to act on

behalf of these low profile devices. JXME uses the JXTA relay as a support to help these devices to become a peer [10].

JXTA relay acts as a junction between JXTA network and peers. The JXTA relay:

- Receives commands from a J2ME client, performs what is necessary on the client's behalf
- Represents a JXME client to the JXTA network
- Provides interoperability with JXTA protocols
- Filters JXTA traffic
- Acts as a proxy and trims the advertisements

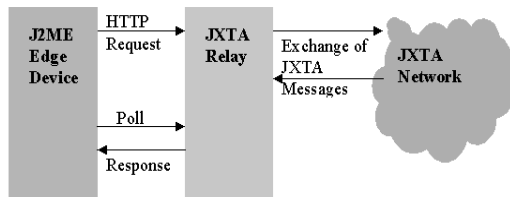


Fig. 1. JXTA Relay

Different JXME peers connected to different JXTA relays can still discover each other and communicate. They are not required to establish and maintain a static relationship with one specific JXTA relay [3]. It is worthwhile to mention that a JXTA relay is not just meant to serve J2ME clients. A relay can serve any client that can communicate according to the specified protocols [10].

**The JXME API:** There are just three classes in the JXME API [3]:

- Message class: Include methods to create and manipulate JXTA messages
- Element class: Include methods to construct and manipulate the basic components of JXTA messages
- PeerNetwork class: Include operations that can be invoked on the JXTA network

### III. IMPLEMENTATION OF MESSAGE PASSING AND FILE SHARING APPLICATIONS

#### A. Components of the Application

The message passing and file sharing application has the following components.

- Graphical User Interface
- The Online Presence Service
- The Message Passing service
- The File Sharing service
- The JXTA Relay Server

To build the file sharing service, JSR 75, an extra API is used. JSR 75 is used to take advantage of features commonly found on PDAs in the J2ME space. Here the

FileConnection Optional Package (FC) API gives J2ME devices access to file systems residing on mobile devices.

#### B. Setting up the Environment

At first we need to install Java 2 runtime environment, and the Java2 software development kit. Then we install the Java 2 Micro Edition Wireless Toolkit. In our research, we worked on Netbeans editor. In this research stable 2.1.3 version of JXME was used. To set up a personal computer as a JXTA Relay, the first thing that must be done is to install the JXTA Shell. The JXTA shell can be downloaded from JXTA's homepages. After the shell is configured we log onto the shell and the JXTA relay we just created is active and ready for use.

#### C. Explanation of Functions

Some important built-in functions are used when developing the applications. Several useful methods in the PeerNetwork class are mentioned below [10]:

- createInstance() is a factory method that returns an instance of PeerNetwork with specified peer name.
- connect() method connects to a relay at a specified HTTP URL.
- create() method creates peers, pipes, groups on the JXTA network through the relay proxy.
- poll() method polls the relay for messages addressed to this mobile peer.
- listen() and close() methods open and closes the pipes respectively.
- send() method sends a message to a specified pipe.

A typical application will perform the following actions in response to calls to its MIDlet methods:

- startApp() is called by the application manager when the MIDlet is started and contains statements that are executed each time the application begins execution.
- The pauseApp() method is called before the application manager temporarily stops the MIDlet.
- destroyApp() method is called prior to the termination of the MIDlet by the application manager.
- commandAction() method is listening when the user makes a choice.
- run() method is started when a peer is created. This is an essential method when implementing the Runnable interface.

#### D. Message Passing

Target of this application is to provide presence service that is to show who are online at a certain period and to provide P2P message passing. A peer sends an "alive message" at a regular interval. When a peer does not get alive message from a particular peer for about 10 seconds, it will consider the peer as dead and remove all its informations. After a while when the peer again receives

alive message from that particular peer, it again adds all relevant informations of that peer. A peer can send personal message to another peer whenever it wishes. Steps of this application are below:

- Connect to server
- Create propagate pipe
- Create own unicast pipe
- Listen to both, poll server
- Send alive and personal messages

Screen shot of this application is show in Figure 2.

#### E. File Sharing

Target of this application is to provide presence service like message passing application and to provide file sharing upon demand of a peer. Two screen shots are shown in Figure 3 and 4. At first a peer should go to online. While online, a user can search for a file finding other peers connected to the JXTA relay. Next few steps are below:

- The client peer sends a query message with a file name to the network.
- All peers on the networks receive the message and search its local file system to see whether it contains the file.
- If it contains, it sends a positive reply to the client; otherwise negative reply is sent.
- After getting the reply, the client will request the peer who positively replied first to send the file.
- That peer will send the requested file to the client peer directly.

#### F. Test Results

##### Message Passing

All the network connections and peer creation were completed successfully. “Alive” and “p2p” messages were sent and received without any problem.

##### File Sharing

Network connections, peer creation all were done without any problem. File Connection Package of JSR75 API worked flawlessly. But sometimes the positive/negative reply of the query message is lost in the network for some unknown reason. This may happen because JXTA/JXME is still a developing project. As the reply of the query is lost, all consecutive messages that were to be propagated after the reply were not generated all the times.

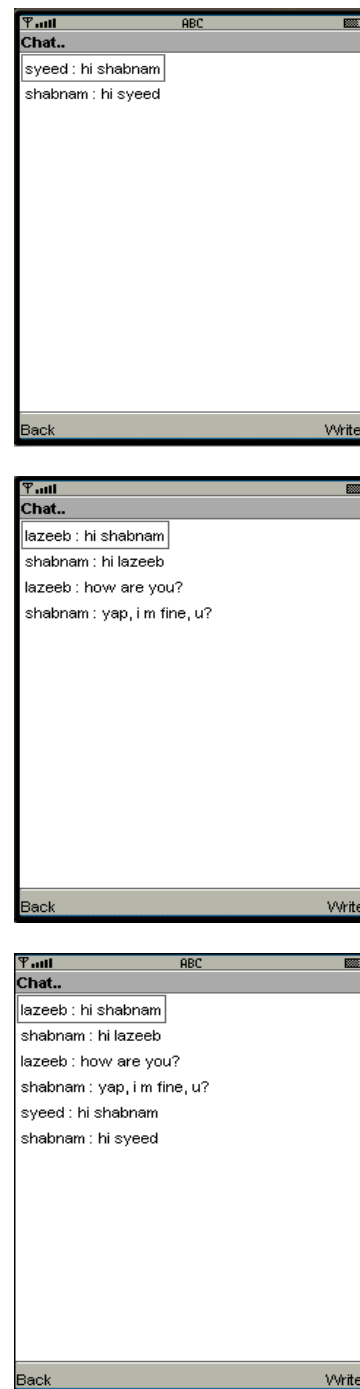


Fig. 2. Chat Screen



Fig. 3. Shared Files



Fig. 4. Search Screen

#### G. Future Developments

The P2P applications developed during this research were developed to show that it is possible to develop P2P applications for mobile devices. The developed applications are very primitive and can easily be improved to have better value to a potential user. P2P Mobile Applications can be developed for transferring not only messages and files but also voice data. So that, for example, if a user of a mobile phone is out of service and needs to make an emergency phone call, a P2P software can search for other mobile phones in the vicinity using Bluetooth or WiFi technology and establish a P2P connection to a nearby device which has an active network coverage. A P2P application like Gnutella or Napster for mobile can also be developed which may take advantage of media sharing via the internet or even Bluetooth or Wifi. The possibilities of P2P applications are endless and with mobile phones increasing in power and memory every day, the future of mobile is perhaps in P2P Mobile Applications.

#### IV. CONCLUSION

As the number of wireless mobile devices used in everyday life increases exponentially, it is essential to explore the feasibility of various mobile applications and their limitations. During this research, two P2P applications for mobile have been partially developed using JXME to show that it is possible to develop useful P2P applications for mobile devices. The first application that was developed, a P2P Message Passing Software for mobile devices, is a practical implementation of chatting software independent of any client-server architecture. This message passing software demonstrates the message passing capability of JXTA and establishes a building block for the File Sharing Software developed later. This application can be practically used today as a replacement or complement to SMS. The second application is a P2P File Sharing Application for mobile devices. This application shows the potential of transferring large files between mobiles using P2P. The aim of this application was to see if it is possible to develop a P2P application for mobile devices capable of searching and transferring large files without going through a central server. With the power and storage capacity of mobile devices increasing exponentially, this software may hold the future for mobile computing. The research demonstrates that with further development and research work, various P2P applications developed using JXTA might become practical applications for everyday use.

#### ACKNOWLEDGEMENT

We would like to thank the JXTA/JXME community to provide us with useful information and for answering our query threads several times. Without their support it would be impossible for us to work in such a new arena. We would also like to thank all our family members, teachers and batch mates for their supports and encouragements.

#### REFERENCES

- [1] Wikipedia-the free encyclopedia, "Peer-to-Peer", <http://en.wikipedia.org/wiki/Peer-to-Peer> (accessed on 10-07-2008)
- [2] "CM Walker, Leslie. Uncle Sam Wants Napster! The Washington Post, Nov 8, 2001", <http://www.washingtonpost.com/ac2/wp-dyn?pagename=article&node=washtech/techthursday/columns/dotcom&contentId=A59099-2001Nov7> (accessed on 10-07-2008)
- [3] Jan-Erik Romøren, "Developing Peer-to-Peer Applications for Mobile Handheld Devices", *Faculty of Engineering and Information Technology Griffith University, Gold Coast Campus*; June 2003
- [4] Wikipedia-the free encyclopedia, "GPRS", <http://en.wikipedia.org/wiki/GPRS> (accessed on 01-07-2008)
- [5] Official JXTA Website <http://jxta.dev.java.net/>

- [6] *JXTA Java™ Standard Edition v2.5: Programmers Guide*; September 10th, 2007
- [7] Overview of JXTA, Sams Publishing, [http://www.developer.com/java/other/article.php/10936\\_1450221\\_1](http://www.developer.com/java/other/article.php/10936_1450221_1) (accessed on 10-07-2008)
- [8] Juan Carlos Soto, *Project JXTA: An Open P2P Applications Platform, Introduction and Update*, Project JXTA
- [9] Brendon J. Wilson, *JXTA*, June 2002
- [10] Troy Espiritu, "Universal Chat Client for J2ME Enabled Mobile Devices Using the JXME Implementation of JXTA", July 2007

# Connecting Telephone Exchange Centers over Metro Ethernet Networks by using Diffserv-MPLS

Ghasem Mirjalily, Parisa Bastani, and S.M.T. Almodarresi

Faculty of Electrical and Computer Engineering, Yazd University, Yazd, Iran

**Abstract** — Recent large deployment of high-speed metro Ethernet networks creates opportunity to connect legacy telephone exchange centers and cellular telephone base stations to the core telephony network. The success of data and legacy voice services convergence shall depend on whether it can provide the customer with the required QoS or not. Packet-base telephony service requires stringent bounds on end-to-end packet delay, jitter, and loss. New protocols such as MPLS and Diffserv are realized due to the major advantages associated with them such as QoS. This paper studies the performance of connecting telephone exchange centers over a metro Ethernet network using MPLS and DiffServ QoS model. We will implement some simulations to evaluate the performance in terms of delay, jitter and loss.

**Index Terms** — Data communication, Metropolitan area networks, Queuing analysis, Telephone exchanges.

## I. INTRODUCTION

Given the emerging deployment of metro Ethernets [1], a major aspect of interest is to employ these networks to connect legacy telephone centers and GSM base stations to the core telephony network. This protects investments in existing infrastructure and creates new revenues for network providers. However, to implement this convergence, these networks have to be configured in a way that customary QoS requirements for telephony, can be met. This is a crucial problem given that the data cross traffic in these networks is bursty. This burstiness can cause excessive queuing delay and frame losses due to buffer overflow.

In a telephone exchange or Base Station, the analog voice signal arriving from a user is converted to a digital signal. Then multiple calls are aggregated using Time division multiplexing (TDM) and transmitted over a single line with sufficient capacity. In Europe and Middle East, usually 32 voice channels are multiplexed into a single carrier, named the E1 standard with a bit rate equal to 2.048 Mbps [2].

During the past years, there is considerable interest in technologies, systems, and architectures supporting voice over packet (VoP) applications. Two factors have held back the deployment of VoP at the technical level on a broad scale: quality of service (QoS) considerations and robust signaling. Sager, Baumann and Fiedler [2]-[3], introduced the QoS problem of carrying TDM telephony

traffic over native Metropolitan Gigabit Ethernets. They showed by simulation and analytic evaluation that QoS requirements defined by the Metro Ethernet Forum cannot be met even when data traffic utilization is as low as 1%. They didn't use any priority or QoS mechanism in your work.

Salah [4], considered the deploying of VoIP in a small enterprise network. He studied the QoS requirements and utilized both analysis and simulation tools to determine the number of VoIP calls that can be supported for a typical native Ethernet network. Furthermore, he discussed some design and engineering issues including QoS requirements. He didn't consider QoS mechanisms in his work.

Wright, in his paper [5], compared the advantages of VoMPLS with VoATM, VoIP, and VoFR, from the viewpoints of bandwidth utilization, implementation issues, and the region of the network (access/backbone) in which implementation takes place. He concluded that VoMPLSoPPP is more efficient than VoMPLSoEthernet, and VoIP is highly inefficient.

In [6], authors proposed two new AF service models for overlay networks to accommodate emerging applications such as VoIP and real-time video streaming over the DiffServ/MPLS networks. In both service models, incoming calls are first serviced with the EF-class up to a certain threshold and after that calls would be served with AF-class. In [7], the issues of carrying voice in MPLS domain are studied and its performance in terms of delay and jitter under varying load conditions is analyzed. As a result, authors proposed Diffserv model as a solution to provide required QoS. Nair, in his thesis [8], proposed some solutions to QoS for Ethernet in the metro area. This involves traffic engineering with load balancing, uniform resource utilization and reliability.

In this paper, we review the problem whether it is reasonable to expect that TDM E1 telephony traffic can be run on Metro Ethernet by using MPLS and Diffserv mechanisms? We conduct a simulation study to investigate whether the minimum QoS requirements (in terms of loss, jitter and delay) defined by the Metro Ethernet Forum can be met when an integrated environment of MPLS and Diffserv is provided. As a simulation environment, we use OPNET modeler [9], which is widely used in both industry and academia.



This paper is organized as follows: Section II gives a brief overview of the preliminary concepts: QoS requirements, Metro Ethernet and MPLS. In Section III, using MPLS to support Diffserv is described and Simulation results are presented in Section IV.

## II. PRELIMINARY CONCEPTS

In this section, we introduce some preliminaries in brief, including: QoS, Metro Ethernet and Multi-Protocol Label Switching (MPLS) concepts. QoS is the quality a user or customer can expect from a given service. When specifying the QoS, a number of factors are taken into account:

- Latency (Delay) - the time from a packet is sent until it is received at another point.
- Jitter – timing variations from an ideal position in time, caused by packets arriving either out of order or at an inconsistent rate.
- Packet Loss - the percentage of packets lost in the transmission.

For telephony service, the first two (delay and Jitter) are very important factors. Quantifying these parameters allows us to find out how efficiently the traffic in different networks is being managed and whether the network is suitable for the data we wish to transmit or not. Different kinds of applications have different requirements for the parameters listed above. Table I summarizes QoS requirements for some common application types.

On the other hand, Ethernet technology has so far been widely accepted in enterprise deployments, and millions of Ethernet ports have already been deployed. It is said that today 90% of data traffic is Ethernet-encapsulated.

Among the attributes that have underpinned this phenomenal success are plug-and-play operation, easy troubleshooting, simple operation, and inexpensive interfaces based on economies of scale and backward compatibility. As data services become predominant, carriers are waking up to the possibility of deriving the same operational benefits by replacing their (FR/ATM/TDM) metro network infrastructures with Ethernet infrastructures. However, metro networks have very different requirements from the early LANs. The major challenge for deploying Ethernet in metro networks and wide area networks is to guarantee the quality of service (QoS) using Diffserv mechanisms [10]. This can be met either by Ethernet evolution or by MPLS encapsulation. The spread of MPLS into metro networks is driven by familiar business goals: cost reduction and revenue growth. MPLS promises to cut metro-area bandwidth costs by consolidating diverse access traffic. It also promises to boost service provider revenue by supporting more robust, scalable Ethernet services [11].

TABLE I  
QOS REQUIREMENTS FOR SOME REAL-TIME APPLICATION

Application Type	QoS requirements		
	Delay	Jitter	Packet Loss
Streaming Media	Sensitive	Sensitive	Sensitive
Video-conference	Critical	Critical	Sensitive
Pcketized Voice	Critical	Critical	Sensitive

The MPLS protocol resides between the middle of the second layer and the third layer, and it inserts a shim header between these layers. The shim header has two fields, label and exp. Labels are applied in MPLS to create virtual lines called LSPs (label-switched paths). An LSP is set up on an appropriate route that can accommodate the required bandwidth.

The IETF has also defined a way to carry L1 and L2 traffic over an MPLS network. This includes carrying TDM services over MPLS. With this type of encapsulation, Pseudowires are constructed by building a unidirectional MPLS virtual connection (VC) LSPs between the two Provider Edge endpoints. The VC-LSPs are identified using MPLS labels that are statically assigned or assigned using the Label Distribution Protocol (LDP).

## III. MPLS SUPPORTING DIFFSERV

The exp field is utilized by MPLS to support Diffserv classes characterized by different per-hop-behaviors (PHBs) [11]. The exp value of each packet is mapped to a PHB, which is implemented at each MPLS router. The Diffserv architecture defines three classes: EF (expedited forwarding), AF (assured forwarding), and BE (best effort). The EF PHB provides low packet loss, low latency, low jitter, and the maximum guaranteed bandwidth. The EF PHB is usually implemented by applying priority queuing in accordance with the order EF > AF > BE at the output queues of the routers. The EF inflow rate at the edge router must be lower than the EF maximum guaranteed bandwidth, which is determined by the contract and is in the range of the LSP required bandwidth.

The AF PHB provides a minimum bandwidth guarantee, as well as efficient utilization of excess bandwidth. The AF minimum guaranteed bandwidth of each LSP is the balance of the LSP required bandwidth that is not consumed by EF traffic. To use this excess bandwidth effectively, AF packets are allowed to inflow at a higher rate than the AF minimum guaranteed bandwidth, as long as they are labeled with an out-of-profile flag specifying a high drop preference at the edge router.

As part of the Diffserv mechanism, each MPLS router must implement some queuing discipline that governs how packets are buffered while waiting to be transmitted. Various queuing disciplines can be used to control which packets get transmitted (bandwidth allocation) and which packets get dropped (buffer space). The queuing discipline also affects the latency experienced by a packet, by determining how long a packet waits to be transmitted. Examples of the common queuing disciplines are first-in first-out (FIFO) queuing, weighted fair queuing (WFQ) and priority queuing (PQ). The idea of FIFO queuing is that the first packet that arrives at a node is the first packet to be transmitted. Given that the amount of buffer space at each node is finite, if a packet arrives and the queue (buffer space) is full, then the router discards (drops) that packet. In PQ queuing, the routers implement multiple FIFO queues, one for each priority class. Within each priority, packets are still managed in a FIFO manner. This queuing discipline allows high priority packets to cut to the front of the line [12].

In the next section, we will set up a network that carries two types of data: Packetized E1 traffic and variable rate data cross traffic. We will study how the choice of the queuing discipline in the nodes can affect the performance of the E1 traffic.

#### IV. MODELING A TYPICAL NETWORK TOPOLOGY

A network model based on a single bottleneck link was developed (Fig. 1).

This model has been simulated using OPNET package. Using this model, both MPLS and Diffserv are implemented together in the same network, providing insight on the performance of the two protocols in terms of delay, jitter and loss.

All links and ports are Ethernet-based. Center\_1 is an E1 traffic generator simulating a telephone center (or a cellular telephone base station), and Center\_2 is its corresponding destination. S1 is a variable bit rate (VBR) data generator and D1 is its corresponding destination. Sources Center\_1 and S1 are connected to the ingress node R0, passing the packets through the network to corresponding destinations, connected to the egress node R2.

Here, E1 traffic is simulated with EF flow and data traffic with BE flow. The two intermediate links between the ingress and egress nodes are assigned as a Diffserv-MPLS domain with a bandwidth of 10Mbps and a length of 10 Km, while the access links to sources and destinations are assigned a bandwidth of 10Mbps and a length of 1 Km. The explicit route of MPLS domain is R0-R1-R2. The MPLS has two LSPs; one carries the E1 traffic, other carries data traffic.

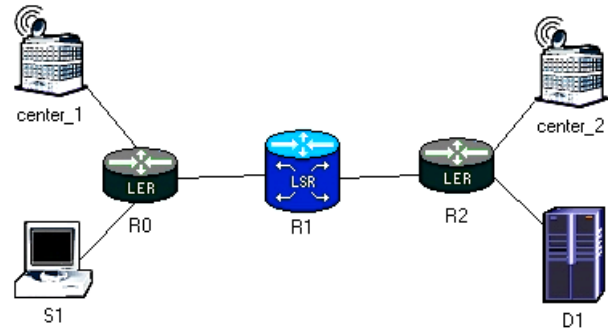


Fig. 1. Modeled Network

In a telephone exchange or Base Station, the analog voice signal arriving from a user is converted to a digital signal with a sample rate of 8 kHz according to Shannon's sampling theorem, resulting in a data rate of 64 Kbits/s for one telephone channel. Usually 32 voice channels are time division multiplexed into a single carrier, named the E1 standard with a bit rate equal to 2.048 Mbps. Actually, there are only 30 of the 32 channels used for voice traffic, the remaining 2 channels are used for synchronization purpose and other control data [2].

Since the delay has to be kept as small as possible, it is certainly better to take each individual E1 carrier packet and send it immediately instead of accumulating them to fill an Ethernet frame up to its maximum payload size of 1500 bytes. Because the E1 packet should also make it through Ethernet layer, a layer 2 header of 26 bytes is attached to the E1 packet of 32 bytes, giving 58 bytes for the Ethernet frame. This leads to a gross data rate of 3.712 Mbps [2]. Therefore, Center\_1 generates the E1 traffic at a constant bit rate of 3.712 Mbps and 800 Packets/Second, then, sends it to the destination node Center\_2. Fig. 2 shows the traffics generated by Center\_1 and S1.

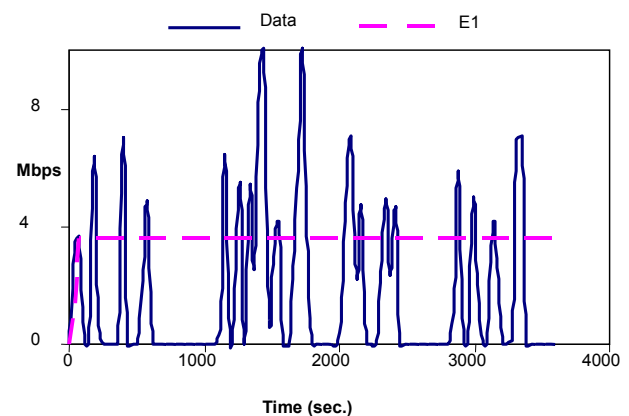


Fig. 2. Generated Traffics by Sources

All simulations are done with traffic in only one direction to keep the simulation effort at a minimum without losing any important model characteristic. As you can see, there is bottleneck on intermediate links, because the aggregated traffic exceeds 10 Mbps. Here, E1 traffic flow between telephone centers is taken as a representation to study the delay, jitter and loss.

The Metro Ethernet Forum (MEF) is a consortium of ISPs and network equipment manufacturers that specify quality of service requirements for telephony circuit emulation over metro Ethernets. These requirements pertain to a requirement for maximum delay, jitter and frame loss rate within an observation interval. Numbers are listed in Table II. Moreover, MEF specifies an additional composite requirement that is the number of frames that violate any of the three above stated requirements. This requirement is called the frame error rate (FER), and is defined as [3]:

$$FER = \frac{\text{number of errored frames}}{\text{total frames sent}}$$

Where an errored frame is a frame delayed more than 25 ms or exceeded the jitter limit of 10 ms or dropped by any network node. MEF specifies that this frame error rate is less than  $8.75 \times 10^{-7}$ .

#### A. QoS Analysis

The total delay in a network topology like in Fig. 1, consists of propagation, transmission and queuing delays. Signal propagation delays over links (due to distance) are an important cause of communication delays. The speed of propagation in optical fibers is  $0.68c$ , where  $c$  is the speed of light in vacuum (300000 Km/s). For the bottleneck links of the network simulation topology in Fig. 1, the propagation delay is  $49 \mu s$  which corresponds to a distance of 10 km for the optical link. For access links, the propagation delays are  $4.9 \mu s$  for 1 km optical links that are very small compared to the bottleneck link's propagation delays. All propagation delays together from the source nodes to the sink nodes are  $107.8 \mu s$  for both types of traffic.

Transmission delay is the second form of delay incurred by each packet as it is sent over a link. An E1 frame consists of 464 bits that are to be transmitted over the 10 Mbps links. The transmission delay at each node is therefore  $46.4 \mu s$ . Any frame sent from source to sink node in the network is processed 4 times (Center\_1, R0, R1 and R2), introducing total  $185.6 \mu s$  of transmission delay.

Queuing delays occurs when data is made to wait for resources to become available, typically a communications link or a processor.

TABLE II  
METRO ETHERNET FORUM'S QoS REQUIREMENTS [3]

Metric	QoS requirement
Jitter	Max. 10ms
Delay	Max. 25 ms
Loss rate	Max. $8.75 \times 10^{-7}$
Frame error rate	Max. $8.75 \times 10^{-7}$

This type of delay can in some cases account for the majority of a frame's end-to-end delay from its source to its destination node. Queuing delays is the only delay that depends on the variable workload traffic. When no data traffic is sent, the total end-to-end delay of an E1 frame is at its minimum of  $420 \mu s$  as shown by simulation results (see Table III and IV). Subtracting the propagation delay ( $107.8 \mu s$ ) and the transmission delay ( $185.6 \mu s$ ) from it, the minimum queuing delay in the network results to be  $126.6 \mu s$ , which accounts for the processing delays of all nodes of the given network. In the simulation setup of this work, E1 and data traffic is sent independently over a 10 Mbps link to the bottleneck entry node R0. As long as the incoming traffic at the R0 is less than the output link's capacity (which is also 10 Mbps), no relevant queues build up in this node and the delay of the E1 traffic is constant at its minimum value of approx.  $420 \mu s$ . When the incoming traffic at the R0 is greater than the output link's capacity, the queue grows at the rate of the output link's exceeding capacity. With growing queue size, also the delay increases.

In networks, packets are sometimes lost or damaged due to buffer overflows or transmission errors. In the simulation setup of this work, losses mainly originate from buffer overflows in the bottleneck entry node R0 (Fig. 1). Losses due to transmission errors are negligible small and buffer overflows in other nodes do not occur.

The point of buffering in the network is to absorb data bursts and to transmit them during the (hopefully) ensuing bursts of silence. This is essential to permit the transmission of bursty data. OPNET Modeler's node queues use first-in-first-out (FIFO) queue scheduling and tail drop queue management algorithms like most of the QoS unaware equipment today.

#### B. Simulation Results

The network was studied with different conditions. At the first scenario, we assumed native Ethernet links and switches without MPLS and without any QoS mechanism. There is naturally, a large FIFO queue at each node. Therefore, packet loss is negligible, but the delay and jitter are not acceptable, as shown in Fig. 3 and Table III.

TABLE III  
QoS Metrics in Implemented Scenarios without MPLS

		FIFO Unlimited	FIFO Size 100
Delay (msec)	Min	0.42	0.42
	Max	22902	173.8
	Avg.	266.8	37.8
Jitter (msec)	Min	1042	64
	Max	8630	86
	Avg.	5340	72
Packet Loss Ratio		$7.53 \times 10^{-7}$	$62.24 \times 10^{-3}$
Frame Error Rate		1	1

In second scenario, we limited the FIFO queue size to 100 packets. As a consequence, delay and jitter are decreased, but packet loss is increased. As shown in Fig. 4 and Table III, QoS parameters are larger than the MEF's acceptable values.

In third scenario, we implemented MPLS without Diffsev. Two LSPs are set, one for E1 traffic and another for data traffic. As you can see, in MPLS without Diffserv model, packet loss is negligible, but the delay and jitter are not acceptable, as shown in Fig. 5 and Table IV.

Last scenario is implemented by using Diffserv-MPLS model with PQ queuing discipline. For PQ, the queue size of BE traffic is set to 60 and the queue size of EF traffic is set to 40 (total queue size is 100).

TABLE IV  
QoS Metrics in Some Implemented Scenarios with MPLS

		MPLS without Diffserv	MPLS with PQ
Delay (msec)	Min	0.42	0.42
	Max	20727	16.2
	Avg.	167.8	4.6
Jitter (msec)	Min	303	4.43
	Max	6740	7.96
	Avg.	4490	5.75
Packet Loss Ratio		$7.53 \times 10^{-7}$	$7.53 \times 10^{-7}$
Frame Error Rate		1	$7.53 \times 10^{-7}$

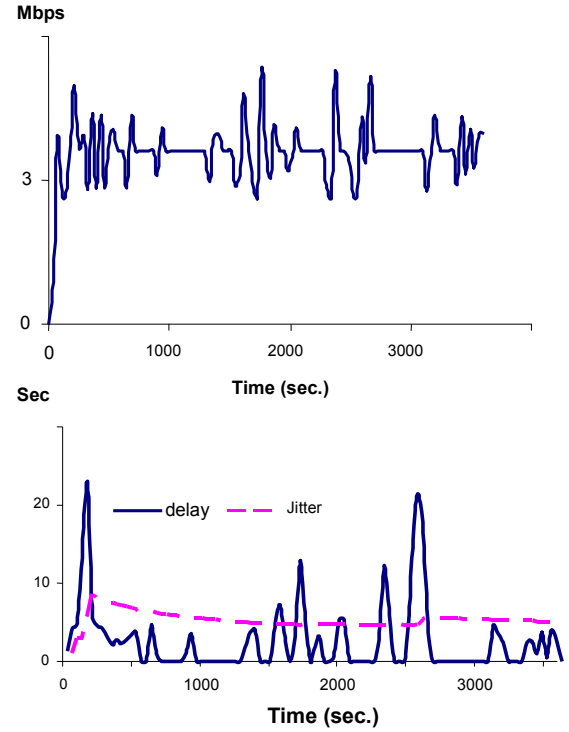


Fig. 3. Delivered E1 traffic, Delay and Jitter (large FIFO queue without MPLS)

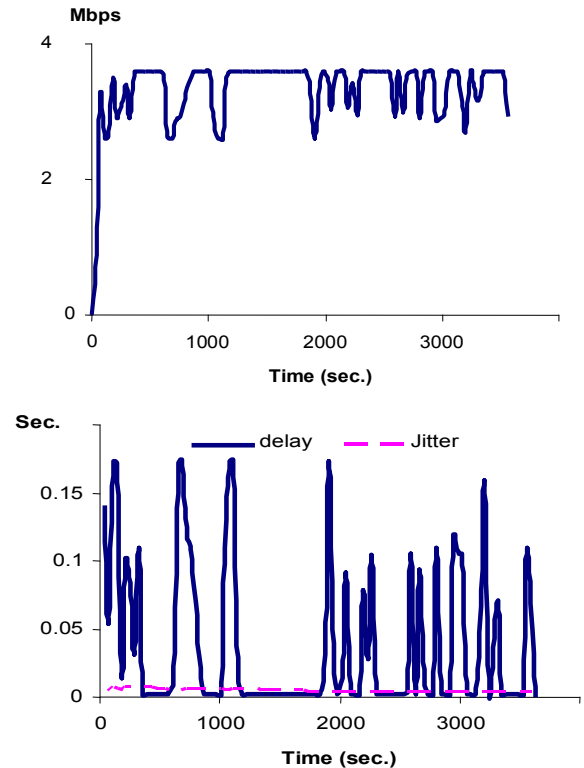


Fig. 4. Received E1 traffic, Delay and Jitter (FIFO queue size 100 without MPLS)

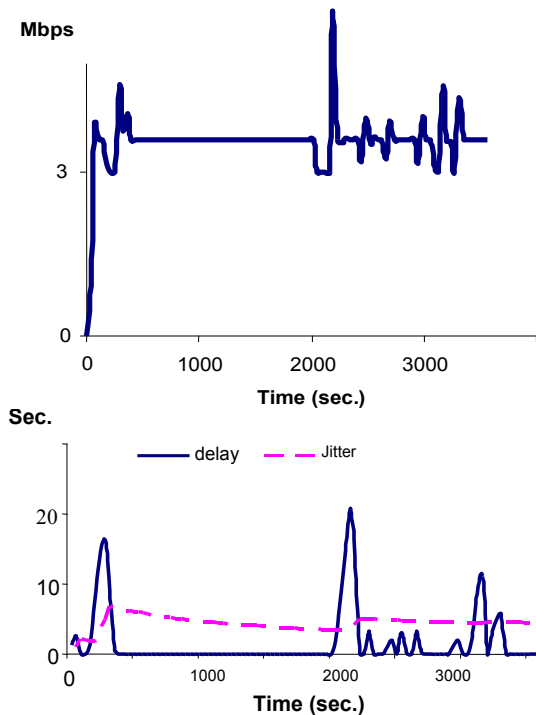


Fig. 5. Received E1 traffic, Delay and Jitter (MPLS without DiffServ)

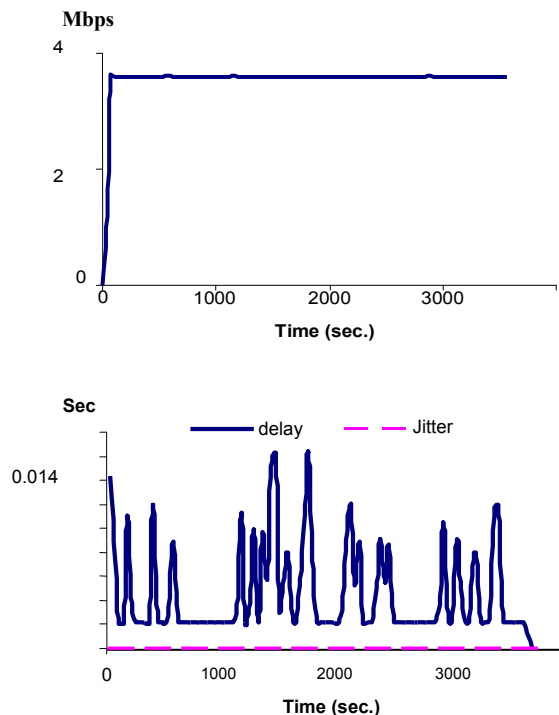


Fig. 6. Received E1 traffic, Delay and Jitter (MPLS with Priority Queue)

For MPLS with PQ model, the network performance reflected acceptable loss, delay and jitter. The delay is very small as shown in Fig. 6 and Table IV. This was expected, as the EF traffic normally arranges to pass through Diffserv domain without much delay, jitter or loss.

## V. CONCLUSION

In this paper, we studied the performance of packet-based E1 traffic transmission over a metro Ethernet network using MPLS and DiffServ QoS model. We implemented some simulations to evaluate the performance of connecting telephone exchange centers in terms of delay, jitter and loss under varying conditions. We showed that without using DiffServ-MPLS, MEF's QoS requirements are not met, but by using MPLS with PQ queuing discipline, these requirements are satisfied.

## ACKNOWLEDGEMENT

This work is supported by Iranian Telecommunication Research Center (ITRC).

## REFERENCES

- [1] M. Huynh, and P. Mohapatra, "Metropolitan Ethernet Network: A move from LAN to MAN", *Computer Networks*, vol. 51, no. 17, pp 4867-4894, 2007.
- [2] P., Sager, *Does circuit emulation in metro gigabit Ethernets require service priority?*, Thesis, ETH Zurich, 2005.
- [3] R. Baumann, and U. Fiedler, "Towards Connecting Base Stations over Metro Gigabit Ethernets", *Proc. of the Int. Conf. on QoS in Heterogeneous Networks*, 2006.
- [4] K. Salah, "On the deployment of VoIP in Ethernet networks: methodology and case study", *Computer Communications*, vol. 29, pp. 1039-1050, 2006.
- [5] D. Wright, "Voice over MPLS compared to voice over other packet transport technologies", *IEEE Communications Magazine*, vol. 40, no. 11, pp. 124-132, 2002.
- [6] H. Chen, Hui-Kai Su and Bo-Chao Cheng, "An objective-oriented service model for VoIP overlay networks over DiffServ/MPLS networks", *Computer Communications*, vol. 30, pp. 1039-1050, 2007.
- [7] S. AL-Irhayim, J. A. Zubairi, Q. Mohammad and S. A. Latif, "Issues in Voice over MPLS and Diffserv Domains", *Proc. of PDCS'2000*, vol. 2, pp. 467-472, 2000.
- [8] P. M. V. Nair, *Quality of Service in Metro Ethernet*, Ph.D. Thesis, Southern Methodist University, 2006.
- [9] OPNET Technologies, [www.opnet.com](http://www.opnet.com)
- [10] Meddeb, A., "Why Ethernet WAN transport?", *IEEE Communications Magazine*, vol. 43, pp.136-141, 2005.
- [11] S. Alvarez, *QoS for IP/MPLS Networks*, Cisco Press, USA, 2006.
- [12] Mirjalily, G., Ahmadian, M., "A New approach in Scheduling for Differentiated Services Networks", as a chapter of the book named: "*Recent Advances in Communication Systems and Electrical Engineering*", Springer, 2008.

# Security Enhancement of MD5 Hashed Passwords by Using the Unused Bits of TCP Header

Md. Didarul Alam Chawdhury, and A.H.M. Ashfak Habib

Chittagong University of Engineering & Technology, Chittagong – 4349, Bangladesh

**Abstract** — When a password is encrypted by a hash algorithm the resultant is called hashed password. In a server client based communication system such as Yahoo Messenger, AIM, passwords of clients are hashed by MD5 and passed to the server for authentication [1]. This type of transmission is always a subject of interception by the hackers. These hashed passwords are passed through the internet as a data packet. TCP header is a most common part of the data packet. In a TCP header there are six reserved bits which remains always unused [2]. In this paper we propose a new approach to enhance the security of hashed passwords by using the six reserved bits of a TCP header. Here we encrypt the hashed password by a random key using simple mathematical function. The information needed to decrypt the encrypted hashed password is carried by the six bits of TCP header.

**Index Terms** — Client-server systems, Cryptography, Computer network security, Data security, Security.

## I. INTRODUCTION

Communication is the most important aspect of this modern age. Computer network i.e. internet is one of the most important part of communication of this modern civilization. Since valuable information is passed through the internet, the question of security arises. In a server client based communication, passwords are top secret information. Therefore passwords are first hashed and then passed through the internet. At the time of transmission the hackers always try to intercept those hashed passwords. So it is important to make the transmission of those passwords secure as much as possible.

A few techniques are developed for securing hashed passwords. Most common of them is salting. Salting means adding some bit pattern to the password and then encrypting it by message digest algorithm [1]. The encrypted password is then passed through the internet. Another technique is double hashing, where the password is hashed twice. Usually data passes through the network as a packet. A data packet consists of a number of bits. There are some bits which always remain unused in that packet. In fact in a TCP header 54 bits remain unused on an average case [2]. Moreover there are six reserved bits in TCP header which are never used. In this paper we tried to improve the security of the hashed passwords by using the unused six bits of TCP header.

## II. SERVER CLIENT SYSTEM

Among various types of communication systems, server & client based systems are one of the most important and widely used communication systems. Fig. 1 illustrates a server client based communication system. Server is a computer which is used to provide a particular service to the clients and clients are a group of computer that receives the service provided by the server [3]. To determine the eligible users the server applies an authentication scheme. Most typical one is user name and password verification.

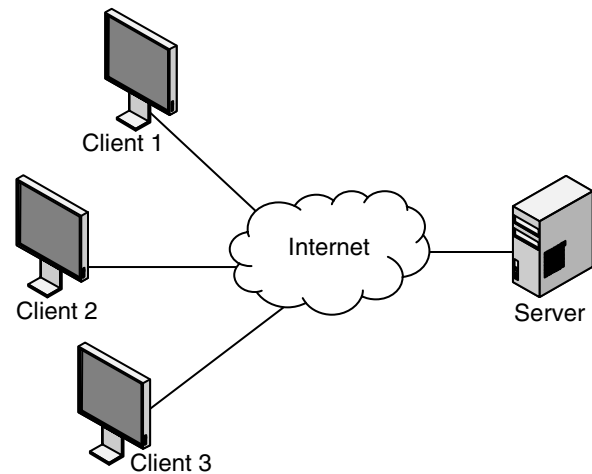


Fig. 1. Server Client Communication System.

Before transmitting through the media the passwords are encrypted. The most commonly used encryption algorithms are hash (message digest) algorithms. The passwords in the server database are stored as hashed password. When a client is challenged for authentication then a login page appears at the client end. The client enters his/her user name and password. The user name is passed as plain text but the password is encrypted by a message digest algorithm. The resultant cipher text i.e. hashed password is passed to the server and the server verifies them. The most widely used hash algorithm is Message Digest 5 (MD5). In this paper we only concentrate on MD5 hash algorithm.

### III. MESSAGE DIGEST 5 (MD5)

The MD5 algorithm is an extension of the MD4 message-digest algorithm. It is an irreversible encryption algorithm. That is, from an encrypted message it is not possible to recover the original message. The algorithm takes a message of arbitrary length as input and produces a 128-bit "fingerprint" or "message digest" of the input as output. It is conjectured that it is computationally infeasible to produce two messages having the same message digest, or to produce any message having a given pre-specified target message digest. The MD5 algorithm is intended for digital signature applications, where a large file must be "compressed" in a secure manner before being encrypted with a private (secret) key under a public-key cryptosystem such as RSA [4].

### IV. ATTACKS ON MD5 HASH

The hashed password can be hacked during transmission. When MD5 hash is hacked then we call it attack on MD5 hash. This hash can be hacked effectively by using two efficient techniques [1]. These are:

- a) Dictionary Attack
- b) Rainbow Table.

#### A. Dictionary Attack

A dictionary attack refers to the general technique of trying to guess some secret by running through a list of likely possibilities, often a list of words from a dictionary [5]. It contrasts to a brute force attack in which all possibilities are tried. The attack works because users often choose easy to guess passwords, even after being exhorted against doing so. Dictionary attacks either pre-compute hash values for a given dictionary file and then compare target password hashes to the pre computed tables for a match, or words in a dictionary file are hashed and compared against a target password hash for a match.

#### B. Rainbow Table

In 2003, Philippe Oechslin [6] introduced Rainbow Tables, a time-memory trade-off used for cryptanalysis of hashed passwords. From his work, several applications have been developed that create rainbow tables.

The entry in a rainbow table is called chain. Chain is pre-computed to hack a password. Rainbow tables use a successive reduction function for every point in the chain. They start with a reduction function 1 and end with a reduction size of  $t-1$  (one less than chain length). The password is first hashed and then the reduction function is applied to it. Only the first and the intermediate hash (at the length of the chain) are stored in the table. To hack a hash, a chain from the hash is generated. For each

password that in the chain, hackers check to see if it is at the end of the chain. When they hit a match at the end of the chain, then it is known that the hash was part of this chain. Thus the complete chain again generated and the password is hacked.

### V. HASHED PASSWORD ENCRYPTION

We name the new approach of encryption which enhances the security of MD5 hashed password as Hashed Password Encryption (HPE). There are three steps in this encryption process which are executed at the client side. The first step is the generation of a random key. By using the key we determine a value from a spiral arrangement of number sequence. Second step is the encryption of the MD5 hashed password. Encryption of the key is the third and final step of this encryption process.

Decryption takes place at the server side and it consists of two steps. First step of decryption is to recover the original key and the final step is to decrypt the encrypted hashed password.

The encryption and decryption processes are explained in details with the help of an example as follow:

Consider a server client based communication system where MD5 is used to encrypt the client password. Suppose the original password of a client is "abc" and the corresponding MD5 hashed password is  
900150983cd24fb0d6963f7d28e17f72

#### A. Encryption Process (Client Side)

At first a random key is generated. For choosing a key the following constraints are followed:

- 1) Key must be a prime number.
- 2) Key length will be 3, 4 or 5. That is the number of digits in the key is either 3 or 4 or 5.
- 3) In case of 3-digit-key the last digit will be greater than or equal to the preceding digits and no digit can be zero. For example 233 can be a key but 253 cannot be a key.
- 4) In case of 4-digit-key the value of the last two-digit-number will be greater than or equal to each of the preceding two digits. Moreover none of the first two digits can be zero. For example 1049 or 2301 cannot be a key but 6871 can be a key.
- 5) In case of 5-digit-key the value of last three-digit-number will be greater than or equal to each of the preceding two digits. In the first two digits no digit can be zero. For example 10049 or 14001 cannot be a key but 14983 can be a key.

The range of the key space of our HPE cryptosystem is 15000. In this range we get over 1300 keys.

We use a spiral arrangement of number sequence. This arrangement can be considered as a number grid. Suppose

the random key generated is 227. The grid for the key is shown in Table I.

TABLE I  
SPIRAL ARRANGEMENT OF NUMBERS

31	30	29	28	27	26	49
32	13	12	11	10	25	48
33	14	3	2	9	24	47
34	15	4	1	8	23	46
35	16	5	6	7	22	45
36	17	18	19	20	21	44
37	38	39	40	41	42	43

The first digit of the key indicates row and the second digit indicates the column. The rest of the digits indicate the size of the grid. The above grid is of size 7.

The value that will be used for encryption is found from the grid. For the key 227 the value is 17.

#### 1. Encryption of the Hashed Password

Now we encrypt the hashed password by using a simple arithmetic addition and mod operation. The encryption is stream ciphering. That is we encrypt the hashed password character by character. The function can be shown as follows:  $C = (A + V) \bmod 256$

Where A is the ASCII value of each character in the hashed password, V is the value found from the key. For our example it is 17, C is the ASCII value of the character that will replace the character in the previously encrypted hashed password.

In our example the hashed password is 900150983cd24fb0d6963f7d28e17f72. For the first character '9' in the hashed password the encryption can be shown as follows:

$$C = (57 + 17) \bmod 256$$

$$C = 74$$

Since the ASCII value of J is 74, '9' is replaced by 'J'. Thus the encrypted hashed password obtained is JAABFAJIDtuCEwsAuGJGDwHuCIvBHwHC. This newly encrypted hashed password is sent to the server.

#### 2. Encryption of the key

We have the key 227. In this step this key is encrypted. First we calculate the sum of ASCII value of all characters in the original hashed password. Then we use the following function to encrypt the key:

$$Q = (P * K) \bmod 61$$

Where, P is the sum of ASCII value of all characters in the original hashed password, K is the key, Q is the encrypted key. In our example the encryption is as follow:

$$Q = (2107 * 227) \bmod 61$$

$$Q = 49$$

The value 49 is passed through the six reserved bits of TCP header.

#### B. Decryption (Server Side)

The newly encrypted password is sent to the server through the data field of the TCP packet. Moreover the value of Q is passed through the reserved six bits of the header.

##### 1. Decryption of the key

Server has original hashed password against the user name in its database. It calculates the sum of ASCII value of all characters in the original hashed password. From the six reserved bits in TCP header of the packet received from the client the server gets the value  $Q = 49$ . After this the server has to do the followings

- 1) Calculates  $M = (P * S) \bmod 61$  for all keys in a range. Where P is the sum of ASCII value of all characters in the original hashed password, S is the possible key.
- 2) If  $M = Q$  then the value of S is kept in a list.

On an average case 20 possible keys are found. In our example the possible keys are: 227, 349, 1447, etc.

##### 2. Decryption of the Encrypted Hashed Password

The server has the encrypted hashed password JAABFAJIDtuCEwsAuGJGDwHuCIvBHwHC. For all possible keys the server does the following:

- 1) Determine the value that is used in encryption from the spiral arrangement of numbers. For the first possible key 227 the value  $H = 17$  is found.
- 2) Determine  $T = H \bmod 256$ . In our example  $T = 17$ .
- 3) Determine  $A = C - 17$ . If A is less than zero then the following operation is done:  $A = A + 256$ .

A is the ASCII value of the corresponding characters in the original hashed password. For the example J is the first character of the encrypted hashed password. For J we get  $A = 74 - 17 = 57 = '9'$

Thus the original password is recovered. Then the server finds match with the password in its database. If a match is found the server stops and the client is authenticated. If no match is found it is considered that the client supplied wrong password.

## VI. PERFORMANCE ANALYSIS

We evaluated our approach by calculating the new computation information for the two hacking techniques illustrated in this paper.



#### A. Computation Information for Dictionary Attack without Using HPE

Computation information for the dictionary attack [1], when HPE algorithm is not used, is as follows:

Character set = [abcdefghijklmnopqrstuvwxyz]  
 Password length = 7  
 Words = 23109  
 Dictionary = English  
 Algorithm = MD5  
 Record Format = \$hash\$password , ordered by \$hash  
 File Size = 231MB  
 Average lookup time =  $2.9 * 10^{11}$  floating seconds

#### B. Computation Information for Dictionary Attack Using HPE

Computation information for dictionary attack taking HPE algorithm into account is as follows:

Character set = [abcdefghijklmnopqrstuvwxyz]  
 Password length = 7  
 Words = 23109  
 Dictionary = English  
 Algorithm = MD5  
 Record Format = \$hash\$password , ordered by \$hash  
 File Size = 947469 bytes  
 Average lookup time = 1134440021.7656 floating seconds

#### C. Pre-computation estimates for Rainbow Table without Using HPE

The following data are the pre-computation estimates for Rainbow Table without using HPE and are calculated by using MATLAB [1].

[key space] 8,031,810,176  
 [chain length] 2100  
 [chain count per table] 8000000  
 [precomputation time(days)] 0.83844

#### D. Pre-computation estimates for Rainbow Table Using HPE

The following data are the pre-computation estimates for Rainbow Table when HPE algorithm is used.

[key space] 8,031,810,176  
 [chain length] 2100  
 [chain count per table] 2048000000  
 [precomputation time(days)] 214

## VII. CONCLUSION

In this paper we proposed a new cryptosystem for further encrypting the MD5 hashed password before transmitting it through the internet. We showed the performance analysis of our cryptosystem. From the performance analysis the benefit of our proposed HPE cryptosystem is clear. If a server client based communication system uses HPE in authenticating users then the possibility of interception of the password will be reduced to a considerable extent. It will be infeasible for the hackers to pre-compute the hash for dictionary and rainbow table when HPE is implemented in the system. Moreover HPE makes a good use of six reserved bits in TCP header which implies an efficient utilization of network resources. Thus the transmission overhead due to transmitting unused bits is reduced.

## REFERENCES

- [1] D. Kaminsky, "MD5 to be considered harmful someday," Doxpara Research, December 2004. Available: [http://www.doxpara.com/md5\\_someday.pdf](http://www.doxpara.com/md5_someday.pdf).
- [2] A. S. Tanenbaum, *Computer Networks (4<sup>th</sup> Edition)*, Prentice Hall, New Jersey, 2004.
- [3] D. E. Comer, *Computer Networks and Internets (4<sup>th</sup> Edition)*, Pearson Education, Inc., New Jersey, 2004.
- [4] R. Rivest, "RFC 1321: The MD5 Message-Digest Algorithm," April 1992.
- [5] B. Pinkas and T. Sander, "Securing passwords against dictionary attacks," *Proceedings of the 9th ACM conference on Computer and communications security*, pp. 161 - 170, 2002.
- [6] P. Oechslin, "Making a Faster Cryptanalytic Time-Memory," *Proceedings of Crypto'03*, 2003. Available: <http://lasecwww.epfl.ch/~oechslin/publications/crypto03.pdf>

# Crosstalk and Bit Error Rate Performance of a Proposed Optical Cross-Connect Topology

Md. Prince Hossen\*, Md. Abubakar Siddique and Md. Rafiqul Islam

Department of Electrical and Electronic Engineering, Khulna University of Engineering and Technology, Khulna-9203, Bangladesh.

\*Corresponding author e-mail: prince\_mhmd@yahoo.com

**Abstract** --- This paper presents an optical wavelength division multiplexed (WDM) cross-connect (OXC) topology, which is the modified form of the existing topologies. The analytical approach for the proposed topology is derived to explore the effect of crosstalk induced in multichannel WDM network due to OXC. The factors that affect the magnitude of crosstalk in the OXC are investigated and identified. The bit error rate (BER) performance of a WDM optical system is evaluated at a bit rate of 10 Gb/s on account of OXC-induced crosstalk and different noises. Power penalty is evaluated at BER of  $10^{-9}$  shows that there is a significant impact of crosstalk on the number of wavelengths and input fibers. The results obtained from the present study demonstrate that the crosstalk performance of the proposed topology can be improved significantly than the existing topologies.

**Index Terms** -- optical cross connect, optical crosstalk, optical filter, optical switches, Wavelength division multiplexing.

## I. INTRODUCTION

The advent of wavelength division multiplexing (WDM) technique provides a dynamic increase in the information carrying capacity of a fiber [1]. WDM corresponds to the scheme in which multiple optical carriers at different wavelengths are modulated by using independent electrical bit streams and are then transmitted through the fiber. WDM networks are very attractive due to high capacity, network security, flexibility and possibility of upgrading the existing optical fiber networks to WDM networks [2]. The optical cross-connect (OXC) is an essential WDM network element which offers a dynamic wavelength routing scheme that can reconfigure the network while maintaining its non blocking (transparent) nature to WDM networks [3].

In practical systems while propagating through the switching element that are the part of OXC many signals and wavelength channels could influence each other and cause significant crosstalk [4]. Crosstalk is one of the major limiting factors that may degrade WDM network performance. The crosstalk sources are related to the different individual components of the OXC's. Up to now several OXC topologies [4] were utilized to evaluate the bit error rate (BER) performance of WDM optical fiber transmission system including the effect of OXC-induced crosstalk [5,6]. Since the performance of WDM networks is

limited by the OXC-induced crosstalk [8-10], it is desirable to develop a suitable OXC topology, so that, it contributes minimum crosstalk due to OXC. It is also desirable to investigate the BER performance of a WDM transmission system including the effect of crosstalk induced by the OXC components of the proposed topology.

In this work we have proposed an OXC topology with modification of the existing topologies. Afterwards the analytical equation for the proposed topology is derived by identifying and quantifying different crosstalk sources. The BER performance for a particular WDM system is evaluated at 10 Gb/s on account of OXC-induced crosstalk and different noises. Finally, with a specific BER, the amount of power penalty as a function of the component parameters of the OXC is evaluated and compared with the existing topologies.

## II. OXC TOPOLOGY

Several topologies exist for all optical WDM cross connects [4,7]. The most suitable topology for an application depends in general on the required functionality and on the cost, capacity and flexibility constraints. In this paper, we have analyzed the performance of OXC by placing filter before and after the switching matrix and selecting of the wavelength channels. The schematic diagram of the proposed OXC topology is shown in fig. 1. This topology is the combination of the first and third topology as defined in Ref. [4]. In the proposed topology, the switching matrix of topology 1 is used but the wavelength channels are selected by the filters (with a fixed centre frequency), before and after being routed to the desired output fiber.

The block diagram of a WDM transmission link with proposed OXC is shown in fig.1. The switching matrix in the OXC shown in fig.1 is based on an array of gates which are implemented by gain-clamped semiconductor optical amplifier (GC-SOA) [11] to route the different wavelength channels. Splitters and combiners are placed in front of and behind the switch matrix and filters are used to select the wavelength channels. The wavelength channel to be transmitted is multiplexed by a WDM MUX and fed to an

## OXC

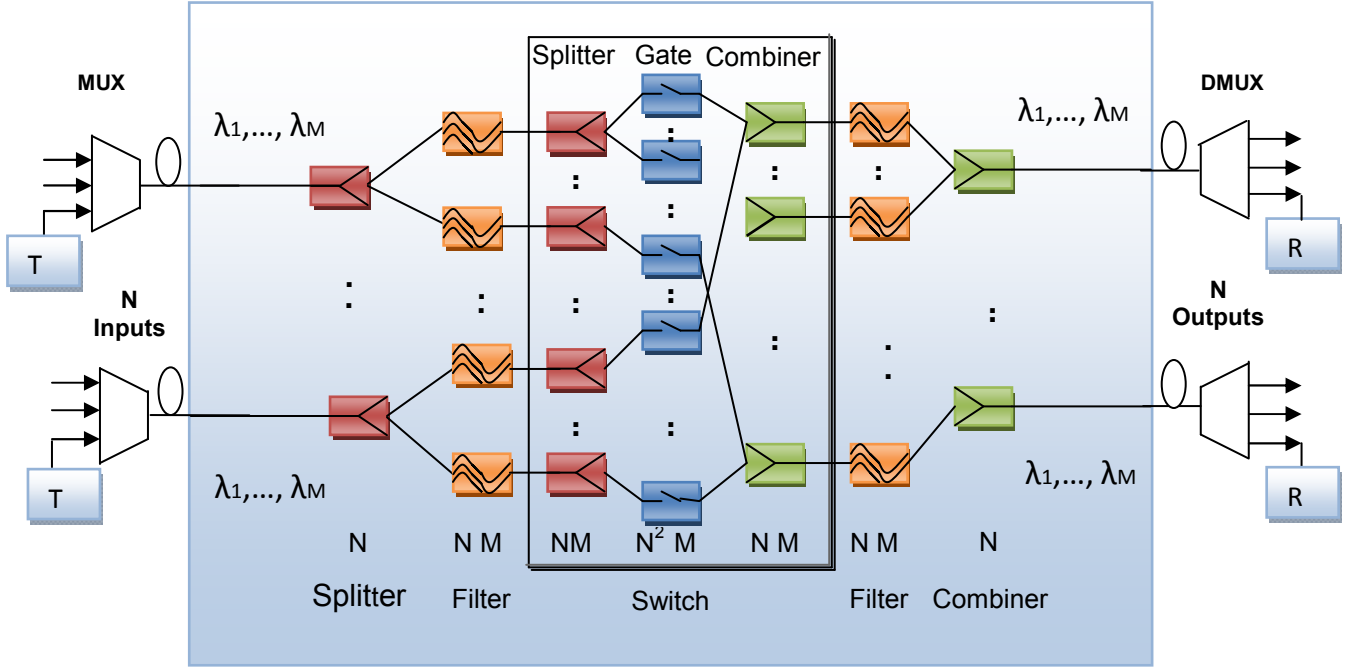


Fig. 1: Block Diagram of a WDM optical system with a new OXC topology.

incoming fiber. At the input of cross-connect, the incoming signals are splitted by a first array of power splitters followed by a array of filters and a second array of power splitters. At the input of GC-SOA gates, all channels are present. The gate selects the wavelength that carries the desired channel. The OXC enables any wavelength channel from any input fiber to be cross-connected to any output fiber, on condition that no two channels in the output fiber have the same wavelength. In fig.1, each fiber carries wavelength channels  $\lambda_1, \lambda_2 \dots \lambda_M$ . Given that the  $N$  input fibers are routed to the desired  $N$  output fibers, each carrying  $M$  wavelength channels, so there are a total of  $N \times M$  wavelength channels. The  $N \times M$  wavelength channels passed through first array of power splitters. There are  $N$  power splitters for all the  $N$  input fibers. All the different wavelength channels appear at the output of the power splitter at lesser power due to power splitting which are fed to  $N$  different filters. Wavelengths  $\lambda_1, \lambda_2 \dots \lambda_M$ , from the output of filters are then fed to another array of  $M$  power splitters, one for each wavelength. So there is a total of  $N \times M$  power splitters at the second array. The output of the second array of power splitters are fed to  $N \times M$  gates, which allows only specific wavelength to pass through. The outputs of the gates are combined by  $N \times M$  combiner at the first array of combiner. Outputs of  $N \times M$  combiners are fed to  $N \times M$  filters. Then,  $N$  times  $M$  outputs with a different central wavelength, are combined into the  $N$  output fibers by the second array of  $N$  combiners. Finally, the desired

wavelength channel of an outgoing fiber is demultiplexed by a WDM DMUX and is received by a direct detection receiver.

### III. CROSSTALK AND BER MODEL

The analytical equation illustrated in this paragraph which describes the output power of the proposed topology in function of the input powers and OXC component parameters. The crosstalk is calculated for a certain wavelength channel, this channel will be called the channel under study. The absolute power levels are only relevant in front of the GC-SOA. Therefore, the power levels at the input of the gate are taken as reference, gains and losses are not relevant since they are assumed uniform for all possible routing ways.

To drive the equations for the proposed topology, the signal power is defined by  $P_i^j$ , where  $i$  designates the wavelength channels and  $j$  the number of the fibers. The fiber which contains the channel under study is indicated by  $j_0$ , and the wavelength under study is  $i_0$ .  $P_{i_0}^{j_0}$  is the input power of a channel,  $P_{i_0}^{out}$  is defined as the output power of wavelength channel  $i_0$  with crosstalk contributions added with all wavelength channels carrying bit 1.  $P_i^{j_0}$  is the signal power at fiber  $j_0$  with another wavelength  $i$ .  $P_{i_0}^j$  is the signal power at another fiber  $j$  that carries the wavelength under consideration,  $i_0$ . We assume that all wavelength channels including  $P_{i_0}^{j_0}$  carry bit 1.  $P_{i_0}^{out}$  at any output fiber

may be just  $1/N$  of the output power, due to division of power before entering the GC-SOA. However, it is also assumed that the GC-SOA is set with gain of  $N$  times to compensate for the division of output optical power.  $N$  is the number of input fibers into OXC,  $M$  is the number of wavelengths per input fiber. Due to the gate's imperfection in gain clamping, it adds crosstalk. The crosstalk power at wavelength  $i$  is then given by [4]

$$P_{cros,i} = X_{gate} P_{in,i} \sum_{k=1}^M P_{in,k} \quad (1)$$

with  $X_{gate}$  is the crosstalk parameter of the GC-SOA,  $P_{in,k}$  is the power at wavelength  $k$  and  $M$  the number of wavelength channels at the input of the gate. The sum is made over all  $M$  wavelength channels.  $T_F$  is derived from the filter suppression of wavelength channel,  $T_F^{-1}$ . This means that  $T_F$  is the transmission of the filter seen by that channel ( $< 1$ ). The ON-OFF ratio of the gate is given by  $R_{gate}^{-1}$ , so that  $R_{gate}$  is the transmission of a gate or gate extinction ratio in the off-state ( $< 1$ ). Extinction ratio is defined as,  $R_{gate} = P_{off}/P_{on}$  where  $P_{on}$  is the input power and  $P_{off}$  is the off state power through each gate. The expression for output power with crosstalk of wavelength channel  $i_0$  can be given by

$$\begin{aligned} P_{i_0 1}^{out} = & P_{i_0}^{j_0} + P_{i_0}^{j_0} \{ X_{gate} ((M-1)T_F P_i^{j_0} + P_{i_0}^{j_0}) \} \\ & + P_{i_0}^j \left\{ \begin{aligned} & (N-1)R_{gate} [1 + X_{gate} T_F M P_i^j] \\ & (M-1)^2 T_F^2 [1 + X_{gate} T_F M P_i^j] \\ & + (M-1)^2 (N-1) T_F^2 R_{gate} \end{aligned} \right\} \\ & - 2\sqrt{P_{i_0}^{j_0}} \sqrt{P_{i_0}^j} \left\{ \begin{aligned} & (N-1)\sqrt{R_{gate}} + (M-1)^2 T_F + (M-1)^2 (N-1) T_F \sqrt{R_{gate}} \\ & (N-1)(M-1)^2 T_F \sqrt{R_{gate}} \\ & + (N-1)^2 (M-1)^2 T_F R_{gate} \\ & + (N-1)(M-1)^4 T_F^2 \sqrt{R_{gate}} \end{aligned} \right\} \\ & - 2P_{i_0}^j \left\{ R_{gate} \sum_{t=1}^{N-2} t + T_F^2 \sum_{t=1}^{(M-1)^2-1} t + T_F^2 R_{gate} \sum_{t=1}^{(M-1)^2(N-1)-1} t \right\} \quad (2) \end{aligned}$$

for the proposed topology under the assumption that all channels carry bit 1. The above equation has six contributions. The first three terms are the non interfering contributions and the last three terms are the contributions due to the interference of different channels (beat terms). In this topology the crosstalk due to the non-perfect gain clamping of the gate is same as the topology 3 [4]. The analytical equations for the proposed topology differ from the topology 1 and 3 due to the filters placed both sides of the switching matrix. It is expected that the crosstalk performance of the proposed topology may improve because

the desired wavelength channels are selected by the filters (with a fixed centre frequency) before and after the channels are routed to the output fiber.  $P_{i_0}^{out(ref)}$  is the output power of wavelength channel  $i_0$  when the OXC carries only this wavelength channel, such as when there is no crosstalk.  $P_{i_0}^{out(ref)}$  is given by [4]

$$P_{i_0}^{out(ref)} = P_{i_0}^{j_0} + X_{gate} (P_{i_0}^{j_0})^2 \quad (3)$$

Since wavelength channel  $i_0$  may carry bit 1 or bit 0 at any instant of time, equation (3) has to be modified. If wavelength channel  $i_0$  carries bit 0, then equation (2) reduces to equation (4).

$$\begin{aligned} P_{i_0 0}^{out} = & P_{i_0}^j \left\{ \begin{aligned} & (N-1)R_{gate} [1 + X_{gate} T_F M P_i^j] \\ & (M-1)^2 T_F^2 [1 + X_{gate} T_F M P_i^j] \\ & + (M-1)^2 (N-1) T_F^2 R_{gate} \end{aligned} \right\} \\ & - 2P_{i_0}^j \left\{ \begin{aligned} & (N-1)(M-1)^2 T_F \sqrt{R_{gate}} \\ & + (N-1)^2 (M-1)^2 T_F R_{gate} \\ & + (N-1)(M-1)^4 T_F^2 \sqrt{R_{gate}} \end{aligned} \right\} \\ & - 2P_{i_0}^j \left\{ R_{gate} \sum_{t=1}^{N-2} t + T_F^2 \sum_{t=1}^{(M-1)^2-1} t + T_F^2 R_{gate} \sum_{t=1}^{(M-1)^2(N-1)-1} t \right\} \quad (4) \end{aligned}$$

$P_{i_0}^{out(ref)}$  when wavelength channel  $i_0$  carries bit 0 can be written as,  $P_{i_0}^{out(ref)} = 0$ .

Following the analytical equations for the OXC topology presented here, the crosstalk contribution from different crosstalk sources are quantified and total crosstalk is calculated. The expression for relative crosstalk for a channel under study can be given by [4]

$$Crosstalk = \frac{P_{i_0}^{out(ref)} - P_{i_0 1}^{out}}{P_{i_0}^{out(ref)}} \quad (5)$$

To evaluate the BER of a WDM system, the crosstalk model presented for the proposed OXC topology is used to derive BER model. The mathematical formulation for BER of the proposed topology in case of IM-DD system can be given by [12].

$$BER_{worstcase} = \frac{1}{8} \left[ \begin{aligned} & \operatorname{erfc} \left( \frac{1}{\sqrt{2}} \frac{i_1 + i_{CT0} - i_D}{\sigma_{1,0}} \right) + \operatorname{erfc} \left( \frac{1}{\sqrt{2}} \frac{i_D - i_0 - i_{CT0}}{\sigma_{0,0}} \right) \\ & + \operatorname{erfc} \left( \frac{1}{\sqrt{2}} \frac{i_1 + i_{CT1} - i_D}{\sigma_{1,1}} \right) + \operatorname{erfc} \left( \frac{1}{\sqrt{2}} \frac{i_D - i_0 - i_{CT1}}{\sigma_{0,1}} \right) \end{aligned} \right] \quad (6)$$

where,  $i_D$  is the threshold current. In equation (6), it is assumed that when a bit 1 is interfered by a crosstalk bit 1, bit 0 is interfered by crosstalk bit 0, bit 1 is interfered by crosstalk bit 0 and bit 0 is interfered by crosstalk bit 1 as reported so far [8]. The variance of the interference is

given in equations (7) to (10)

$$\sigma_{1_0}^2 = \sigma_{th}^2 + 2qR_d(P_s + P_{sp} + P_{CT0})B + \sigma_{s-sp}^2 + \sigma_{sp-sp}^2 + \sigma_{CT0-sp}^2 + \sigma_{s-CT0}^2 \quad (7)$$

$$\sigma_{1_1}^2 = \sigma_{th}^2 + 2qR_d(P_s + P_{sp} + P_{CT1})B + \sigma_{s-sp}^2 + \sigma_{sp-sp}^2 + \sigma_{CT1-sp}^2 + \sigma_{s-CT1}^2 \quad (8)$$

$$\sigma_{0_1}^2 = \sigma_{th}^2 + 2qR_d(P_{sp} + P_{CT1})B + \sigma_{CT1-sp}^2 + \sigma_{sp-sp}^2 \quad (9)$$

$$\sigma_{0_0}^2 = \sigma_{th}^2 + 2qR_d(P_{sp} + P_{CT0})B + \sigma_{CT0-sp}^2 + \sigma_{sp-sp}^2 \quad (10)$$

where  $\sigma_{1_0}^2$  is the noise variance when signal bit 1 is interfered by crosstalk due to bit 0 and  $\sigma_{0_0}^2$  is the noise variance when signal bit 0 is interfered by crosstalk due to bit 0.  $\sigma_{1_1}^2$  is the noise variance when signal bit 1 is interfered by crosstalk due to bit 1 and  $\sigma_{0_1}^2$  is the noise variance when signal bit 0 is interfered by crosstalk due to bit 1.  $\sigma_{th}^2$  is variance of thermal noise,  $q$  is the electronic charge,  $R_d$  is the receiver responsivity,  $B$  is the bandwidth of the receiver low-pass filter,  $P_s$  is signal power.  $i_1$  is the photocurrent for a transmitted bit 1 and is given by  $i_1 = 2R_d P_s$ , The photocurrent  $i_0 = 0$  for a transmitted bit 0, with  $P_s$  assumed to be zero. The expression for amplified spontaneous emission (ASE) power,  $P_{sp}$  is given by  $P_{sp} = hf\eta_{sp}(G-1)\Delta\nu_{sp}$ , where  $h$  is the plank's constant,  $f$  is the optical frequency,  $\eta_{sp}$  is spontaneous emission factor,  $G$  is the optical amplifier gain, and  $\Delta\nu_{sp}$  is the effective bandwidth of spontaneous emission. The crosstalk power due to bit 1 is  $P_{CT1} = P_{i0_1}^{out(ref)} - P_{i0_1}^{out}$  and due to bit 0 is  $P_{CT0} = -P_{i0_0}^{out}$ . The corresponding crosstalk currents are  $i_{CT1} = R_d P_{CT1}$  and  $i_{CT0} = R_d P_{CT0}$  respectively.

#### IV. RESULTS AND DISCUSSION

Using the analytical approach presented in section III, the

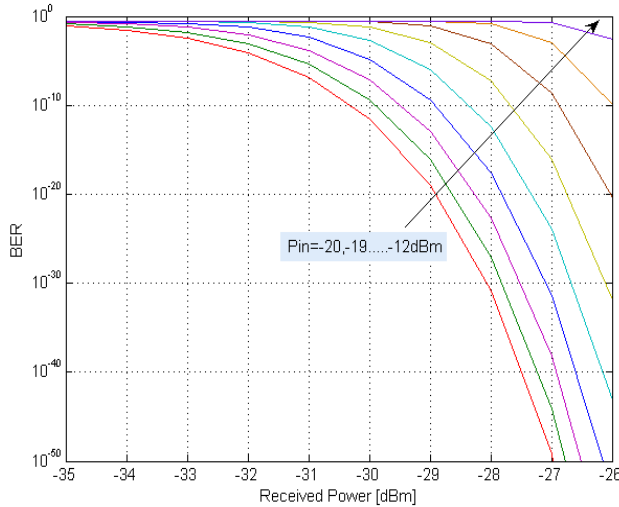


Fig. 2 : Plot of BER as a function of received power for 10 Gb/s optical WDM system with proposed OXC topology

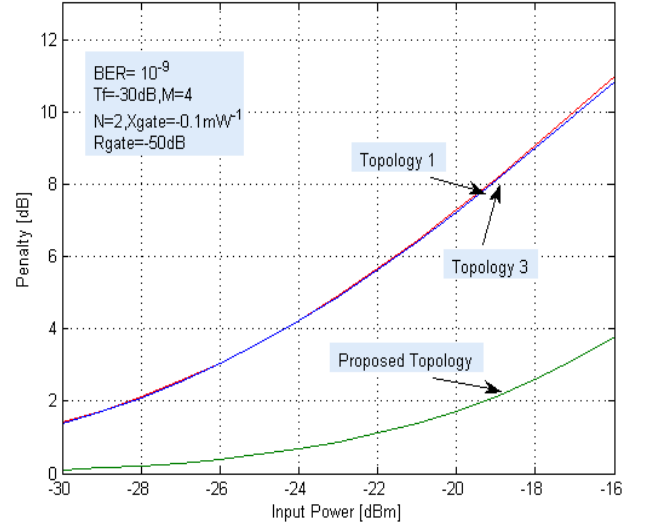


Fig. 3: Comparison of power penalty for the proposed topology and the existing topologies 1 and 3. The penalty is calculated at 10Gb/s for  $M = 4$ ,  $N = 2$ ,  $T_f = -30\text{dB}$ ,  $X_{gate} = 0.1\text{mW}^{-1}$  and  $R_{gate} = -50\text{dB}$ .

BER performance is evaluated for the proposed system at 10 Gb/s taking into account of OXC-induced crosstalk, ASE, and receiver noises. The plots of BER vs received power is shown in fig 2 with different values of input powers. It is observed from fig. 2 that the BER decreases with increasing the received power for a particular input power. However, for a particular received power, the BER is found to increase with increasing the input power. The power penalties due to crosstalk are determined for the topology 1 and 3 [4] and the proposed topology at BER of  $10^{-9}$ . The results obtained for  $M = 4$ ,  $N = 2$ ,  $T_f = -30\text{ dBm}$ , and  $R_{gate} = -50\text{ dBm}$  are shown in fig. 3. Form the plots, it is observed that the system suffers a power penalty due to OXC-induced crosstalk and the penalty increases with increasing input power for a given number of wavelengths and fibers. As seen in fig. 3 the power penalty for the topology 1 and 3 is almost the same. But, for the proposed topology the penalty is found to be lower than the topology 1 and 3. For a 2 dB penalty, the allowable input power for the topology 1 and 3 is less than -29 dBm. On the other hand, for the proposed topology it is less than -19 dBm. The increase in penalty at higher input power can be compensated if the performance of the switching matrix as well as that of filter is improved and wavelength converter is used.

The plots of penalty versus the number of wavelengths and the number of fibers for different topologies are shown in figs. 4 and 5, respectively. It is noticed from figs. 4 and 5 that there is an increase in penalty with increasing number of wavelengths as well as the number of fibers, at a given input power. The proposed topology shows lower crosstalk level and lower power penalty than the topology 1 and 3.

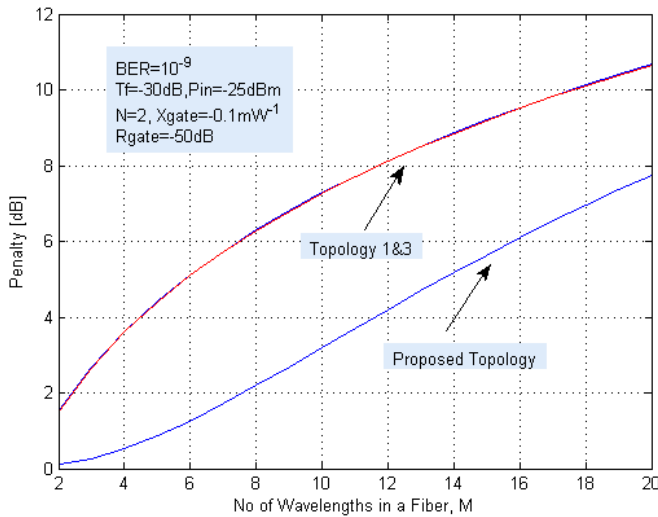


Fig. 4: Power penalty vs number of wavelengths in a fiber for different topologies. Data for topology 1 and 3 are taken from Ref. [4].

From the above discussion it is concluded that OXC-induced crosstalk limits the number of wavelengths, the number of input fibers as well as allowable input power per fiber in a WDM system. Considering the system impairments, the performance of the proposed topology is found to improve significantly than the existing topology 1 and 3.

## V. CONCLUSION

In this paper, crosstalk and BER performances of a WDM network are evaluated analytically for a proposed OXC topology and then compared with the existing topologies 1 and 3. Different factors that affect magnitude of crosstalk

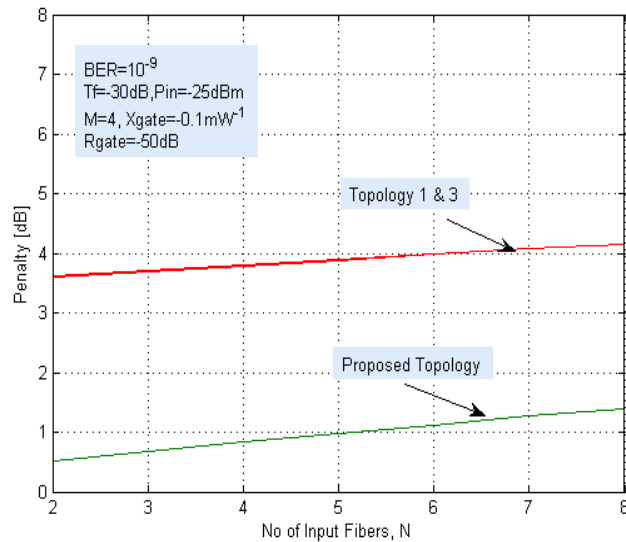


Fig. 5 : Power penalty vs number of input fibers for different topologies. Data for topologies 1 and 3 are taken from Ref. [4].

and BER in the OXC are also investigated in details. The system performance is evaluated in terms of power penalty at BER of  $10^{-9}$  with different OXC parameters and noises. It is found that the allowable input power for the proposed topology is higher than the exiting topologies, which means that the influence of OXC-induced crosstalk is lower in the proposed topology than the existing topologies. The results obtained from the present study indicate that for an allowable power penalty more number of wavelengths or fibers can be used in the proposed OXC topology than the existing topologies 1 and 3. Therefore, system throughput ( $N$  multiplied with  $M$ ) can be increased by the proposed topology with the suffering of minimum power penalty.

## REFERENCES

- [1] C. A. Brackett, "Is there an emerging consensus on WDM networking," *J. Lightwave Technol.*, pp. 936-941, vol. 14, June 1996
- [2] D. J. Blumenthal, M. Shell, and M. D. Vaughn, "Physical limitations to scalability of WDM all-optical networks," *Opt. Photon. News*, Feb. 1997.
- [3] G. Soulage et al., " $4 \times 4$  space-switch based on clamped-gain semiconductor optical amplifiers in a  $16 \times 10$  Gbit/s WDM experiment," in Proc. ECOC 1996, ThD 2.1, pp. 4.145-4.148.
- [4] Tim Gyselings, G. Morthier and R. Baets, "Crosstalk analysis of multiwavelength optical cross connect", *Journal of Lightwave Technology*, pp. 1273-1283, vol. 17, no. 8, August 1999.
- [5] S.P Majumder, Syeda Faria Mahmud, Md.Hossam-e-Haider, "Crosstalk Limitation of a WDM System due to Optical Cross-Connect (OXC)," Proc. in ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh, pp. 159-162.
- [6] Ngee Thiam Sim, S.P Majumder, M S Islam, "Bit Error Rate and Crosstalk Performance in Optical Cross-Connect with Wavelength Converter" Proc. In ICECE 2006, IEEE Catalog No.: 06EX1362, 19-21 December 2006, Dhaka, Bangladesh pp. 278-281.
- [7] E. Iannone and R. Sabella, "Optical path technologies: A comparison among different cross-connect architectures", *Journal of Lightwave Technology*, vol. 14, no. 10, pp. 2184 - 2195, October 1996.
- [8] T. Y. Chai, H. Chen, S. K. Bose and C. Lu, "Crosstalk analysis for limited wavelength interchanging cross connects", *IEEE Photonics Technology Letters*, vol.14, no.5, pp. 696-698, May 2002.
- [9] Y. Shen, K. Lu and W. Gu, "Coherent and incoherent crosstalk in WDM optical networks", *Journal of Lightwave Technology*, vol. 17, no. 5, pp. 759-764, May 1999.
- [10] J. Zhou, E. Casaccia Cavazzoni and M. J. O'Mahony, "Crosstalk in multiwavelength OXC networks", *Journal of Lightwave Technology*, vol. 14, no. 6 pp. 1423-1435, June 1996.
- [11] J. Sun, G. Morthier, and R. Baets, "Numerical and theoretical study of the crosstalk in gain clamped semiconductor optical amplifiers," *IEEE J. Select. Topics Quantum Electron.*, vol. 3, pp. 1162-1167, Oct. 1997.
- [12] Y. S. Jang, C.-H. Lee, and Y. C. Chung, "Effects of crosstalk in WDM systems using spectrum-sliced light sources" *IEEE Photonic Technology Letters*, vol. 11, no. 6, pp. 715-717, June 1999.



# Solar Cell Efficiency Improvement Using Compound Parabolic Concentrator and an Implementation of Sun Tracking System

Eklas Hossain, Riza Muhida, Ahmad Faris Dzulkpli, and Khairul Azizi Abdul Rahman

Department of Mchatronics Engineering, International Islamic University Malaysia, Jalan Gombak,  
53100 Kuala Lumpur, Malaysia

**Abstract** — This paper presents how the efficiency of solar cell can be increased with the design of compound parabolic concentrator and the implementation of sun tracking system. Sun tracking systems is an application of the Machine Vision (MV) and collaboration with Data Acquisition (DAQ) to systems by using a webcam as a sensor and sound card as an output channel to drive a motor. The motors react as the mechanism of the camera to make sure it always focuses on the target of sun. By using a MatLab software programming environment, closes loop control is implemented to provide interaction between camera and sound card as a device to control the motor to track the sun.

**Index Terms** — Solar cell, concentrator, machine vision, close loop control, sun tracking.

## I. INTRODUCTION

Sunlight is a source of energy has been regarded as an alternative source of energy along with other alternative source such as hydrogen and natural gas. Solar energy is a vital that can make environment friendly energy more flexible, cost effective and commercially widespread. Photovoltaic source are widely used today in many applications such as battery charging, water heating system, satellite power system, and others. Although the installation cost is high but the advantage of photovoltaic source is free population.

The problem identification for the PV solar cell is the solar cell has low efficiency if the position of the solar cell in static location. It needs maximum heat sun light to produce the highest efficiency. Therefore, the compound parabolic concentrator is used to solve this problem. When the tracking system is combined with the compound parabolic, it can produce the voltage in the maximum efficiency. Compound parabolic concentrator is a significant of photovoltaic source because it is the combination of parabolic concentrator and the solar panel. The purpose of using parabolic concentrator is to get the ray reflection of light at the highest efficiency and to get the source of energy at the maximum value. The development of this system is as an alternative of power source and also to increase the source of the electricity at the highest efficiency from morning until afternoon.

Solar radiation can be converted into useful energy directly. In order to get the maximum efficiency of solar

cell, the sun has to be targeted at focal point. Since sun is changing position all the time, it should be tracked. Currently there are some researches on sun tracking systems, but almost all of them were using photosensitivity sensors as devices. The mechanism which follows the sun is not a new problem. In 1962 in Chile, Finster first attempted was completely mechanical [1]. Later Saavedra [2] presented a mechanism with automatic electronic control to orient an Eppley pyrheliometer. A recent review of the existing systems has been published by Myers et al. [3]. Roth, Georgiev, and Boudinov [4] presented a design of a system for sun-tracking. In this paper, the webcam and sound card are chosen as alternative devices to get the best performance of the tracking system with machine vision method and data acquisition (DAQ). By using the webcam as a sensor, the machine vision techniques are used to implement to make the systems efficient.

## II. DESIGN CONFIGURATION AND SELECTION

Figure 1 illustrates the design of the project starts with the base design by using CATIA programs and AUTOCAD programming. The design is used to measure and plot the dimension of the prototype of the project. It can be used as a guideline before start doing assembly process of the prototype. There are two types of DC motor that uses in this project, the 6W DC motor for pan rotation of the CPC, and 40W DC motor for tilt rotation of the base. The worm gear is combined with the 6W DC motor to make the CPC rotate and hold its position when the motor stop. The worm gear is combining with spur gear with ratio reduction 10:1. The DC motor also is combined with the gear head. The gear head is used to reduce the speed from high rotation to slow rotation with reduction ratio 50:1. Therefore, the combination of both reduction gears will produce 500:1 of the reduction speed. The slow rotation of the CPC is important because the movement of the sun is slow and it is easy to control the position of the CPC when doing the tracking process. The reduction gear for base rotation is also same as 6W DC motor.

Figure 2 illustrates the dimension of the CPC. The material that use in designing compound parabolic concentrator are aluminum beam for supporting base,

stainless steel for reflection, and camera as a sensor. The most important part in designing CPC is the focus point location of the solar cell. The focus point is the place where the location of the sunray is reflected on that point. Therefore, the calculation of the focus point of CPC is based on the standard parabolic formula.



Fig. 1. The base design and full design

Figure 3 illustrates the curve and calculation of the CPC. From the calculation shows that the focus point of the reflection is 15cm. therefore, the solar cell should be located at this position.

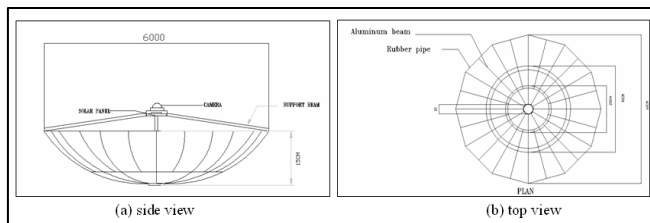


Fig.2. The dimension of the compound parabolic concentrator

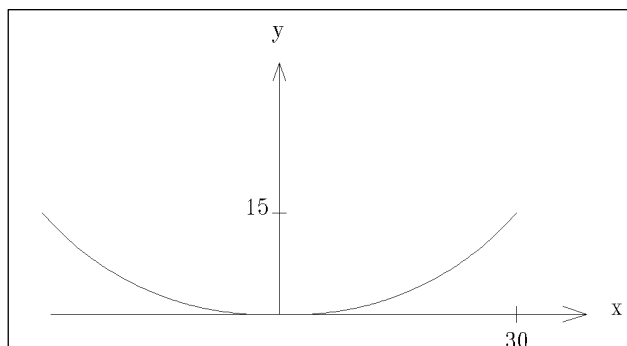


Fig.3: The curve of the parabola and focus point

### III. FLOW OF THE SYSTEM

Figure 4 shows the system of this project uses web camera as a sensor detector of the sun. Then it will connect with computer and process by using matlab

programming. In this program the image will process by using simulink image processing tool. From the matlab programming, it will be used to determine the center position of the sun. The signal will be created from the matlab programming to create the PWM (Pulse Width Modulation) to control the speed of motor. This system uses two types of signal, sound card from motherboard and external sound card. The signal from the sound card produce negative and positive signal. The signal positive signal will generate the forward rotation and the negative signal will generate backward rotation. The output signal is applied at the H bridge circuit and amplified with Opamp. The opamp is used to amplify the signal to increase the value of the output voltage. This two sound card will control the motor at tilt rotation and pan rotation. The rotation of the motor is used to make the position of the camera center with the sun position.

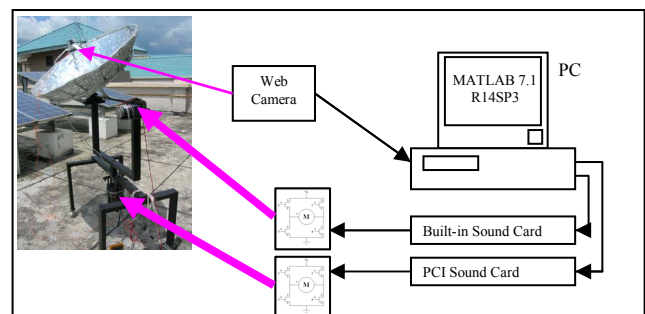


Fig. 4. Schematic diagram for the sun tracking system

Figure 5 shows the controlling system for the sun tracking system uses MATLAB simulink programming and M file to control the speed of motor rotation. Basically, the MATLAB simulink is used to process the data and image processing of the sun. The process that includes in the image processing step is thresholding, median filter, image enhancement, centroid and quadrant of the image. All the steps are important because it is the way to determine the center position of the sun with camera. Part A is included the process of the image processing. It is also include the pixel adjustment of the camera. The adjustment of the pixel is very important because it can protect the camera from any damage. From the image processing, the position of the camera will be controlled by using M file coding to generate the PWM signal. The PWM signal is produced at the sound card. All the controlling motor rotation is at part B. The motor of tilt and pan rotation is controlled at this part and it will default condition when the position of the camera and sun is centered. The centroid position of the camera is displayed at the centroid display. The pulse generator is used to generate the controlling system according



whatever time we want. In this system, the system is set to do the process for every 30 second. Therefore, the process of the controlling system will be generating in every 30 second.

All the four images is the output from the image processing toolbox in the simulink programming. The original image is the image that produces form the camera. The camera is filtered with film. The purpose of using film

is to filter the sunray and to get only the picture of the sun. The enhancement of the image is used to reduce the pixel of the sun resolution. It is also to reduce the noise of the sun image. Form the enhancement process, the sun image is thresholding to get black and white of the image. The black and white is used to get the quadrant position of the centroid of the image. The centroid is the center position of the object.

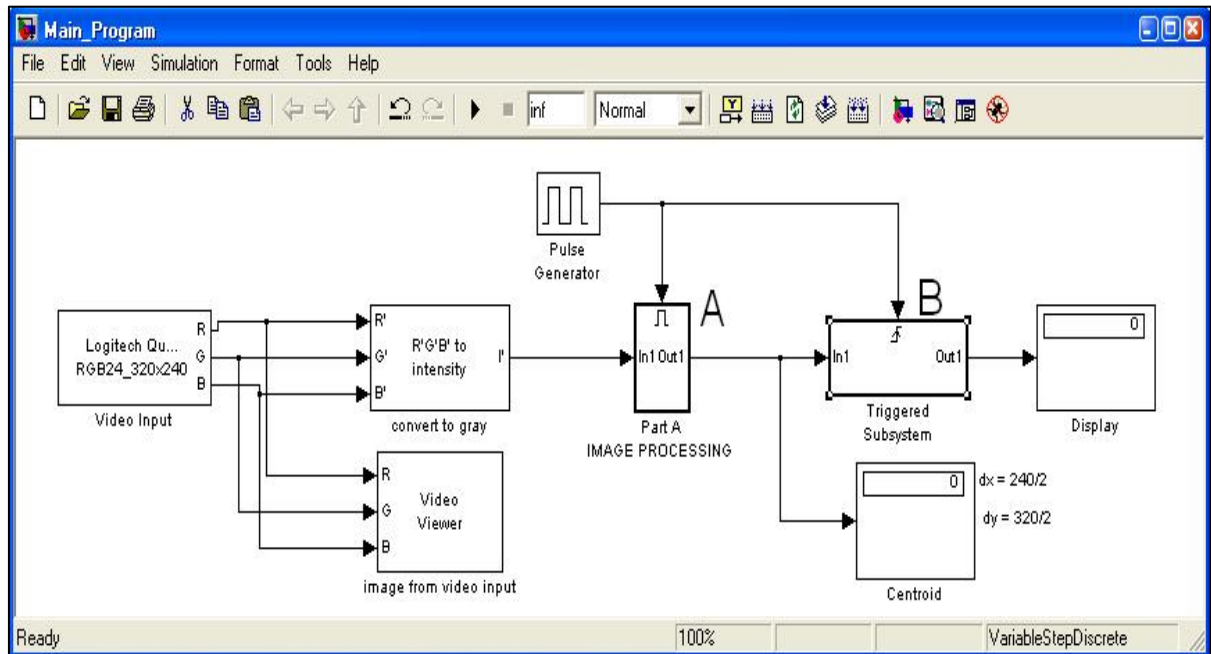


Fig. 5. The whole controlling system

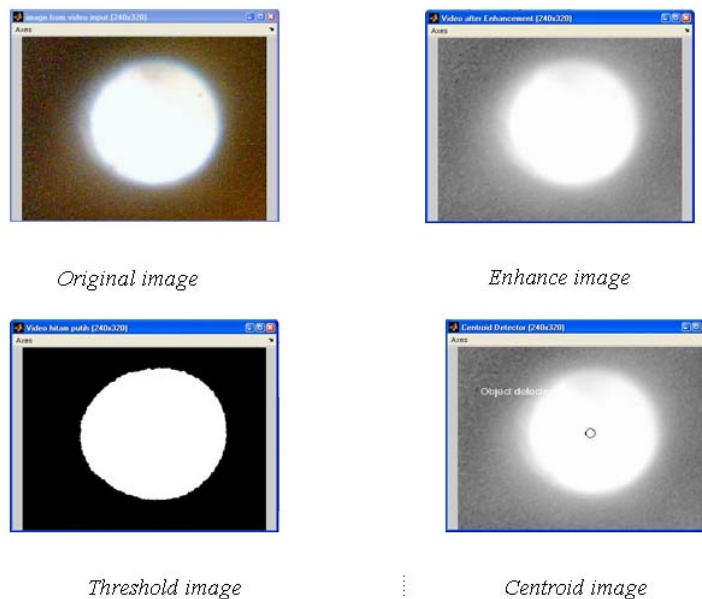


Fig. 6. The image processing

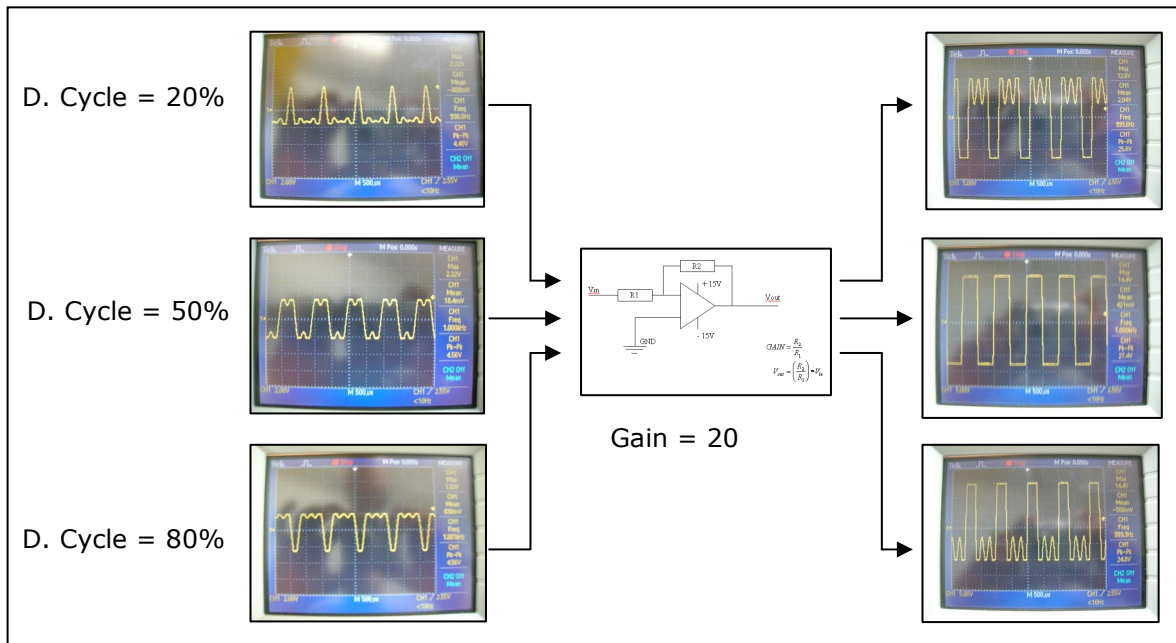


Fig. 7. Pulse Width Modulation (PWM)

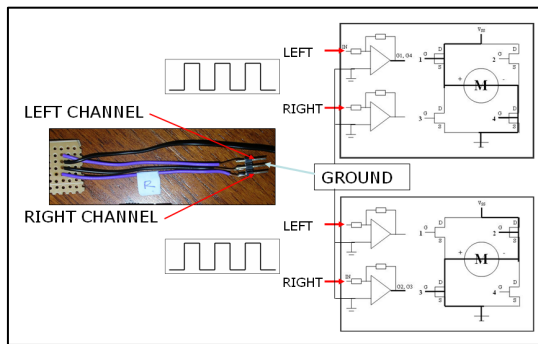


Fig. 8. h-bridge circuit

Figure 7 illustrates the matlab command can generate PWM signal from the sound card. The PWM is very important because it is used to control the motor speed rotation. The speed of motor is depending on the duty cycle of the PWM signal. The motor will rotate at high speed rotation if the duty cycle is high. The impulse signal from the sound card is small and it is not enough to generate the motor. Therefore, the signal needs to amplify at the opamp. The opamp is used to increase the impulse of the signal. The impulse of the signal is depend on the gain of the resistor that uses in the opamp. From the figure

shows that, the signal that produce from different types of duty cycle and the output that produces when it is amplified at the opamp with gain 20. When the signal is amplified, the value of the impulse voltage is increased and it is enough to make the motor move. From the opamp, it will go to the H bridge circuit. Figure 8 shows the H bridge circuit that controls the movement of the DC motor. The H bridge circuit is used to rotate the motor move forward and backward.

#### IV. EXPERIMENTAL RESULTS

Figure 9 illustrates the result of the experimental graph that plot during the experiment test. The result of the experiment is the comparison between solar cell at CPC and non CPC. From the experiment, the result of efficiency of solar cell at the CPC is better than non CPC. The experiment shows that the power of solar cell can successfully produce when it is located at the CPC. The range of Current (I) vs Temperature ( $^{\circ}\text{C}$ ) is shown in Figure 10. It is evident that CPC system's temperature increase in much higher than non CPC. Our proposed design and system could provide better performance improvement for solar cell energy improvement.

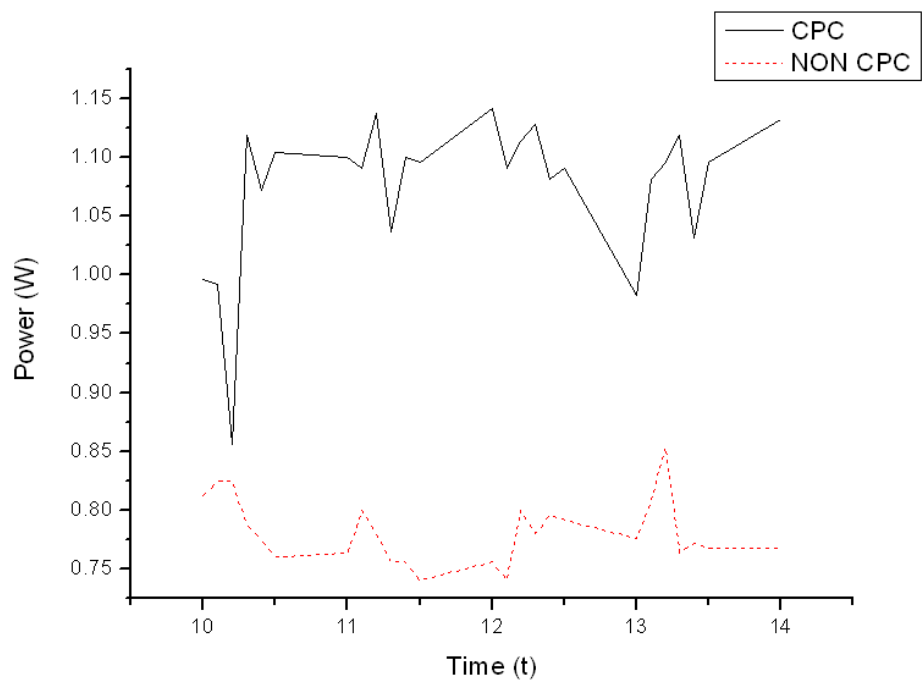


Fig. 9. Graph Power (W) vs Time (h)

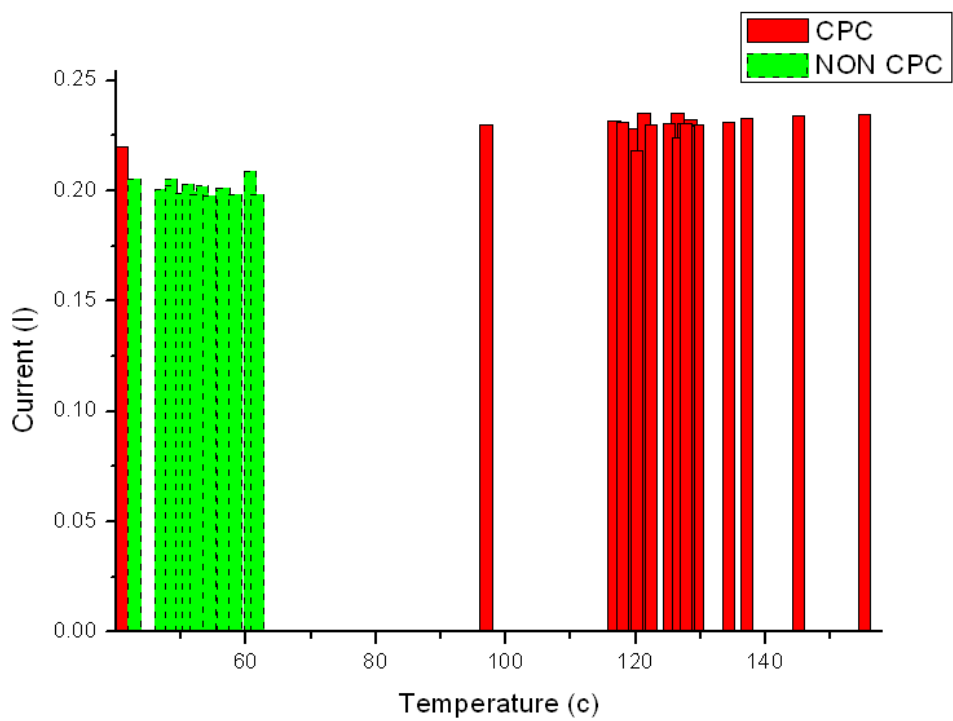


Fig. 10. Graph for the range Current (I) vs Temperature (°C)

## V. CONCLUSIONS

Camera is the sensor of the system, be an eye of the human to take care and observe the sun. As soon the sun can be tracked with the CPC, the voltage, current and power received will always be at the maximum of the efficiency. If the efficiency becomes the matter of the improvement of the solar cells, now with this project we already surpass the problems. This project should be continued and upgrade from time to time to get the higher efficiency to the maximum.

## REFERENCES

- [1] C. Finster, "El heliostato de la Universidad Santa Maria," *Scientia*, 119, pp. 5–20, 1962.
- [2] A. S. Saavedra, "Diseño de un servomecanismo seguidor solar para un instrumento registrador de la irradiación solar directa," *Memoria*, Universidad Técnica Federico Santa Maria, Valparaiso, Chile, 1963.
- [3] D. R. Myers, T.L. Stoffel, I. Reda, S.M. Wilcox and A.M. Andreas, "Recent progress in reducing the uncertainty in and improving pyranometer calibrations," *Journal of Solar Energy Engineering*, vol. 124, pp. 44–50, 2002.
- [4] P. Roth, A. Georgiev, and H. Boudinov, "Design and construction of a system for sun-tracking," *Renewable Energy*, vol. 29, no. 3, pp. 393–402, 2004.
- [5] T. Pavlidis, "Computers versus Humans", November 2002, <http://www.theopavlidis.com/comphumans/comphuman.htm>.
- [6] M. H. Rashid, *Power Electronics Circuits, Devices, and Applications*, 2004.
- [7] K. Sandberg, *Introduction to image processing in Matlab 1*, Department of Applied Mathematics, University of Colorado at Boulder, 2007.
- [8] Machine Vision, [//en.wikipedia.org/wiki/Machine\\_vision](http://en.wikipedia.org/wiki/Machine_vision).
- [9] A. Basid, and N. Azah, *Design and Development of Automated Visual Servoing System (AVSS)*, IIUM, Malaysia, November 2004.
- [4] W. R. Norshuhada, *Design and Development of Camera Controlled by Using X-Y Cartesian Robot*, December 2003.
- [4] R. Jain, R. Kasturi, and B. G. Schunk, *Machine Vision*, McGraw-Hill International Editions, 1995.

# Comparison of GARCH and Neural Network Methods in Financial Time Series Prediction

Altaf Hossain and M. Nasser

Department of Statistics, Rajshahi University, Rajshahi-6205, Bangladesh

E-mails: rasel\_stat71@yahoo.com; mnasser.ru@gmail.com

**Abstract** — Many researchers have already published huge number of papers comparing Autoregressive (AR) model, a model based on Box-Jenkins methodology, and Back Propagation Artificial Neural Network (BPANN) in financial time-series forecasting. Among them, some compared SVMs and BPs taking AR as a benchmark in forecasting the six major Asian stock markets. They showed that both the SVMs and BPs outperform the traditional models, ARs. They did prediction of transformed data, but not level data. They did not take account of GARCH model, specially developed to model financial time series. Generalized Autoregressive Conditional Heteroskedastic (GARCH) model is needed to capture high volatility for better forecasts. This article applies GARCH model instead AR or ARMA model to compare with standard BP in forecasting of the four international including two Asian stock markets indices. Our fitted GARCH models give better forecasts than the fitted standard BP models in forecasting of the four international markets indices except one market.

**Index Terms** — Autoregressive Moving Average (ARMA); Generalized Autoregressive Conditional Heteroskedastic (GARCH) and Back Propagation Artificial Neural Network (BPANN).

## I. INTRODUCTION

A time series  $\{Y_t\}$  is a discrete time, continuous state, process where  $t: t=1, 2, \dots, T$  are certain discrete points. Usually time is taken at equal spaced intervals and the time increment may be from everything from seconds to years. The study of time series is concerned with time correlation structures. It has diverse applications ranging from oceanography to finance. In the business and economic environment, it is very important to predict various kinds of financial variables accurately in order to develop proper strategies and avoid the risk of potentially large losses. There are some classical methods that have been developed in predicting financial time series. Among them, there is a powerful method, available in the literature for univariate time series forecasting, known as Box-Jenkins [2] ARMA (Autoregressive Moving Average) approach on stationary time-series. In conventional econometric models, the variance of the disturbance is assumed to be constant. But many economic and financial time series such as exchange rates, indices, market returns, inflation rate etc exhibit periods of unusual large volatility, followed by periods of relative

tranquility (i.e., time series exhibits clustering of large and small disturbances). Such circumstances suggest a form of heteroskedasticity in which the variance of the disturbance depends on the size of preceding disturbance and hence the conditional variance is non-constant over the sample period. Engle [5] showed that it is possible to model the mean and the variance of a series simultaneously by a model named Autoregressive Conditional Heteroskedastic (ARCH); whereas Box-Jenkins [2] modeled only the mean series. ARCH model allows the conditional variance to be an AR process. In latter, Bollerslev [1] extended Engle's original work by developing a technique that allows the conditional variance to be an ARMA process and that extended process is known as the GARCH process.

The classical methods are based on some specific assumptions, such as linearity; or on error distributions, such as normality. In this circumstance, ANNs were developed to meet the increasing demand that can predict, detect, classify and summarize the structure of variables and define the relationships between them – without relying too much on such assumptions. Not all relationships in economics and finance are direct. But the hidden layer of an ANN can capture all non-direct relationships between input and output variables. ANNs have been successfully applied in different areas, including marketing, retail, banking and finance, insurance telecommunications and operations management (Smith and Gupta [11]). The reason for its popularity among econometricians can be attributed to the ability to approximate any nonlinear relationship with a reasonable degree of accuracy. A number of research articles used Artificial Intelligence (AI) methods to forecast the price behavior of the financial commodities and compare the results with the traditional statistical methodologies. Refenes et al. [10], for example, modeled stock returns with ANNs in the framework of Arbitrage Pricing Theory (APT) and compared their study with regression models. Steiner and Wittkemper [12] investigated the performance of several ANN models that forecast the return of a single stock. Kamruzzaman and Sarker [7] modeled and forecasted currency of exchange rates using three ANNs and a comparison was made with traditional ARIMA model. In most of studies like the above, they showed that even the simple neural learning

procedures, such as BP algorithms easily outperform the best practice of traditional statistical models. They reported that the relationship between the return of a stock and factors of the capital market is usually non-linear and may contain dependencies in time. Chen and Wu [3] compared SVMs and BPs taking AR as a benchmark in forecasting the six major Asian stock markets. They showed that both the SVMs and BPs outperform the traditional models, ARs. They did prediction of transformed data, but not level data.

Although the superiority of the AI models has been shown comparing with traditional statistical methods like ARs in financial markets, as far we know, no research has been carried out comparison between the recently developed statistical method, GARCH and AI methods. So far the GARCH model has been developing theoretically. Few applied works regarding GARCH, especially, in forecasting financial time series are available in the literature.

The data source and nature are introduced in Section II. The methodology including two prediction methods in this paper is discussed in Section III. The results and discussion are presented in Section IV. Finally, our conclusion is in Section V.

## II. DATA SOURCE AND NATURE

For analysis we have used the monthly data of the four international stock market indices including two Asian markets, collected from the *Financial Forecast Center Home Page*. The names of the stock markets are Japan (Nikkei 225 from January, 1984 to January, 2006), Hong Kong (Hang Seng from December, 1986 to January, 2006), U.K. (FTSE 100 from April, 1984 to January, 2006), and Germany (DAX from November, 1990 to January, 2006).

## III. METHODOLOGY

A time-series is a dynamic consequence of events over time. It can be interpreted as a stochastic process. A time series of successive observations is regarded as a sample realization from an infinite population which has been generated by a process. A single equation ARMA model states how any value in a single time series is linearly related to its past values. In theory, if the “correct” model is chosen, the ARMA forecasts are “best”. See Box & Jenkins [2] and Pankratz [9]. Standard ARMA analysis rests on the simplifying assumption that the mean and unconditional variance of time series is independent of time, i.e., the series is stationary. The plot of ACF (Autocorrelation Function) and PACF as well some conventional tests like ADF (Augmented Dickey Fuller) test are used to decide whether a data series is stationary

or not. For non stationary time series, the data is transformed to be stationary; see Pankratz [9] for details. In the article, the most recent developed model, GARCH, has been used to overcome high volatility problem in forecasting financial time series. In the GARCH process, we additionally model the volatility (i.e., the variance of the disturbance) and most of the tests like stationarity are similar to that of the ARMA process.

### A. The ARCH and GARCH Modeling

Engle [5] showed that the serial correlation in squared returns, or conditional heteroskedasticity, can be modeled using an autoregressive conditional heteroskedasticity (ARCH) model of the form

$$y_t = E_{t-1}[y_t] + \epsilon_t \quad (2)$$

$$\epsilon_t = z_t \sigma_t \quad (3)$$

$$\sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + a_2 \epsilon_{t-2}^2 + \dots + a_q \epsilon_{t-q}^2 \quad (4)$$

where  $E_{t-1}[\cdot]$  represents expectation conditional on information available at time  $t-1$  and  $z_t$  is a sequence of iid random variables with mean zero and unit variance. In the ARCH model  $z_t$  is assumed to be iid standard normal. The restrictions  $a_0 > 0$  and  $a_i \geq 0$  ( $i = 1, \dots, q$ ) are required for  $\sigma_t^2 > 0$ . The representation (2) - (4) is convenient for deriving properties of the model as well as for specifying the likelihood function for estimation.

An important extension of the ARCH model proposed by Bollerslev [1] replaces the AR (p) representation in (4) with an ARMA (p,q) formulation

$$\sigma_t^2 = a_0 + \sum_{j=1}^p b_j \sigma_{t-j}^2 + \sum_{i=1}^q a_i \epsilon_{t-i}^2, \quad (5)$$

where the coefficients  $a_i$  ( $i = 0, \dots, q$ ) and  $b_j$  ( $j = 1, \dots, p$ ) are all assumed to be positive to ensure that the conditional variance  $\sigma_t^2$  is always positive. The model in (5) together with (2)-(3) is known as the generalized ARCH or GARCH (p,q) model. When  $p=0$ , the GARCH model reduces to the ARCH model.

The GARCH model reduces the number of estimated parameters from infinitely many to only just a few. GARCH has gained fast acceptance and popularity in the financial world. This can be explained by various arguments: the GARCH process has a close relation to ARMA processes. This suggests that the theory behind the GARCH process might be closely related to the theory of ARMA processes, which is well studied and widely known. That is, there is no logical difference between ARMA and GARCH modeling. In the GARCH modeling, we model and forecast the mean series additionally with the conditional variance; as a result, GARCH is more able than ARMA to capture the unusual volatility. For testing ARCH/GARCH effects, assessing the fit, diagnostic checks for model adequacy and estimation see Enders [4].

### B. Parsimonious Property of GARCH over ARCH

A high-order ARCH model may have a more parsimonious GARCH representation that is much easier to identify and estimate. Usually a GARCH (1,1) model with only three parameters in the conditional variance equation is adequate to obtain a good model fit for financial time series. Indeed, Hansen and Lunde [6], provided compelling evidence that it is difficult to find a volatility model that outperforms the simple GARCH (1,1).

### C. ANNs for Time Series Forecasting

An ANN is a biologically inspired form of distributed computation. It simulates the functions of the nervous system by a composition of interconnected simple elements (artificial neurons) operating in parallel, as shown in Figure 1. An element is a simple structure that performs three basic functions: input, processing and output. ANNs can be organized into several different connection topologies and learning algorithms (Lippmann [8]). In case of time-series, the series at different lags act as inputs of ANNs. The number of inputs to the network is constrained by the problem type, whereas the number of neurons in the output layer is constrained by the number of outputs required by the problem type. However, the number of hidden layers and the size of the layers are decided by the designer.

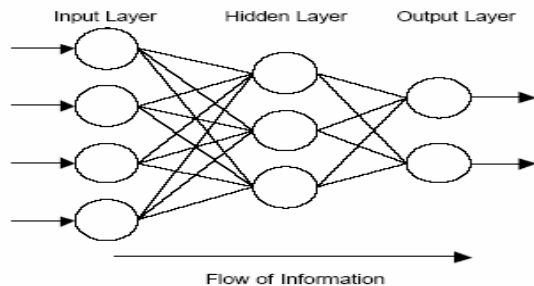


Figure 1 A neural network with one hidden layer

BP trains multilayer feedforward networks with differentiable transfer functions to perform the function approximation, pattern association and pattern classification. It is the process by which the derivatives of network error, with respect to network weights and biases, are computed to perform computation backwards through the network. Computations are derived by the chain rule of calculus.

There are several different BP training algorithms with a variety of different computation and storage requirements. Kamruzzaman and Sarker [7] used three algorithms, Standard Backpropagation (SBP), Scaled Conjugate Gradient (SCG) and Backpropagation with Bayesian

Regularization (BPR), in training the ANN. It was found that SCG based model performs best when measured on the two most commonly used metrics and showed competitive results when compared with BPR based model on other three metrics. Chen and Wu [3] used only the SBP and mentioned, "No single algorithm is best suited to all the problems." So we consider SBP for training the ANN models in our paper. All the algorithms use the gradient of the performance function to determine how to adjust the weights to minimize the performance. For example, the performance function of the feed forward networks is the Mean Square Error (MSE). The training process requires a set of examples of the proper network inputs and target outputs. During the training, the weights and biases of the network are iteratively adjusted to minimize the network performance function.

### D. Performance Criteria

Although the MSE is a perfectly acceptable measure of performance, in practice the ultimate goal of any testing strategy is to confirm that the results of models are robust and capable of measuring the profitability of a system. It is important, therefore, to design a test from the outset. According to Tay and Cao [13] and Thomason [14], [15], the prediction performance is evaluated using the following statistics: MSE, Normalized Mean Squared Error (NMSE), Mean Absolute Error (MAE), Directional Symmetry (DS) and Weighted Directional Symmetry (WDS). These criteria are defined in the Table 1. See Chen and Wu [3] for more details.

## IV. RESULTS AND DISCUSSION

### A. Explorative Data Analysis (EDA)

We examined the four international stock market indices. The time periods used and the indices' statistical data are listed in Table 2, and the monthly closing prices used as the data sets are plotted in Figure 2. The time periods cover many important economic events, which we believe are sufficient for the training models.

### B. Preprocessing and Design of The Models

We transform the original closing price into white noise series differencing two times at lag one for each series except the market NK, where we take one time difference. White noise series is the first requirement for a data series to which we can apply the classical statistical techniques as well as the AI models for better fitting. According to Thomason [14], [15] and Tay and Cao [13], this transformation has many advantages. The most important is that the distribution of the transformed data becomes more symmetrical so that it follows a normal distribution

more closely. This also improves the predictive power of neural network.

TABLE 1  
PERFORMANCE METRICS AND THEIR  
CALCULATION

METRICS	CALCULATION
MSE	$MSE = \frac{1}{n} \sum_{i=1}^n (a_i - p_i)^2$
NMSE	$NMSE = \frac{1}{\delta^2 n} \sum_{i=1}^n (a_i - p_i)^2$ $\delta^2 = \frac{1}{n-1} \sum_{i=1}^n (a_i - \bar{a})^2$
MAE	$MAE = \frac{1}{n} \sum_{i=1}^n  a_i - p_i $
DS	$DS = \frac{100}{n} \sum_{i=1}^n d_i$ $d_i = \begin{cases} 1 & (a_i - a_{i-1})(p_i - p_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$
WDS	$WDS = \frac{\sum_{i=1}^n d_i  a_i - p_i }{\sum_{i=1}^n d'_i  a_i - p_i }$ $d_i = \begin{cases} 0 & (a_i - a_{i-1})(p_i - p_{i-1}) \geq 0 \\ 1 & \text{otherwise} \end{cases}$ $d'_i = \begin{cases} 1 & (a_i - a_{i-1})(p_i - p_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$

TABLE 2  
DESCRIPTION OF DATA SETS

INDEX	TIME PERIOD	N	HIGH	LOW	MEAN	SD
Nikkei 225 (NK)	1984.01-2006.01	265	38920	7837	18020	6643.7
Hang Seng (HS)	1986.12-2006.01	230	17410	2138	8977	4411.6
FTSE 100 (FT)	1984.04-2006.01	262	6930	1009	3559	1612.3
DAX (DX)	1990.11-2006.01	183	7645	1398	3637	1690.4
Here N indicates the number of realizations for each market						

The modification is known, so we can reverse it to compare the forecasts with the original series. Each of the

four data sets is partitioned into three subsets according to the time sequence in the ratio 95: 2.5: 2.5. The first part is used for training; the second is a validating set and the last is used for testing. In Standard BP models, we use one hidden layer of different units for each market. We select the number of inputs for the ANNs, and units for the hidden layer of that ANNs which give minimum MSE on the validation sets. The used activate functions are sigmoid in the hidden layer and linear in the output layer.

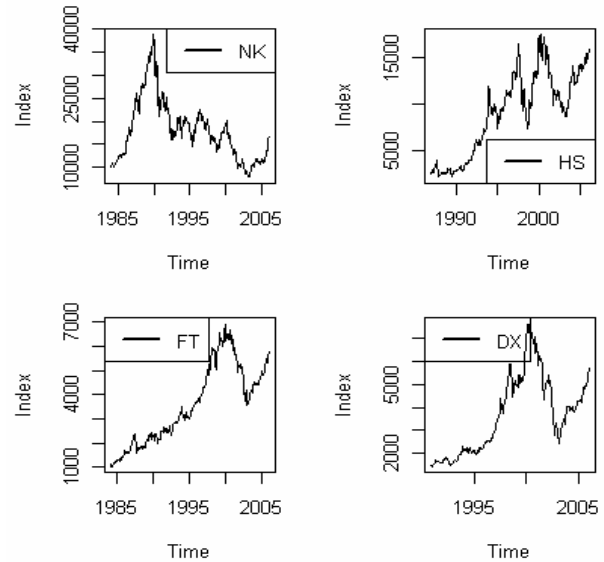


Figure 2 Indices of the four International stock markets

### C. Comparison and Discussion

For each market, we also design different ordered ARMA models as to show the necessity of GARCH model. The ACF of squared residuals for each fitted ARMA indicates the presence of GARCH effects. The forecasting results of the ARMA, GARCH and BP for the test set are collated in Table 3. In general, GARCH and BP models outperform the ARMA model in the deviation performance criteria. However, in the either direction and weighted direction performance criteria, the ARMA model shows better results than BP model for the four markets, possibly because BPs are trained in terms deviation performance. With regard to the direction criteria, the time-series model has an intrinsic merit because it emphasizes the trend of the series. Therefore, the BP model, even GARCH model cannot outperform ARMA in the direction criteria in the four international markets under study. Figure 3 shows the fitted lines of the GARCH, BP models and real value line where the fitted line of the GARCH model seems closer to the real value line than the BP model for different markets except the market FT for the period 1991-1998.



TABLE 3  
COMPARISON OF THE RESULTS OF NN AND  
GARCH MODELS ON THE TEST SET

		SQRT MSE	MAE	NMSE	DS	WDS
N K	ARMA (3,4)	3172.5	2706.4	3.12	57.1	0.66
	GARCH (1,1)	824.82	686.75	0.21	85.7	0.00
	BP (2,12)	900.91	765.64	0.25	71.4	0.42
H S	ARMA (3,4)	545.12	459.55	1.30	16.6	2.07
	GARCH (1,1)	656.85	525.59	1.88	0.00	-----
	BP (4,8)	778.94	684.11	2.65	33.3	1.18
F T	ARMA (4,3)	521.08	497.87	7.33	71.4	0.12
	GARCH (1,1)	196.91	170.64	1.04	42.8	1.23
	BP (7,4)	236.81	220.12	1.51	28.5	5.87
D X	ARMA (7,0)	961.35	921.55	9.23	60.0	0.38
	GARCH (1,1)	234.99	229.32	0.55	40.0	0.66
	BP (4,5)	191.05	156.14	0.36	60.0	0.79

The first order of BP ( , , ) indicates the number of inputs and the second order indicates the number of units used in the hidden layer.

Table 4 gives the comparison of the result of the GARCH and BP models. In the NK, HS and FT data sets, the GARCH models perform better than the BP models. On the other hand, the BP method performs only a little better than GARCH in only one market, DX. If we consider the ARMA models to compare with GARCH and BP models, both seem to perform better for three markets in terms of the deviation measurements. Only in the market HS, the ARMA performs better than both. For the other one directional criteria (DS), BPs work better than GARCHs in the NK and FT indices. The performance of GARCHs is not always better than the BPs in these performance criteria. Table 4 summarizes the two models. Generally speaking, based on the five performance criteria, GARCHs fit the NK, HS and FT indices well, whereas BP fits only DX indices.

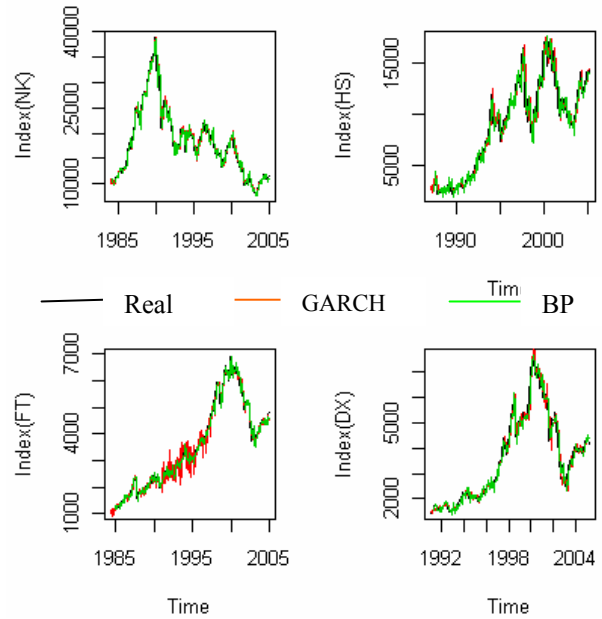


Figure 3 The values of indices fitted by GARCH and NN models and their real values

TABLE 4  
COMPARISON OF NN AND GARCH ON THE  
TEST SET

	SQRT MSE	MAE	NMSE	DS	WDS	SUMMARY
NK	+	+	+	-	+	+
HS	+	+	+	+	-	+
FT	+	+	+	-	+	+
DX	-	-	-	+	+	-

GARCH, '+' represents GARCHs that perform well and '-' represents NNs that perform well.

## V. CONCLUSION

In this article, we examine the feasibility of applying two models (GARCH and standard BP) to financial time-series forecasting for the four international stock markets including two Asian markets. Our experiments demonstrate that both models perform better than the ARMA model in the deviation measurement criteria. Previous research claims that the NN method is superior to the classical statistical method, ARMA or AR. If we consider the classical statistical method GARCH (just extended version of ARMA process), it is superior to the NN method except in the only one market DX of indices. The statistical methods, ARMA and GARCH, require large number of sample size for better forecasting and these drastically reduce the original sample size when the

high order model is fitted. From Figure 2, it is evident that DX data set is less volatile. Perhaps, due to this reason in this case standard BPANN (wins in three metrics out of five) slightly outperforms GARCH. Again if we look at the validation (short-time) errors, the ANNs perform better than the GARCHs – which indicates the generalization power of NN method is not as good as that of GARCH. If we investigate ANN with other training methods and compare with GARCH, what would be happened – it is the matter of study. ANN as well GARCH has several improved versions. The GARCH models are more parsimonious than the NN models since the GARCH models take smaller number of parameters for better forecasting. The GARCH models are also more interpretable than that of NN because the ANN structures contain hidden layers which are not interpretable. In general, the GARCH and the BP models perform well in the prediction of indices behavior in terms of the deviation criteria. Our experiments also show that both the statistical and the AI techniques can assist the stock market trading and the development of the financial decision support systems. But if we consider predictability along with interpretability and parsimonious property, then we recommend the GARCH models in financial time series prediction.

Our future study is to apply more complex GARCH methods and ANN with other training methods to the forecasting of Bangladesh stock markets and compare it with also the SVM methods. In our current GARCH models, we only select the return of index price information as input variables. But we want to enhance the performance of prediction GARCH models by including other efficient input variables (such as some macroeconomic variables).

#### ACKNOWLEDGEMENT

First I wish to my thanks and gratitude to my elder brother Faisal Zaman for his cordial guidance. I acknowledge the valuable suggestion of reviewer. I am also grateful to the core team of R language; using their packages our output is found.

#### REFERENCES

- [1] Tim. Bollerslev, “Generalized autoregressive conditional heteroscedasticity,” *Journal of Econometric* 31, 307-27, (1986).
- [2] G. E. P. Box, and G. M. Jenkins, *Time Series Analysis Forecasting and Control*, San Francisco: Holden-Day, (1976).
- [3] W-H. Chen, J-Y. Shih, and S. Wu, “Comparison of support-vector machines and back propagation neural networks in forecasting the six major Asian stock markets,” *Int. J. Electronic Finance*, Vol. 1, No. 1, pp.49-67, (2006).
- [4] Walter Enders, *Applied Econometric Time Series*, John Wiley & Sons, Inc., (2004).
- [5] Robert F Engle, “Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation,” *Econometrica* 50, 987-1007, (July 1982).
- [6] P. Hansen, and A. Lunde, A “Forecast comparison of volatility models: does anything beat a GARCH (1,1) model?,” *Journal of Applied Econometrics*, 20, 873-889, (2004).
- [7] J. Kamruzzaman, and R. Sarker, “Forecasting of currency exchange rates using ANN: a case study,” 2003 *Proc. IEEE Intl. Conf. on Neur. Net. & Sign. Process. (ICNNSP03)*, pp. 793-797, China, (2003).
- [8] R.P. Lippmann, “An introduction to computing with neural nets,” *IEEE ASSP Magazine*, April, pp.36-54, (1987).
- [9] A. Pankratz, *Forecasting with Dynamic Regression Models*, John Wiley & Sons Inc., (1991).
- [10] A-P. Refenes, A.D. Zaprakis, and G. Francis, “Modeling stock returns in the framework of APT: a comparative study with regression models,” *Neural Networks in the Capital Markets*, pp.101-125, (1995).
- [11] K.A. Smith, and J.N.D. Gupta, “Neural networks in business: techniques and applications for the operations research,” *Computers and Operations Research*, Vol. 27, pp.1023-1044, (2000).
- [12] M. Steiner, and H-G. Wittkemper, “Neural Networks as an alternative stock market model,” *Neural Networks in the Capital Markets*, pp.135-147, (1995).
- [13] F.E.H. Tay, and L. Cao, *Application of support vector machines in financial time-series forecasting*, Omega, Vol. 29, pp.309-317, (2001).
- [14] M. Thomason, “The practitioner method and tools: a basic neural network-based trading system project revisited (parts 1 and 2),” *Journal of Computational Intelligence in Finance*, Vol. 7, No. 3, pp. 36-45, (1999a).
- [15] M. Thomason, “The practitioner method and tools: a basic neural network-based trading system project revisited (parts 3 and 4),” *Journal of Computational Intelligence in Finance*, Vol. 7, No. 3, pp. 35-48, (1999b).

# Neural Network and Genetic Algorithm Approaches for Forecasting Bangladeshi Monsoon Rainfall

Shipra Banik, Farah Habib Chanchary, Khodadad Khan, Rifat Ara Rouf and Mohammed Anwer

School of Engineering and Computer Science, Independent University, Bangladesh, Dhaka, Bangladesh  
email: [banik@secs.iub.edu.bd](mailto:banik@secs.iub.edu.bd), [farah@secs.iub.edu.bd](mailto:farah@secs.iub.edu.bd), [khan@secs.iub.edu.bd](mailto:khan@secs.iub.edu.bd),  
[rifatar@secs.iub.edu.bd](mailto:rifatara@secs.iub.edu.bd), [manwer@secs.iub.edu.bd](mailto:manwer@secs.iub.edu.bd)

**Abstract-** True information about rainfall is crucial for human activities such as the use and the management of water resources, hydroelectric power projects, warning to impend droughts or floods, urban areas sewer systems and many others. This paper investigates the development of an efficient model to forecast monthly monsoon rainfall for a number of stations, namely Barishal, Chittagong, Dhaka, Khulna, Rajshahi and Sylhet. It is believed that rainfall forecasting is difficult and also a challenging task for anyone because rainfall data are multi-dimensional and nonlinear. Therefore, to model rainfall data, the AI models, namely artificial neural network (ANN), adaptive neuro fuzzy inference system (ANFIS) and genetic algorithm (GA) have been used. Results found by the AI models are also compared to the linear regression model to show advantages of selecting these models. Findings suggest that ANFIS and GA approaches could be used more accurately than the other two selected approaches to forecast the Bangladeshi monsoon rainfall.

**Index Terms-** Atmospheric dynamics, forecasting, artificial intelligence, neural network, fuzzy inference system, genetic algorithm

## I. INTRODUCTION

The important hydrological event rainfall is the quantity of water falling in drops from vapor condensed in the atmosphere. When water droplets in clouds become too heavy to stay in the air, they fall out towards the ground. This paper proposes an appropriate model to predict monsoon rainfall in different stations of Bangladesh (see section II, for details). Making reliable prediction about rainfall is very important in many areas of human activity, indicated as follows: Rainfall mainly controls our water supplies, which are the basis for crops production. It is especially important for rainfed (or dry) land agriculture. For example, crops plant use a huge amount of water and rains, to a major extent, provide this water through the soil, for the development and growth of plants. Since Bangladesh is an agro-based country, accurate prediction of rainfall will be useful for proper plannings of cultivation. That's why those who are involved in agriculture, they will be interested to know whether the next days (or months) will be rainy/non-rainy. Although water is vital to life, yet water can be extremely destructive. Thus, rainfall forecasting can warn of happening flood or drought so that one can save their lives and properties. Rainfall forecasting is also

important for engineering applications, mainly for the design of hydroelectric power projects, because this system requires prior information about average rainfall, maximum/minimum rainfall, maximum intensity, duration etc. for a year/each month. In urban areas, rainfall also has a strong influence on traffic control, operations of sewer systems and others. Thus, we believe that precise rainfall prediction is important for practitioners who are interested to make wise policies related to this event.

True quantitative rainfall forecasting is generally difficult and also a challenging task for anyone because of our complex atmospheric processes. Thus, rainfall is treated as one of the most complex and difficult events among other hydrological events. Forecasting techniques, namely the statistical methods (ARIMA, regression, exponential smoothing etc.) and the AI methods (neural networks (artificial neural network (ANN), adaptive network based fuzzy inference system (ANFIS) etc.), fuzzy inference system (FIS), genetic algorithm (GA) and others) are proposed in literature. Note that statistical techniques have been developed for many years but have proven to be incapable to handle the non-linear series [1-2]. Hence, the AI methods have been proposed and engaged successfully in series forecasting, where the only goal is to minimize predictive error for the considered model. Thus, these methods have been applied in solving numerous forecasting problems (e.g. [3-6] and many others). We believe rainfall data are multi-dimensional, non-linear and dynamic, therefore to search for an appropriate model, the powerful techniques ANN, ANFIS and GA have been chosen. Results obtained by these models are also compared to the linear multiple regression (MR) model to show selected AI models performances. Empirical works (e.g. [7-13] and many others) have been carried out to forecast rainfall in context of various countries. Although studies have been conducted to predict rainfall for Bangladesh (e.g. [11-13] and others), to our knowledge, no comparative work is available to predict rainfall under the AI models for various Bangladeshi stations. We have considered this issue in this paper, which is planned as follows: Rainfall model and data sets description are explained in the next section. Considered techniques to predict rainfall are reviewed in the section III. Computational settings of the

TABLE I.  
NUMERICAL SUMMARY STATISTICS FOR  $R_t$

Stations	Time Periods	Mean	SD	Skewness
Barishal	Jan-53:Dec-02	285.43	176.38	0.79
Chittagong	Jan-49:Dec-02	387.81	285.10	1.06
Dhaka	Jan-53:Dec-02	274.24	154.35	0.73
Khulna	Jan-53:Dec-02	234.09	150.81	0.91
Rajshahi	Jan-64:Dec-02	210.16	147.09	0.82
Sylhet	Jan-56:Dec-02	547.55	291.42	0.38

selected techniques are given in the section IV. The relative performances of the considered methods are discussed in the section V. The final section concludes the paper with some future research plans.

## II. MODEL AND DATA SET

### A. Rainfall Model

The following functional relation is considered to predict monsoon rainfall:

$$R_t = f(\text{Intercept}, T_t, RH_t, CL_t), t=1,2,\dots,n \quad (1)$$

where  $R_t$  = rainfall (mm),  $T_t$  = temperature ( $^{\circ}\text{C}$ ),  $RH_t$  = relative humidity (%) and  $CL_t$  = cloud coverage (octas). Our target is to estimate model (1) parameters using the techniques of ANN, ANFIS, GA and MR.

### B. Data Set

For different stations of Bangladesh (see Table I), monthly  $R_t$ ,  $T_t$ ,  $RH_t$  and  $CL_t$  for monsoon periods (April to October) are used, which are extracted from <http://www.barc.gov.bd/database>. To understand behaviors of the considered variables, summary statistics are reported in Tables I-II. It can be seen that the highest mean rainfall is observed for the Sylhet region (547.55 mm) and the least mean amount for the Rajshahi region (210.16 mm). Standard deviation measures, e.g., for the Sylhet region, show that all times rainfall was not observed 547.55 mm (expected range: 256.13 to 838.97). Skewed measures indicate that rainfall patterns do not follow the normal distribution. For example, skewness 0.38 mm for the Sylhet region indicates that most of months rainfall amounts are observed below 547.55 mm. From Table II, it is observed that mean of  $T_t$  and mean of  $RH_t$  are almost same for all regions. But records of  $CL_t$

TABLE II.  
NUMERICAL SUMMARY STATISTICS FOR  $T_t$ ,  $RH_t$  AND  $CL_t$

Sta.	Mean			SD			Skewness			CC		
	$T_t$	$RH_t$	$CL_t$	$T_t$	$RH_t$	$CL_t$	$T_t$	$RH_t$	$CL_t$	$T_t$	$RH_t$	$CL_t$
Bar.	29	85	5.2	0.7	4.5	1.3	0.1	-0.9	-0.5	-0.3	0.6	0.64
Chi.	28	83	5.2	0.6	3.7	1.3	-0.1	-0.3	-0.5	-0.2	0.7	0.70
Dha.	29	81	5.2	0.9	5.6	1.3	-0.2	-1.1	-0.6	-0.1	0.6	0.60
Khu.	29	83	4.6	0.9	5.4	1.4	-1.7	-1.1	-0.4	-0.2	0.6	0.66
Raj.	30	79	4.6	0.9	9.6	1.5	-0.4	-1.4	-0.4	-0.1	0.6	0.71
Syl.	27	84	5.9	1.1	4.8	1.3	-0.6	-1.1	-0.8	0.1	0.6	0.72

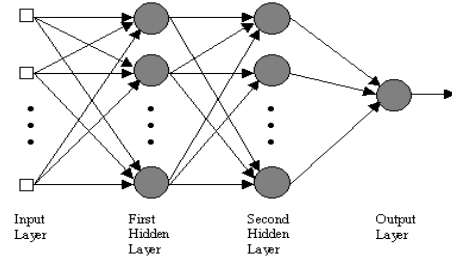


Fig. 1. An ANN architecture

show that like rainfall, Sylhet region experiences the highest mean amount (5.9 octas), can be a cause of occurring highest rain. Skewed measures (like  $R_t$ ) show that  $T_t$ ,  $RH_t$  and  $CL_t$  variables do not follow normal distribution. Correlation coefficient (CC) results indicate that there is a negative relationship exists between  $T_t$  and  $R_t$  and a positive relationship between  $RH_t$  and  $CL_t$ . We believe these findings will be useful for individuals who are interested about the events  $R_t$ ,  $T_t$ ,  $RH_t$  and  $CL_t$ .

## III. TECHNIQUES USED TO PREDICT RAINFALL

### A. ANN

Culloch and Pitts [14] introduced this system. A basic ANN structure (see Fig.1, details [5]) contains 1 input layer, number of hidden layers (with sufficient no. of neurons) and 1 output layer. Each neuron receives an input  $P_{Q \times 1}$  ( $T_t$ ,  $RH_t$  and  $CL_t$ ) and produces the net output as  $\hat{R}_t = W_{N \times Q} + b$ , where  $W_{N \times Q}$  are the weights,  $b_{N \times 1}$  is the bias,  $N$  are the number of neurons in the hidden layers and  $Q$  are the number of elements in  $P$ . The training algorithm is the standard BP, which uses the GD technique to minimize error. During training, each desired  $\hat{R}_t$  is compared with the actual  $R_t$  and calculates error at the output layer. When the training process is completed, then the network with adjusted  $W_{N \times Q}$  and  $b_{N \times 1}$  (estimated parameters) is used to test a set of data, which is different from the training set of data.

### B. ANFIS

This system is developed by Jang [3]. It is a combination of ANN and FIS, in such a way that the ANN learning algorithm is used to determine parameters of FIS. An ANFIS system (see Fig.2, details [3] or [5]) has 5 fixed layers: 1 input layer, 3 hidden layers and 1 output layer. The learning algorithm is a hybrid system, which uses two types of error estimation techniques, the least square method (LSE) and the GD method. The hidden layer is computed by the GD method of the feedback structure and the final output ( $\hat{R}_t$ ) is estimated by LSE. Initially, one solution is calculated and then it is optimized by readjustment of parameters by repeated application of the hybrid system.

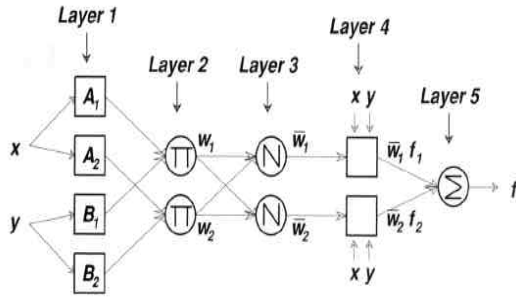


Fig.2. An ANFIS architecture

### C. GA

This technique is proposed by Holland [15]. Later Koza used this approach (called his method genetic programming (GP)) to evolve programs to perform certain tasks (details, see [16]). By definition, GA is a technique based on the ‘Darwin’s Principle of Natural Selection’ and is used to solve optimization problems. The basic idea is to select the best, discard the rest and differs from other systems in the following way. GA does not require any assumptions such as linearity, stationarity, homogeneity or chaotic. Thus, to handle the complex multi-dimensional behaviors of a system, this approach has been used efficiently in literature (e.g. [10, 15-17] and others). Fig.3 flowchart shows the basic steps involved in a GA process (see Table III, for the standard GA options). Each of steps brief explanation of Fig.3 (or Table III) are as follows:

**Step-1:** Create an initial population consisting of random chromosomes. To understand the GA process, e.g., for the model (1), consider the following random population of 6 chromosomes with 4 parameters each: (0.13, 0.01, 0.84, 0.68), (0.20, 0.74, 0.52, 0.37), (0.19, 0.44, 0.20, 0.83), (0.60, 0.93, 0.67, 0.50), (0.27, 0.46, 0.83, 0.70), (0.19, 0.44, 0.01, 0.42). Population size is chosen usually from 100-500. Larger population may produce more robust solution.

**Step-2:** Fitness scaling is used to provide a measure of how selected chromosomes perform in the problem

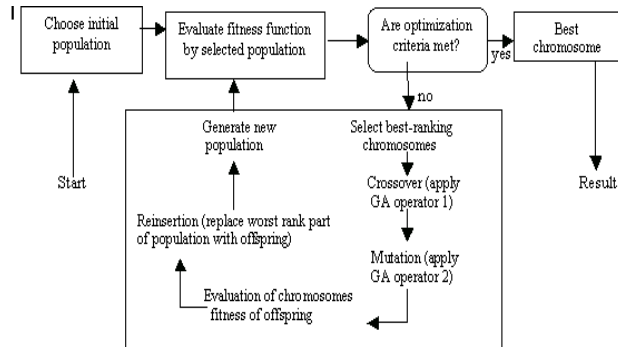


Fig. 3. A GA structure

TABLE III.  
COMMON GA OPTIONS

Step	Algorithm Option
1. Creation	Uniform, normal
2. Fitness scaling	Rank, proportional, top (truncation), linear scaling, shift
3. Selection	Roulette, stochastic uniform selection, tournament, uniform
4. Crossover	Binary-valued (single-point, two-point, n-point, uniform), real-valued (intermediate, line)
5. Mutation	Gaussian, uniform
6. Reinsertion	Pure, uniform, elitist, fitness-based

domain. Model (1) fitness is evaluated through a criterion like RMSE (CC, MAE can be also used). For (1), we get RMSE: 26.19, 32.09, 53.75, 20.18, 18.67 and 66.64. Using the linear-ranking process (e.g.) (details, see [18]), Fit(RMSE): 1.2, 0.8, 0.4, 1.6, 2.0, 0.

**Step-3:** Based on the step-2 results, choose parents for the next generation. To understand it, consider the following distribution:

RMSE:	26.19	32.09	53.75	20.18	18.67	66.64
Fit(RMSE):	1.2	0.8	0.4	1.6	2.0	0
pdf of RMSE:	0.2	0.13	0.06	0.26	0.33	0
cdf of RMSE:	0.2	0.33	0.39	0.65	0.98	1.0

It is observed that chromosome 5 is the fittest chromosome, because it occupies the largest interval, whereas chromosome 4 is the second least fit chromosome as the smallest interval. Chromosome 6 is the least fit interval that has a fitness value of 0 and get no chance for reproduction. Using (e.g.) the roulette wheel method (purpose is to eliminate the worst chromosomes and to regenerate better substitutes), selected 4 parents are: (0.20, 0.74, 0.52, 0.37), (0.13, 0.01, 0.84, 0.68), (0.27, 0.46, 0.83, 0.70), (0.13, 0.01, 0.84, 0.68). Next step is to produce offspring from selected parents by combining entries of a pair of parents (known as crossover) and also by making random changes to a single parent (known as mutation).

**Step-4 (GA operator-1):** Basic operator for producing new (improved) chromosomes is known as crossover (a version of artificial mating). It produces offspring that have some parts of both parents genetic material. Offspring are produced using the intermediate crossover method, because this is a method proposed to recombine of parents with real valued chromosomes (details, [18]). Thus, crossover offspring are: (0.16, 0.16, 0.85, 0.57), (0.13, 0.22, 0.76, 0.43), (0.13, 0.15, 0.83, 0.69), (0.26, 0.45, 0.84, 0.68).

**Step-5** (GA operator-2): Offspring are mutated after being produced crossover offspring and this GA operator increases the chance that the algorithm will generate better fittest RMSE than the step-4. GA creates 3 types of offspring: These are elite (number of best RMSE values in the current generation that are guaranteed to survive to the next generation), crossover and mutation. To understand it, consider an example: Suppose population size is 20 and elite count is 2. If crossover fraction is 0.8, then the distribution of offspring: 2 elites, 14 ( $18 \times 0.8$ ) are crossover offspring and the remaining 4 are mutation offspring. Just to know, a crossover fraction of 1 means, all offspring other than elite are crossover offspring, while a crossover fraction of 0 means that all offspring are mutation offspring. How offspring are produced under the mutation process, see [18]. The mutation offspring are found (0.16, 0.17, 0.85, 0.56), (0.13, 0.22, 0.76, 0.43), (0.13, 0.14, 0.83, 0.69), (0.26, 0.45, 0.84, 0.68) respectively.

**Step-6:** Once offspring have been produced using steps 4-5, offspring fitness (i.e. RMSE values) must be determined (procedure similar to step-2). We get improved RMSE: 23.37, 28.13, 24.11 and 18.62. If less offspring are produced than the original population size, then to maintain size, offspring have to be reinserted into the old population. By this step, it is determined which chromosomes will be replaced by offspring. Using (e.g.) the fitness-based reinsertion method, the following RMSE are found: 20.18, 18.67, 18.62, 23.37, 24.11 and 28.13. If termination criteria are not defined, GA returns to the step-3 and continues up to the step 6. It is satisfied when either maximum number of generations is achieved or when all chromosomes in the population are identical (i.e. converge). The creator sets this number before running GA, which ensures that GA does not continue indefinitely.

#### D. MR

This method is developed based on the assumption of linearity and its main aim is to minimize error. Model (1) parameters are estimated by  $\hat{\alpha}_i = (X'X)^{-1}(X'R_t)$ ,  $i=1,2,3$ , where  $X$  includes  $T_t$ ,  $RH_t$  and  $CL_t$  and  $\hat{\alpha}_0 = \bar{R}_t - \hat{\alpha}_1 \bar{T}_t - \hat{\alpha}_2 \bar{RH}_t - \hat{\alpha}_3 \bar{CL}_t$ .

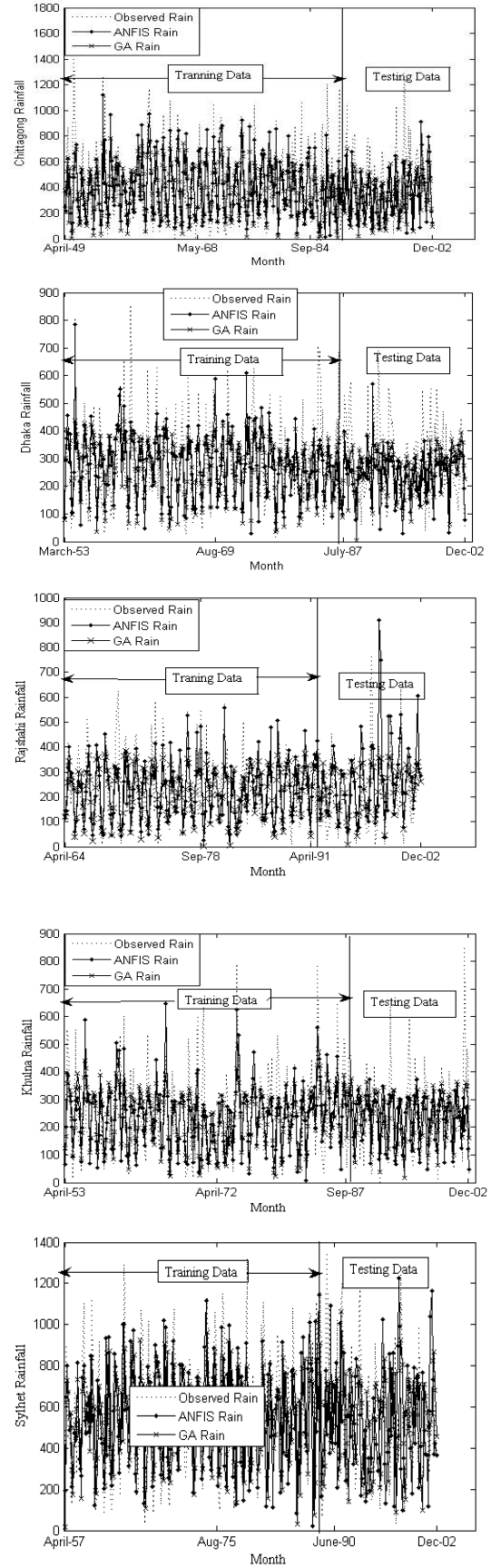
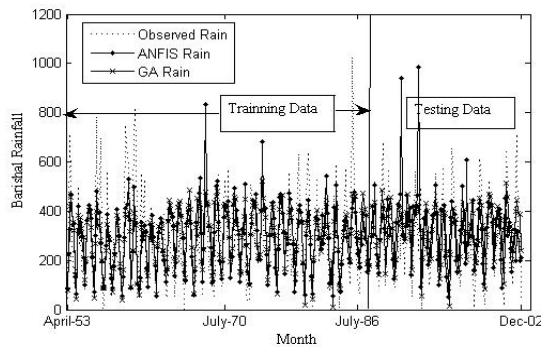


Fig. 4. Training and testing results of the considered stations

TABLE IV.  
COMPUTATIONAL SETTINGS FOR ANN, ANFIS AND GA NETWORKS

ANN	ANFIS	GA
1) <i>Topology</i> -3:10:7:1 i.e. 3 input nodes and two hidden layers consisting of 10 and 7 hidden neurons respectively.  2) <i>Learning rate</i> -0.15.  3) <i>Momentum parameter</i> -0.90.  4) <i>Activation function</i> for the hidden neurons- <i>Purelin</i> .	1) <i>Input- <math>T_i</math>, <math>RH_i</math> and <math>CL_i</math></i> . 2) 2 mfs (type Gaussian) for each of input variables. Thus, 8 if-then fuzzy rules were learned. Total parameters (44) = premise parameters (12) + consequent parameters (32), where premise parameters is calculated by # of inputs * # of mfs * # of parameters of Gaussian distribution and consequent parameters is calculated by # of mfs * # of parameters of Gaussian distribution * # of fuzzy rules.	1) <i>Population size</i> -1500, generated at random from the uniform probability distribution. 2) <i>Fitness function</i> -RMSE of model (1). 3) <i>Fitness scaling</i> -Linear ranking process. 4) <i>Selection</i> -Roulette wheel. 5) <i>Crossover</i> -fraction: 0.90; Intermediate method. 6) <i>Mutation</i> -fraction: 0.05; Uniform method. 7) <i>Reinsertion</i> -The fitness-based reinsertion method. 8) <i>Termination Criteria</i> - Maximum # of generations, assumed 300.

#### IV. DESIGN OF CONSIDERED TECHNIQUES

The first 70% data for each selected stations are used as the training period and the rest 30% as the testing period (see Fig.4). To reduce error, an error and trial approach is used to find the computational settings for each selected AI techniques, which are reported in the Table IV.

#### V. RESULTS AND DISCUSSION

Systems performances have been measured by statistical measures, RMSE and coefficient of determination ( $R^2$ ). Note that RMSE is a measure of the accuracy of prediction through representing the degree of scatter and  $R^2$  is a measure of the accuracy of prediction of the trained network models. Higher  $R^2$  value indicates better prediction. All computational works were carried

TABLE V.  
PERFORMANCES OF DIFFERENT CONSIDERED MODELS

Sta.	Mea.	ANN		ANFIS		GA		MR	
		Train	Test	Train	Test	Train	Test	Train	Test
Bar	RMSE	135	111	123	105	135	103	138	112
	$R^2$	0.84	0.85	0.87	0.88	0.84	0.88	0.82	0.83
Chi.	RMSE	193	193	176	169	193	186	196	193
	$R^2$	0.82	0.82	0.88	0.89	0.84	0.83	0.80	0.81
Dha	RMSE	120	111	107	105	120	103	121	111
	$R^2$	0.85	0.86	0.89	0.86	0.86	0.87	0.83	0.84
Khu	RMSE	112	105	97	94	112	98	115	107
	$R^2$	0.84	0.85	0.88	0.85	0.84	0.85	0.82	0.83
Raj.	RMSE	98	107	82	95	98	87	99	108
	$R^2$	0.85	0.82	0.90	0.86	0.86	0.82	0.81	0.83
Syl.	RMSE	185	170	161	128	185	151	187	172
	$R^2$	0.91	0.92	0.93	0.91	0.91	0.93	0.89	0.90

out using the programming code of MATLAB (version 7.0). Table V summarizes training and testing results achieved for selected approaches, which clearly reveal efficiency of the proposed techniques. In terms of both measures RMSE and  $R^2$ , for all stations, ANFIS performed better (i.e. good match between actual and predicted) than other models, followed by GA. It is found that ANN and MR models performances are almost same for all considered cases. See the Fig.4, to understand our findings graphically, where only the ANFIS model and the GA model performances (to avoid clumsiness in figure) are shown with observed  $R_i$ .

#### VI. CONCLUSION

Rainfall forecasting is important for many areas of human activity such as agriculture, water resources, hydroelectric power projects, happening of droughts or floods and others. This paper modelled the complex multi-dimensional behaviors of monthly monsoon rainfall for a number of Bangladeshi stations, namely Barishal, Chittagong, Dhaka, Khulna, Rajshahi and Sylhet using several soft computing techniques. We considered ANN, ANFIS and GA processes and results obtained by these models are compared to the linear multiple regression model to show advantages of proposing the AI models. Findings suggest that the ANFIS model and the GA model can be used to forecast monthly monsoon rainfall more accurately than the ANN model and the statistical model. We believe the findings of this paper will be useful to the individuals for formulating wise policies related rainfall in the context of our country, Bangladesh. Our next plan is to model the daily monsoon rainfall data, which have significance in the field of agriculture, transportation, sports, tourism activities and others. This is left for future research.

#### REFERENCES

- [1] J. Kamruzzaman and R.A. Sarker, "Comparing ANN based models with ARIMA for prediction of Forex rates", *ASOR Bulletin*, pp.1-10, 2003.
- [2] J.M. Kihoro, R.O. Otieno and C. Wafula, "Seasonal time series forecasting: A comparative study of ARIMA and ANN models", *African Journal of Science and Technology*, pp.41-49, 2000.
- [3] J.S.R. Jang, "ANFIS: Adaptive-network-based fuzzy inference systems", *IEEE Transactions on Systems, Man, and Cybernetics*, pp.665-685, 1993.
- [4] S.B.Charhate and M.C. Deo, "Storm surge prediction using NN and GP", *Paper Presented in the 6<sup>th</sup> Conference on AI application to Environmental Science AMS Annual Meeting*, 20-24 January, New Orleans Louisiana, USA. 2007.
- [5] S.Banik, F.H.Chanchary, R.A.Rouf and K. Khan, "Modeling chaotic behavior of Dhaka stock market index values using the neuro-fuzzy model", *Proceedings of 10<sup>th</sup> ICCIT*, pp.80-85, 2007.
- [6] A. Kumar and T. Debroy, "Tailoring fillet weld geometry using a genetic algorithm and a neural network trained with convective heat flow calculations", *Welding Journal*, pp.26-33, 2007.
- [7] P.A. Whigham and P.F. Crapper, "Modeling rainfall-runoff



- using genetic programming”, *Mathematical and Computer Modeling*, Canberra, Australia, pp.707-721, 2001.
- [8] H.D.Navone and H.A.Ceccatto, “Predicting Indian monsoon rainfall: A neural network approach”, *Climate Dynamics*, pp.305-312, 1994.
- [9] Z. Sen and A. Oztopal, “Genetic algorithms for the classification and prediction of precipitation occurrence,” *Hydrological Sciences*, pp.255-267, 2001.
- [10] Y.H. Lee, S.K. Park and D.E. Chang, “Parameter estimation using the genetic algorithm and its impact on quantitative precipitation forecast”, *Annales Geophysicae*, pp.3185-3189, 2006.
- [11] R.U.Mondol and D.S. Shahid, “A neural network approach for the prediction of monthly rainfall”, *Jahangirnagar University of Science*, pp.87-95, 2004.
- [12] M.N. Islam, S. Islam, T. Hayashi, T. Terao and U. Hiroshi, “Application of a method to estimate rainfall in Bangladesh using GMS-5 data”, *Journal of Natural Disaster Science*, pp. 83-89, 2002.
- [13] C.M. Wahid and M.N. Islam, “Patterns of rainfall in the northern part of Bangladesh”, *Journal of Science Research*, pp.115-120, 1999.
- [14] W.S. Culloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity”, *B. Math Bioohys*, pp.115-133, 1943.
- [15] J.N.Holland, *Adaptation in Natural and Artificial Systems*, Ann Arbor, The University of Michigan Press, 1975.
- [16] J.R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, Cambridge: MIT Press, 1992.
- [17] Z. Wei, W.U. Zhi-ming and Y. Gen-Ke, “Genetic programming based chaotic time series modeling”, *Journal of Zhejjang University Science*, pp.1432-1439, 2004.
- [18] H. Pohlheim, *Documentation for Genetic and Evolutionary Algorithm Toolbox for Use with MATLAB*.



# On the Multi-Channel Noise Cancellation Using Beamforming Algorithm

Md. Zahangir Alam<sup>\*</sup>, Md. Zahangir Alam<sup>\*\*</sup>, M Abdus Sobhan<sup>†</sup>

<sup>\*</sup>Dept. of Computer Science & Engineering, World University of Bangladesh, Dhaka.

<sup>\*\*</sup>Dept. of Electronics and Telecommunication Engineering, Prime University, Bangladesh.

<sup>†</sup>SECS, Independent University, Bangladesh (IUB), Dhaka, Bangladesh

Email: {zahan\_cse, zahan\_cse, sobhan30}@yahoo.com.

**Abstract** — This paper presents performance analysis of beamforming algorithm for canceling multiple channel noise depending on variation of hidden layer of the multilayer feedforward network and the number of epoch (also known as number of iteration). We consider the learning rate variation of the multi layer perception (MLP) to determine the adaptive learning rate of the network. An MLP has been considered to perform beam-forming, which is an array of sensors connected to the MLP inputs. We use the Backpropagation algorithm as the learning rule for MLP and improving the signal quality. This involves a desired signal whilst removing any noise or interference signals which may come from different sources [1].

**Index Terms** — DSPs,  $\mu$  LMS, MLP, AM, BSS, Nyquist Sampling Rate.

## I. INTRODUCTION

Various signal processing techniques have been proposed over the years for noise reduction in the environment. The explosive growth of digital signal processing algorithms and technologies has given an impetus to the application of these techniques to the real world. Digital signal processors (DSPs) have shrunk tremendously in size while their processing capabilities have grown exponentially. This has enabled the use of DSPs in a variety of portable hearing enhancement devices such as hearing aids, headsets, hearing protectors, etc. There is an approach for electrical noise reduction. These electrical noise reduction techniques are applied in hearing aids, cochlear implants, etc. where the signal and ambient noise are recorded using microphone. Noise reduction techniques such as spectral subtraction and the  $\mu$  LMS algorithm are applied and the listener hears only the signal [2]. One of the important assumptions of this technique is that the listener is acoustically isolated from the environment. In this research, beam-forming technique is used for noise cancellation and performance is analyzed for various hidden layers in feedforward neural network and number of iteration

(epoch). Backpropagation algorithm is used to improve the signal quality and to train the data [3].

## II. ARCHITECTURE OF THE PROPOSED SYSTEM

The arrangement for the neural network based on beamforming is depicted in Fig. 1. From the Fig. 1 it is clear that there will be a neuron for each array element (AE). A beamforming is an array of sensors connected to the MLP inputs. The electromagnetic signal impinges each antenna at different time. This time difference is effectively a delay on the signal and the physical spacing of the antenna elements is used to calculate the required delays. In this case the linear array will consist of 6 elements which process the desired signal has been contaminated with noise.

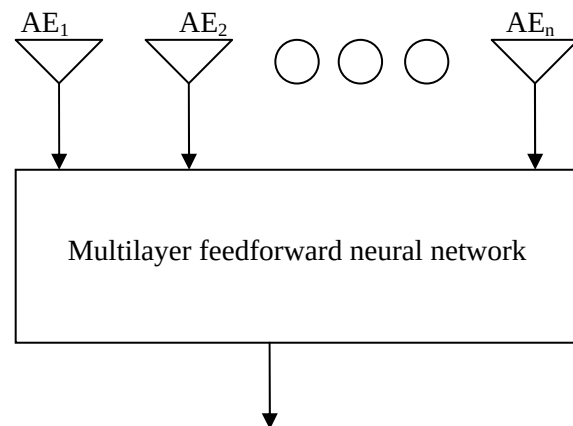


Fig. 1 Neural network for beamforming

### A. Multilayer Feedforward Neural Network

Fig. 2 represents the fully connected feedforward three layers MLP. Fig. 2 is known as a 7-4-3 architecture because it has 7 input nodes, 4 hidden nodes, and 3 output nodes. In general, the input layer has not performed any computation; it just acts as a buffer. The neural network architecture will be

optimized by trial and error. In this system, 6 neurons have considered in the input layer, 3 neurons in the hidden layer and 1 neuron in the output layer. A bipolar sigmoid transfer function is used in each layer due to its ability to solve nonlinear problems. The short-term representation of a signal is obtained by windowing (Hamming window). Sampling rate of this signal is higher than the Nyquist sampling rate.

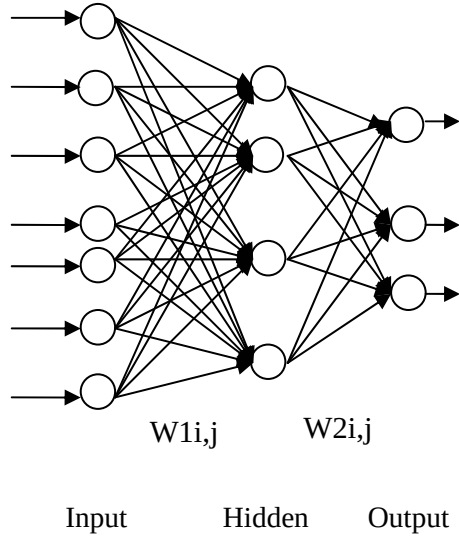


Fig. 2. Three layer MLP

### B. Beamforming Technique

Beamforming technique is currently used in numerous radio communication systems, radar and sonar to eliminate interference and jamming. A Beam-former is another name in spatial filtering where an array of sensors together with appropriate signal processing can either direct or block the radiation or the reception of signals in specified directions. Fig. 3 represents the block diagram of a simple beamformer. At first, it samples the propagating wave field in space, is typically used for processing narrowband signals [4].

The output of the sensors are weighted and then summed to produce the output of the overall beam-former. The beam-former output at time  $k$ ,  $y(k)$  is given by a linear combination of the data at the  $M$  sensors at time  $k$ .

$$y(k) = \sum_{i=1}^m (W_i u_i(k))^* \quad \text{where } * \text{ represents complex}$$

conjugate. It is conventional to multiply the data by conjugate of the weights to simplify notation.

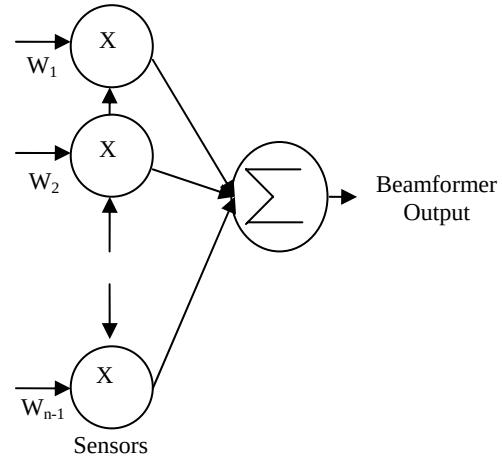


Fig. 3. A general beamforming system

## II. SYSTEM IMPLEMENTATION

Different steps to implement the system are given below.

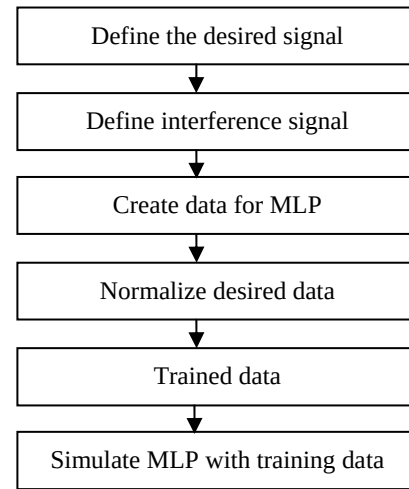


Fig. 4. Implementation steps

According to the Fig. 4 to create a desired signal AM wave is selected. This signal arrives at each antenna at given time intervals. A random noise is added to this signal when it arrived at each antenna. Desired signal created is represented by the following equation.

$$d\_signal = (A \times (1 + m \times \cos(\omega_m t)) \times (\cos(\omega_c t))) \quad (1)$$

Then the noisy signal is represented by the following equation.

$$d\_signal = (A \times (1 + m \times \cos \omega_m t \times \cos \omega_c t)) + (rand() \% 1000) / 1000 \quad (2)$$

where  $A$  = carrier amplitude of the desired signal  
 $m$  = modulation index of the desired signal.  
 $\omega_m$  = angular modulation frequency of the desired signal.  
 $\omega_c$  = angular carrier frequency of the desired signal

The difference between these two signals is the angle of arrival (i.e. the delay of the signal between antennas). The equation which generate interference signal is shown below.

$$i\_signal = i_A \times (1 + i_m \times \cos j \omega_m t \times \cos j \omega_c t) \quad (3)$$

where,  $i_A$  = carrier amplitude of the interference signal.

$i_m$  = modulation index of the interference signal.

When the interference signal arrived at the antenna, end noise is added to the interference signal. Thus interference equation is same as (2). The signal to impinge on the antenna is composed of both the desired signal and the interference signal. The input signal to arrive at each antenna will be the superposition of the two signals. This was created by summing the desired signal and the interference signal. The input signal of each antenna is given by the following equation.

$$Input = d\_signal + i\_signal \quad (4)$$

The amplitude of the desired signal is not known but due to the 'sigmoid' transfer function the output amplitude level lies between -1 to 1. The normalized data transfer equation

$$Target(t) = 2 \times (d\_signal(t) - minn) / (maxn - minn) - 1 \quad (5)$$

where  $d\_signal$  is the current value,  $minn$  and  $maxn$  is the minimum and maximum value of the desired signal respectively. Finally Backpropagation algorithm is used to train the desired data [5]. We have considered learning rate 0.01 and the mean square error (MSE) 0.001.

#### IV. RESULT AND DISCUSSION

The trained network is simulated with only one input signal to produce an output. The variables that were changed for the AM signal were the carrier frequency. For each of the test wave the modulation

frequency was set between 20 Hz to 20KHz. The carrier frequency was set between 535 KHz and 1605 KHz (i.e. typical range of radio station frequency). The modulation depth had a range of 0 to 1, and the carrier amplitude has no specific range.

##### A. Sample Input

In Table I, five signals are considered as input signal.

##### B. Correlation Value Calculation and its Graphical Representation with Respect to Feedforward Network

Table II contains the correlation values of the five input signals. In feedforward network, we have considered 6 nodes in the input layer and 1 node in the output layer. But hidden layer nodes vary from 1 to 7. With respect to this variation, we have calculated correlation values for five input signals.

##### C. Correlation Value Calculation and Graphical Representation with Respect to Epoch

On the other hand we conducted experiment with respect to the epoch/number of iteration. One complete pass through the training set of data is known as one epoch (also known as iteration). Table III presents the correlation values of five input signals and Fig. 6 shows the graphical representation of correlation values with respect to the number of epoch and input signals.

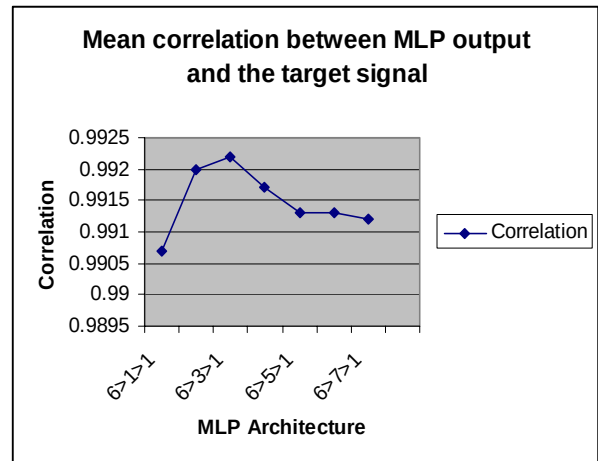


Fig. 5. Correlation values for variation of hidden layer nodes

##### D. Correlation value calculation and graphical representation with respect to learning rate.

We have tested the performance with respect to the different learning rates. One complete pass through the training set of data is known as an epoch (also known as number of iteration). Table-IV presents the correlation values of five input signals and Fig.7 presents the graphical representation of correlation values with respect to the different learning rates. From the Table II, it is clear that the 6-3-1 combination produced the best correlation value of 0.9922. Its graphical representation with respect to the input signals and hidden layers is shown below.

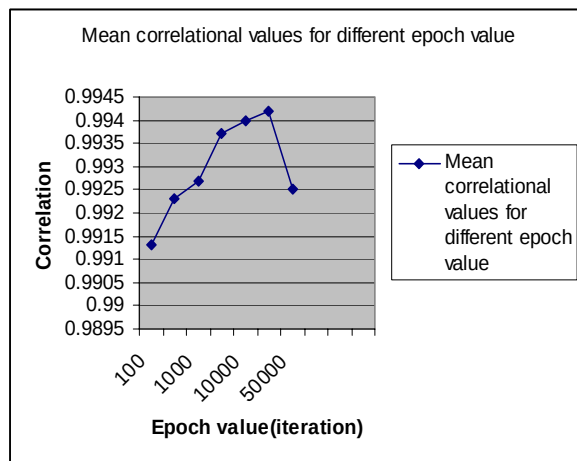


Fig. 6. Correlation values for different epoch

Fig. 6 shows the correlation value for different epoch values and also shows that if the number of epoch increases then the performance of the MLP also increases respectively. For less than 100 epochs, correlation value between the desired signal and output signal is not satisfactory.

Fig-7 shows the correlation value maximum for learning rate 3. After the learning rate of 3 has been reached, the accuracy of the MLP rapidly decreases as the learning rate increases. In general, the preference of the network gradually decreases as the learning rate is reduced from the optimum value of 3. At 300000 epochs, a correlation value 0.9942 is obtained and that is the average highest result achieved in this experiment. If we increase the number of iteration, but carrier amplitude of the interference signal is twice than the carrier amplitude of the desired signal, then it returns low correlation value. In Test-4, 50000 epoch returns low correlation value of 0.9857.

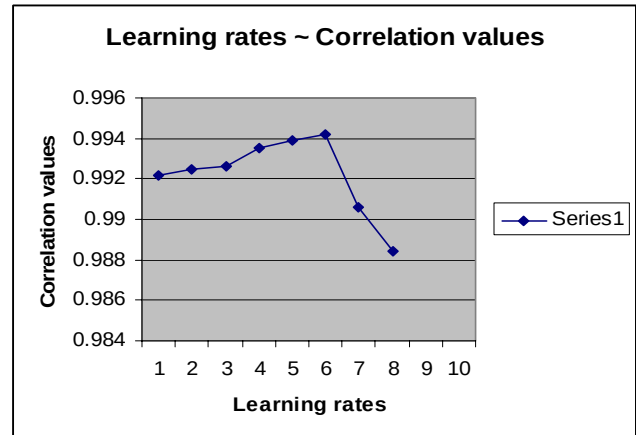


Fig.7. A graph to show the correlation values for different learning rate values.

## V. CONCLUSION

According to our observation, it is possible to say that an MLP with 6 input neurons, 3 middle neurons and 1 output neuron should be used with a learning rate of 3, and an epoch value of 30,000. With this arrangement a normalized correlation value of approximately 0.9945 can be obtained between the transmitted desired signal and the MLP output for Test-2 signal. This can be a much higher correlation than the correlation between the signal which arrives at the array and the output signal. For the learning rate 3, the network represents the highest correlation value. So, we conclude that learning rate 3 is the adaptive learning rate. Therefore this report has shown that it is possible to use an MLP to produce exceptional results for beam-forming algorithm. The optimized network produces exceptional results not only for the data that was trained with but also for the whole range of signals used in AM radio broadcasting. The MLP has shown that it has the ability to reproduce the desired signal by attenuating any signals arriving at different angles. Unfortunately, the research was subject to time limitations and further work would be able to enhance the work. By using alternative method like blind source separation (BSS), it would be possible to compare the various methods and find a better solution.

TABLE I  
SAMPLE INPUTS

Test No	Carrier Amp. of signal (A)	Modulation frequency ( $f_m$ ) of the signal	Carrier Amp. of Intf. signal ( $i_A$ )	Modulation Index of Interference signal( $i_m$ )	Carrier frequency of Intf. signal( $f_c$ )	Modulation Frequency of Intf. signal( $f_m$ )
1	1	$3 \times 10^3$ Hz	1	0.5	$1605 \times 10^3$ Hz	$3 \times 10^3$ Hz
2	1	20	0.8	0.9	$535 \times 10^3$ Hz	100 Hz
3	1	$5 \times 10^3$ Hz	0.5	0.8	$1000 \times 10^3$ Hz	$2 \times 10^3$ Hz
4	1	1000 Hz	2	0.9	$700 \times 10^3$ Hz	1000 Hz
5	1	100 Hz	0.2	1	$1500 \times 10^3$ Hz	5000 Hz

TABLE II.  
THE CORRELATION VALUE FOR THE INPUT SIGNALS

Architecture	Test-1	Test-2	Test-3	Test-4	Test-5	Mean Value
6-1-1	0.9711	0.9707	0.9709	0.9709	0.9707	0.9708
6-2-1	0.9921	0.9921	0.9912	0.9925	0.9921	0.9920
6-3-1	0.9921	0.9923	0.9915	0.9928	0.9923	0.9922
6-4-1	0.9921	0.9917	0.9908	0.9923	0.9917	0.9917
6-5-1	0.9913	0.9916	0.9906	0.9916	0.9916	0.9913
6-6-1	0.9915	0.9915	0.9905	0.9915	0.9915	0.9912
6-7-1	0.9914	0.9915	0.9905	0.9915	0.9915	0.9912

TABLE III.  
THE CORRELATION VALUES FOR DIFFERENT EPOCH

Number of Epoch	Test-1	Test-2	Test-3	Test-4	Test-5	Mean Value
100	0.9913	0.9918	0.9908	0.9912	0.9918	0.9913
500	0.9924	0.9925	0.9914	0.9930	0.9925	0.9923
1000	0.9928	0.9929	0.9918	0.9935	0.9929	0.9927
5000	0.9937	0.9939	0.9928	0.9942	0.9939	0.9937
10000	0.9940	0.9942	0.9931	0.9945	0.9942	0.9940
30000	0.9943	0.9945	0.9935	0.9945	0.9945	0.9942
50000	0.9944	0.9946	0.9936	.09857	0.9946	0.9925

TABLE- IV.  
A TABLE TO SHOW THE CORRELATION VALUES FOR DIFFERENT LEARNING RATES.

Learning Rate	Test-1	Test-2	Test-3	Test-4	Test-5	Mean Value
0.01	0.9921	0.9923	0.9915	0.9928	0.9923	0.9922
0.05	0.9927	0.9927	0.9917	0.9931	0.9927	0.9925
0.1	0.9927	0.9928	0.9918	0.9932	0.9928	0.9926
0.5	0.9935	0.9938	0.9925	0.9941	0.9938	0.9935
1	0.9940	0.9941	0.9930	0.9945	0.9941	0.9939
3	0.9943	0.9945	0.9935	0.9945	0.9945	0.9942
5	0.9944	0.9896	0.9848	0.9947	0.9896	0.9906
8	0.9885	0.9896	0.9804	0.9940	0.9898	0.9884

#### REFERENCES

- [1] Zahangir Alom, Iqbal Aziz khan, Md Zahangir alam, Mst Shamima Nasrin "Performance Analysis of Beamforming Algorithm for Noise Cancellation.", International Conference on Computer and Information Technology 2007 (ICCIT 2007), 27-29 December, 2007.
- [2] L. C. Godara, "Improved LMS algorithm for adaptive beamforming", *IEEE Trans. Antenna Propagate*, vol. 38, no. 10, pp. 1631-1635, October 1990.
- [3] J. A. Freeman and D. M. Skapura, "*Neural Networks*", Addison-Wesley Longman Inc, California, pp. 89-124, 1991.
- [4] B.D. Van Veen and K.M. Buckley, "Beamforming: A versatile approach to spatial filtering", *IEEE ASSP Mag.*, vol. 5, pp. 4-24.
- [5] C A Jacobson, C R Johnson Jr., D C McCormick and W A Sethares, "Stability of Active Noise Control Algorithms", *IEEE Signal Processing Letters*, Vol. 8, Number 3, March 2001.
- [6] Caudill and Bulter, 1992, "Understanding neural networks", *MIT press*, Vol. 1 & 2.
- [7] Winder, R.O. (1964) "Threshold Logic in Artificial Intelligence", *IEEE Publication*, S-142, pp. 107-128.

# Recurrent Neural Network Classifier for Three Layer Conceptual Network and Performance Evaluation

Md. Khalilur Rhaman and Tsutomu Endo

Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan.  
rhaman@dumbo.ai.kyutech.ac.jp, endo@pluto.ai.kyutech.ac.jp

**Abstract**—Contextual analysis in dialog is a hard problem. In this paper three layers memory structure is adopted to address the challenge which we refer to as Three Layer Conceptual Network (TLCN). This highly efficient network simulates the human brain by episodic memory, discourse memory and ground memory. An extended case structure framework is used to represent the knowledge. The knowledge database is constructed by the collection of target system information and utterances. This knowledge is updated after every dialog conversation. A Recurrent Neural Network classifier is also introduced for classifying the knowledge for the target system. This system prototype is based on doctor-patients dialogs. 78% disease classification accuracy is observed by this system prototype. Disease identification accuracy is depending on number of disease and number of utterances. This performance evaluation is also discussed in details.

**Index Terms**—Three Layer Conceptual Network, knowledge representation, Neural Network.

## I. INTRODUCTION

This paper presents a dialog processing architecture. It utilizes three layer conceptual network (TLCN) instead of traditional natural language processing architecture. This architecture is developed by RHAMAN [1] from the idea of memory model of NOMURA[2]. NOMURA's model was a framework for memory management but RHAMAN developed architecture for a real time system and introduced a discourse algorithm which was related to [3], [4], [5]. It is a very efficient and unique model which can simulate the human brain. We used the traditional modules for input and response generation but developed an extended case frame model of BRUCE [6] and SHIMAZU & NOMURA [7] for handling the knowledge database. Case frame network is very useful for unstructured and ambiguous languages like Japanese. However, most systems deal with rather simple sentences, in which the analysis is a relatively easy task. Our extension of case frame architecture reduced the complexity using formalism. In this paper we propose a Neural Network Classifier to classify the diseases of our system prototype. ELMAN [8] and LAWRENCE [9] are considered to apply in this application. This is an initiative to introduce Neural Network in Three Layer Conceptual Network.

## II. THREE LAYER CONCEPTUAL NETWORK

TLCN is the three layer memory model proposed by RHAMAN[1]. It is a concept to use the knowledge like a human brain. The layers are: 1) Ground Memory, 2) Discourse Memory and 3) Episodic memory. In the network all the knowledge and their relationships are represented in Ground Memory. Discourse memory provides the most basic, useful and recently used knowledge. Episodic Memory is a User interface. It contains the knowledge to understand the input dialogs and generate response dialogs. However, each memory is not simply an accumulation of the contents of other memories. It is a structured memory coherently organized by merging old and new knowledge. The process relates to what is called knowledge learning. Fig. 1 describes the processing of the TLCN.

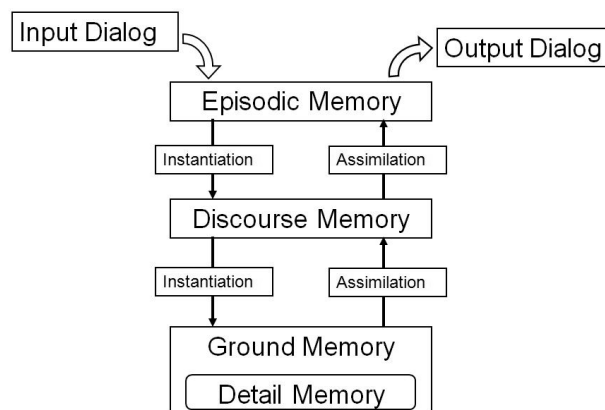


Fig. 1. Processing of TLCN

## III. MEMORY MODULES

### A. Ground Memory Module

The ground memory of GMM contains all kind of knowledge. This is the main data source of the system. It stores knowledge related to linguistic knowledge such as dictionary, and grammar, and non-linguistic knowledge such as disease information, previous dialog information, and procedural knowledge which relates inferring a fact from a collection of facts. In Fig. 2. the nodes represent the case structures and the edges represent their relationship.

The knowledge is classified in predefined classes. For this system prototype the nodes are classified on each disease. A class can be a property of other classes. The GMM manager is an algorithm that searches the disease information in Ground memory when it requested by the Dialog Manager and also finds the discourse knowledge requested by Discourse memory module. Ground memory is always updated after every conversation but never delete anything.

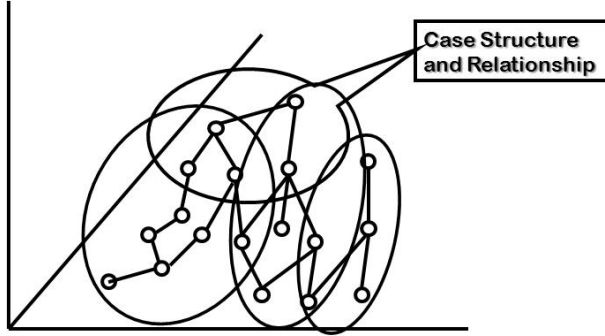


Fig. 2. Case structure and their relationship

### B. Discourse Memory Module

Discourse memory provides situational and contextual information for an utterance environment. It contains the nodes with different frequency. The frequency is calculated depending on the uses of knowledge. In our proposed system we used three frequencies: medium, high and very high. We standardize the frequency three after examining with other frequencies. We found better response using the frequency three. The very high frequency is determined by the mostly used classes, recently used classes and seasonal disease classes. The high frequency nodes are determined by the basic information of the patient and his past medical record. Medium frequency nodes are determined by the medium uses of classes. The knowledge and their frequency of discourse memory are different for different patients. The Fig. 3 represents the frequency of nodes.

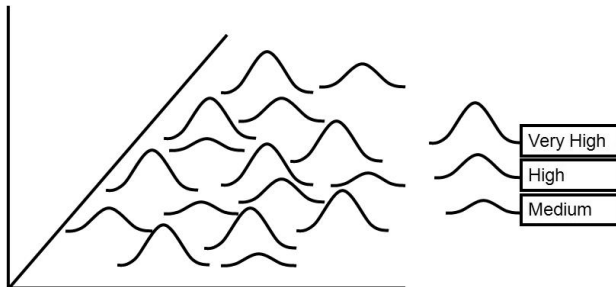


Fig. 3. Frequency of nodes

### C. Episodic Memory Module

It is the module that receives input from the user and provides output to the user. The input processor receives the input text dialog and understands. Episodic memory stores the meaning of ongoing segment of utterances. As understanding proceeds, the essence of episodic memory assimilated into discourse memory and the essence of discourse memory assimilated into ground memory. In the same way the output processor receives case frames from dialog manager and generates the appropriate text for user. To understand the dialog the input processor used some grammatical modules, text normalization unit, Semantic classifier and a converter for case structured knowledge. Episodic memory contains all the linguistic information and the recent conversation with present user. EMM manager includes the entire previous dialog, responses with the case structure of new dialog and classifies in the pre-defined groups. It sends these case structures to the discourse memory module. Episodic memory is a temporary memory. It updates only for the ongoing process requirement. After every conversation with a user, it refreshes the memory.

## IV. TLCN-DP ARCHITECTURE

The TLCN dialog processing architecture system is illustrated in Fig. 4. This architecture provides efficient integration of component modules. Text input utterances are sent to dialog understanding unit named episodic memory module (EMM) which first normalizes the sentences and produces some case frames then these case frames are sent to EMM manager which integrates the previous dialogs case frames and structures by the assimilation and instantiation of discourse memory and ground memory to understand the dialog. Then the input dialog is classified in some classes and is sent to Dialog Manager (DM). DM searches this knowledge in discourse memory with the help of discourse memory module (DMM) manager and finds the disease information. DM will make the decision whether the disease is identified or not enough information in discourse memory or have to ask more question for confirming the disease. If the information is not enough in discourse memory then DM sends the input knowledge to ground memory module (GMM) manager. GMM manager does the same job as DMM manager and sends the disease information to DM. If DM finds that the patients information is not enough for matching a disease for the Recurrent Neural Network it will generate a question type response case structure for the dialog generator. Dialog generator will ask the question to the patient. If the dialog manager found the goal it will give advice or suggested medicine name to the dialog generator. Then the dialog generator generates a text dialog for the patient.



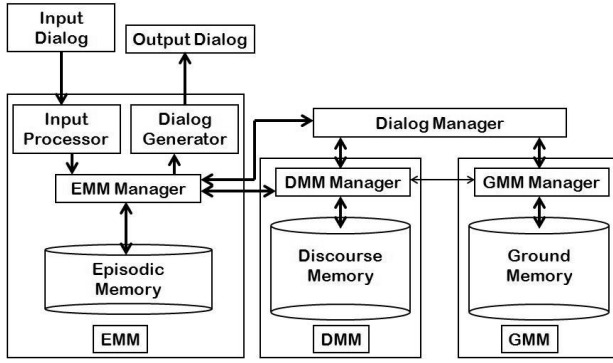


Fig. 4. TLCN Dialog Processing System

## V. TLCN-DP: COMPONENT TECHNOLOGIES

The TLCN-DP system was based on doctor-patient dialog, the statistics of which are given in Table I. These diseases information have been collected from some medical web link [10] [11] [12].

TABLE I  
TRAINING CORPUS STATISTICS

	Number
Diseases	25
Sentences for each disease	80
Normalized Sentences for each disease	180
Vocabulary Size	2000

### A. Text Understanding

The text understanding component of the TLCN-DP consists of language understanding module and the language generation module. The understanding module normalizes the input text and converts the output into a language-neutral meaning representation a case frame. The language generation module produces the response text from the meaning representation. In this section we discuss four component of language understanding technology: i) text normalization, ii) case frame representation, and iii) classification.

### B. Text normalization

Text normalization is an essential step for minimizing noise variations among words and utterances. The text normalization component is essentially based on using synonyms and other forms of syntactic normalization. The main steps include stemming, using a synonyms dictionary, and removal of confusions, non-alphanumeric and non-white space characters.

### C. Case frame Model

The extended case frame model we used is an understanding model consisting of predicates, semantic case relations (roles), modalities, and conjunctive relations [7]:

*sentence* = *simple\_sentence*|*sentence* +  
*conjunctive\_relation* + *sentence*  
*simple\_sentence* = *predicate* + *case\_relations* +  
*modalities*  
*case\_relations* = *object*|*method*|*direction*|*timeand*  
*space*|*supplement*|*modification*  
*modalities* = *tense*|*aspect*|*manner*|*intention*|*guess*  
|*attitude*|*negation*|*ability*|*necessity*|...  
*conjunctive\_relation* = *time*|*cause*|*reason*|*result*|  
*contrast*|*goal*|*assumption*|

- 1) Case Relations: One characteristic of this model is the relatively large collection of case relations. The collection includes; 1) basic cases like Fillmore's case system, 2) case relations such as extent, manner, and degree, which appear as adverbial phrases, and 3) additional cases represented by the inflection of a certain verb.
- 2) Modalities: Semantic structures must include information on modality such as tense, aspect, intention, manner, attitude and assumption. Real (not toy) sentences convey much of this type of information. Unfortunately, however, most understanding systems and linguistic theories have not dealt sufficiently with modalities.
- 3) Conjunctive Relations: A sentence generally consists of several sub-sentences, each of which represents a unit event. In the model, event relations are expressed by conjunctive relations such as time, cause, reason, result, goal, assumption, contrast and circumstance. These are essential for understanding a sentence, and must be organized into semantic structures.

### D. Extended Case frame Representation

An extension of case structure model is proposed as a linguistic model for representing the text meaning structure. Thus the case frames represent a representation scheme for episodic memory. The traditional case structure is a structure for a unit sentence which consists mainly in relations between noun and verb. This is not sufficient to represent structures of real sentences which sometimes have complex noun phrase and compound sentences. Also our proposed case structure has to have facilities for representing other structures involving relations between two nouns and two verbs etc. It has been designed to integrate those structures into one linguistic model. Its nature is hierarchical with respect to the way constituents are connected, iterative with respect to conjunction and recursive with respect to embedding. Using this formalism, the syntactic and semantic structures of sentences can be represented uniformly.

In our proposal every noun has six properties: adjective, delimiter, preposition, auxiliary verb, adverb and verb.

Noun can be related by either verb or preposition. This case frame representation makes the matching algorithm simple and efficient. It's depicted in Fig. 5.

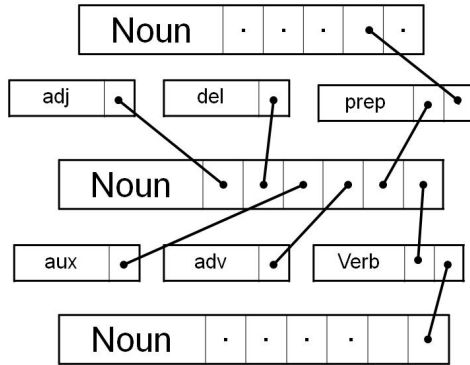


Fig. 5. Extended case frame structure

### E. Knowledge Representation

In TLCN-DP the nodes are classified in 9 properties: 1) Disease, 2) Symptom, 3) Environment Cause, 4) Physical Cause, 5) Other Cause, 6) Tests, 7) Treatment & Medicine, 8) Future Effect and 9) Prevention.

Disease is the head node. All the other nodes are related with disease. Fig. 6. explains symptom and cause used to find the goal disease. Then pharmaceutical tests confirm the disease (if any tests exist). After that the TLCN-DP will give detail to the patient detail including treatment, medicine, effect, carefulness & prevention.

If any utterance from patients dialog contains any data related to body then it will be treated as Symptom. If any utterance contains any data about treatment or Medicine then it will be treated as Treatment. If any utterance contains any data about test then it will be treated as Test oriented. Otherwise it will be treated as Cause oriented. Classifier sends the classified utterances to the DM manager.

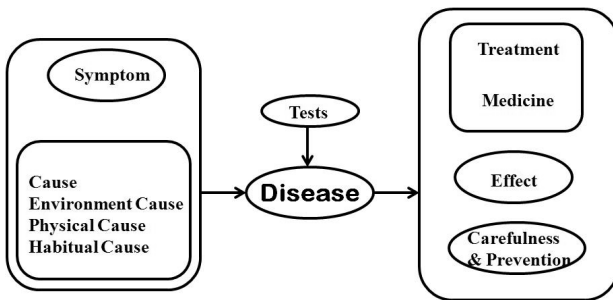


Fig. 6. Relationship of knowledge

Fig. 7. shows the knowledge representation where the disease is the head classes and others are the daughter

classes. This also explains how case frames can be related with each other.

### F. Classifier

Recurrent Neural Network Classifier is used to identify a disease. The inputs of this classifier are the Indexes of classified knowledge properties and their amplitudes.

Recurrent Neural Network: Recurrent Neural Network is a class of neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic temporal behavior.

Elman Recurrent Neural Network: A recurrent network with feedback from each hidden node to all hidden nodes. When training the Elman network back propagation-through-time is used rather than the truncated version used by Elman, i.e. in this paper Elman network refers to the architecture used by Elman but not the training algorithm. An overview of Elman Network can be seen in the Fig 8.

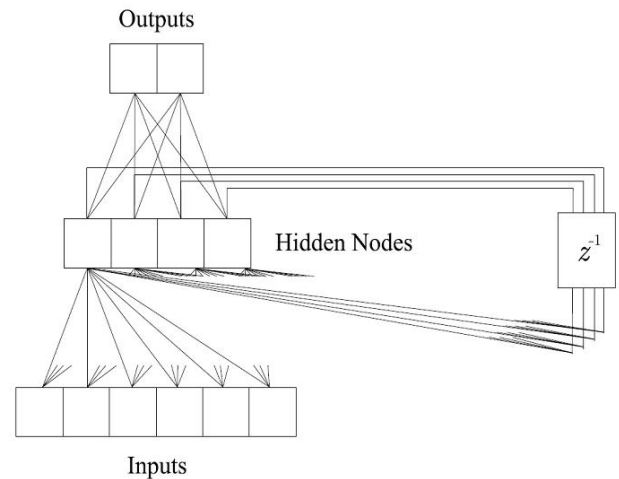


Fig. 8. Elman Recurrent Neural Network

Back-propagation-through-time has been used to train the recurrent networks. Elman networks using: four word inputs, 7 hidden nodes, the quadratic cost function, the learning rate schedule shown below, an initial learning rate of 0.2 and we use the random weight initialization strategy. The Recurrent Network we used for this application is depicted in Fig. 9. Where S indicates Symptom, EC indicates Environment Cause, PC indicates Physical Cause, OC indicates Other Cause, T indicates Tests, H indicates Hidden node, D indicates Disease and NI indicates Not Identified.

- 1) Target Output: Targets outputs were 0.1 and 0.9 using the logistic activation function. This helps avoid saturating the sigmoid function.

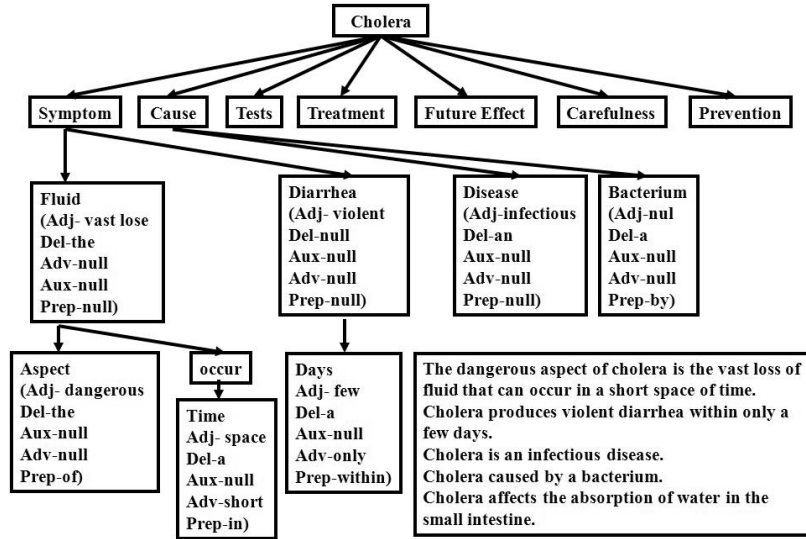


Fig. 7. Knowledge representation

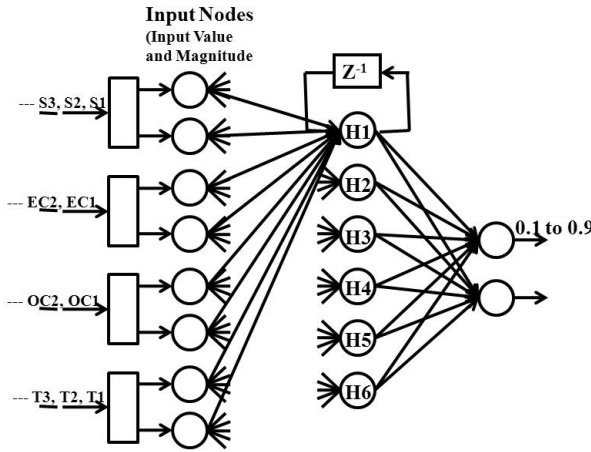


Fig. 9. Disease Classifier

- 2) Weight Initialization: Random weights are initialized with the goal of ensuring that the sigmoids do not start out in saturation but are not very small (corresponding to a flat part of the error surface). In addition, several (20) sets of random weights are tested and the set which provides the best performance on the training data is chosen. In our experiments on the current problem, it was found that these techniques do not make a significant difference.
- 3) Learning Rate Schedules: Relatively high learning rates are typically used in order to help avoiding slow convergence and local minima.
- 4) Activation Function: Symmetric sigmoid functions often improve convergence over the standard logistic function. For our particular problem we found that the difference was minor and that the logistic

function resulted in better performance.

- 5) Cost Function: We used quadratic cost function.

$$E = \frac{1}{2} \sum_k (y_k - d_k)^2 \quad (1)$$

Where  $y$  and  $d$  correspond to the actual and desired output values,  $k$  ranges over the outputs.

- 6) Sectioning of the Training Data: We investigated dividing the training data into subsets. Initially, only one of these subsets was used for training. After 100% correct classification was obtained or a pre-specified time limit expired, an additional subset was added to the working set. This continued until the working set contained the entire training set. The data was ordered in terms of sentence length with the shortest sentences first. This enabled the networks to focus on the simpler data first. Elman suggests that the initial training constrains later training in a useful way.

#### G. Dialog Manager

In our TLCN-DP architecture, we have incorporated a response module to help patients to find the goal disease. Dialog manager provides case frame response to the language generator. It analyzes the result and selects the type of system response. The system responses are divided in mainly three types: 1) response before disease identification, 2) after disease identification and 3) failed to identify.

- 1) Response before disease identification: In this type of response, first the dialog consider the symptom and the causes of height match disease which are the minimum similar to the other selected diseases. Response module selects case frames for symptoms and

cause. After 10 conversations it select case frames for Pharmaceutical tests (if there any for height match disease) and also the effect and prevention oriented case frames for height 3 matching diseases.

- 2) After Disease Identification: After Disease Identification the dialog manager provides case frames for treatment including Medicine, Effect, and Prevention for the specific disease to the language generator. It can also provide more case frames for detail of this disease by the requirement of patients.
- 3) Failed to identify: If the dialog manager doesn't get any information about any diseases, it generates failure notification.

#### H. Language Generator

It generates the sentences by arranging the parts of speeches of the case frames. And send the output to the user.

### VI. SYSTEM

We developed a prototype system for real-time medical advices application, called TLCN-DP, to help the people who have to go to the hospital frequently. First we selected some keyword for classification of the utterances. Then we have trained the semantic classifier using a corpus. We collected the disease information from some medical books and some medical web link. By the Question Answering information we trained the Recurrent Neural Network Classifier. We built the ground memory with case frame by this information.

#### A. PERFORMANCE EVALUATION

Evaluation of dialog system performance is a complex task and depends on the purpose of the desired dialog metric. We have trained our system with mostly common 25 diseases from [10], [11], [12] where 9 are mainly related with age, 8 are related with season and the other of 8 diseases are related with food, accident, habitual, health condition and location. The real doctor-patient dialog conversation is collected for this evaluation. The dialogues have been collected from various books on medical interviews and some frequently asked questions- answering medical web links. 80 patients dialogs were simulated with this system prototype where 10 patients were tested 3 times and 15 patients 2 times and 20 patients only once. Table 2 presents the existing disease identification accuracy with the number of dialog conversation. According to the result analysis table; we found that the system achieved a very satisfactory result for 10 to 20 diseases. For 15 diseases we found 78% accuracy of disease identification. For less than 10 diseases the system can identify with very few dialog conversation. 69% accuracy achieved only for 6 utterances. The success of recurrent neural network is always more than 90% for existing disease.

TABLE II  
DISEASE IDENTIFICATION EVALUATION

Diseases	Utterances					
	5	6	7	8	9	10
10	51%	69%	82%			
15	33%	38%	58%	70%	78%	
20	15%	22%	34%	49%	59%	65%
25			20%	26%	42%	51%

### VII. CONCLUSION

In this paper we presented some architecture and technologies. The main underlying strategy is to adopt 1) Three Layer Conceptual Network, 2) uses of case frame, 3) dialog matching, 4) disease identification and 5) response generation. In addition of these fundamental strategies, we investigated a variety of linguistic constraints.

Still, there is a big problem to understand the dialog and it takes a plenty of time so there are some scope to improve this section. We can also use Neural Network to improve efficiency of the conceptual network, adopt some module to train this system by the real time conversation of a doctor and patient. And in future we plan to use voice dialog rather than text dialog.

This is only beginning of Neural Network with Three Layer Conceptual Network. Our goal is make human-computer successful interaction. We strongly believe this research is a new breed of natural-language dialog processing.

### REFERENCES

- [1] M. K. Rhaman and T. Endo, "Three layer conceptual network dialog processor," in *The Twelfth IASTED International Conference on AI and Soft Computing*, Sept. 2008, pp. 628-802.
- [2] H. Nomura, "Modeling and representative framework for linguistic and non-linguistic knowledge in natural language understanding," in *Germany-Japan Science Seminar*, 1986, pp. 1-10.
- [3] M. A. Walker, "Evaluating discourse processing algorithms," in *Proceedings of the ACL*, 1989.
- [4] H. Nomura, "Experimental machine translation systems lute," in *Second Joint European-Japanese Workshop on Machine Translation*, 1985, pp. 1-2.
- [5] D. Marcu and A. Echihiabi, "An unsupervised approach to recognizing discourse relations," in *The 40th Annual Meeting of the Association for Computational Linguistics*, July 2002, pp. 368-375.
- [6] B. B., "Case systems for natural language," *Artificial Intelligence*, vol. 6, no. 4, pp. 327-360, 1975.
- [7] H. N. A. Shimazu, S. Naito, "Japanese language semantic analyzer based on an extended case frame model," in *International Joint Conference on Artificial Intelligence*, 1983, pp. 717-720.
- [8] J. Elman, "Distributed representations, simple recurrent networks, and grammatical structure," *Machine Learning*, no. 7, p. 195226, 1991.
- [9] C. L. G. Steve Lawrence and S. Fong, "Natural language grammatical inference with recurrent neural networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 12, p. 126140, 2000.
- [10] "http://www.netdoctor.co.uk/," Weblink, Online Medical Solution.
- [11] "http://www.doctoronline.nhs.uk/ver02/index.asp," Weblink, Online Doctor.
- [12] "http://www.nzdoctor.co.nz/default.aspx," Weblink, online Medical Solution.

# An Energy Saving Program for Bangladesh, for Reducing Load Shedding, and for Continuity of Power for IT Sector

Shahriar Khan

School of Engineering and Computer Science, Independent University, Bangladesh, 14/2 A,  
Baridhara North Road, Baridhara, Dhaka - 1212, email: khandhaka@gmail.com

**Abstract** — The success of computers, IT, and telecommunications hinges greatly on the inexpensive and uninterrupted availability of power. Even with the gradual depletion of fossil fuel, and the rise in fuel prices, there has been insufficient effort towards saving electricity and making the most of remaining fuel reserves. It is time to adopt energy saving programs in Bangladesh, creating awareness that extravagant consumption is detrimental to national, global, and long-term interests. Such energy saving programs have been implemented to various extents in countries like Brazil, Norway, New Zealand, and the USA. With its high population density, low development, poor economic infrastructure, and dynamic river network, Bangladesh provides a unique backdrop to an Energy Saving programs. Such a program should reduce load-shedding, distribute electricity among more users, and divert electricity to productive uses like irrigating fields and preventing manufacturing stoppage. Continuity of electricity would also save our huge investment in generators and UPS equipment. An Energy Saving program should be well-understood and researched by policymakers and experts. An important part of such a program should be changing consumer behavior, such as turning off appliances when not in use. Power-hungry air-coolers develop acclimatization and dependence in users, who then perceive normal outside temperatures as uncomfortable. Low-power alternatives to air-coolers are ceiling fans and ventilator fans on walls. LED lights, compact fluorescent lamps, and IT and telecommunications are to be promoted as efficient uses of electricity. Urban sprawl to distant suburbs should be contained in favor of compact townships, which are better adapted to buses. Buses should be promoted over smaller cars, which should be promoted over larger hungry cars and SUVs. This paper pursues such common sense approaches, together with recommendations by experts, and implementation by policymakers.

**Index Terms** — Energy, Power, Load shedding, Power Outage, Air conditioner, Fossil Fuel, Conservation.

## I. INTRODUCTION

The success of computers, IT, and telecommunications depends greatly on the inexpensive and uninterrupted availability of power. The recent increases in fuel prices and load sheddings have been an awakening that the days of inexpensive fossil fuel are on the way to being over. In response to the energy crises, countries such as Brazil, New Zealand, and Norway, have implemented energy savings programs with various degrees of success.

Hungary has documented significant health and environmental benefits from its energy savings program [1]. In Canada, the power generation companies have become important participants in their energy savings programs [2].

With the decline of local and global fuel reserves, actions to save energy and maximize fuel utilization are long overdue for Bangladesh. Such a program is to have a two-pronged approach. Firstly, there should be understanding and research among planners of an energy saving program. Secondly, there should be implementation of the program by policymakers, taking concrete steps, such as enacting legislation and conducting media campaigns.

## II. LIMITED FOSSIL FUEL

The limited nature of fossil fuel reserves (oil, gas, coal), and the need to ultimately transition to alternative sources is well recognized. However, in Bangladesh and worldwide, there has been insufficient effort to conserve energy, and make the most of remaining fuel reserves [3].

The price of oil is dictated both by geopolitical conditions, and by scarcity of reserves [4]. Most oil and gas reserves of the world have already been discovered, and there has been a shift to deeper oil wells, in our search for economically viable reserves. Even with the latest technology such as Directional Drilling, and Logging While Drilling, and underground sensors like Radio frequency and Magnetic Resonance Imaging, we can economically extract at best about two-thirds of oil in a reserve [5,6]. About a decade ago, natural gas was burnt as a by-product of oil-production, in the US and Canada, because of its lack of economic viability. Amidst the increasing scarcity and demand, natural gas itself has become widely viable economically, even in the US and Canada. In Bangladesh, the energy resources advisor recently declared that no further gas-powered power stations would be constructed, because of depleting gas reserves [7].

### A. Alternative Energy

There is no source that quite provides the ease and energy density of fossil fuel. Common-sense approaches dictate that we maximize utilization of our remaining reserves, as most alternative energy sources cannot come close to fossil fuel in economic viability. Solar energy requires major investment in large areas of silicon panels, and the laws of physics dictate that no more than about a third of solar energy incident on a solar panel is converted to electricity. Solar panels covering a household rooftop would be barely enough to supply the household requirements (without air conditioning), and would need batteries for use at night. Wind energy requires large investment and maintenance for turbines. Air flow over Bangladesh is obstructed by the Himalayas in the North, and wind speeds here are too low to give any viability to wind turbines. Airplanes running on solar energy, wind energy, or batteries are almost inconceivable. Nuclear power and its lead shielding would be too heavy for flight, and would have to run on the principle of propellers rather than jet propulsion.

### B. Biodiesel

Recently, there has been an initiative in the West, to produce biodiesel from grains. In some countries, gasoline must be mixed with 2.5 % biodiesel by law. This has been cited as the primary cause of the dramatic increases in the price of rice and other foods. In Indonesia thousands of square kilometers of forests have been clear-cut to make way for growing grains for biodiesel.

It is evident that production of biodiesel uses up land that could have been otherwise used for edible grains. In addition clear-cutting of forests for biodiesel is detrimental to the environment. A Saving Energy program would ease the pressure on biodiesel, and its negative consequences.

## III. BACKDROP OF BANGLADESH

With its high population density, low development, and poor economic infrastructure, Bangladesh provides a unique backdrop to energy and electricity saving programs. Electricity is vital for economic development, but only about 20 percent of the population has access to it (25 percent of urban population and 10 percent of rural) [8,9,10]. Other problems in the power sector include high system losses, delays in completion of new plants, low plant efficiencies, shortage of natural gas, erratic power supply, electricity theft, blackouts, etc. These factors contribute to Bangladesh's per capita electricity consumption being among the lowest in the world at about

110 kWh per person (Table 1, and Figure 1). Shortfalls sometimes go up to 500 MW (as of August 2005), requiring new plants and fossil fuel the country can barely afford [11, 12]. The availability of gas is already in decline, and has been predicted to last until about 2011. Although power generation capacities cannot be increased overnight, energy saving programs can be initiated to distribute the electricity among more people.

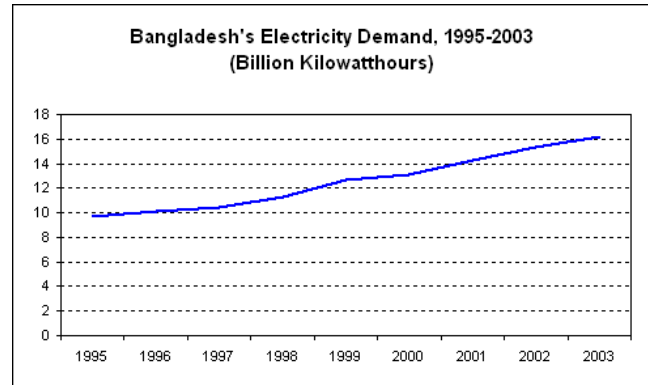


Fig 1. Electricity Demand from 1995 to 2006 for Bangladesh

TABLE 1  
ELECTRICITY CONSUMPTION PER CAPITA IN NEIGHBOURING COUNTRIES (KWH).

India	474
Pakistan	431
Sri Lanka	336
Bangladesh	110
Nepal	65

### A. Load-shedding and Back-up Power

Insufficient supply of electricity, such as during peak demand, is normally met by disconnecting selected areas, otherwise known as "load-shedding." This is understandably unpopular, and a socio-political factor during important events and sports [8,9]. Affluent areas are less subjected to load-shedding, and public protests for lack of electricity turned violent in the mid-2006 [10].

### B. Back-up Power

Load-shedding requires heavy investment in generators and UPS equipment for both domestic and industrial consumers. Without such equipment, there is significant loss in commercial and industrial productivity, such as for irrigation and manufacturing. Such investments in back-up power are almost unheard of in the West.

#### IV. ENERGY EFFICIENCY PROGRAMS

The discrepancy between electricity production and demand, coupled with dwindling global fossil fuel reserves and global warming, strongly favor new policies for reducing electricity consumption. In response to the energy crises, many countries have implemented energy savings programs with varying degrees of success. Hungary has experienced significant health and environmental benefits from its energy savings program [1]. In Canada, the power generation companies have become important participants in their energy savings programs [2].

Bangladesh should have an energetic Energy Saving programs designed by scientists and policymakers. There should be well-founded policies, and promoting awareness that extravagant electricity consumption is detrimental to national, popular, global, and long-term interests.

Such an Energy Saving program should consist of two parts: (a) research and investigation by researchers (b) planning, policy-making, and implementation by policymakers.

##### *A. Research and Planning*

For the success of any Energy saving program, there should be an appropriate understanding of the issues by planners and policymakers. Research may be conducted on issues which require clarification. In the context of Bangladesh, the foremost planner and implementer of an energy savings program is the government. It is also possible for the government to subcontract the planning of the program to others, such as a non-profit organization, or academia.

##### *B. Changing Consumer Behavior*

The essence of a Save Energy program is changing consumer behavior, such as turning off lights and appliances when not in use. Such programs have been successfully implemented in Brazil, California, New Zealand, Norway, Canada, and the U.S.A., where demand was reduced by 3-20%. Numerous websites have come up promoting saving of electricity [13,14].

Lighting and air conditioning can be practiced in moderation. Lights can be turned off when away for 5 minutes or more. Stickers can be posted near switches saying, "Turn off Lights when Out for 15 Min"

Air conditioners can be turned off, when away for an 15 or more. Increasing the thermostat by just one degree results in significant savings. Often, much of a house or office is cooled or lit, when only one room is in use. These behaviors may be targeted for change, thorough user-friendly messages in the media.

#### V. AIR CONDITIONING AND ACCLIMATIZATION

Recent worldwide and local warming of the weather, has been met with by increased usage of air conditioners, ACs. Air conditioners are very power-hungry by nature, and may comprise much or most of a household bill. In the US, 5 % of all electric energy goes to ACs [15]. Although similar statistics for Bangladesh are not readily available, warming of the weather, and increased per-capita income has increased the popularity of ACs in recent years. The case against air-coolers is that the body becomes dependent and "acclimatized" to the colder temperatures, with inability to tolerate normal outside warmth and humidity [10]. Those who spend much or most of the time under an air-cooler, begin to perceive outside heat and humidity as uncomfortable.

Acclimatization produces the desire to be under an AC at other times, such as inside a car (up to several hours a day, in our sprawling cities and traffic jams). Lowering dependence on the car AC, and measures such as blocking sunlight at the windows are major potentials for energy reduction.

Humans are capable of surviving in extreme temperatures, ranging from the cold of the Arctic, to the 40s °C in the tropics. From anecdotal evidence, when moving from a cold country to a warm country or vice versa, the body adapts in about 2 weeks, if not less. These are other examples of acclimatization.

ACs inherently operate in closed rooms with little or no ventilation. In the daytime and with high human occupancy, closed rooms tend to be a few degrees warmer than the outside air, putting the burden on the AC to first cool the room to the outside temperature. This equalization of inside and outside temperatures and humidity could otherwise be efficiently done by a ventilating fan at the windows or near the ceiling.

Ceiling fans and table fans provide circulation of air and evaporation of sweat, hence cooling the body, with much lower power consumption (Table 2).

TABLE 2  
COMPARISON OF COOLING METHODS [17,18,19]

	Size	Power
Air cooler,	10,000 BTU	1300 watts
Ceiling fan	48" fan	75 watts
Floor fan	Various sizes	15-95 watts

With the multi-storied buildings coming up in large cities, there is less wind movement near the lower floors. Ventilator fans at walls would be very useful in expelling the warm and humid air in people-inhabited rooms. An

energy saving program should promote ventilator fans at walls in preference to ACs.

#### A. The "Heat" of Direct Sunlight

High temperatures and humidity are not to be confused with the perception of "heat" from direct sunlight. Direct sunlight can be uncomfortable and dangerous, whereas one can be acclimatized to normal outside temperatures and humidity of the shade. A body receives its full requirement of Vitamin D from sunlight from only about 5 minutes of direct exposure per day.

There is good reason to be concerned about the thinning Ozone layer, and the increasing ultraviolet rays we are exposed to nowadays. Direct sunlight can be associated with nausea and discomfort, among many other effects, and should be blocked to counter discomfort. The ozone layer over certain places like New Zealand has deteriorated to the point that even Bangladeshis get sunburnt from half an hour's exposure to the sun.

Sunlight in a car may be blocked with reflectors and shades at the windows.

### VI. LIGHTING

Much of the lighting locally and globally is with energy-inefficient incandescent light bulbs of 60 to 100 watts. Light Emitting Diodes (LEDs) have been introduced in pilot projects in Bangladesh and India [20]. LEDs have very low power consumption, and better adaptability to batteries but have much higher initial cost (Table 2) [21]. There is little or no medical evidence that low lighting cause deterioration of vision.

TABLE 2  
WATTAGE AND COSTS OF DIFFERENT LIGHTS

	Wattage	Cost
Bulb	60 to 200	Tk 20-40
Fluorescent	7 to 20	Tk 170-250
LED	(18 LEDs) 2 W	US\$ 27

An Energy Saving program can replace incandescent bulbs with Compact Fluorescent Lamps, CFLs, which use 50 – 80 % less energy [22]. Although, the CFL is more expensive initially, they last about 10 times longer than incandescent bulbs, and are less expensive in the long run. A single 18 watt CFL used in place of a 75 watt incandescent will save about 570 kWh over its lifetime. In California, consumers installed nearly eight million CFLs during their crisis period, resulting in demand reduction of almost 500 MW.

Lights activated by motion and infra-red sensors are household items in the West, but at about \$20 a piece, they are not yet affordable to Bangladeshi customers.

Prices may be brought down further, spurred by rising electricity prices, and widespread manufacture in Asia.

#### A. Festival lights

Display and festival lights, consisting of thousands of little bulbs, are another example of conspicuous and avoidable consumption. The government of Bangladesh had discouraged such festival or display lighting in 2007, but their use has continued ever since. Power consumptions of a single lighting can be about 10 KW, assuming 15,000 bulbs of 0.6 watt each.

### VII. TELECOMMUNICATION

Applications of telecommunication, such as telephones (mobile and fixed), telemedicine, the internet, television, radio, etc. are efficient uses of electricity, as they reduce our dependence on fuel-dependent traveling [23,24,25]. Over the next few years, as traveling continues to be expensive in terms of fossil fuel, telecommunication will be an increasingly affordable substitute to traveling.

### VIII. CONTAINING URBAN SPRAWL

Dhaka has sprawled over distant suburbs like Uttara, Bashundhara, Purbachol, Gazipur, and Tongi. This type of urban sprawl promotes the need for cars, rather than buses. Carrying numerous passengers, buses are efficient modes of transportation, but are only adapted to compact cities, rather than spread-out cities with distant suburbs. An Energy Saving program should discourage and contain the urban sprawl taking place in Dhaka and other suburbs.

The popularity of heavier and fuel-hungry Sports Utility Vehicles has continued for decades in the West. The price increases of fuel in 2008 has forced consumers to step back from such conspicuous consumption. An energy saving program should encourage buses in preference to cars.

### VIII. PROPOSED SAVING ELECTRICITY PROGRAM

Summarizing, a proposed Saving Energy program may have the following.

- Awareness that fossil fuels are a limited resource, to be used with care.
- Media campaigns (TV, radio, newspapers) creating awareness among the masses that extravagant consumption are detrimental to national, global, and long-term interests.
- Campaign to "switch off" appliances that are not being used. Post stickers near switches saying, "Turn off Lights when Out for 5 Min"



- Awareness that extravagant consumption may prevent field irrigation, or cause manufacturing disruptions in the under-developed country of Bangladesh.
- Reduce or Avoid usage of Festival Lights
- Discourage air conditioners, or increase AC thermostat temperatures. Realization that the human body is easily adaptable to heat, and easily acclimatizes itself to its surroundings.
- Ceiling and Ventilator fans are to be used in preference to ACs, which lead to temperature acclimatization, and dependence.
- Direct sunlight, such as in a car, should be blocked with reflectors and shades at windows
- Urban sprawl should be contained, and compact cities, favorable to buses should be encouraged.
- Large cars and SUVs should be discouraged in favor of smaller cars, which in turn should be discouraged in favor of buses.
- Media campaigns should create awareness that our children and our children's children will be the beneficiaries or victims of our actions today.
- Recognition that one should not misuse energy and electricity just if one can afford it.

## VIII. CONCLUSION

The success of computers, IT, and telecommunications hinges greatly on the inexpensive and uninterrupted availability of power. Although wastages of electricity and energy are obvious to most of us, there has not been sufficient common-sense approach to the problem. Conspicuous consumption has continued unabated in Bangladesh and worldwide. Energy saving policies must be envisioned and implemented, as a common sense response to the continuing depletion of fossil fuel.

## REFERENCES

- [1] K. Aunan, G. Patzay, et. al. "Health and Environmental Benefits from Implementation of an Energy Saving Program in Hungary," [www.cicero.uio.no/media/92.pdf](http://www.cicero.uio.no/media/92.pdf), University of Budapest, ISSN: 0804-4562, January 1997.
- [2] Rouse, S., Ontario Power Generation's green response program, IEEE Power Engineering Society Winter Meeting, Vol. 1, pp. 113, 27-31, Jan 2002.
- [3] S. Khan, F. Ndiaye, "A New Approach and Proposal for Saving Electricity, so as to Minimize Load-shedding, and Make Most of Remaining Fossil Fuel Reserves," Proceedings of National Conference on Electronics, Information, and Telecommunication," pp. 1218, Bangladesh Electronics Society, Rajshahi University, Rajshahi - 6205, Bangladesh, 29-30 June, 2007
- [4] Crude Oil Prices, Oil Prices History and Analysis, [www.wtrg.com/prices.htm](http://www.wtrg.com/prices.htm), accessed July 24, 2008
- [5] Schlumberger Oilfield Services, [www.slb.com](http://www.slb.com), accessed July 23, 2008.
- [6] Halliburton Oil Technologies and Services, [www.halliburton.com](http://www.halliburton.com), accessed July 22, 2008
- [7] The Daily Star, July 22, 2008, [www.dailystar.net](http://www.dailystar.net)
- [8] "City goes haywire over power demand, Youths clash with cops". Dhaka, Bangladesh: May 10. (<http://independent-bangladesh.com>)
- [9] The New Age. 2006. "Power outages on as Meghnaghat plant shut partially". Dhaka, Bangladesh:
- [10] The Daily Star. 2006. "Kansat agitators attack cops, damage vehicles". Dhaka, Bangladesh : April 9.
- [11] CIA. 2006. "Electricity Consumption". [www.cia.gov](http://www.cia.gov) : CIA, The World Fact Book.
- [12] EIA. 2005. "Bangladesh Country Analysis Brief". [www.eia.doe.gov](http://www.eia.doe.gov) : Energy Information Administration.
- [13] Meier, Alan. 2004. "Saving Electricity Quickly". [www.iea.org](http://www.iea.org). Paris: International Energy Agency.
- [14] Power Scorecard TM. 2005, <http://powerscorecard.org>
- [15] AC Doctor, [www.acdoctor.com](http://www.acdoctor.com)
- [16] Acclimatization; Adjusting to Temperature, <http://www.uihealthcare.com/topics/medicaldepartments/internalmedicine/acclimatization/index.htm>
- [17] Briggs and Stratton, The Power Within, [www.briggsandstratton.com](http://www.briggsandstratton.com), accessed July 24, 2008.
- [18] Progress Energy Website, [www.progress-energy.com/custservice/carres/energytips/ceilingfans.asp](http://www.progress-energy.com/custservice/carres/energytips/ceilingfans.asp)
- [19] Saving Electricity, Calculate your Energy Cost, <http://michaelbluejay.com/electricity/howmuch.html>, accessed July 24, 2008.
- [20] Surya Bijlee, Solar Energy Lighting, [www.suryabijlee.com/program.htm](http://www.suryabijlee.com/program.htm), accessed July 24, 2008.
- [21] C. Crane, LED Lights, [www.ccrane.com/lights](http://www.ccrane.com/lights), accessed July 24, 2008.
- [22] Energy Star. 2006. "Compact Fluorescent Light Bulbs". <http://www.energystar.gov> : Energy Star.
- [23] S. A. Wadud, S. Khan, "Overview of Telecommunication Sector in Bangladesh," Proc. of Prospects and Problems of Mobile and Land Phones in Bangladesh, pp 103-109, Dec. 2005.
- [24] M. R. Siddiqui, S. Khan, "Transmission Network Planning Strategies for Cellular Mobile Networks of Bangladesh to Facilitate Transition to Third Generation," Prospects and Problems of Mobile and Land Phones in Bangladesh, Dhaka, pp. 27-44, 2005.
- [25] I. Al-Mamoon, S. Khan, "Telemedicine Services at the Diabetic Association of Bangladesh with the Technical Collaboration of Grameen Telecom," Prospects and Problems of Mobile and Land Phones in Bangladesh, Dhaka, Bangladesh, December 8th, 2006.
- [26] H. Herring, "Does energy efficiency save energy? The debate and its consequences," Journal of Applied Energy Volume 63, Issue 3, Pages 209-226, July 1999.
- [27] Van Gorp, J.C, "Maximizing energy savings with enterprise energy management systems," Pulp and Paper Industry Technical Conference, 2004, pp. 175 - 181, 27 June -1 July 2004.

# Analysis of Broadband Slotted Microstrip Patch Antenna

Mohammad Tariqul Islam<sup>1</sup>, Mohammed Nazmus Shakib<sup>2</sup>, Norbahiah Misran<sup>1,2</sup>, and Baharudin Yatim<sup>1</sup>

<sup>1</sup>Institute of Space Science (ANGKASA), <sup>2</sup>Dept. of Electrical, Electronic & Systems Engineering,  
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

Email: titareq@yahoo.com, engmdns@yahoo.com

**ABSTRACT** — In this paper, a novel multiple slot microstrip patch antenna with enhanced bandwidth is presented. The design adopts contemporary techniques namely; probe feeding, inverted patch structure and multiple slotted patch. The composite effect of integrating these techniques and by introducing the novel multiple shaped patch, offer a low profile, wide bandwidth, high gain and compact antenna element. The proposed patch has a compact dimension of  $0.707 \lambda_0 \times 0.354 \lambda_0$  (where  $\lambda_0$  is the guided wavelength of the center operation frequency). With the proposed concept, an antenna prototype is simulated and analyzed. The proposed antenna achieves a fractional bandwidth of 27.15% (1.91 to 2.51 GHz) at 10 dB return loss while maintaining a maximum gain of 10.5 dBi. The design is suitable for array applications especially for base station.

**Index Terms** — Broadband antenna, microstrip patch antenna, probe fed.

## I. INTRODUCTION

The current demand of wireless communication systems and their miniaturization, antenna design becomes more challenging in present days. The main reason of microstrip patch antenna to attract wide interest is because of its inherent characteristics such as light weight, low profile, low cost, mechanically robust, simple to manufacture, easy to be integrated with MIC/MMIC and others RF devices [1], [2]. In spite of its several advantages, microstrip patch antennas suffer from low gain and a narrow impedance bandwidth. To overcome its limitation of narrow impedance bandwidth and low gain, many techniques have been suggested e.g., for probe fed stacked antenna, microstrip patch antennas on electrically thick substrate, slotted patch antenna and stacked shorted patches have been proposed and investigated [3]-[5].

The size of microstrip patch antenna is basically determined by its resonance length and width. The reduction of the patch size can be achieved by using patch substrate material with very high permittivity and small substrate height [6]. But, in this case, the low radiation efficiency will reduce the antenna gain. By using two stacked patches with the walls at the edges between the two patches, one can obtain high gain [7].

Recently, aperture-coupled fed stacked patch antenna [8] have been investigated and bandwidths up to 69% have been reported, however, the major drawbacks are the level of back radiation due to the use of a resonant aperture and the surface wave excitation. The use of L-shaped or F-shaped probes have also been proposed yielding to wide impedance bandwidths [9], [10]. In [9], an L-probe fed stacked U-Slot patch antenna was proposed with a bandwidth up to 44.4% being achieved. V-slotted rectangular microstrip antenna with a stacked patch has been shown that it is able to achieve bandwidths as high as 47% [11]. However, feeding techniques in these designs especially of the probes have increased complexity of the design and fabrication process. In this paper, a novel multiple slotted patch is investigated for enhancing the impedance bandwidth and gain.

## II. ANTENNA LAYOUT AND STRUCTURE

The geometry of the proposed antenna structure is shown in Fig. 1. The proposed multiple slotted patch antenna structure is composed of rectangular patch, air substrate and a vertical probe connected to the patch. For a microstrip antenna, the patch width  $w$  and length  $l$  that support the operation at the required resonant frequency can be designed using the formulas as given in [12] as

$$w = \frac{\lambda_0}{2\sqrt{\frac{\epsilon_r + 1}{2}}} \quad (1)$$

$$l = \frac{\lambda_0}{2\sqrt{\epsilon_r}} - 2\Delta l \quad (2)$$

where

$$\Delta l = 0.412h \frac{(\epsilon_r + 0.3)\left(\frac{w}{h} + 0.264\right)}{(\epsilon_r - 0.258)\left(\frac{w}{h} + 0.8\right)} \quad (3)$$

$$\epsilon_e = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2\sqrt{1 + 12\frac{h}{w}}} \quad (4)$$

The location of the probe point example, if  $(x_1, y_1)$  for matching to 50 ohm can be determined as

$$x_1 = \frac{w}{2} \quad (5)$$

$$y_1 = \frac{l}{\pi} \cos^{-1} \frac{w\sqrt{5(\epsilon_r - 1)}}{3\epsilon_r l} \quad (6)$$

where  $0 \leq y_1 < l/2$  [13].

While for a microstrip antenna, it is found that the patch length can still be designed using formulas (1) to (6) for the required resonant frequency. However, the patch width calculated by formula (1) usually needs to be widened to ensure the excitation and matching using the probe feed.

In this design, the inverted rectangular patch, with dimension of  $0.707 \lambda_0 \times 0.354 \lambda_0$  is supported by a low dielectric superstrate with dielectric permittivity  $\epsilon_r$  and thickness  $h_1$ . An air-filled substrate with thickness  $h$  is sandwiched between the superstrate and a ground plane.

The proposed patch integrates with multi slots on the same radiating element are embedded in parallel on the radiating edge of the patch symmetrically with respect to the centerline of the patch. The multiple slots of the patch are shown in Fig. 1(a), where,  $l_1$  and  $w_1$  are the length and width of the slots. The patch is fed by a direct connected probe along the centerline at a distance  $f_p$  from the edge of the patch as shown in Fig. 1(b). Table 1 shows the optimized design parameters obtained for the proposed patch antenna. A Rogers RT 5880 Duroid™ dielectric substrate with dielectric permittivity,  $\epsilon_r$  of 2.2 and thickness,  $h_1$  of 1.5748mm has been used in this research. An Aluminum plate with dimensions of  $1.473 \lambda_0 \times 1.326 \lambda_0$  and thickness of 1 mm is used as the ground plane.

Reducing the size of the antenna is one of the key factors to miniaturize the wireless communication devices. However, reducing the antenna size will usually reduce its impedance bandwidth as well. Therefore designing a small antenna with a wide impedance bandwidth which satisfies future generation wireless application is a challenging work, especially having stable radiation patterns across the operating frequency band. In this paper, the use of probe feeding technique, multiple slotted on the patch with thick air-filled substrates provide the wide bandwidth and gain enhancement and the use of parallel slots reduce the crosspolarization. The multi slots of the proposed patch

introduce a capacitive reactance, which counteracts the inductive reactance of the probe pin and produces second resonance of the patch to achieve a wide bandwidth.

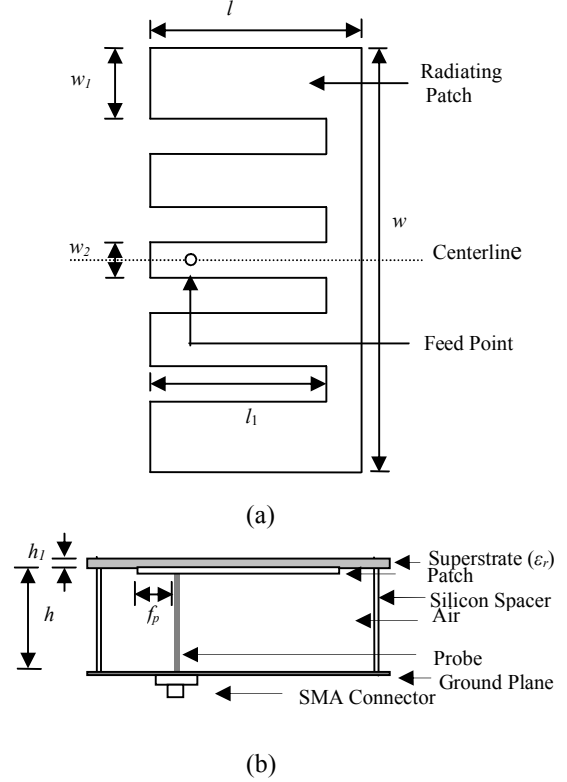


Fig. 1. Geometry of the proposed patch antenna: (a) Top view and (b) side view.

TABLE I  
PROPOSED PATCH ANTENNA DESIGN PARAMETERS IN  
MILLIMETERS (MM)

Dimensions	$w$	$l$	$w_1$	$l_1$
(mm)	96	48	14.5	42
Dimensions	$w_2$	$h$	$h_1$	$f_p$
(mm)	7	12.5	1.5748	8.5

### III. RESULTS AND DISCUSSION

A simulator “Ansoft HFSS” based on finite-element method (FEM) has been used to calculate return loss, impedance bandwidth, radiation pattern and gains. Fig. 2 shows the simulated result of the return loss of the proposed antenna. The two closely excited resonant frequencies at 2.0 GHz and at 2.35 GHz as shown in the figure gives the measure of the wideband characteristic of the patch antenna. The simulated impedance bandwidth of

27.15% from 1.91 GHz to 2.51 GHz is achieved at 10dB return loss ( $VSWR \leq 2$ ).

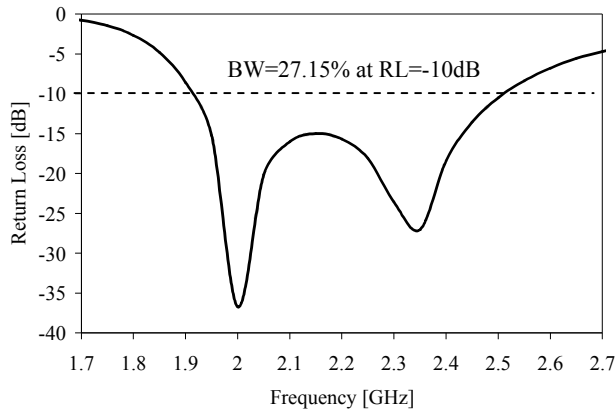


Fig. 2. Simulated return loss of the proposed patch antenna.

The simulated radiation patterns at the second resonant frequency in the  $xz$ -plane and  $yz$ -plane are plotted in Fig. 3. For the sake of brevity, only simulated radiation pattern for second resonant frequency is given in this paper.

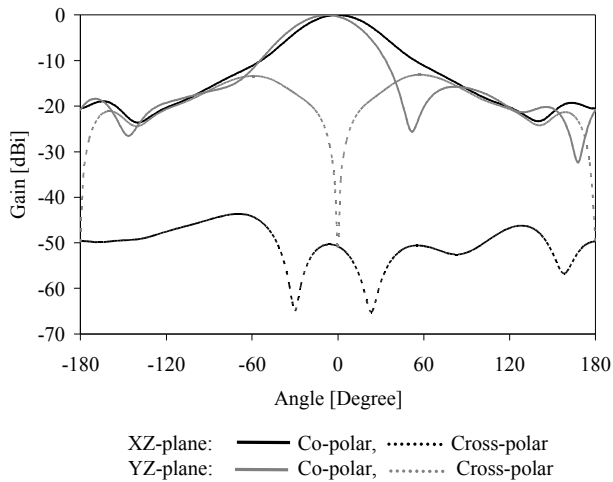


Fig. 3. Normalized radiation pattern of proposed patch antenna at resonance frequency 2.35 GHz.

As shown in Fig. 3, the designed antenna displays good broadside radiation patterns in the  $xz$ -plane and  $yz$ -plane at second resonant frequency. It can be seen that 3-dB beamwidth of  $52^\circ$  and  $46^\circ$  for  $xz$ -plane and  $yz$ -plane respectively at 2.35 GHz. The crosspolarization level is lower than about -43dB in  $xz$ -plane. The proposed antenna exhibits better crosspolarization than the design reported in [14]. Notable, the radiation characteristics of the

proposed antenna are better to those of the conventional patch antenna.

The simulated gain of the proposed patch antenna at various frequencies is shown in Fig. 4. As shown in the figure, the maximum achievable gain is 10.5 dBi at the frequency of 2.2GHz and the gain variation is 1.57dBi between the frequency ranges of 1.91GHz to 2.51GHz (i.e. a gain bandwidth of 27.15%).

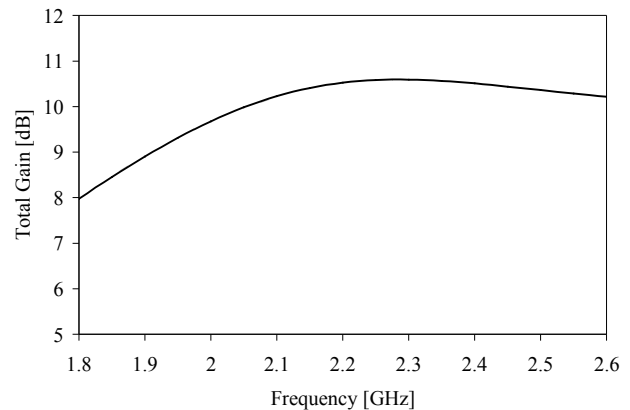


Fig. 4. Simulated gain of proposed patch antennas at different frequencies.

#### IV. CONCLUSION

A multiple slotted patch antenna for enhancing bandwidth and gain was successfully designed in this paper. Simulation results of a wideband microstrip patch antenna covering 1.91 to 2.51 GHz frequency have been presented. Techniques for microstrip broadbanding, gain enhancement, and crosspolarization reduction are applied with significant improvement in the design by employing proposed slotted patch shaped design, inverted patch, and probe feeding.

The proposed microstrip patch antenna achieves a fractional bandwidth of 27.15% (1.91 to 2.51 GHz) at 10 dB return loss. The maximum achievable gain of the antenna is 10.5 dBi. The proposed patch has a compact dimension of  $0.707 \lambda_0 \times 0.354 \lambda_0$ . The design has demonstrated that patch with multiple slots and probe fed can be used to form an antenna with the broad bandwidth of 27.15%, furthermore due to its high gain, and broad bandwidth more applications can be anticipated.

#### ACKNOWLEDGEMENT

The authors would like to thank the MOSTI Secretariat, Ministry of Science, Technology and Innovation of Malaysia, e- Science fund: 01-01-02-SF0376, and Institute

of Space Science (ANGKASA), UKM for sponsoring this work.

#### REFERENCES

- [1] S.L.S. Yang, A.A. Kishk, and K. F. Lee, "Frequency Reconfigurable U-Slot Microstrip Patch Antenna," *IEEE Antennas and Wireless Propagation Letters*, vol. 7, pp. 127-129, 2008.
- [2] S.L.S. Yang and K.M. Luk, "A Wideband L-Probes Fed Circularly-Polarized Reconfigurable Microstrip Patch Antenna," *IEEE Transactions on Antennas and Propagation*, vol. 56, no.2, pp. 581-584, February 2008.
- [3] D.M. Pozar and D.H. Schaubert, *Microstrip antennas, the analysis and design of microstrip antennas and arrays*. New York: IEEE press, 1995.
- [4] D. Sanchez-Hernandez. and L.D. Robertson, "A Survey of Broadband Microstrip Patch Antennas," *Microwave Journal*, pp. 60-84, September 1996.
- [5] K. Chang, "RF and Microwave Systems," Texas A & M University, 2000.
- [6] J.R. James, P. S. Hall, and C. Wood, *Microstrip Antenna Theory and Design*. Peter Perigrinus, London, 1981.
- [7] S. Egashira and E. Nishiyama, "Stacked microstrip antenna with wide bandwidth and high gain," *IEEE Transactions on Antennas and Propagation*, vol. 44, pp. 1533-1534, November 1996.
- [8] S.D. Targonski, R.B. Waterhouse, and D.M. Pozar, "Design of Wide-Band Aperture Stacked Patch Microstrip Antennas," *IEEE Transactions on Antennas and Propagation*, vol. 46, no. 9, pp. 1245-1251, September 1998.
- [9] B.L. Ooi and C.L. Lee, "Broadband Air-Filled Stacked U-Slot Patch Antenna," *Electronics Letters*, vol. 35, no. 7, pp. 515-517, April 1999.
- [10] B.L. Ooi, C.L. Lee, P.S. Kooi, and S.T. Chew, "A novel F-probe fed broadband patch antenna", in *IEEE Antennas and Propagation Society International Symposium*, vol. 4, pp. 474-477, July 2001.
- [11] G.Z. Rafi, and L. Shafai, "V-slotted rectangular microstrip antenna with a stacked patch," in *IEEE International Symposium on Antennas and Propagation Society*, vol. 2, pp. 264-267, June 2003.
- [12] K. Hirasawa, *Analysis, Design, and Measurement of Small and Low- Profile Antennas*. Norwood, MA: Artech House, 1991.
- [13] W.L. Stutzman and G.A. Thiele, *Antenna Theory and Design*. New York: Wiley, 1998.
- [14] M. Tariqul Islam, N. Misran, and K.J. Ng, "A 4×1 L-probe fed Inverted Hybrid E-H Microstrip Patch Antenna Array for 3G Application," *American Journal of Applied Sciences*, vol. 4, no.11, pp. 897-901, 2007.

# Analysis of Antennas for the Wireless Communication and Radar Applications

Md. Osman Goni, Abu Md. Numan-Al-Mobin, Md. Hasanur Rahman

Department of Electronics and Communication Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh

**Abstract** — This paper signifies the distributed electromagnetic simulation for different thin wired antennas or structures mounted on earth or vehicles or space shuttle for communication purposes. A lot of electromagnetic analysis softwares are available commercially. This paper presents some simulation results of different antennas using Wire-MoM based on thin-wire approximation. The method gives frequency domain results of currents and fields which is further converted into time domain for better understanding.

**Index Terms** — Antenna feeds, Antenna measurements, Antenna proximity factors, Antenna radiation patterns, Antenna transient analysis.

## I. INTRODUCTION

Electromagnetic fields can be predicted by experiment, analysis and computation. The last mentioned technique, otherwise known as numerical simulation is the most recent and rapidly advancing. The modeling of electromagnetic processes on computer can help us understand how products work and hopefully lead to better products as in the philosophy of computer-aided design (CAD).

The numerical computation for distributed electromagnetic simulation and optimization is the very cost effective than the direct experimental method.

For idealized geometry, the calculation of electromagnetic fields generated by power line conductors at each specific location is a trivial algebraic task. However, the presence of man made and natural objects, uneven profile of ground surface, and non-homogeneous conductivity of the ground at different depths affect the resulting distribution of magnetic and electric fields.

One of the earliest applications of a numerical electromagnetic (EM) analysis to the antenna design and lightning study [1] modeled a lightning channel attached to a tall structure. A time-domain code was used in that study. A similar time-domain code was also used in a study on induced voltages on a distribution line [2] over perfectly conducting ground. A recent study also employs a similar time-domain code [3] to model the wired antennas with induced current and voltage in each wire and a lightning channel. Moini *et al.* applied the time-domain approach utilizing the Thin-Wire Time-Domain code (TWTD) [2],[4].

However, more authors have applied the Numerical Electromagnetic Code (NEC-2) [5], a frequency-domain code to lightning electromagnetic pulse (LEMP) studies or lightning surge analysis. Heidler *et al.* [6] used NEC-2 to

analyze induced voltages on conductor loops illuminated by LEMP.

Cristina and Orlandi [7] employed NEC-2 to evaluate the EM field inside a building struck by lightning. Chai *et al.* [8] employed this code to study field inside a wire-array lightning protection system for a lunch vehicle when it is struck by lightning. Baba and Ishii applied the frequency-domain approach utilizing the Numerical Electromagnetic Code (NEC-2) [9],[10].

For practical purposes, one of the classes of models applied recently in lightning studies is the so-called electromagnetic models [11]. They are usually based on thin-wire antenna approximation and involve a numerical solution of Maxwell's equations using the method of moments [12] to find the current distribution along the lightning channel, from which electromagnetic fields can be computed.

## II. NUMERICAL MODEL

Different types of return-stroke current models are used to determine the return-stroke current. As a part of a general investigation into the physics of lightning, as a mechanism by which return strokes currents at ground can be determined from remotely measured electric and magnetic fields, and hence by which either currents of individual strokes or statistical distributions of the stroke currents can be obtained, and as a mechanism for calculating realistic fields to be used in "coupling" calculations such as to determine the lightning-induced voltages appearing on electric utility or telecommunication lines when lightning occurs near those lines.

1) The first defined class of models, gas dynamic or "physical" models, is primarily concerned with the radial evolution of a short segment of the lightning channel and its associated shockwave. Principal model's outputs include temperature, pressure, and mass density as a function of the radial coordinate and time.

2) Electromagnetic Field Theory Approach models, or second class models, are usually based on the so-called lossy thin wire antenna approximation of the lightning channel.

3) The third class of models is the distributed circuit models, also called RLC transmission line models. These models are used to determine the channel current versus

time and height and can therefore also be used for the computation of remote electric and magnetic fields.

4) The last class is the engineering models in which a spatial and temporal distribution of the channel current (or the channel line charge density) is specified based on such observed lightning return-stroke characteristics as current at the channel base, the speed of the upward propagating wave front, and the channel luminosity profile.

In this paper Field Theory Approach is adopted for the evaluation of the effects. For numerical solution computer program “Wire-MoM” is used which is based on Method of Moments (MoM). Here the computed results are obtained in frequency domain for a step current input.

In “Wire-MoM” two methods are used for the solution of the numerical analysis. They are:

1) LU decomposition.

In linear algebra, the LU decomposition is a matrix decomposition which writes a matrix as the product of a lower and upper triangular matrix. The product sometimes includes a permutation matrix as well. This decomposition is used in numerical analysis to solve systems of linear equations or calculate the determinant.

2) Conjugate gradient method.

We have used here the LU decomposition method.

### III. NUMERICAL ELECTROMAGNETIC SIMULATION

#### A. Simulation-1: Dipole Antenna

For the dipole antenna, fed with the ideal 5V voltage source at the frequency range 200MHz – 1200MHz. Finding of the current distribution along the antenna at the frequency at which the antenna would radiate most efficiently. Plotting of 3D antenna radiation pattern at that frequency and the graphical representation on the calculation of the power budget. Here the structure frequency limit is approximately 660 MHz.

TABLE I  
SIMULATION PARAMETERS FOR 1M DIPOLE ANTENNA

Structure	Excitation	Output
No. of Wires=1 No. of Connections=0 No. of Basis id.=21 Ground plane at z=0 No. of Impedance elements= 0	No. of Frequencies=15 No. of Voltage sources= 1 No. of Current sources= 0 Plane wave excitation= Yes	No. of Far-field directions=7381 No. of Near-field points=100 Saving Current for No. of wires=1 Saving S-parameters=Yes

Here the 1m dipole antenna made of a perfect conductor with the radius of 1mm was represented by a vertical cylindrical conductor. The model simulating the current and frequency response measurement of the single vertical conductor representing the dipole is shown in Fig. 1.

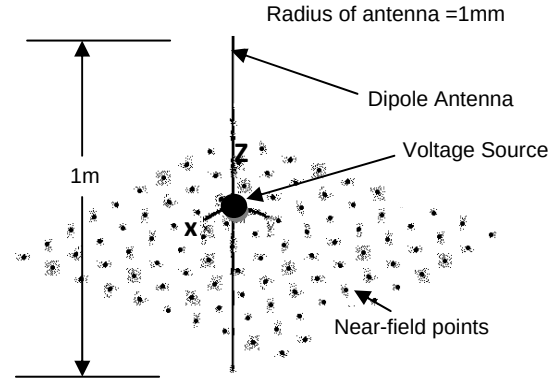


Fig. 1. Wired-Structure of the dipole antenna with the voltage source at the middle of the wire.

The simulation was run over the frequency range from 200 MHz to 1200 MHz with a step of 71.429 MHz. Computation is done in 2 minutes and 20 seconds.

#### 1. Result for Simulation-1

The frequency domain plot of  $I(f)$  is done under the 11th basis function corresponding to the feeding point on the antenna. The total current is computed in absolute value of the current. The plot of total current in time domain is shown in Fig. 2. Now the current on each segment of the wire is available at the frequency 258.56 MHz.

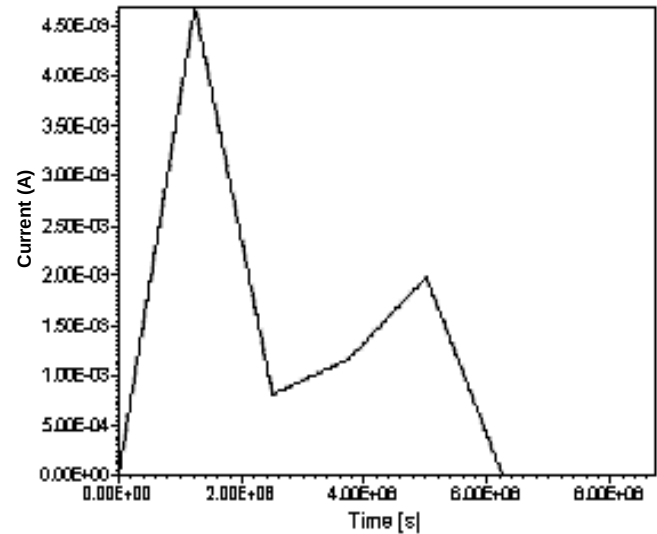


Fig. 2. Plot of current in time domain for dipole antenna.

The 3D radiation pattern for the dipole antenna is shown in Fig. 3. Here the dipole antenna is approximated with the thin wire.

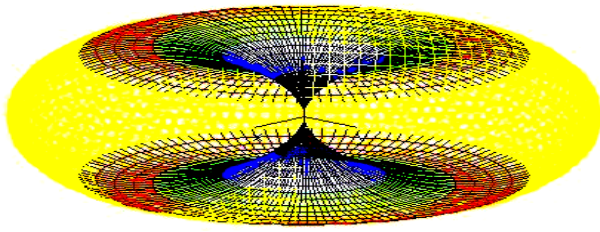


Fig. 3. 3D Radiation pattern of the dipole antenna

## 2. Power Budget Computation for Dipole Antenna

The graphical representation for the input power, radiated power and dissipated power versus frequency is shown in Fig. 4(a). The input power and radiated power have the same magnitudes. And the plotting of the efficiency is shown in Fig. 4(b).

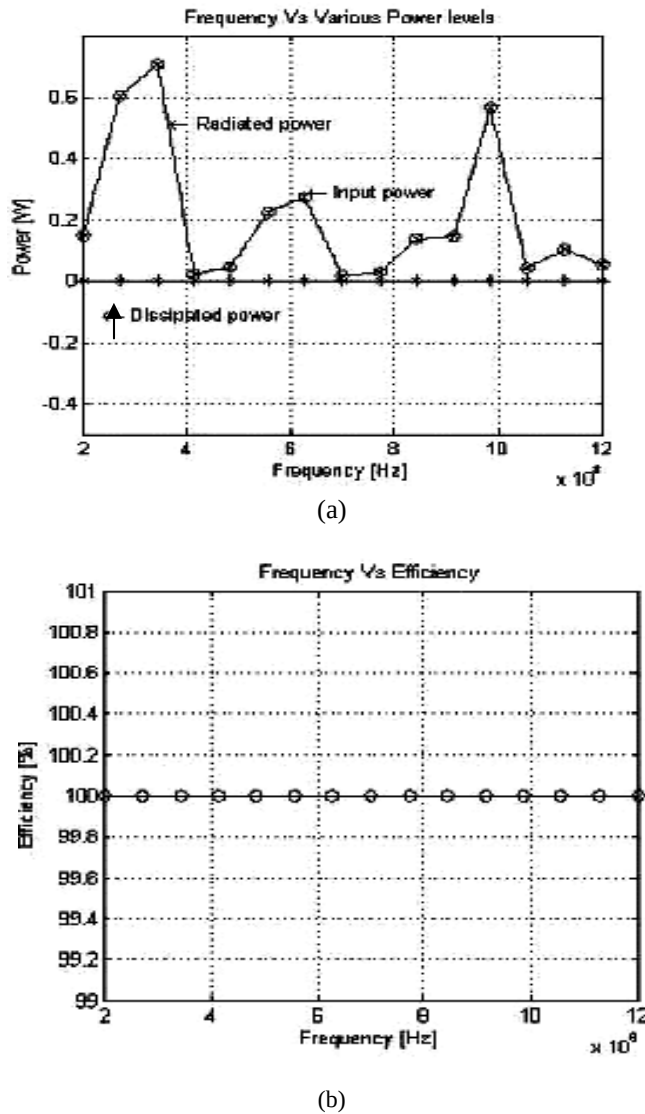


Fig. 4. Dipole antenna analysis (a) frequency vs power, (b) frequency vs efficiency.

## B. Simulation-2: Grounded Monopole Antenna

For the grounded monopole antenna, fed with the ideal 5V voltage source at the frequency range 200 MHz – 1200 MHz. Plotting of the total current curve in the time domain. Finding of the current distribution along the antenna at the frequency at which the antenna would radiate most efficiently. Here the structure frequency limit is approximately 480 MHz.

TABLE II  
SIMULATION PARAMETERS FOR GROUNDED MONOPOLE ANTENNA

Structure	Excitation	Output
No. of Wires=1 No. of Connections=1 No. of Basis id.=16 Ground plane at z=0 PEC No. of impedance elements =1	No. of Frequencies=30 No. of Voltage sources= 1 No. of Current sources= 0 Plane wave excitation=Yes	No. of Far-field directions=3751 No. of Near-field points=36 Saving Current for No. of wires=1 Saving S-parameters= Yes

Here the 1m long grounded monopole antenna made of a perfect conductor with the radius of 1mm was represented by a vertical cylindrical conductor.

The model simulating the current and frequency response measurement of the single vertical conductor representing the grounded monopole is shown in Fig. 5.

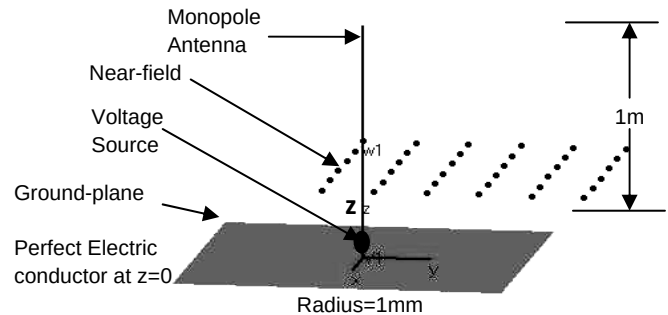


Fig. 5. Grounded monopole antenna with the voltage source at the bottom of the wire.

The simulation was run over the frequency range from 50MHz to 1500MHz with a step of 50 MHz. Computation is done in 2 minutes and 23 seconds.

## Result for Grounded Monopole Antenna

The frequency domain plot of current is done under the 11th basis function corresponding to the feeding point on the antenna. The total current is computed in absolute value of the current. The plot of total current in time domain is shown in Fig. 6. This current on each segment of the wire is available at the frequency 153.1 MHz.



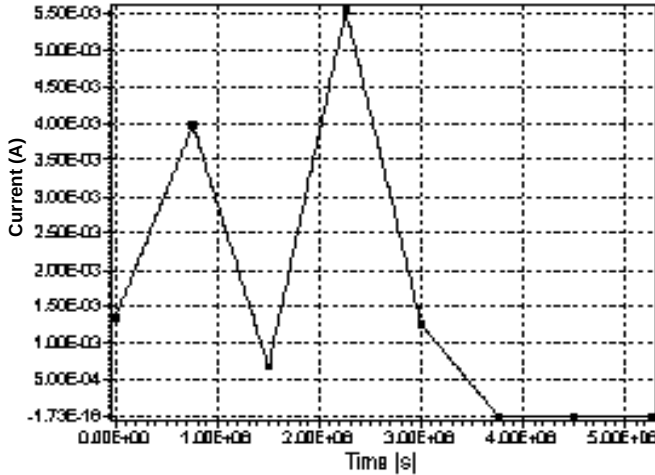


Fig. 6. Plot of current in time domain for grounded monopole antenna.

### C. Simulation-3: Monopole Antenna Mounted on a Car at Low Excitation

For the monopole antenna mounted on a car, fed with the ideal 5V voltage source at the frequency range 1MHz – 150MHz. Plotting of the total current curve in the time domain. Finding of the current distribution along the antenna at the frequency at which the antenna would radiate most efficiently and also the graphical representation on the calculation of the power budget. Here the structure frequency limit is approximately 10 MHz.

TABLE III  
SIMULATION PARAMETERS FOR 2M MONOPOLE ANTENNA MOUNTED ON A CAR AT LOW EXCITATION

Structure	Excitation	Output
No. of Wires=80	No. of Frequencies=8	No. of Far-field directions =3751
No. of Connections=113	No. of Voltage sources= 1	No. of Near-field points=144
No. of Basis id.=513	No. of Current sources= 0	Saving Current for No. of wires=80
Ground plane at z=0 PEC	Plane wave excitation= Yes	Saving S-parameters = Yes
No of Impedance elements=1		

Here 2m long monopole antenna made of a perfect conductor with the radius of 0.005mm was represented by a vertical cylindrical conductor. The model simulating the current and frequency response measurement of the 2m long monopole antenna mounted on a car at low excitation is shown in Fig. 7.

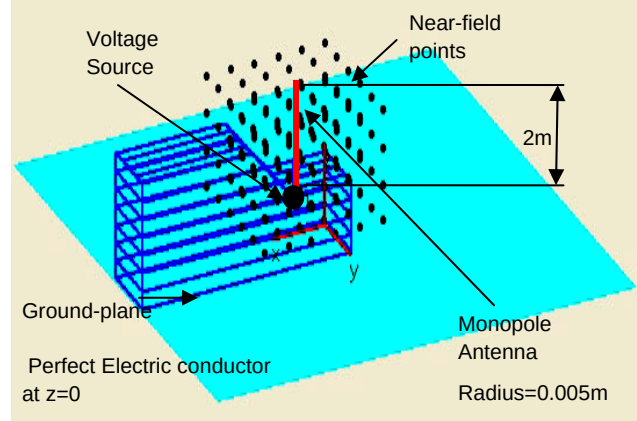


Fig. 7. Wired-Structure of the monopole antenna mounted on vehicle fed with the ideal 5V voltage source.

The simulation was run over the frequency range from 1 MHz to 150 MHz with a step of 50 MHz. Computation is done in 2 minutes and 43 seconds. The simulation was run over the frequency range from 1 MHz to 150 MHz with a step of 50 MHz. Computation is done in 22 minutes and 11 seconds.

#### 1. Result for Simulation-3

The frequency domain plot current under the 640th basis function corresponding to the feeding point on the antenna. The total current is computed in absolute value of the current. The plot of total current in time domain is shown in Fig. 8.

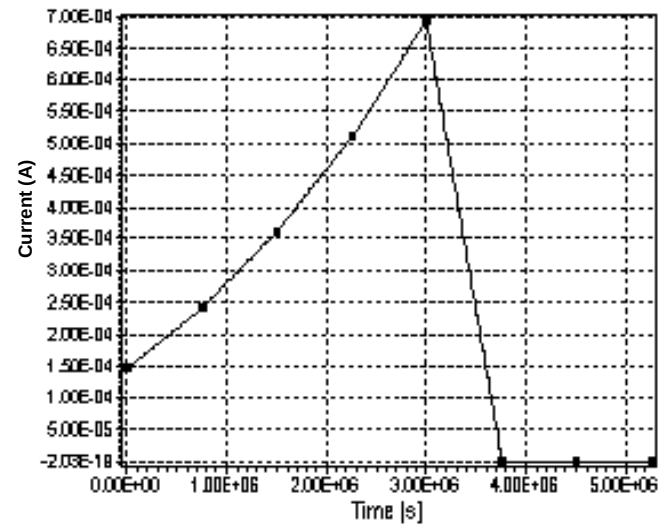


Fig. 8. Plot of current in time domain for low voltage excitation.

## 2. Power Budget Computation for Monopole Antenna Mounted on a Car

The computation of input power, radiated power and dissipated power versus frequency is shown in Fig. 9(a) and the plotting of efficiency in frequency domain is shown in Fig. 9(b).

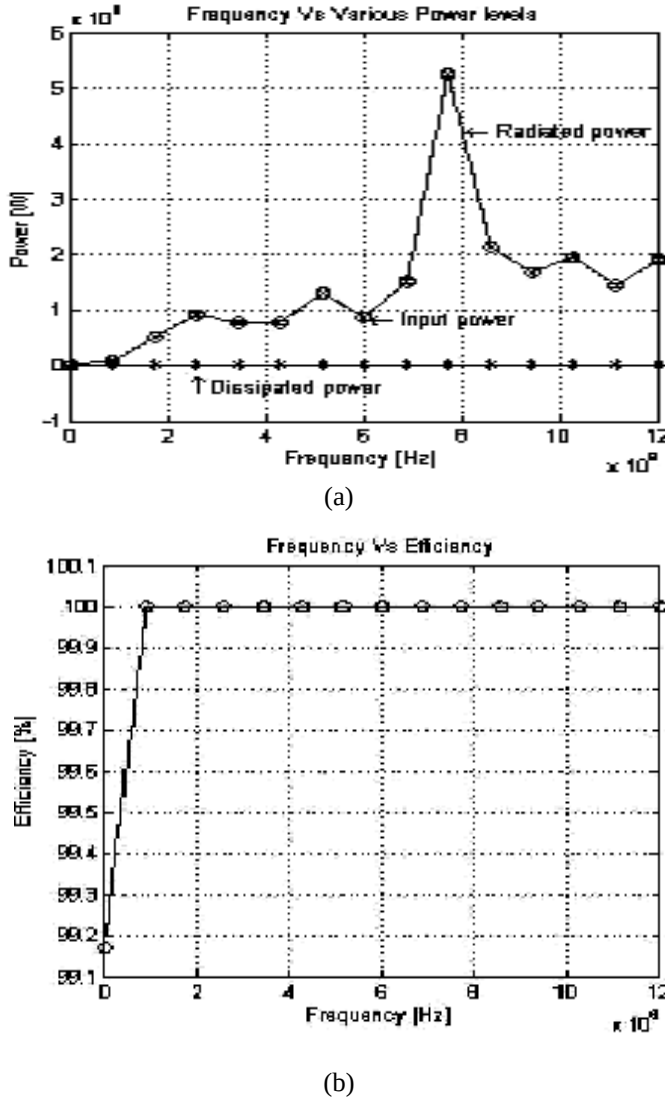


Fig. 9. Monopole antenna analysis (a) power vs frequency, and (b) efficiency vs frequency.

### D. Simulation-4: Patch Antenna Mounted on a Rocket at Low Excitation

For the patch antenna mounted on a rocket, fed with the ideal 5V voltage source at the frequency range 1 MHz – 300 MHz. Plotting of the total current curve in the frequency domain. Finding of the current distribution along the antenna at the frequency at which the antenna would radiate most efficiently and also the graphical representation on the calculation of the power budget. Here the structure frequency limit is approximately 89 MHz.

TABLE IV  
SIMULATION PARAMETERS FOR PATCH ANTENNA MOUNTED ON A ROCKET AT LOW EXCITATION

Structure	Excitation	Output
No. of Wires=368	No. of Frequencies= 8	No. of Far-field directions=3751
No. of Connections=521	No. of Voltage sources= 1	No. of Near-field points= 27
No. of Basis id.=889	No. of Current sources= 0	Saving Current for No. of wires=368
Ground plane at z=0	Plane wave excitation=Yes	Saving S-parameters= Yes
No. of Impedance elements=0		

Here 0.237m wide and 0.67m long patch antenna was represented by the horizontally aligned cylindrical conductors. And the height of Rocket is taken as 3m long. The wire radius is 0.001m and antenna thickness is 0.002m which is shown in Fig. 10.

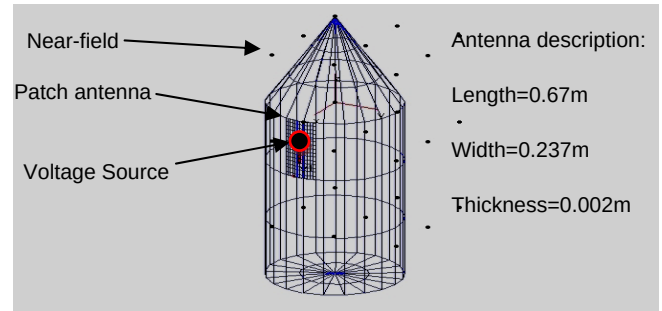


Fig. 10. Wired-Structure of the patch antenna mounted on rocket fed with the ideal 5V voltage source.

The simulation was run over the frequency range from 1 MHz to 300 MHz with a step of 42.714 MHz. Computation is done in 6 minutes and 15 seconds.

### 1. Result for Simulation-4

The frequency domain plot current under the basis function corresponding to the feeding point on the antenna. Here the total current is computed in the absolute value of the current which is shown in Fig. 11.

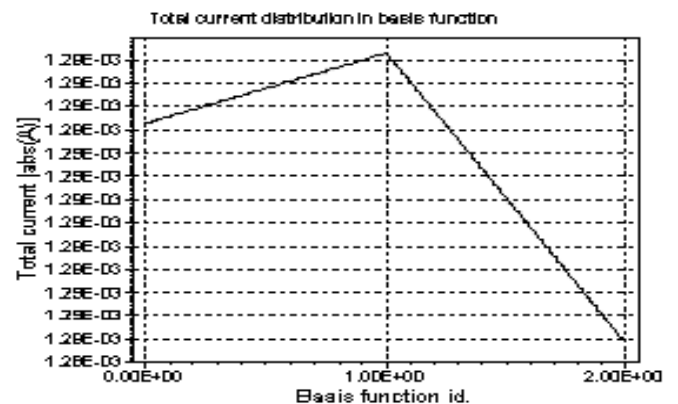


Fig. 11. Total current distribution along the wire.

## 2. Power Budget Computation for Patch Antenna Mounted on a Rocket

The computation of the input power, radiated power and the dissipated power versus frequency is shown in Fig. 12(a) and the plotting of efficiency is shown in Fig. 12(b).

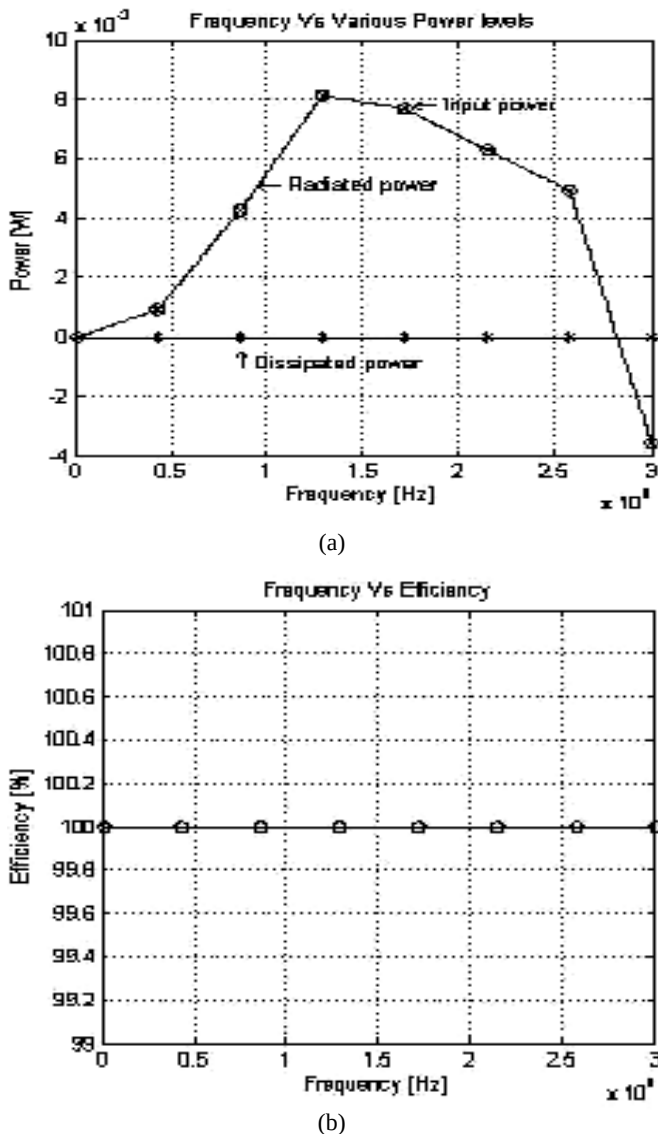


Fig. 12. Patch antenna analysis (a) frequency vs power, (b) frequency vs efficiency.

## IV. CONCLUSION

During the past one year, we have gained a great deal of experience and insight into current and voltage calculations on different antennas and structures, the electromagnetic field calculations and fields of lightning return stroke current modeling. The experimental evaluation for the lightning transient study is much expensive than the numerical solutions. The authors of this paper carried out the numerical electromagnetic analysis of various antennas

used in wireless communication and radar. As the MoM automatically computes the electromagnetic coupling between the conductors so no coupling model was required and the solution by MoM is more accurate than the time domain EMTP solution. The analysis shows that the amount of current and voltage induced on the antenna elements are also of large magnitude for the degradation and disruption of the transmission information. The result and analysis also provides that a large amount of current and voltage are generated by the Lightning Over-Voltage which implies the installation of surge arrester of lightning current arrester for the protection of the base station.

## REFERENCES

- [1] S. Podgorski and J. A. Landt, "Three dimensional time domain modeling of lightning," *IEEE Trans. Power Delivery*, vol. 2, pp. 931–938, July 1987.
- [2] R. Moini, B. Kordi, and M. Abedi, "Evaluation of LEMP effects on complex wire structures located above a perfectly conducting ground using electric field integral equation in time domain," *IEEE Trans. Electromagn. Compat.*, vol. 40, pp. 154–162, May 1998.
- [3] E. K. Miller, A. J. Poggio, and G. J. Burke, "An integro-differential technique for time-domain analysis of thin wire structures," *J. Comput. Phys.*, vol. 12, pp. 24–48, 1973.
- [4] R. Moini, B. Kordi, G. Z. Rafi, and V. A. Rakov, "A new lightning return stroke model based on antenna theory," *J. Geophys. Res.*, vol. 105, no. D24, pp. 29 693–29 702, 2000.
- [5] G. J. Burke and A. J. Poggio, "Numerical electromagnetic code (NEC)—method of moments," in *Technical Document 116*. San Diego: Naval Ocean Systems Center, 1980.
- [6] F. Heidler and K. Muller, "LEMP calculations with the traveling current source model," in *Proc. Int. Conf. Lightning Static Electricity*, Bath, U.K., Sept. 1989.
- [7] S. Cristina and A. Orlandi, "Lightning channel's influence on currents and electromagnetic fields in a building struck by lightning," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, Washington, D.C., Aug. 1990, pp. 338–342.
- [8] J. C. Chai, H. A. Heritage, and R. Briet, "Electromagnetic effects of the four-tower supported catenary wires array lightning protection system," in *Proc. 16th Int. Aerospace Ground Conf. Lightning Static Electricity*, Mannheim, CA, May 1994, pp. 377–386.
- [9] M. Ishii and Y. Baba, "Numerical electromagnetic field analysis of tower surge response," *IEEE Trans. Power Delivery*, vol. 12, pp. 483–488, Jan. 1997.
- [10] Y. Baba and M. Ishii, "Numerical electromagnetic field analysis on lightning surge response of tower with shield wire," *IEEE Trans. Power Delivery*, vol. 15, pp. 1010–1016, Apr. 2000.
- [11] V.A. Rakov, "Transient response of a tall object to lightning," *IEEE Trans. Electromagn. Compat.*, vol. 43, no. 4, pp. 654–661, Nov. 2001.
- [12] R. F. Harrington, *Field Computation by Moment Methods*, New York: McMillan, 1968.

# Amplitude Banded Technique and Parallel Structure for Godard and Sato Algorithms of Blind Channel Equalization

Muhammad Lutfur Rahman Khan, Mohammed H. Wondimagegnehu and Tetsuya Shimamura

Graduate School of Science and Engineering, Saitama University  
255 Shimo-okubo, Sakura-ku, Saitama 338-8570, Japan.  
Email: khan@sie.ics.saitama-u.ac.jp

**Abstract**—The least-mean-squares (LMS) algorithm which updates the filter coefficients by gradient-based method is the most popular adaptive filtering one. In this paper we propose a novel amplitude banded (AB) technique with LMS on Godard (ABGodard) and Sato (ABSato) algorithms for the equalization of communication channels. The non-linear properties of the AB technique with LMS algorithm are inherited into the ABGodard and ABSato algorithms, resulting in an improvement of equalization performance in favor of the ABSato algorithm. Mean square error (MSE) as well as bit error rate (BER) are investigated on a simple communication channel model. Observations on simulations show that ABGodard and ABSato algorithms provide better performance than the standard Godard and Sato algorithms, respectively. The parallel structure of the Sato and ABSato algorithms provides a further performance improvement. This is also observed from MSE and BER performances.

**Index Terms**—Least-Mean-Square, Non-linear Adaptive Algorithm, ABGodard Algorithm, ABSato Algorithm, Blind Equalization.

## I. INTRODUCTION

The physical channel introduce a distortion to the transmitted signal. To recover the original signal, channel equalization plays an important rule in digital communication systems. To reduce or ideally to eliminate completely the inter symbol interference (ISI), adaptive equalization [1] is required. Although conventional equalization techniques rely on training sequence based equalization, they suffer from the trade-off between the sequence length and the capacity of the link. To avoid this problem and when the training sequence is not available, blind equalization technique [2] can be used. Thus the blind channel equalization is the great deal of attention for its importance in digital communication systems. The Godard [3] and Sato [4] adaptive algorithms are widely used blind algorithms for equalization of a channel. They are commonly derived based on measuring the output of the channel in case of lacking explicit knowledge of its input.

Blind adaptive equalizers are often composed of two distinct sections: (i) an adaptive filter adapted by linear adaptive algorithm (ii) followed by a non-linear estimator to improve the filter outputs, the improved output of the filter is taken to improve the estimator output in the adaptation process at the next iteration.

In the adaptation process, a linear adaptive LMS algorithm is used for the Godard and Sato equalizers. Recently, Shimamura et.al [5] derived a new non-linear adaptive algorithm, called amplitude banded LMS (ABLMS) algorithm. The adaptation of the ABLMS algorithm considers the amplitude information of the channel output to select the coefficients of the equalizer as non-linear switching pattern. The ABLMS algorithm exhibits better performance than the conventional LMS algorithm. In this paper, we set out to apply the AB technique [6] in blind adaptive algorithm to obtain better performance. Actually, we propose the AB version of the Godard and Sato algorithms. In the proposed ABGodard and ABSato algorithms, the amplitude information is deployed to select the coefficients to be updated. Based on the amplitude level of the received sequence, the equalizer coefficients are (for each iteration) selected from the elements of the coefficient matrix, and then updated. Since the AB technique itself provides the capability of non-linear classification [7] and the increase of division number of the AB technique enhances its non-linearity, thus it is expected that the proposed ABGodard and ABSato algorithms provide better performance than the Godard and Sato algorithms, respectively and the increase of division number provides a further improved performance.

## II. CHANNEL MODEL AND BLIND EQUALIZER

In this paper, it is assumed that the channel model is given by

$$u(k) = \sum_{i=0}^L h_i x(k-i) + v(k) \quad (1)$$

where  $h_0, h_1, \dots, h_L$  are the channel coefficients,  $u(k)$  is the transmitted sequence, and  $v(k)$  is a white Gaussian noise uncorrelated with  $x(k)$ . The channel output  $u(k)$  becomes the input for the equalizer. Figure 1 shows a block diagram of the blind equalizer [8]. In Figure 1,  $y(k)$  is the transversal filter output to the equalizer input  $u(k)$ .  $y(k)$  plays a key rule in the cost function of blind adaptation. For the Sato algorithm,  $y(k)$  is applied to estimate the transmitted sequence  $x(k)$  as  $\hat{x}(k)$ . For most of the blind algorithms, the LMS algorithm can be commonly utilized

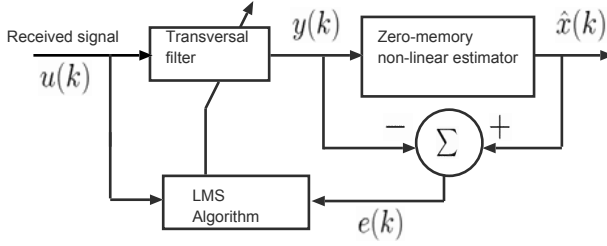


Fig. 1. Block diagram of the blind equalizer

to update the coefficient vector with the information of the estimation error,  $e(k)$ .

### III. AB BLIND ALGORITHM

The two proposed algorithms, ABGodard and ABSato, are described in this section.

#### A. ABGodard Algorithm

In implementing the ABGodard algorithm, the cost function minimization is considered as the same as that of the Godard algorithm. The cost function of the ABGodard algorithm is

$$J(k) = E[|y(k)|^p - R_p]^2 \quad (2)$$

where  $y(k)$  is the transversal filter output,  $p$  is a positive integer and  $R_p$  is a positive real constant defined by

$$R_p = \frac{E[|x(k)|^{2p}]}{E[|x(k)|^p]} \quad (3)$$

where  $E$  denotes expectation. The error signal can be calculated as

$$e(k) = y(k) |y(k)|^{p-2} (R_p - |y(k)|^p) \quad (4)$$

For the adaptation of the ABGodard algorithm, a  $Q \times M$  coefficient matrix  $\mathbf{Z}_{\text{abg}}(k)$  is considered, the elements of which are given by  $e_{i,j}(k)$ ,  $i = 1, 2, \dots, Q$ ,  $j = 1, 2, \dots, M$ .  $\mathbf{Z}_{\text{abg}}(k)$  is initialized at  $k = 0$  and all the elements are set to zero at initial condition. In the adaptation process of the algorithm, the elements of  $\mathbf{Z}_{\text{abg}}(k)$  are updated as the pattern of switching of the elements to be updated. Among the  $Q \times M$  elements of  $\mathbf{Z}_{\text{abg}}(k)$ , only  $M$  elements,  $e_{r(s),s}(k)$ ,  $s = 1, 2, \dots, M$  are selected for each iteration and a coefficient vector,  $\mathbf{w}_{\text{abg}}(k) = (e_{r(1)1}(k), e_{r(2)2}(k), \dots, e_{r(M)M}(k))^T$ , where  $r(s)$  is an integer and determined on the basis of the amplitude level of each element  $u_{k-s+1}$  of the input vector  $\mathbf{u}(k)$ , is formed for  $s = 1, 2, \dots, M$  as follows:

- if  $|u_{k-s+1}| \leq A_{\text{max}}/Q$ , then  $r(s) = 1$ .
- if  $A_{\text{max}}/Q < |u_{k-s+1}| \leq 2A_{\text{max}}/Q$ , then  $r(s) = 2$ .
- if  $2A_{\text{max}}/Q < |u_{k-s+1}| \leq 3A_{\text{max}}/Q$ , then  $r(s) = 3$ .

• . . .

- . . .
- if  $(Q-1)A_{\text{max}}/Q < |u_{k-s+1}|$ , then  $r(s) = Q$ .

Here,  $A_{\text{max}}$  denotes the maximum amplitude of the channel output and  $Q$  corresponds to the number of divisions used to partition the amplitude of the channel output.  $A_{\text{max}}$  should be measured from the received sequence before the equalizer is implemented. Accurate estimation of  $A_{\text{max}}$  is desired, but slightly inaccurate estimation may be also acceptable. This is because the range of the amplitude corresponding to  $r(s) = Q$  is not severely restricted, and occurs with the lowest probability compared with the other range cases.

As an example, consider that the equalizer length is  $M=5$  and the input vector is given by

$$\mathbf{u}(k) = [0.35 \quad -1.7 \quad 0.70 \quad 0.55 \quad 0.15]^T$$

If we consider the division number  $Q=4$ , then the 4 by 5 coefficient matrix is prepared as follows

$$\mathbf{Z}_{\text{abg}}(k) = \begin{bmatrix} e_{11}(k) & e_{12}(k) & e_{13}(k) & e_{14}(k) & e_{15}(k) \\ e_{21}(k) & e_{22}(k) & e_{23}(k) & e_{24}(k) & e_{25}(k) \\ e_{31}(k) & e_{32}(k) & e_{33}(k) & e_{34}(k) & e_{35}(k) \\ e_{41}(k) & e_{42}(k) & e_{43}(k) & e_{44}(k) & e_{45}(k) \end{bmatrix}$$

The value of  $A_{\text{max}}$  is assumed to be 0.8. Since each element of  $\mathbf{u}(k)$  produces  $(r(1) = 2, r(2) = 1, r(3) = 4, r(4) = 3, r(5) = 1)$ , the coefficient vector  $\mathbf{w}_{\text{abg}}(k)$  becomes  $\mathbf{w}_{\text{abg}}(k) = [e_{21}(k), e_{12}(k), e_{43}(k), e_{34}(k), e_{15}(k)]^T$ . This vector is updated by the Godard algorithm, and then the updated coefficients  $e_{21}(k+1), e_{12}(k+1), e_{43}(k+1), e_{34}(k+1), e_{15}(k+1)$  are inserted into the coefficient matrix  $\mathbf{Z}_{\text{abg}}(k+1)$ . For the next iteration, a coefficient vector is again prepared based on the elements of the input and then updated by the Godard algorithm. In such a way all the elements of  $\mathbf{Z}_{\text{abg}}(k)$  are updated for all input data. For selection of the filter coefficients in such a way at each iteration, the filter provides better output, which helps to estimate more accurate desired signal than that at the previous iteration. At the  $k^{\text{th}}$  iteration, the adaptive equations for the ABGodard algorithm are given by

$$y(k) = \mathbf{u}(k)^T \mathbf{w}_{\text{abg}}(k) \quad (5)$$

$$e(k) = y(k) |y(k)|^{p-2} (R_p - |y(k)|^p) \quad (6)$$

$$\mathbf{w}_{\text{abg}}(k+1) = \mathbf{w}_{\text{abg}}(k) + \mu e(k) \mathbf{u}(k) \quad (7)$$

where  $\mu$  is the step size and  $\mathbf{w}_{\text{abg}}(k+1)$  is the coefficient vector. From the definition of the cost function  $J(k)$  in (2) and (3), we see that the equalizer adaptation according to the ABGodard algorithm does not require carrier phase recovery. The algorithm therefore tends to converge slowly.

#### B. ABSato Algorithm

In implementing the ABSato algorithm, the cost function minimization is considered as the same as that of the

Sato algorithm. That is, the ABSato algorithm consists of minimizing a nonconvex cost function

$$J(k) = E[(\hat{x}(k) - y(k))^2] \quad (8)$$

where  $y(k)$  is the transversal filter output and  $\hat{x}(k)$  is an estimate of the transmitted sequence.  $\hat{x}(k)$  is given by

$$\hat{x}(k) = \gamma \text{sgn}[y(k)] \quad (9)$$

where the function  $\text{sgn}()$  is the signum function which returns the sign of the argument.  $\text{sgn}[y(k)]$  is 1 if  $y(k)$  is positive, 0 if  $y(k)$  is zero and (-1) if  $y(k)$  is negative. The constant  $\gamma$  sets the gain of the equalizer, which is defined by

$$\gamma = \frac{E[x^2(k)]}{E[|x(k)|]} \quad (10)$$

and the adaptation is taken in the same way as the ABGodard algorithm.

#### IV. PERFORMANCE EVALUATION

The performances were experimented on a channel model whose transfer function is given by

$$H(z) = 1.0 + 0.5z^{-1}. \quad (11)$$

Figure 2 shows the mean square error (MSE) conver-

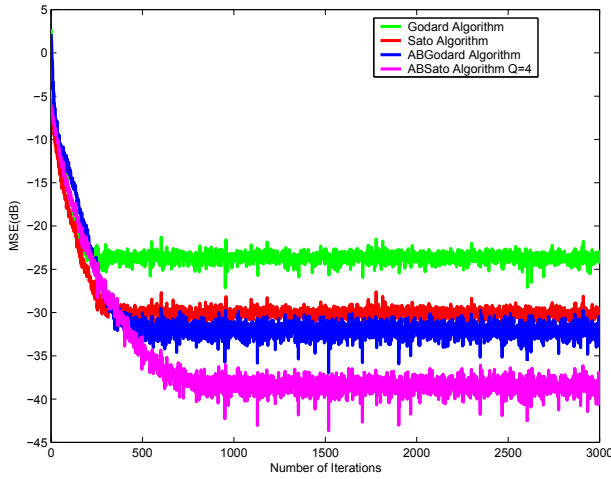


Fig. 2. MSE performance of Godard, Sato, ABGodard and ABSato on channel (11) with SNR=40dB

gence plots of the Godard, Sato, proposed ABGodard and proposed ABSato algorithms on channel (11) with SNR=40dB, step size  $\mu=0.025$ , numbers of divisions  $Q=4$ , filter order  $M=4$  and number of input data  $N=3000$ . Figure 2 indicates that commonly the AB versions provide better performance, while the ABSato algorithm results in a smaller MSE level. According to this result, we decided to select the ABSato algorithm and investigated the performance of the ABSato algorithm furthermore. Figure 3 illustrates the BER performance of the Sato and

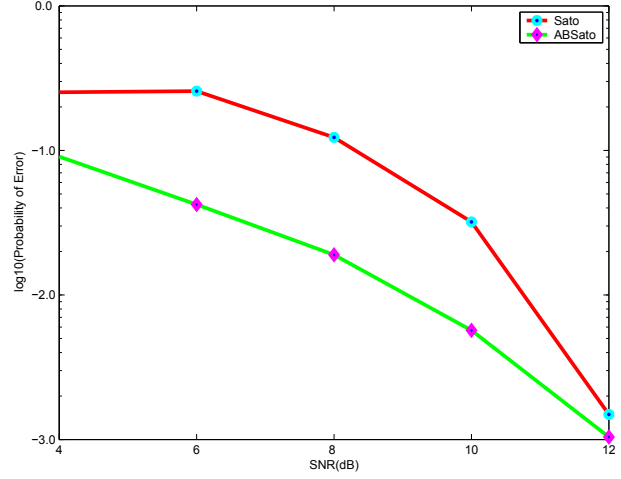


Fig. 3. BER performance of Sato and ABSato on channel (11)

ABSato algorithms on channel (11), where the step size  $\mu=0.031$ , number of division  $Q=2$ , number of input data  $N=100000$  and filter order  $M=6$  were used. Figure 3 also suggests that the ABSato algorithm is better than the Sato algorithm even for BER performance. Figure 4 shows the

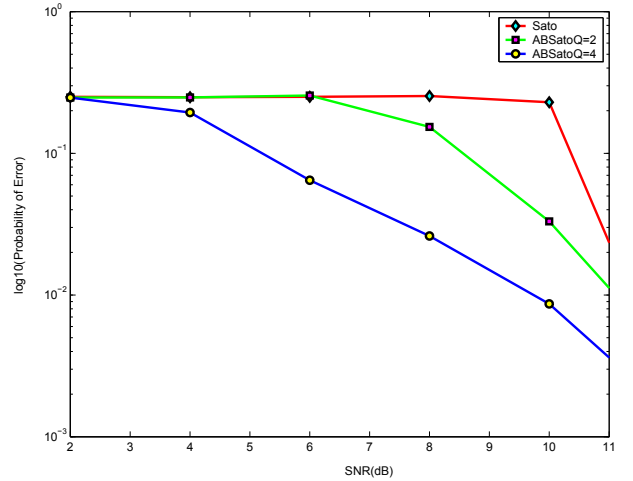


Fig. 4. BER performance of Sato and ABSato on channel (11) with division number  $Q=2,4$

BER performance of the Sato and ABSato algorithms on the same channel based on division numbers  $Q=2$  and  $Q=4$ , where the step size  $\mu=0.031$ , number of input data  $N=1000000$  and filter order  $M=6$  were used. Figure 4 suggests that as the division number  $Q$  is increased, the performance of the ABSato algorithm is improved.

#### V. FILTER STRUCTURE

To improve the performance of the ABSato algorithm, another filter structure is considered in this section.

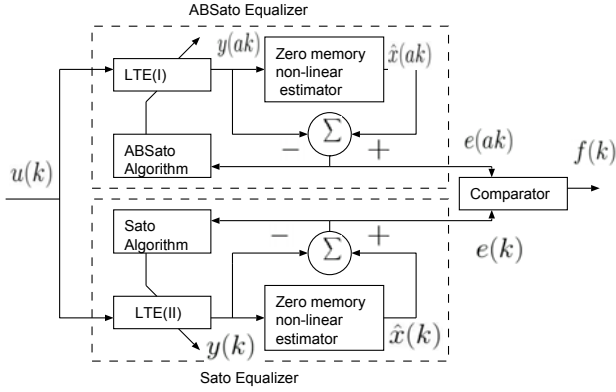


Fig. 5. Parallel Sato-ABSato equalizer

#### A. Parallel Structure

Figure 5 depicts the parallel structure of the Sato and ABSato equalizers. In the parallel structure, two linear transversal equalizers; LTE(I) and LTE(II), are updated by the ABSato and Sato algorithms with  $e(ak)$  and  $e(k)$ , respectively. The comparator provides  $f(k)=e(ak)$  if  $(e(ak))^2 \leq (e(k))^2$  otherwise  $f(k)=e(k)$ . The parallel equalizer provides output  $y(ak)$  when  $f(k)=e(ak)$  and  $y(k)$  when  $f(k) = e(k)$ .

#### B. Performance Evaluation

The performances of the parallel Sato-ABSato equalizer were investigated on the same channel. Figure 6 shows the MSE convergence plots of the Sato, ABSato and parallel Sato-ABSato equalizers with the SNR=40dB,  $M=8$ , step size  $\mu=0.025$  and  $Q=4$ . Figure 6 clarifies that the parallel structure of the Sato and ABSato algorithms provides better performance than non-parallel structure algorithms.

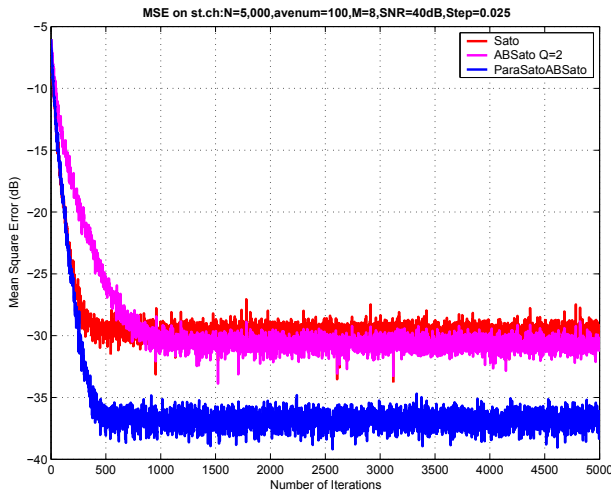


Fig. 6. MSE performance of Sato, ABSato and Parallel Sato-ABSato on channel (11) with SNR=40dB

Figure 7 shows the BER performance of the Sato, ABSato

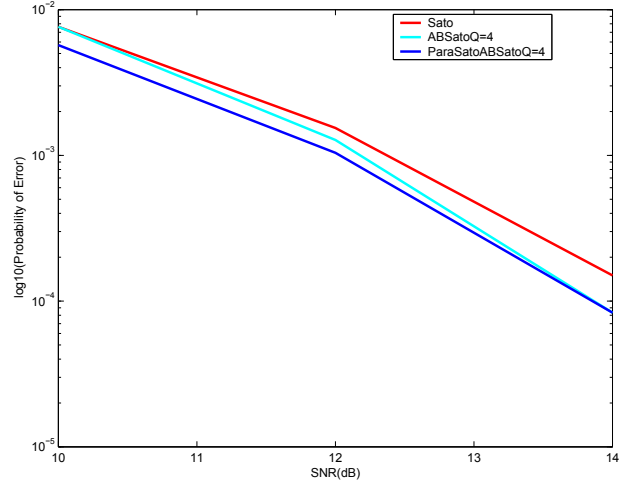


Fig. 7. BER performance of Sato, ABSato and Parallel Sato-ABSato on channel (11)

and parallel Sato-ABSato algorithms with  $M=8$ , step size  $\mu=0.025$ ,  $Q=4$  and  $N=100000$ . Figure 7 suggests that even for BER, the parallel combination of the Sato and ABSato algorithms improves the performance. From figure 6 and 7 we can confirm that the parallel structure enhances the performance of the ABSato algorithm with the support of the Sato algorithm. Additionally we checked the performance of the parallel equalizer with the Godard and ABGodard algorithms, which is constructed in same way as in Figure 5. The performance of the parallel Godard-ABGodard equalizer was, however, worse than that of the parallel Sato-ABSato equalizer.

#### VI. CONCLUSION

In this paper, ABGodard and ABSato algorithms have been proposed for blind channel equalization. Simulation results have demonstrated that the ABGodard and ABSato algorithms perform better than the Godard and Sato algorithms, respectively, for simple communication channel equalization. The ABSato algorithm behaves more accurately than the ABGodard algorithm, and the increased division number and the use of parallel structure enhance the performance of the ABSato algorithm further.

#### REFERENCES

- [1] S. U. H. Qureshi, "Adaptive equalization", Proc. IEEE, Vol.73, No.9, pp.1349-1387, 1982.
- [2] A. Benveniste and M. Goursat, "Blind equalizers", IEEE Trans. Commun., Vol. COM-32, pp.871-882, Aug. 1982.
- [3] Dominique N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems", IEEE Trans. Commun., Vol Com-28, No.11, pp.1867-1875, Nov. 1980.
- [4] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation", IEEE Trans. on Commun., Vol. COM-23, pp.679-682, June 1975.
- [5] T. Shimamura and C. F. N. Cowan, "Equalization of time-variant multipath channels using amplitude banded LMS algorithms", IEICE Trans. Fundamentals on Electronics, Communications and Computer Sciences, Vol.E84-A, No.3, pp.802-812, 2001.

- [6] T. Shimamura and C. F. N. Cowan, "Amplitude banded technique for time variant channel equalisation", Proc. IMA Int. Conf. Mathematics in Signal Processing, Dec. 1996.
- [7] T. Shimamura, "Performance of the amplitude banded LMS equalizer on stationary channels", Proc. IEEE International Workshop on Nonlinear Dynamics of Electronic Systems, pp.289-292, 2007.
- [8] S. Haykin, "Adaptive Filter Theory", Prentice-Hall, 1996.



# Performance Investigation of Earth-to-Satellite Microwave Link Due to Rain Fade in Bangladesh

<sup>1</sup>Md. Rafiqul Islam, <sup>1</sup>Md. Arafatur Rahman, <sup>1</sup>Farhat Anwar and <sup>2</sup>Md. Mahfujur Rashid

<sup>1</sup>Kulliyyah of Engineering, International Islamic University Malaysia  
Jalan Gombak, 53100 Kuala Lumpur, Malaysia

<sup>2</sup>Dept. Of Electrical & Electronic Engineering, Ahsanullah University Of Science & Technology, Dhaka  
T/N : +603-6196 4572 E-mail: rafiq@iiu.edu.my

**Abstract** — The fast growth in communication systems has resulted congestion in lower frequency bands and system designers are forced to explore higher and higher frequencies. Rain is a dominant source of attenuation at higher frequencies and consequently degrades the system performance in tropical and subtropical regions. The knowledge of rain fade and its performance is essential in order to optimize system capacity and meet quality and reliability. This paper has investigated performance of earth-to-satellite link due to rain fades operating in a subtropical country Bangladesh. The rain intensity data are derived from forty years measured annual rainfall data. The converted rain intensity data is used to estimate rain fades at C, Ku and Ka-bands. The rain fade is also estimated using ITU-R recommended rain intensity. Noise generated during rains is predicted for all three bands and carrier-to-noise ratios are estimated to compare the performances of earth-to-satellite link at different frequency bands in a subtropical region.

**Index Terms** — Rainfall Data, Rain Fade, Link Budget, Earth-to-satellite, CNR.

## I. INTRODUCTION

The very fast growth in communication systems has brought saturation to the most desirable frequency band (1 to 10 GHz). This fact has led to the utilization of higher frequencies extending the radio frequency spectrum into the millimeter wavelength region. Rain is a dominant source of attenuation at higher frequencies. Attenuation due to rain at frequencies above 10 GHz, mainly leads to outages that compromise the availability and quality of service, making this one of the most critical factors in satellite link design. The design of new telecommunication systems requires the knowledge of rain fade in order to optimize system capacity and meet quality and reliability criteria [1].

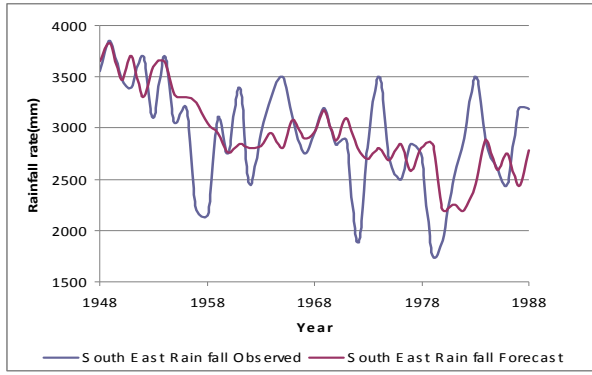
In the design of a radio communication system, one of the major concerns is to assess system unavailability also called outage time. Unavailability or outage time is the amount of time during which the system's performance will be below some threshold value or that it will not be usable. For designing a reliable system, the amount of outage time has to be kept below some objective. In microwave systems, outages can occur either due to equipment failure or it can be propagation outage. In modern systems, the equipment outage time can be made negligibly small by using standby equipment and

automatic protection switching systems. However, at higher frequencies, attenuation by rain can cause an outage, which can not be easily protected. Therefore the practical way of achieving a reliable radio system at those frequencies where there is substantial rain attenuation is to design the system in such a way that the expected amount of rain outage is below some objective [2].

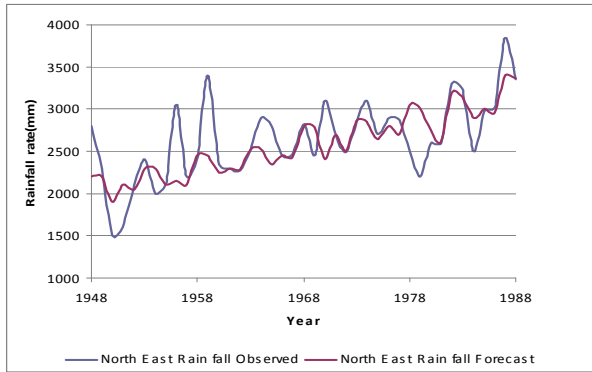
For a reliable communication system, unavailability time during a year has to be kept at 0.01 percent [6]. This corresponds to availability time of 99.99 percent during a year. Therefore rainfall with one-minute integration time is very important parameter to predict attenuation at 0.01% of time availability. The long term cumulative annual rainfall data are available for most of the countries of the world. This paper has presented the cumulative rainfall data collected for forty years in four different parts of Bangladesh. Using appropriate conversion model, the long-term annual rainfall data has been converted to rain intensity data. The rain intensity proposed by International Telecommunication Union (ITU-R) as well as converted data are used to predict the rain fade for earth-to-satellite at C, Ku and Ka-Bands. This paper has investigated performance of earth-to-satellite link due to rain fades at different frequency bands operating in Bangladesh.

## II. ANNUAL RAIN STATISTICS

Bangladesh has a subtropical monsoon climate characterized by wide seasonal variations in rainfall, moderately warm temperatures, and high humidity. Most parts of the country receive at least 2000 mm of rainfall per year. It locates at 88° to 93° longitude (East) and 20° to 27° Latitude (North). Because of its location just south of the foothills of the Himalayas, where monsoon winds turn west and northwest, Bangladesh receives the heavy average precipitation which classify into four distinct regions. They are as North-West, North-East, South-West and South-East. The annual rainfall data collected from 1948 to 1988 for two regions are shown in from Figure 1 [7]. The mean annual rainfall of North-East is 2635 mm, North-West is 1821 mm, South-East is 2937 mm and South-West is 1768 mm[7].



(a)



(b)

Figure 1: Variations of annual rainfall data measured and forecasted for forty years in (a) South East region and (b) in North East region of Bangladesh.

### III. CONVERSION OF RAIN STATISTICS TO RAIN INTENSITIES

Common rain attenuation prediction methods require 1-min rain rate data, which is scarce in the tropical and subtropical region [2]. However, yearly rainfall data are available at many meteorological stations. A method for converting the available rainfall data to the equivalent 1 min rain rate cumulative distribution (CD) would be very useful for radiowave engineers. For this reason 1 min rain rate CD can be estimated by the use of Chebil model and long-term mean annual rainfall data. His model is as follows:

$$R_{0.01} = M \quad (1)$$

where  $a$  and  $b$  are regression coefficients. In Chebil model the regression coefficients  $a$  and  $b$  are defined as [3]  
 $a = 12.2903$  and  $b = 0.2973$  (2)

Chebil has developed his model based on measured values of  $M$  in Malaysia, Indonesia, Singapore, Brazil and Vietnam. He showed that his model is the best estimate of

the measured data [3,4]. Using Chebil model, long-term mean annual rainfall data has been converted to 1 min rain rate data and are presented in Table 1.

Table 1: Measured mean annual rainfall and converted corresponding rain intensity at four regions in Bangladesh.

Name of Region	Measured Mean Annual Rainfall, mm	Converted Rainfall rate $R_{0.01}$ , mm/hr
South East	2937	132
South West	1768	114
North East	2635	128
North West	1821	115
ITU-R Map [6]	-	95

The highest rain intensity is observed at south east region of Bangladesh at 132 mm/hr and lowest at south west is 113 mm/hr. The rain intensity recommended by ITU-R map is found 95 mm/hr for Bangladesh[6] which is far lower than converted rain intensity from measured long term annual rainfall.

### IV. RAIN ATTENUATION STATISTIC FROM RAINFALL RATE

The following procedure provides estimates of the long-term statistics of the slant-path rain attenuation at a given location for frequencies up to 55 GHz. The following parameters are required:

- $R_{0.01}$  : point rainfall rate for the location for 0.01% of an average year (mm/h)
  - $h_s$  : height above mean sea level of the earth station (km)
  - $\theta$  : elevation angle (degrees)
  - $\phi$  : latitude of the earth station (degrees)
  - $f$  : frequency (GHz)
  - $R_e$  : effective radius of the Earth (8 500 km).
- The geometry is illustrated in Figure 2 where  
 A = frozen precipitation  
 B = rain height  
 C = liquid precipitation  
 D = Earth-space path

Step 1: Determine the rain height,  $h_R$ , as given in Recommendation ITU-R P.839.

Step 2: For  $\theta \geq 5^\circ$  compute the slant-path length,  $L_s$ , below the rain height from:

$$L_s = \frac{(h_R - h_s)}{\sin \theta} \quad \text{km} \quad (3)$$

For  $\theta < 5^\circ$ , the following formula is used:

$$L_s = \frac{2(h_R - h_s)}{\left(\sin^2 \theta + \frac{2(h_R - h_s)}{R_e}\right)^{1/2} + \sin \theta} \quad \text{km} \quad (4)$$

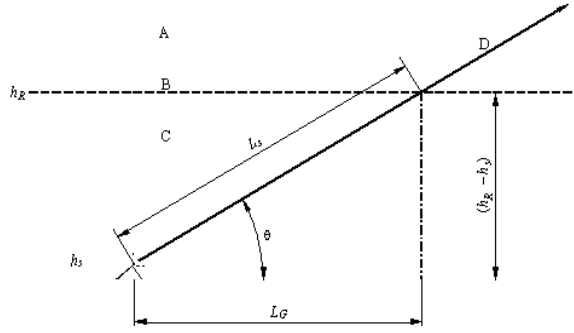


Figure 2: Schematic presentation of an earth-space path given the parameters to be input to the attenuation prediction process.

If  $h_R - h_s$  is less than or equal to zero, the predicted rain attenuation for any time percentage is zero and the following steps are not required.

Step 3: Calculate the horizontal projection,  $L_G$ , of the slant-path length from:

$$L_G = L_s \cos \theta \quad \text{km} \quad (5)$$

Step 4: Obtain the rainfall rate,  $R_{0.01}$ , exceeded for 0.01% of an average year (with an integration time of 1 min). If this long-term statistic cannot be obtained from local data sources, an estimate can be obtained from the maps of rainfall rate given in Recommendation ITU-R P.837. If  $R_{0.01}$  is equal to zero, the predicted rain attenuation is zero for any time percentage and the following steps are not required.

Step 5: Obtain the specific attenuation,  $\gamma_R$ , using the frequency-dependent coefficients given in Recommendation ITU-R P.838 and the rainfall rate,  $R_{0.01}$ , determined from Step 4, by using:

$$\gamma_R = k (R_{0.01})^\alpha \quad \text{dB/km} \quad (6)$$

Step 6: Calculate the horizontal reduction factor,  $r_{0.01}$ , for 0.01% of the time:

$$r_{0.01} = \frac{1}{1 + 0.78 \sqrt{\frac{L_G \gamma_R}{f}} - 0.38 (1 - e^{-2L_G})} \quad (7)$$

Step 7: Calculate the vertical adjustment factor,  $v_{0.01}$ , for 0.01% of the time:

$$\zeta = \tan^{-1} \left( \frac{h_R - h_s}{L_G r_{0.01}} \right) \quad \text{degrees}$$

$$\text{For } \zeta > \theta, \quad L_R = \frac{L_G r_{0.01}}{\cos \theta} \quad \text{km}$$

$$\text{Else,} \quad L_R = \frac{(h_R - h_s)}{\sin \theta} \quad \text{km}$$

$$\text{If } |\varphi| < 36^\circ, \chi = 36 - |\varphi| \quad \text{degrees}$$

$$\text{Else, } \chi = 0 \quad \text{degrees}$$

$$v_{0.01} = \frac{1}{1 + \sqrt{\sin \theta} \left( 31 \left( 1 - e^{-(\theta/(1+\chi))} \right) \sqrt{\frac{L_R \gamma_R}{f^2}} - 0.45 \right)}$$

(8)

Step 8: The effective path length is:

$$L_E = L_R v_{0.01} \quad \text{km} \quad (9)$$

Step 9: The predicted attenuation exceeded for 0.01% of an average year is obtained from:

$$A_{0.01} = \gamma_R L_E \quad \text{dB} \quad (10)$$

Step 10: The estimated attenuation to be exceeded for other percentages of an average year, in the range 0.001% to 5%, is determined from the attenuation to be exceeded for 0.01% for an average year:

$$\text{If } p \geq 1\% \text{ or } |\varphi| \geq 36^\circ: \quad \beta = 0$$

$$\text{If } p < 1\% \text{ and } |\varphi| < 36^\circ$$

$$\text{and } \theta \geq 25^\circ: \beta = -0.005(|\varphi| - 36)$$

$$\text{Otherwise: } \beta = -0.005(|\varphi| - 36) + 1.8 - 4.25 \sin \theta$$

$$A_p = A_{0.01} \left( \frac{p}{0.01} \right)^{[-(0.65 + 0.03 \ln(p) - 0.04 \ln(A_{0.01}) - \beta(1-p) \sin \theta)]} \quad \text{dB} \quad (11)$$

This method provides an estimate of the long term statistics of attenuation due to rain. Considering Singapore Satellite ST1 located at 88° East longitude and earth station at Dhaka at 90° longitude (East) and 24° Latitude (North), the elevation angle is 61.8°. Using above methodology from step1 to step10 and the rain rates derived in Table 1, rain fades have been estimated for South-East region and other region as well as those recommended by ITU-R and presented in Fig. 3 and Fig. 4. All cases the signals are assumed as vertically polarized and the regression coefficients are given as table Table 2.

Table2: Regression Coefficients for Estimating Specification Attenuation.

vertical	C-band (6GHz)	Ku-band (14GHz)	Ka-band (20GHz)
$\alpha$	1.265	1.128	1.065
$k$	0.00155	0.0168	0.0691

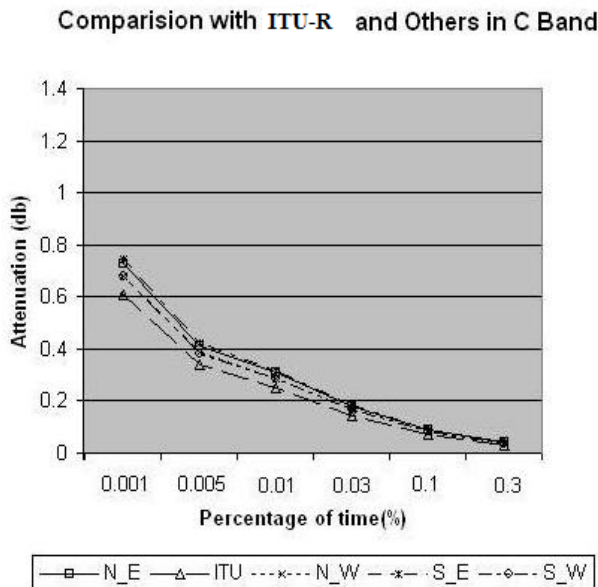


Figure 3: Variations of predicted rain fade for C-Band at different percentage of time of the year.

At C-band, attenuation due to rain is predicted 0.3 dB for 0.01% of the time of the year. But the same at Ku-band and Ka-band are predicted as 13.8 dB and 39.2 dB respectively. The difference between C and Ku-band rain fade is almost 13.5 dB and C and Ka-band is about 38.8 dB for the design of earth-to-satellite microwave link with 99.99% reliability. Hence to design reliable earth-to-satellite microwave link is very critical at Ku and Ka-bands and needs careful and accurate estimation of rain attenuation in Bangladesh.

**Comparison with ITU-R and Worst case in Different Band**

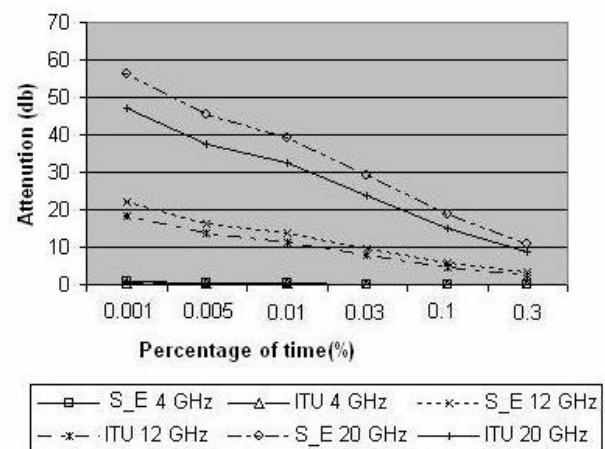


Figure 4: Variations of predicted rain fade for all three frequency bands at different percentage of time of the year.

All three bands, rain attenuation are also estimated based on ITU-R recommended rain. All rain fades at 0.01% are presented in Table 3. It is obvious that ITU-R predicted rain attenuation are much lower than those predicted using measured rain rate in all cases. The differences are 0.1 dB at C-band, 2.5 dB at Ku-band and 6.9 dB at Ka-band.

Table 3: Estimated rain fades for all three frequency bands.

	4 GHz	12 GHz	20 GHz
$A_{0.01\%}$	dB	dB	dB
ITU-R	0.25	11.3	32.3
Measured	0.31	13.8	39.2

## V. LINK BUDGET AT CLEAR AIR AND DURING RAIN

The atmospheric absorption has been estimated using ITU-R proposed graph for elevation angle  $61.8^\circ$  using the equation (12) and presented in Table 4.

$$\text{Atmospheric Absorption, [AA]} = [\text{AA}]_{900} * \text{cosec}(\text{El}) \quad (12)$$

Link Budget has been calculated for Earth-to-satellite downlink at three bands using ITU-R proposed rain rate and converted measured rain rate for Bangladesh. The receive signal level as well as Carrier-to-noise ratio has

Table 4: Atmospheric absorption losses at three frequency bands

Frequency	[AA]
4 GHz	0075 B
12 GHz	011 dB
20 GHz	0.40 dB

been estimated both clear air condition and raining conditions and are tabulated in Table 5. The noise power is calculated using equation (13).

$$N = K T_{\text{sky}} B_w \quad (13)$$

K is Boltzmann constant

TS is sky temperature and is assumed as  $140^\circ$ .

Bw is noise band width and 36 MHz at C, Ku and Ka-bands.

Noise temperature increased during rain is estimated equation (14) as follows:

Sky noise temperature due to rain

$$T_{\text{sky.rain}} = 273(1 - 10^{-A/10}) \text{ K} \quad (14)$$

Hence, total system noise temperature in rain,

$$T_{\text{s.rain}} = T_{\text{LNA}} + T_{\text{sky.rain}} \quad (15)$$

Finally, increase of noise temperature due to rain in different polarization,

$$N = 10 \log (T_{\text{s.rain}} \cdot T_{\text{s}}) \quad (16)$$

It is observed that ITU-R recommendation underestimates the rain rate measured in Bangladesh and consequently the

link budget estimation introduces significant errors specially at Ku and Ka-bands. Even though C/N ratio is 3-4 dB higher in Ku and Ka-bands than C-band during clear air condition. But during the rains C/N ratios falls to 8.6 dB at C-band and -4.36 dB at Ku and -24.13 dB at Ka-bands. These estimation are based on ITU-R recommended rain rates. But the above estimation is done based on converted measured rain rate and C/N ratio is fallen to -6.8 dB at Ku-band and -31.1 dB at Ku-band. These prediction shows that using of Ku and Ka-bands is very challenging and critical in this region. However the measured rain rate data is converted from measured long term annual rainfall data, it is too preliminary to comment correctly. Therefore it is recommended to measure rain intensity and rain drop size distribution urgently for the design of reliable microwave link in Bangladesh.

## VI. CONCLUSION

Rain is a dominant source of attenuation at higher frequencies in tropical and subtropical regions. Therefore accurate estimation of rain fade is very essential in order to design reliable microwave links in such regions. This paper has predicted rain fades in a subtropical country Bangladesh using rain intensity data derived from forty years measured annual rainfall data. It is observed that rain fade is not significant at C-band, but is very critical at Ku and Ka-bands. This paper has also investigated performance of earth-to-satellite link due to rain fades. Noise generated during rains is predicted for all three bands and carrier-to-noise ratios are estimated to compare the performances of earth-to-satellite link at different frequency bands. From observation, it is found that C/N ratio is 3-4 dB higher in Ku and Ka-bands than C-band during clear air condition. But during the rains C/N ratios falls to 8.6 dB at C-band and -4.36 dB and -24.13 dB at Ku and at Ka-band respectively. These estimation are based on ITU-R recommended rain rates. But the C/N ratio falls to -6.8 dB at Ku-band and -31.1 dB at Ka-band based on converted measured rain rate. These prediction shows that using of Ku and Ka-bands is very challenging and critical in this region. However the measured rain rate data is converted from measured long term annual rainfall data, it is too preliminary to comment correctly. Therefore it is recommended to measure rain intensity and rain drop size distribution urgently for the design of reliable microwave link in Bangladesh.

Table 5: Link Budget Calculation for Earth-to-satellite downlink at three bands using ITU-R proposed and measured rain rate for Bangladesh

	ITU			S_E region (Worst case)		
Dounlink Power Budget	4GHz	12GHz	20GHz	4GHz	12GHz	20GHz
EIRP	49	49	49	49	49	49
$B_0$ = Transponder output backoff (typically)	-2	-2	-2	-2	-2	-2
Gr = Recieving antenna gain	33.93	43.47	47.91	33.93	43.47	47.91
[FSL] = Free space path loss	-199.7	-206.6	-209.7	-199.7	-206.6	-209.7
$L_{ant}$ = Edge of beam loss for satellite antenna (typically)	-3	-3	-3	-3	-3	-3
[AA] = Clear air atmospheric loss	-0.075	-0.11	-0.4	-0.075	-0.11	-0.4
$L_m$ = Other losses (typically)	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
$P_r$ = Receive power at receive station (Clear Air)	-122.20	-119.56	-117.93	-122.2	-119.6	-117.9
Rain Attenuation $A_{rain}$	0.25	11.28	32.31	0.31	13.76	39.17
$P_r$ (due to rain)	-122.45	-130.84	-150.24	-122.52	-133.3	-157.1
Dounlink Noise Budget						
$k$ = Boltzmann's constant in dBW/K/Hz	-228.6	-228.6	-228.6	-228.6	-228.6	-228.6
$T_s$ = System noise temperature, 140 K	21.46	21.46	21.46	21.46	21.46	21.46
$B_n$ = Noise bandwidth, 36 MHz	75.56	75.56	75.56	75.56	75.56	75.56
$N$ = Receiver noise power (Clear Air)	-131.58	-131.58	-131.58	-131.58	-131.6	-131.58
Increased due to rain ( $\Delta N$ )	0.534	4.56	4.93	0.66	4.56	5.03
$N$ = Receiver noise power (due to rain)	-131.04	-126.48	-126.11	-130.92	-126.5	-126.01
C/N Ratio in Receiver	4GHz	12GHz	20GHz	4GHz	12GHz	20GHz
C/N = (Clear Air)	9.38	12.02	13.65	9.38	12	13.68
C/N = (due to rain)	8.59	-4.36	-24.13	8.403	-6.8	-31.1

#### REFERENCES

- [1] ITU-R P.618-8, "Propagation data and prediction methods required for the design of Earth-space telecommunication systems", ITU, Geneva, Switzerland, 2003.
- [2] ONG, J.T., and ZHU, C.N.: 'Rain rate measurements by rain gauge network in Singapore', Electron. Lett., 1997, 33, (3), pp. 240-242.
- [3] CHEBIL, J., and RAHMAN, T.A.: 'Rain rate statistical conversion for the prediction of rain attenuation in Malaysia', Electron. Lett., 1999, 35, (12), pp. 1019-1021.
- [4] J. Chebil and T.A. Rahman, "Development of 1 min rain rate contour maps for microwave applications in Malaysian Peninsula", ELECTRONICS LETTERS 30th September 1999 Vol. 35 No. 20, pp.1172-1174.
- [5] MOUPFOUMA, F., and MARTIN, L.: 'Modelling of the rainfall rate cumulative distribution for the design of satellite and terrestrial communication systems', Int. J. Sat. Commun., 1995, 13, pp. 105- 115.
- [6] ITU-R, P. RECOMMENDATION ITU-R P.837-4, "Characteristics of precipitation for propagation modeling", ITU, Geneva, Switzerland, 2003.
- [7] <http://www.sdnpsd.org/sdi/internationaldays/wed/2006/bangladesh/image/rainfal>

## Author Index

### A

Abbas, Majed	341
Abdullah, Raja S. A. R.	104
Abdullah-Al-Mahmud, Md.	150
Acharya, Ayan	110, 470
Acharyya, Sriyankar	132
Ahad, Md. Atiqur Rahman	201, 213
Ahamad, Md Mokaddem	183
Ahmad, Mohiuddin	323, 521
Ahmad, Shamim	52
Ahmed, Abu Shohel	371
Ahmed, Asif	171
Ahmed, Emdad	56
Ahmed, Farruk	166
Ahmed, K.	504
Ahmed, Md.Razu	94
Ahmed, Mohiuddin	341
Ahmed, Reaz	127
Ahmed, Syed Ishtiaque	1
Ahn, Kyoung Kwan	282
Akter, Nasreen	177
Al-Adwany, Maan A. S.	27
Alam, Md. Ashad	399
Alam, Md. Jahangir	73
Alam, Md. Zahangir	646, 741
Alam, Muhammad Raisul	264, 189
Ali, Borhanuddin M.	104, 492
Ali, Md. Sadek	138
Al-Mamun, S.M. Abdullah	554
Almodarresi, S.M.T.	708
Alom, Md. Zahangir	646, 741
Amin, M. R.	85, 161
Amin, Md. Ruhul	116
Anjum, Md. Nashid	150
Anwar, Farhat	773
Anwar, Shahed	669
Anwer, Mohammed	735
Ara, Jannatul Ferdousi	79
Arefin, Md. Shamsul	387
Arefin, Mohammad Shamsul	40, 183, 306

Arif, Abu Shamim Mohammad	405
Aski, Baharak Shakeri	219
Avi, Tamkin Khan	436
Azim, Akramul	581
Aziz, Faisal Aman	681
Aziz, Syed Mahfuzul	276

### B

Babu, Hafiz Md. Hasan	7
Banik, Shipra	735
Bashar, M.A.	481
Bastani, Parisa	708
Begum, Nasima	415
Bennamoun, Mohammed	329
Bhuiyan, Md. Alamin	510
Bhuiyan, Md. Al-Amin	415
Bhuyan, Md. Shariful Islam	25, 127
Billah, M. M. Rajib	312
Bindal, Aman	560
Biradar, S. R.	98
Biswas, Gautam Kumar	405

### C

Chakma, Namvi	288
Chakraborty, Mithun	470
Chanchary, Farah Habib	735
Chattopadhyay, Kaushik	110
Chawdhury, Md. Didarul Alam	714
Choudhury, Lazeeb Faiz	702
Chowdhury, A.K.M	
Rasheduzzaman	248
Chowdhury, Ahsan Raja	7
Chowdhury, Fahim Md.	548
Chowdhury, Gita	629
Chowdhury, M. A.	481

### D

Das, J. K.	161
Das, S. K.	481
Datta, Alak K.	19
Debnath, Mitu Kumar	381
Deldari, Hossein	300
Dewan, Ahsanullah M	611

## Author Index

Dewan, Saifullah M	611	Hasan, Md. Nazmul	581
Didar-Al- Alam, S. M	581	Hasan, Md. Rashidul	548
Djeraba, Chabane	194	Hasan, Mohammad Khalad	548
Dzulkipli, Ahmad Faris	723	Hasan, Muhammad Mehedi Al	
<b>E</b>		Emran	253
Easmin, Nilufar	536	Hashem, M. M. A.	288, 347
Eftakhar, S. M. Ashik	207	Hashimoto, Michio	144
El-Nasr, M. Abou	498	Hirose, Keikichi	294, 335
Elsayed, Adel	242	Hoque, Md. Moshiul	40
Endo, Tsutomu	747	Hoque, Md. Shadiul	571
<b>F</b>		Hoque, Mohammad Moinul	554
Farid, Dewan Md.	652	Hoque, Mohammed Moshiul	183, 306
Farook, Abdullah Al	40	Horiguchi, Susumu	359, 365
Forghani, Yahya	89	Hossain, Altaf	729
Foroughi, Homa	219, 300	Hossain, Eklas	723
Fukumizu, Kenji	399	Hossain, M. Tanvir	669
Fukuzawa, M.	516	Hossain, Md. Delowar	138
<b>G</b>		Hossain, Md. Gahangir	183
Gani, MD. Osman	691	Hossain, Md. Tofazzal	486
Ghods, Leila.	442	Hossain, Md. Zakir	548
Ghosh, Dibyajyoti	353	Hossain, Md. Zulfikar	521
Giasuddin, Shafaat Sheik	542	Hossain, Mohammad Akram	376
Goni, Md. Osman	762	Hossain, Mohammad Kabir	317
Goyal, Akhil	560	Hossain, Mohammad Monoar	459
Gupta, Anand	560	Hossain, Mohammad Shahid	681
Gupta, Ankuj	560	Hossain, Saima	177
<b>H</b>		Hossain, Syed Akhter	166
Habib, A.H.M. Ashfak	714	Hossen, Md. Prince	718
Habib, Sheikh Mahbub	393	Huq, Mohammad Rezwanul	436
Hama, H.	516	<b>I</b>	
Haque, Md. Dulal	94	Ibrahim, Muhammad	7
Haque, Mohammed Mezbaul	686	Ihaddadene, Nacim	194
Harada, Yuichi	144	Imam, Shams M	317
Hasan, K. M. Azharul	536	Inoguchi, Yasushi	359, 365
Hasan, Khondker Shajadul	317, 36	Ishikawa, S.	201, 213
Hasan, Masud	459, 1	Ishikawa, Seiji	207
Hasan, Md. Al Mehedi	52	Islam, A.H.M. Saiful	312
Hasan, Md. Kamrul	465	Islam, Chowdhury Sajadul	646
Hasan, Md. Mehedi	259	Islam, Gazi Zahirul	587
Hasan, Md. Monjur Ul	306	Islam, Md. Ariful	1
		Islam, Md. Ezharul	415



## Author Index

Islam, Md. Imdadul	85, 161	Kaykobad, Mohammad	465
Islam, Md. Mahfuzul	581	Khan, Khodadad	735
Islam, Md. Nazrul	593	Khan, Md. Ali Ashik	436
Islam, Md. Nazrul	448	Khan, Md. Ibrahim	587
Islam, Md. Rabiul	116	Khan, Mozammel H. A.	270
Islam, Md. Rafiqul	411	Khan, Muhammad Lutfor	
Islam, Md. Rafiqul	453,225,236	Rahman	768
Islam, Md. Rafiqul	718	Khan, Shahriar	753, 46
Islam, Md. Rafiqul	773	Khan, Tanjila	566
Islam, Md. Rakibul	225, 411	Khandker, M. R.	635, 640
Islam, Md. Shariful	138	Khandker, Md. Mamun-ur-Rashid	623
Islam, Md. Tajul	177	Khanom, Rashida	270
Islam, Mohammad Rakibul	387	Khatun, Sabira	104, 492
Islam, Mohammad Tariqul	758	Khayat, Saied Hossaini	62
Islam, R.	481	Kim, H.S.	201, 213
Islam, Saiful	571	Kim, Hyoungseop	207
Islam, Syed Foyzol	312	Kim, Jinsang	387
Islam, Syed Nurul	405	Kolahi, Samad S.	663
Islama, Md. Moinul	138	Konar, Amit	110, 470
Israt, Presila	288	Koyama, Osanori	144
<b>J</b>		Kundu , Debojit	19
Jacob , Cyril	393	<b>L</b>	
Jahan, Moslema	347	L'opez-Ortiz, Alejandro	459
Jahanshahi, H.	605	Laila, Runa	85
Jamal, Lafifa	371	Lata, Lutfunnesa	323
Jamil, Md. Subrun	599	Lee, Sungyoung	465
Jewel, Faruq Ahmed	533	Lee, Young-Koo	465
Juthy, Sabiha Rahman	85	<b>M</b>	
<b>K</b>		Mahbub, Md. Shahriar	554
Kader, Md. Abdul	448	Mahdi, M. Adzir	492
Kadir, Mohammad Ismat	571	Mahmood, A.S. M. Ashique	681
Kamal, Abu Raihan Mostofa	248	Mahmud, Mehedi	566
Kamal, K. Raiyan	527	Maiti, Deepyaman	74, 470
Kamal, Mohammad Sarwar	306	Majumder, Koushik	98
Kamal, Mustafa A.	31	Majumder, S. P.	657
Kamal,Tahseen	270	Mandal, Tarak N.	19
Kamruzzaman, Mohammed	533	Mani, Paul	663
Kandeepan, Sithamparanathan	486	Masud, Abdullah Al	73
Karim, Md. Shahriar	161	Masum, Md.	264, 189
Katsuyama, Yutaka	144	Mazady, Md. Anas Boksh	587
Kaykobad, M.	527	Mehedi, Iftekhar Jahan	691

## Author Index

Mirjalily, Ghasem	708
Misbah-Uddin, Abu Hamed	
Mohammad	476
Misran, Norbahiah	758
Miura, Yasuyuki	359, 365
Moghanjoughi, Ayyoub Akbari	104
Mohammadi, S.	605
Molla, Md. Khademul Islam	294, 335
Mondal, M. R. H.	657
Mondal, Md. Nazrul Islam	73
Monir, Mohammad Iftekhar	476
Monir, Sadia	566
Monjur-Ul-Hasan, Md.	183
Mousumi, Fouzia Ashraf	599
Mukaddes, A.M.M.	617
Mukherjee, Nandini	353
Mukhopadhyay, Rupam	353
Munir, M. S.	504
Mustafa, Mohammad Abdur	
Rahim	566
Muhida, Riza	723

## N

Naik, K. Krishna	156
Nakamori, N.	516
Narayan, Shaneel	663
Nashiry, Md. Asif	150
Nasrin, Wahida	576
Nasser, M.	729
Nasser, Mohammed	399, 67
Nazmin, Shams Ara	611
Nguyen, Du D. T.	663
Nilufar, Sharmin	79
Nizamuddin, S. M.	504
Nooruzzaman, MD.	144
Numan, Mostafa Wasiuddin	189
Numan-Al-Mobin, Abu Md.	762
Nur-A-Alam, Mohammad	686
Nurunnabi, A. A. M.	67
Nusrat, Elhum	371
Nusrat, Sabrina	459

## O

Olovsson, Tomas	393
-----------------	-----

## P

Pal, Achira	19
Pal, Rajat K.	19
Parsa, Dr. Saeed.	442
Parvin, Nilufa	536
Perrizo, William	317
Phadikar, Santanu	420
Podder, Anonda	453
Pourreza, Hamidreza	219
Prasad, M.N. Giri	156
Priantoro, Akhmad Unggul	675

## Q

Quazi, Rumana	121
---------------	-----

## R

Rahaman, G. M. Atiqur	405, 171
Rahaman, Mohammad Zahidur	376
Rahat, H. M.	85
Rahman, Md. Anisur	381
Rahman, Rashedur	542
Rahman, Chowdhury Mofizur	13, 691
Rahman, Daud Md. Jamilur	248
Rahman, Dewan Md. Fayzur	13
Rahman, Gazi Mushfiqur	371
Rahman, Khairul Azizi Abdul	723
Rahman, M Lutfar	166
Rahman, M. Atiqur	669
Rahman, M. M. Hafizur	359, 629, 365
Rahman, M. S.	481
Rahman, Md. Arafatur	773
Rahman, Md. Arifur	264
Rahman, Md. Ashiqur	542
Rahman, Md. Asikur	248
Rahman, Md. Fayzur	116
Rahman, Md. Hasanur	762
Rahman, Md. Lutfur	702
Rahman, Md. Mahbubur	138, 150
Rahman, Md. Mostafizur	264
Rahman, Md. Rezwanaur	436

## Author Index

Rahman, Md. Saidur	25	Sil, Jaya	420
Rahman, Md. Saidur	171	Singha, Biswajit	448
Rahman, Md. Tajmilur	510	Sobhan, M Abdus	646, 741
Rahman, Mohammad Zahidur	593, 652	Sohel, Ferdous Ahmed	329
Rahman, Muhammad Arifur	230	Solouk, Vahid	492
Rahman, S. M. Mahfuzur	282	Sultana , Basra	635, 640
Rahman, Tazilur	36	Sultana, Most. Arjuman	629
Rajib, Md. Majharul Islam	576	Sunarto, Yonathan	663
Rajon, S. A. Ahsan	453, 430, 236	<b>T</b>	
Rashid, Md. Mahfujur	773	Tahsin, Tahmina	702
Rezvanian, Alireza	300	Tan, J.K.	201, 213
Rhaman, Md. Khalilur	747	Tan, Joo Kooi	207
Rouf, Rifat Ara	735	Tareq, Shafiul Hasan Md.	189
<b>S</b>		Taslima, Taohida	323
Saha, Sumanta	121	Toaha, Mabruka	46
Sajal, Raihan Ferdous	533	<b>U</b>	
Salahuddin, G. M.	171	Uddin, M. Jasim	675
Sami, Saumen Seal	36	Ullah, M. Habib	675
Sarbishei, Ghazaleh	62	Ullah, Shaikh Enayet	94
1Sarkar, Ajanta De	353	Uragami, Atsushi	617
Sarkar, Subir Kumar	98	<b>W</b>	
Sarker, Badal Chandra	253	Wondimagegnehu,	
Sarker, M. J.	504	Mohammed H.	768
Sarwar, Hasan	566, 177, 691	Wong, Daniel	492
Sato, Yukinori	359, 365	<b>Y</b>	
Sayeed, Taufique	476	Yamada, M.	516
Scurtu, Vitalie	230	Yamada, Masayoshi	323
Seraj, Mehnaz	691	Yatim, Baharudin	758
Shah, Peer Azmat	696	Yousaf, Muhammad	696
Shahriar, Gazi Abdullah	347	<b>Z</b>	
Shaikh, Mostafa Al Masum	294, 335	Zahda, Showayb	341
Shakeri, Moein	300	Zaman, Md. Shahid Uz	116, 73
Shakib, Mohammed Nazmus	758	Zaman, SK. Ali	430
Shams, Rushdi	242	Zia, Md. Ziaul Karim	13
Sharif, Md. Haidar	194		
Sharma, Sunil	276		
Shihavuddin, A.S.M.	504		
Shimamura, Tetsuya	768		
Siddique, Md. Abubakar	718		
Siddique, Javed	36		
Sikder, Iftikhar U.	424		

# **IDCS'08**

## **IEEE International Workshop On Internet and Distributed Computing Systems**

24 December, 2008  
Khulna University of Engineering and Technology (KUET), Bangladesh

## **WORKSHOP PROCEEDINGS**



**Organized by**

Faculty of Electrical and Electronic Engineering (EEE)  
**Khulna University of Engineering and Technology (KUET)**  
Khulna-9203, Bangladesh  
<http://www.kuet.ac.bd/>

# **IDCS'08**

## **IEEE International Workshop On Internet and Distributed Computing Systems**

### **Editors**

Prof. Dr. M.M.A Hashem  
Dr.M.M. Hafizur Rahman  
Pintu Chandra Shill  
Rushdi Shams  
Md. Jahidul Islam

### **Published by**

Faculty of Electrical and Electronic Engineering (EEE)  
Khulna University of Engineering and Technology (KUET)  
Khulna-9203, Bangladesh

### **Additional Copies of the Proceedings can be obtained from**

ICCIT 2008 Secretariate  
Department of Computer Science & Engineering (CSE)  
Khulna University of Engineering and Technology (KUET)  
Khulna-9203, Bangladesh

**IEEE Catalog Number: CFP0817D-PRT**

**ISBN: 978-1-4244-2135-0**

**Library of Congress: 2008900703**

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Operations Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved. Copyright ©2008 by the Institute of Electrical and Electronics Engineers.

# A Trust-Based Distributed Data Fault Detection Algorithm for Wireless Sensor Networks

Zahra Taghikhaki

Computer Engineering Department  
Iran University of Science and Technology  
[Zahra\\_taghikhaki@comp.iust.ac.ir](mailto:Zahra_taghikhaki@comp.iust.ac.ir)

Mohsen Sharifi

Computer Engineering Department  
Iran University of Science and Technology  
[msharifi@iust.ac.ir](mailto:msharifi@iust.ac.ir)

## Abstract

*Fault detection is a difficult and complex task in WSN because of there are many factors that influence data and could cause faults. Large-scale sensor networks impose energy and communication constraints, thus it is difficult to collect data from each individual sensor node and process it at the sink to detect faulty sensors. The proposed approach saves energy and improves network lifetime by detecting data faults locally in cluster head and therefore reducing the number of transmissions required to convey relevant information to the sink. This paper presents a novel approach for detecting sensors which produce faulty data in a distributed way as well as identifying the type of data faults using trust concepts to gain a high degree of confidence. We validate our method with simulations results.*

**Keywords:** Wireless Sensor Networks, Distributed Fault Detection, Trust, Reputation.

## 1. Introduction

The main task of each sensor node in a Wireless Sensor Network (WSN) is to monitor the environment and report the current status of the environment to a center called base station. If a sensor becomes faulty for any reason, then false and unreal data are reported to the center by the sensor. Based on false data, the base station may make inaccurate and faulty decisions leading to serious inefficiencies of the whole network..

Fault diagnosis is one of the main challenges and building blocks of many dependable protocols. In fact, once faulty units are identified, the remaining units can be isolated from the rest of the system, and the computation can proceed, although with reduced performance [1].

Sensor nodes are powered by batteries, which are considered as constrained resources. It is very expensive for the base station to collect information from every sensor and identify faulty sensors in a centralized manner. Different applications may require the fault detection to be conducted in a real-time mode

with low latency or high throughput. Therefore, a localized and distributed generic algorithm for each node is highly preferred in wireless sensor networks.

This paper presents a novel approach for detecting sensors which produce faulty data in a distributed way as well as identifying the type of data faults with a high degree of confidence using reputation concepts.

The rest of paper is organized as follows. Section 2 presents the most related works. Section 3 introduces reputation and trust concepts. In Section 4 faults are modeled and the proposed algorithm to model faults is discussed. Section 5 explains the suggested fault detection algorithm in detail. Section 6 demonstrates evaluations on injected faults in a data set. In Section 7 the proposed algorithm is applied to a real world data set. The complexity of the proposed algorithm is discussed in Section 8 and finally in Section 9 the paper is concluded.

## 2. Related Work

Methods of fault detection based on simplified algorithms such that they could be implemented on a single sensor node are investigated in [2][3]. Both have approaches based on two relationships of correlation: the correlation of a node's measurement and its neighbor's measurements, and the correlation of a node's measurement and its own previous measurement. In [2], a naive bayes algorithm employed which maintains counters of the number of times a particular pair occurs over the history of the sensor network. This approach requires large numbers counters and large amounts of training data. On the other hand, [3] keeps track of the same relationships by only maintaining counters of the differences between the pairs of values. Their claim is that the differences are sufficiently representative of the interesting behavior from the sensors.

However, there are some kinds of faults which require cooperative-diagnosis among a set of sensor nodes. A large portion of faults in WSNs are in this category. For example, Detection method proposed in [4] is to identify faulty sensor nodes in event detection

application. The detection method is based on the assumption that sensor nodes in the same region should have similar sensed value unless a node is at the boundary of the event region. It takes measurements of all neighbors of a node and uses the results to compute the probability of the node being faulty.

Zhang et als. [6] present a different technique for detection of faulty nodes. Each sensor node generates an alert when it observes an inconsistency between itself reading and neighbor reading and then sends alert to sink. Sink makes a graph such that there is an edge  $xy$  in the graph iff  $x$  and node  $y$  are suspicious of each other. Then to each node faulty/non-faulty flag is assigned such that this assignment doesn't affect the consistency of the graph. All of valid assignments is created and the nodes that are marked faulty in all these assignments are labeled as faulty nodes. This approach works in centralized manner then it can not scalable very well.

Chen et als. [7] present a distributed scheme for fault detection. Their scheme have some phases. In the first phase, each sensor compares its reading with neighbors' values and declares itself *locally good* if its value matches with the majority of its neighbors. In the second phase, those nodes that are *locally good* and have a majority of its matching neighbors also *locally good* are declared *globally good*. In the last phase, neighbors of *globally good* nodes that do not agree with those nodes are labeled *faulty*. This scheme is not easily scalable and it has considerable overhead due to high exchange of information between neighbors.

Ding et als. [8] introduce a faulty sensor identification algorithm. Their algorithm is localized but it has too communication overhead. In the algorithm, the reading at a sensor is compared with its neighbors' median reading. If the difference is large, the sensor is very likely to be faulty. Although this algorithm works for large size of sensor networks, the probability of sensor faults needs to be small. If half of the sensor neighbors are faulty and the number of neighbors is even, the algorithm cannot detect the faults well. The location of sensors is needed that it requires GPS or complex techniques for this purpose.

### 3. Trust Concept

The problem of producing reliable information can be diminished to a basic question: "how the sensor nodes trust each other?". Trust is the expectation of one entity about the actions of another [9].

There is much confusion between trust and reputation. When entities face uncertainty, they tend to trust to entities that have high reputation.

Reputation is not a physical quantity but it is a belief; it can only be used to statistically predict the future behavior of other nodes and cannot define deterministically the actual action performed by them.

Trust is a subjective expectation a node has about another node's future behavior. This can be obtained by taking the statistical expectation of the probability distribution representing the reputation between the two nodes. Note that, unlike reputation, the trust metric is simply a number.

We have used the RFSN method for calculating the reputation; more details on RFSN can be found in [9].

### 4. Fault Models

A fault can be introduced into sensor at every point in the sensor network: from failures in the sensor itself, to software bugs and computational errors, to lossy communication [10]. Faults are application and sensor type dependent. There is no significant list of features to consider for fault modeling. The set of models to be used for fault detection contain models for both good and faulty data. Ni et als. [10] categorize the existed faults as: outlier, stuck-at, spike, excessive noise, calibration fault, hardware fault, low battery, and clipping. Sharma et als. [11] categorize faults differently as: NOISE, SHORT and CONSTANT. These three faults have been observed in the four data sets from WSN deployments by others as well [11][12][13]. We start by focusing on a small set of sensor data faults that have been observed in real deployments. We select categorization introduced in [11] to compare our results with their results has achieved on the same dataset [14].

The three types of faults that we intend to detect, as well as how they can be detected, are as follows.

- 1) **SHORT:** A sharp change in the measured value between two successive data points.

*Detection method:* we set an acceptable range for every hours of day and every season. If the reading value is outside this range and the difference between current and previous reading value is above a threshold, it is labeled as outlier.

**NOISE:** The variance of the sensor readings increases. Unlike SHORT faults that affect a single sample at a time, NOISE faults affect a number of successive samples.

*Detection method:* Compute the standard deviation of sample readings within a window  $w$ . If it is above a certain threshold, the samples are affected by the NOISE fault.

- 2) **CONSTANT:** The sensor reports a constant value for a large number of successive samples. The reported constant value is either very high or very

low compared to the “normal” sensor readings.

- 2) *Detection method*: if for several consecutive time intervals the read value by a sensor is invariant then we must check the read values by the sensor’s neighbors. If most of the fault free sensors have variant values in these time intervals then this sensor has constant fault.

## 5. The Proposed Fault Detection Algorithm

First we assume all sensor nodes are divided to some clusters that each cluster has a cluster head.

For fault detection, we have two types of nodes, some sensor nodes (SN) and a cluster head node (CN) in our algorithm as it is shown in Figure 1.

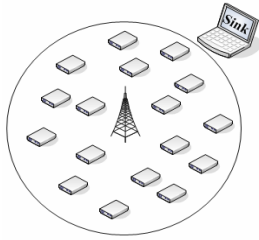


Figure 1. A cluster of SNs and CN

Sensor nodes perform environment sensing and send reading values to the cluster node. Cluster node sends the fusion of all SN data in the cluster to the sink. We assume the cluster node is in the radio range of sensors in the same cluster, meaning that if a sensor broadcasts a message, it can be received by CN in the cluster. SNs are also equipped with the fault detection algorithm and are responsible for detecting faults and reporting them to the CN.

Each SN maintains 3 tables:

1. *Neighbor Data Table (NDT)*: contains data readings of neighbor sensors.
2. *Fault Table (FT)*: contains the neighbor sensors’ faults that detected by the SN. Entry of fault free nodes in this table is “HEALTHY” flag.
3. *Neighbor Reputation Table (NRT)*: contains reputation of neighbor sensors from perspective of the sensor which maintain this table.

The following actions are performed in each cluster. Each sensor broadcast its reading to all neighbors. Each SN receives these data and puts the neighbors reading in its *NDT* then updates its *NRT* considering the difference between its reading and neighbors reading and with comparing the threshold (this threshold is calculated regard to the acceptable difference between values that is application-dependent). SN calculates the reputation for each its neighbors. Actually, each SN puts the number of difference and the number of similarity values

in  $NVote_j$  and  $PVote_j$  respectively ( $j$  is the neighbor  $id$ ), and then it calculates the reputation (or trust) value using statistical expectation and Beta distribution. Now each neighbor has a reputation value which is built by this SN. Based on the fault modeling algorithm described in the previous section, SN detects the fault and puts it in the corresponding fault part in *FT*. In the initially phase all of entry of *FT* is fulfilled by “HEALTHY” word. After  $T$  timeslot each sensor sends *FT* and *RT* to the CN. To detect probably fault of  $S_i$ , CN applies majority voting algorithm on the opinions of some  $S_i$  neighbors (which have reputation value more than threshold) about  $S_i$ .

In reality, *FT* of each neighbor of  $S_i$  that has reputation value more than threshold in the  $RT_i$  (Reputation Table of sensor  $i$ ) is checked by CN to detect final fault detection for  $S_i$ . The final detected fault for  $S_i$  is the majority of trusted neighbors opinions. For example if the result of these opinions be fault type NOISE about fault of  $S_i$ , then the detected fault for this sensor labeled as NOISE.

The final fault table that is constructed by CN is sent to the sink every  $T$  time slot which is specified regard to the network policy and needed accuracy. If we want to monitor fault of sensors frequently then we must set  $T$  to a low value but in this case the generated traffic for sending the *FT* and *RT* to the CN and final *FT* to the sink is increased.

## 6. Evaluation of our method

We analyzed a data set from real-world deployments to detect data faults in sensor traces. The sensor traces contain measurements from a variety of phenomena – temperature, humidity, light, pressure. Authors in [11] used these data set and others to faults prevalence in sensor traces. Also, they detect SHORT, NOISE and CONSTANT faults in these data sets. Then we can compare our results with this reference.

Before we can evaluate the faults detection in real world dataset using the methods discussed in the previous section, we need to characterize the accuracy and robustness of these methods. To do this, we injected faults of the types discussed in section 4 into a real-world data set. We can specify the injected fault with a flag in the data set. By this action we can check what fraction of the injected faults is detected by our method. Injecting faults into a data set gives us an accurate “ground truth” that helps us better understand the performance of a detection method [11].



Similar to [11] we use three metrics to understand the performance of various methods: the number of faults detected, false negatives, and false positives.

### 6.1. Fault injection and validate our method

We want to compare our results with [11], then we use the way for injecting faults that introduced in [11]. To inject SHORT faults, we change value of a sensor  $i$  and replace it by:

$$Val^{new}(S_i) = Val^{old}(S_i) + Val^{old}(S_i) \times C$$

where parameter  $C \in N$  can be a constant or variant.

To inject NOISE faults, we pick a set of successive samples  $W$  and add a random value drawn from a normal distribution,  $N(0, \sigma^2)$ , to each sample in  $W$ . We vary the intensity of NOISE faults by choosing different values for  $\sigma$ .

To inject constant fault we change value of a sensor  $i$  in some successive samples to a constant value.

There are some detection methods that are discussed in [11] and we need them for the purpose of comparing their results to ours. In [11], authors explore three qualitatively different detection methods Linear Least-Squares estimation (LLSE), Hidden Markov Models (HMM), and a rule-based method that leverages domain knowledge to develop heuristic rules for detecting faults. Finally they combine the above methods to build a hybrid method namely Hybrid(I).

Figure 2 compares the accuracy of the SHORT rule, LLSE, HMM, and Hybrid(I) and our method(T) for SHORT faults. The labels used for different detection methods are: **R**: Rule-based, **L**: LLSE, **H**: HMM, **U**: Hybrid(U), **I**: Hybrid(I), **T**: Our method.

As it is shown in Figure 2, our proposed method has detected SHORT faults accurately. Two reasons can create false negatives and false positives. First, when there is inconsistency between the sensed data and reported data to the neighbors by a faulty sensor. In this case, it is possible that the fault or type of fault cannot be detected correctly, leading to false negative.

Second, when the sensors that their reputation are very high (much more above threshold), work badly or have some faults. In this case although the reputations of these sensors are reduced but their reputations are still above the threshold for some periods and they can participate in voting. If there are a lot of these sensors then the final decision can affect them leading to some false negatives and false positives. This condition rarely happens and we have very few false positives or false negatives. To overcome this problem, we can give weight ( $w$ ) to the reputations of sensors. The recent reputation value of a sensor is more important than its old reputation value. We apply this as follows:

$$Re p^{new}(S_i) = w \times Re p^{current}(S_i) + (1 - w) \times Re p^{old}(S_i)$$

Some simulations and their results showed that the proper value for  $w$  is 0.8. With this solution, the results are shown in Figure 3. The two above reasons are used for other faults i.e. NOISE and CONSTANT.

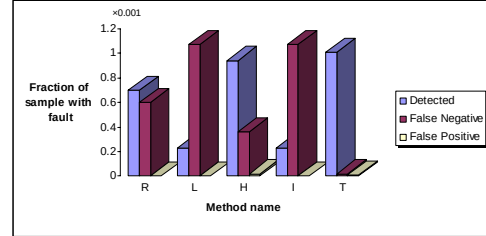


Figure 2. Comparison of methods for injected SHORT faults

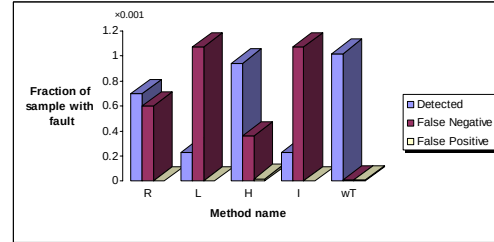


Figure 3. Comparison of methods with our weighted method (wT) for injected SHORT faults

The appearance of NOISE faults are shown in the Figures 4. Figure 4 is a plot of comparison between discussed methods and our weighted and simple methods. As you can see the wT method is more accurate and we can trust to it more than T method because the current reputations considered more than old reputations.

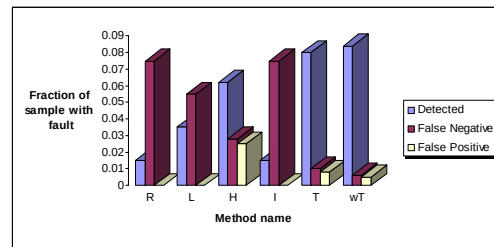


Figure 4. Comparison of methods for injected NOISE faults

The proposed method can also detect CONSTANT faults as well as two previous faults. This is shown in Figure 5.

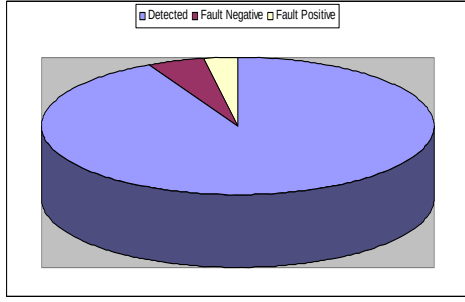


Figure 5. Detection of injected CONSTANT faults

## 7. Faults in a Real world data set

Now we apply our algorithm in a real world data set.

We verified that our proposed method works accurately and showed it in the above figures.

### 7.1. Great Duck Island (GDI) Data Set

We looked at data collected using 30 weather motes on the Great Duck Island over a period of 3 months [14].

We apply our algorithm on the temperature, light, and pressure of this data set and plot results in the following figures. Also, in these figures we compare our method with other detection methods which are applied in this data set and are discussed in the [11].

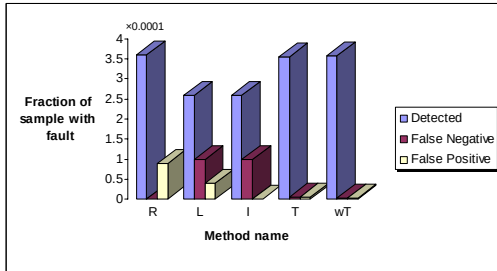


Figure 6. SHORT faults in GDI data set on humidity

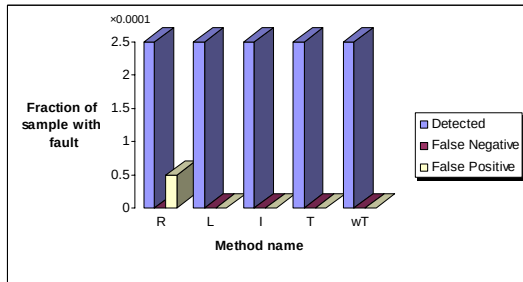


Figure 7. SHORT faults in GDI data set on pressure

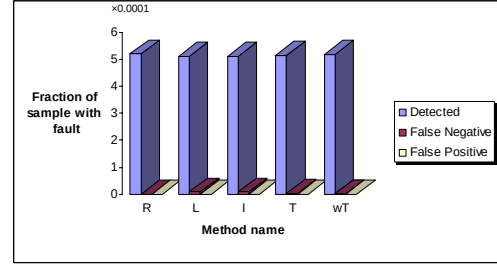


Figure 8. SHORT faults in GDI data set on light field

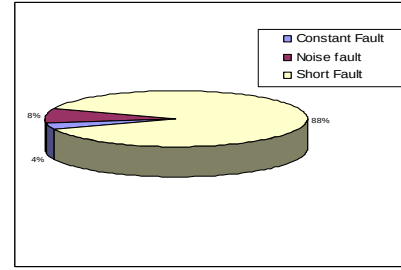


Figure 9. Percentage of faults in GDI data set

All of the above diagrams provide evidence that we can trust the results and faults that are created by our proposed method. This is a major factor to the special sensor networks which do vital and important actions such as health monitoring and battlefield and alike.

## 8. Complexity of Algorithm

Having presented our algorithms for modeling and detecting data faults based on reputation and trust, the complexity of the proposed algorithms are presented here.

A common benchmark for fault detection algorithms is the amount of communications each algorithm creates in the network. Higher communication traffic is unacceptable because it reduce the lifetime of sensor nodes and thus the lifetime of the network. We have measured the generated traffic in the network by our algorithm and discuss about it for evaluating our algorithm. To begin with, let us assume that:

$n$  is the total number of sensors,  $m$  is the number of clusters,  $B$  is the number of neighbors which are located in the radio range of a sensor, considered as a fraction  $n/m$ .

$T_{total}$  is the total time from start till now. This time is a formal unit of measurement to estimate the number of times that for example a sensor sends its data to the cluster head.

The complexity of our method for one cluster is:

$$[(\frac{n}{m} \times B) + (\frac{n}{m} \times \frac{1}{w}) + \frac{1}{w}] \times T_{total}$$

Since  $B$  can be a fraction of  $n/m$  then the total complexity is: (Although  $B$  can also be considered a constant between [3,9] for small and medium size clusters but we discuss the worst case.)

$$[(\frac{n}{m} \times B) + (\frac{n}{m} \times \frac{1}{w}) + \frac{1}{w}] \times T_{total} \approx O(n^2)$$

The communication overhead generated by SNs that send their reading data to each other in a cluster is  $\frac{n}{m} \times B$ . Sending  $FT$  and  $RT$  from every SN to the CN

need a communication thereby the needed

communication for a cluster is  $(\frac{n}{m} \times \frac{T_{total}}{w})$  such that

$w$  is a predefined period of time to sending  $RT$  and  $FT$  to the cluster head. And in  $w$  the cluster head must send their  $FT$  to the sink.

In contrast to other methods the complexity of our algorithm is independent of the failure probability of a sensor; this is an advantage in fault prone environment and fault prone equipments because in these cases this probability is high.

Our algorithm is scalable because fault detection in the first stage is done locally in SNs and in the next stage the final fault detection based on reputations is done in the cluster head.

However our algorithm has some overheads due to maintaining three tables in SN. Other discussed algorithms have overhead too. But our goal was to detect faults, their types and locations with a high degree of confidence using the reputation and trust concept, and our goal may justify the probably additional imposed overheads. In our algorithm one of the main tasks are done by cluster head and we assume these nodes which number of them is constant and low have more resource such as memory and power.

## 9. Conclusion

Fault diagnosis is an important challenge in WSNs. This paper presented how faulty sensors are detected and their types of fault are identified with a high degree of confidence using trust and reputation concepts. All SNs and CN in a cluster are principally responsible for fault detection in a distributed manner in a cluster. With simulations we showed the superiority of our algorithm to related algorithms with respect to accuracy and trustworthiness of results. The complexity of our algorithm was shown to be acceptable with considering main task of our algorithm that is fault detection with high degree of trust. It was also discussed that our algorithm is scalable thereby it can be used in vast environments with a lot of sensors.

## References

- [1] S. Chessa, P. Santi, "Comparison-Based System-Level Fault Diagnosis in Ad Hoc Networks", Proceedings of twentieth IEEE Symposium on Reliable Distributed Systems, 2001, pp. 257–266.
- [2] E. Elnahrawy and B. Nath, "Context-aware sensors", In Proceedings of the European Conference on Wireless Sensor Networks, pages 77–93, 2004.
- [3] L. B. Larkey, L.M.A. Bettencourt, and A.A. Hagberg, "In-situ data quality assurance for environmental applications of wireless sensor networks", Technical Report Unclassified Report LA-UR-06-1117, Los Alamos Laboratory, 2006.
- [4] Krishnamachari and Iyengar, "Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks", IEEE Trans. Comput. Vol.53, no.3, pp. 241–250, March 2004.
- [5] S. Chessa, P. Santi, "Crash Faults Identification in Wireless Sensor Networks", Computer Communications 25 (14) (2002) 1273–1282.
- [6] Q. Zhang, T. Yu, and P. Ning, "A Framework for Identifying Compromised Nodes in Sensor Networks", Proceedings of 2nd IEEE Communication society/ SecureComm2006, 2006.
- [7] J. Chen, S. Kher, and A. Somani, "Distributed Fault Detection of Wireless Sensor Networks", Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks, 2006.
- [8] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized Fault-Tolerant Event Boundary Detection in Sensor Networks", Proceedings of IEEE INFOCOM 2005, Miami, March 2005.
- [9] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, "Reputation-Based Framework for High Integrity Sensor Networks", ACM Transactions on Sensor Networks, Vol. V, No. N, March, 2007.
- [10] Ni, K., et als., "Sensor Network Data Fault Types", ACM Transactions on Sensor Networks, in review.
- [11] A. Sharma, L. Golubchik, R. Govindan, "On the Prevalence of Sensor Faults in Real-World Deployments", Proceedings of the forth Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Networks, pp. 213-222, 18-21 June 2007.
- [12] N. Ramanathan, L. Balzano, M. Burt, D. Estrin, E. Kohler, T. Harmon, C. Harvey, J. Jay, S. Rothenberg, and M. Srivastava, "Rapid Deployment with Confidence: Calibration and Fault Detection in Environmental Sensor Networks", CENS, Tech.Rep. 62, April 2006.
- [13] G. Tolle, J. Polastre, R. Szewczyk, D. Culler, N. Turner, K. Tu, S. Burgess, T. Dawson, P. Buonadonna, D. Gay, and W. Hong, "A Macroscopic in the Redwoods", in SenSys '05: Proceedings of the 2nd international conference on Embedded networked sensor systems. New York, NY, USA: ACM Press, pp.51–63, 2005.
- [14] A. Mainwaring, J. Polastre, R. Szewczyk, and D. C. J. Anderson, "Wireless Sensor Networks for Habitat Monitoring", in the ACM International Workshop on Wireless Sensor Networks and Applications, 2002.

# Variability Model and Management for Web Service: a Petri Net Based Approach

Li Xingyu, Hu Hao, Lü Jian  
StateKey Laboratory for Novel Software Technology,  
Nanjing University, Nanjing, Jiangsu, China  
{lixu,myou,jianlu}@ics.nju.edu.cn

## Abstract

*Variability management is an efficient solution to support service's reuse, but most of the existing variability models pay attention to service's functions and care little about the service's behavior that how these functions are invoked. In this paper, we provide a variability model based on Petri net that both the variable functions and the invocation constraint of these functions are handled. We use General Service to refer the service developed with variability for reuse; by variability binding service instance will be generated for use. Variable functions are modeled by transitions of Petri net; the service behavior is shown by the firing of transitions.*

## 1. Introduction

Variability is the ability of a system to be efficiently extended, changed, customized or configured. Building variability in web service can make the service developed with more adaptability and reusability. The existing variability models pay attention to service's functions, the variability and commonality are depicted, and static dependencies between variable functions are analyzed. But these models care little about the service behavior which is the invocation orders of these functions at runtime. At the same time, Petri net has already been used to model web service from the behavior point of view [8]. To strengthen the variability management of web service, we try using Petri net to model service with variability in this paper; both the variable functions and the invocation sequences of these functions are handled.

In this paper, we actually choose WF-net [1] which is a special case of Petri net to model the life cycle of service, it has only one input place  $i$  and only one output place  $o$ . The only one input place  $i$  corresponds to the call of service while the outplace  $o$  corresponds

to the exiting from service. Transitions of WF-net are extended to model both the service common functions and variable functions; service behavior is shown by the firing transitions; the sequence and choice structures are used to model the invocation order constraint of variable functions. The graphical expression of Petri net gives both the service developer and consumer an intuitionist view of service.

*General Service* is used to mention the service developed with variability for reuse, and *service instance* is the concrete service used by consumer.

The rest of this paper is organized as follows. Section 2 introduces variability issues and formal definition of General Service based on Petri net; Section 3 gives the variability management details that how to generate a service instance from the general service due to consumer's requirements; a case study is learned in Section 4; related work is discussed in Section 5; Section 6 is the conclusion of this work.

## 2. General Service with variability

### 2.1. Variability issues

When a variability model is given, there are some basic elements that a model should contain: they are variation point, variants, and dependencies. The variation point and variants capture the variability of service; and the dependencies constraint these variability.

**2.1.1. Variation point.** In this paper variable functions of service are modeled as variation points. The "variable function" has two meanings in web service: the function can appear in a service instance or not; the function appears in every service instance but there are various ways to implement such function. Since in the service realization, functions are represented by methods, we use service method directly instead of service function in the model.

To facilitate the service development, a classification to variation point is given. By whether a variation point is visible to end user or not there are external type and internal type; by whether a variation point is essential for the service there are mandatory type and optional type. Such definition follows Klaus Pohl [9].

Then variation point can be classified as: external and mandatory type, external and optional type, internal and mandatory type, internal and optional type.

Petri net transitions are extended to model these types' variation points. Seen as Figure 1. Then original transitions of Petri net correspond to common service methods.

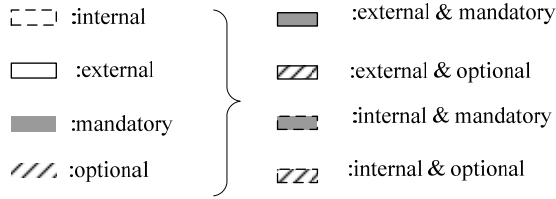


Figure 1. four types variation points

**2.1.2. Variant.** For every variation point, there can be one or more corresponding variants to show how the variability varies. For the mandatory type variation point, there are two or more corresponding variants; and for the optional type variation point, there can be one or more variants.

**2.1.3. Dependencies.** There are three type dependencies between variability issues; they are dependencies between variation points, between variation point and its variants, and between variants of the same VP. In the existed variability models, the dependency between variation points are modeled from the static point of view that can only constraint the selection of variation point. In this paper we use Petri net to constraint both the execution dependencies and the selection dependency of variation points.

*Dependency between variation points.* Require dependency, which means one variation point's execution depends on another one's successful execution, it is modeled by the sequence structure of Petri net. Exclude dependency, the execution of one variation point at runtime excludes another one's execution, if variation point A excludes variation point B, it also means B excludes A, it is modeled by the choice structure of Petri net.

*Dependency between variation point and its variants.* For a variation point, there can be one or more variants corresponding for it. These variants are classified to mandatory type and optional type, that mandatory variants are bound once its variation point is selected while the *optional* ones are not.

*Dependency between variants that belong to same variation point.* For variants that belong to one variation point, they can be mutually exclusive or coexist with each other at runtime. We use *1to1* constraint to mention the previous situation that only one variant can be chosen at runtime, and *1to m* constraint to mention the latter situation that most *m* variants can be chosen at the same time.

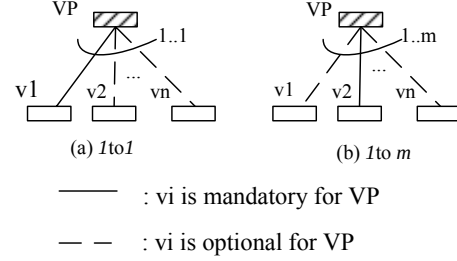


Figure 2. VP is a variation point, v1 to vn are its variants

## 2.2. Formal definition of General Service

After the prepare work, here we give the formal definition of general service and the dependencies of variability issues based on Petri net.

**Definition 1 (General Service):** A general service  $GS$  is a tuple  $(I_{GS}, P_{GS}, M_{GS}, P, T, F, i, o, TM)$  with

- $I_{GS}$  is the identification of  $GS$ ,
- $P_{GS}$  is a set of properties related to  $GS$ ,
- $M_{GS} = M_C \cup M_V$  is a set of public available methods of  $GS$ ,  $M_C$  is the set of common methods and  $M_V$  is the set of variable methods;
- $P$  is a finite set of places,
- $T = T_C \cup T_V$  is a finite set of transitions, such that  $T_C$  is the set of common transitions,  $T_V$  is the set of four types variable transitions, and  $P \cap T = \emptyset$ ,
- $F$  is a set of directed arcs and  $F \in (P \times T) \cup (T \times P)$ ,
- $i$  is the input place of  $GS$ , where for no  $t \in T$ , such that  $(t, i) \in F$ ,
- $o$  is the output place of  $GS$ , where for no  $t \in T$ , such that  $(o, t) \in F$ ,
- $TM = TM_C \cup TM_V$  is a labeling function, such that  $TM_C : T_C \rightarrow M_C \cup \{\varepsilon\}$  label a common transition with a service method's name, or  $\varepsilon$  when there is no corresponding method;  $TM_V : T_V \rightarrow M_V$  label variable point to a variable method's name,

Let  $t^*$  be a fresh identifier not in  $GS$ , the new labeled Petri net

$GS^* = (P, T \cup \{t^*\}, F \cup \{(o, t^*), (t^*, i)\}, i, o, TM \cup \{t^* \rightarrow \tau\},)$  is Strongly Connected.

Definition 4 to 7 are the prepare definitions of the Definition 8.

**Definition 4 (Service State)** The service state  $r$  of the general service  $GS$  is a set denoting the marking of the net. For every service behavior model,  $\{i\}$  is the initial marking  $M_0$  denoting service creation while  $\{o\}$  is the state of service completion.

**Definition 5 (Firing transitions)** A transition in the general service  $GS$  can be fired under Life Cycle State  $r$  when the transition is enabled. A transition  $t$  is enabled if 1) each input place of  $t$  has a token; 2) the method  $m$  associated with  $t$  is invoked. After the performance of  $t$  which is in fact the execution of the method  $m$ , all the input places are removed and the output places are added to the new Life Cycle State  $r'$ . This can be represented by  $r|t\rangle r'$ , and  $r'$  is directly reachable from  $r$ .

**Definition 6 (Reachability Graph)** A Reachability Graph  $G$  is a tuple  $(\bar{r}_0, E)$ , in which  $E \subseteq \bar{r}_0 \times \bar{r}_0 \times TM(T)$ ,  $(r_i, r_j, TM(t)) \in E$ , if and only if  $r_i|t\rangle r_j$ .

**Definition 7 (Method Sequence)** A legal Method Sequence  $\langle m_1, \dots, m_n \rangle$  can be derived from its Reachability Graph of  $GS$  by an algorithm based on depth-first-search framework. The set of all the legal method sequences of  $GS$  is denoted as  $LMS_{GS}$ .

**Definition 8.1 (Dependency between variation point):** for variation points  $VP_A$  and  $VP_B$ , their corresponding methods in variability model are  $m_A$  and  $m_B$ ,

- $VP_A$  Requires  $VP_B$ , means for any  $lms \in LMS_{GS}$ , if  $m_A \in lms$ , then  $m_B \in lms$ , and  $lms$  always shown as  $\langle \dots, m_B, \dots, m_A, \dots \rangle$ ;  $VP_A$  Excludes  $VP_B$ , means for any  $lms \in LMS_{GS}$ , if  $m_A \in lms$ , then  $m_B \notin lms$ ; and if  $m_B \in lms$ , then  $m_A \notin lms$ .

**Definition 8.2 (Dependency between variants and its variation point):** for mandatory type variant  $v$ , after the variants binding of the variation point it belongs to, there will always:

- $\exists lms \in LMS_{GS}$ , that  $v \in lms (1 \leq i \leq n)$

**Definition 8.3 (Dependency between variants):** for variants  $v_1, v_2, \dots, v_n$  that belong to one variation point, after the variants binding replace the variation point with its corresponding variants,

- The dependency between them is *1to1*, then for any  $lms \in LMS_{GS}$ , if  $v_i \in lms (1 \leq i \leq n)$ , then  $v_j \notin lms (j \neq i)$ ; the dependency between them is *1tom*, then for any  $lms \in LMS_{GS}$ , then most  $m (1 < m \leq n)$  of these variants can appear in  $lms$ .

### 3. Variability binding management

When consumer needs a concrete service to use, it can generate a service instance from the general service by variability binding. To do the binding work, a Petri net based model is given for the consumer to describe its requirements which include such information: the functions needed, and the invocation trend of these functions.

**Definition 9 (Service Request):** A service request  $SR$  is a tuple  $(I_{SR}, P_{SR}, M_{SR}, P, T, F, i, o, TM)$  with

- $I_{SR}$  is the service identification that the consumer requires,
- $P_{SR}$  is a set of properties of the required service,
- $M_{SR}$  is the set of methods that this  $SR$  requires,
- $P$  is a finite set of places,
- $T$  is a finite set of transitions, and  $P \cap T = \emptyset$ ,
- $F$  is a set of directed arcs and  $F \in (P \times T) \cup (T \times P)$ ,
- $i$  is the input place of  $SR$ , where for no  $t \in T$ , such that  $(t, i) \in F$ ,
- $o$  is the output place of  $SR$ , where for no  $t \in T$ , such that  $(o, t) \in F$ ,
- $TM : T \rightarrow M \cup \{\varepsilon\}$  is a labeling function, label transitions with methods' name of  $SR$ ,  $\varepsilon$  shows there is no corresponding method for this transition;

Let  $t^*$  be a fresh identifier not in  $SR$ , the new labeled Petri net  $SR^* = (P, T \cup \{t^*\}, F \cup \{(o, t^*), (t^*, i)\}, i, o, TM \cup \{t^* \rightarrow \tau\})$  is strongly connected.

The related definition of service request, e.g. Reachability graph and Method sequence are same to the general service.

The binding work includes three phases: check whether all the mandatory type variable points are contained in the service request; verify whether the variation points chosen by consumer follows the dependencies constraint defined in general service; reply these chosen variation points with their variants.

### 4. Case study

In this section, we use a travel assistant service example to show how this approach works.

Suppose the scenario is that: an online travel assistant can provide functions such as ticket book and hotel book to facilitate travelers. So several travel companies decide to supply such service to their customers, but these companies have different requirements for the service due to their financial constraints or management strategies. And for the same company there may be different types of users such as common users or VIP users that different quality services are needed for them. In such situation, service with variability is needed to depict and handle these companies' requirements instead of develop individual services for every company.

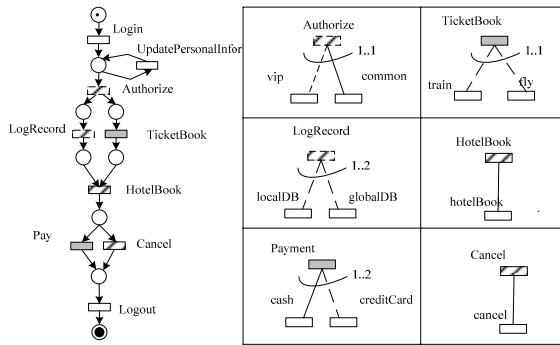
First, by collect different companies' requirements or by the experience, the developer can get the

knowledge that what functions the travel assistant service should have, then by analysis, these functions are divided to common ones and variable ones. For a travel assistant, the common functions are: login, user's personal information update, and logout; the other six functions are variable ones.

How to get the variable functions is not the attention of this paper, here we just suppose the travel assistant should have such functions: login, authorize, update user's personal information, ticket book, hotel book, pay for the book, cancel book, and make log record for the user's operation. The variation points are seen in table 1, the general service of travel assistant is shown as Figure 3.

**Table 1.** variation points of travel assistant

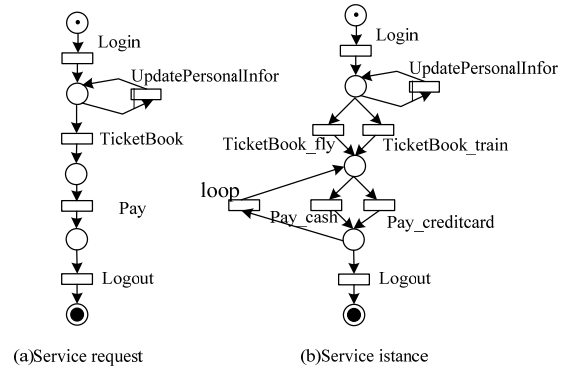
VP	Type	Variants			Dependencies between VP	
		variant		Dependency between variants	Require	Exclude
		mandatory	optional			
Authorize	IO	common user	vip	1to1		
TicketBook	EM		train, fly	1to1		
LogRecord	IO	localDataBase	globalDataBase	1to2		
HotelBook	EO	hotelBook			TicketBook	
Pay	EM	cash	creditCard	1to2	TicketBook, HotelBook	Cancel
Cancel	EO	cancel			TicketBook, HotelBook	Pay



**Figure 3.** general service of travel assistant

Then travel companies can generate service instance from the general one by their own requirements. For example a service request is given as Figure 4(a) shown, it contains the mandatory variation point “TicketBook” in general service; in the general service, “Pay” requires “TicketBook”, since  $LMS_{SR} =$

{<Login, Ticketbook, Pay, Logout>, <Login, updatePersonalInfor, Ticketbook, Pay, logout >}, the service request follows such constraint; then by replacing variation points with their variants, the service instance used by consumer is shown as Figure 4(b).



**Figure 4.** service request and the service instance

## 5. Related Work

We do the comparison work from two aspects: the model point of view and the variability management of web service.

### *Variability model.*

Variability model and management has been well researched and successfully used in Software Product Line. There are several variability modeling techniques have been developed, e.g. FODA, featureRSEB, Cardinality-Based Feature Modeling, COVAMOF, OVM [2]. Among these models, Feature Diagram[10] is a mostly used foundation to describe variability. These model techniques model variability issues following function decomposition fashion.

Comparing to these models, we use Petri net as a formal basic to model variability, the model proposed in this paper can express both functional and behavior information that handles what is variable and constraints on how to vary.

### *Variability management of web service.*

Since web services refer to the networked and modular applications as well as a set of enabling technologies such as SOAP, WSDL, UDDI and emerging Web services flow specifications like BPEL4WS, there are existing work treat variability management of web service. e.g.. [3] [4] [5][6][7].

[3] analyzes three sources of variability in web service that are WSDL documents, web service endpoints, and web service business logic. In [4] variabilities are divided to two perspectives: an architecture-oriented to describe those aspects related to protocols, environments, distribution, and a task-oriented to describe the functions of a web service. For the architecture-oriented variability, WSDL elements, UDDI items, and protocol type are modeled as variation point; for the task-oriented variability service functions are modeled as variation point. In [7], types of variability are given as: workflow variability, composition variability, interface variability and logic variability. A systematic service oriented analyzes and design process is presented for developing highly adaptable services.

These works try to handle all variability involved with web service from service entity to related technologies. Compares to these works, we pay attention to individual service entity, a formal model is given base on Petri net that variability of the service's function and behavior are represented and managed.

[5] and [6] are about the service entity, they treat web service as member of a product family, feature diagram is used to describe service component with

variability, service behavior information are not mentioned.

[3] mentions service business logic variability like service behavior in our work but didn't illustrate details; in [4] service function are modeled to show if a function exist or not, but don't show how these functions are invoked; [7] gives a process to solve the logic variability of web service but lacks the model to capture such variability.

## 6. Conclusion

This paper gives a Petri net based approach to model and manage variability in web service. General Service is given to represent service with variability; for the service consumer, it gives its requirements shown as service request to describe what functions the consumer needs, and how these functions will be used (behavior); after the variability binding, a service instance is generated to service consumer for use.

Related tool will be developed next to support this approach.

## References

- [1] W.M.P. van der Aalst, and T. Basten, "Inheritance of Workflows: An Approach to Tackling Problems Related to Change", *Theoretical Computer Science*, 270(1-2), 2002, pp. 125-203.
- [2] Marco Sinnema, and Sybren Deelstra, "Classifying variability modeling techniques ". *Information and Software Technology*, Volume 49, Issue 7, July 2007, pp.717-739
- [3] Juanjuan Jiang, Anna Ruokonen, and Tarja Systä, "Pattern-based Variability Management in Web Service Development", *ECOWS*, 2005, pp.83-94.
- [4] N. Yasemin Topaloglu, and Rafael Capilla, "Modeling the Variability of Web Services from a Pattern Point of View", *ECOWS*, 2004, pp.128-138.
- [5] Luca Balzerani, Davide Di Ruscio, Alfonso Pierantonio, and Guglielmo De Angelis, "A product line architecture for web applications", *SAC*, 2005, pp.1689-1693
- [6] Silva Robak, and Bogdan Franczyk, "Modeling Web Services Variability with Feature Diagrams", *Web, Web-Services, and Database Systems*, 2002, pp.120-128
- [7] Soo Ho Chang, and Soo Dong Kim, "A Service-Oriented Analysis and Design Approach to Developing Adaptable Services", *IEEE SCC*, 2007, pp. 204-211



- [8] Yin Qin, Hu Hao, Li Jun, Ge Jidong, and Lü Jian, "An Approach to Ensure Service Behavior Consistency in OSGi", *APSEC*, IEEE Computer Society, 2005, pp. 185-192.
- [9] Klaus Pohl, and Andreas Metzger, "Variability management in software product line engineering". *ICSE*, 2006, pp.1049-1050
- [10] Pierre-Yves Schobbens, Patrick Heymans, and Jean-Christophe Trigaux. "Feature Diagrams: A Survey and a Formal Semantics". *RE* 2006, pp.136-145.

# Feasibility of PKC in Resource-Constrained Wireless Sensor Networks

Al-Sakib Khan Pathan and Choong Seon Hong

Department of Computer Engineering, Kyung Hee University, South Korea

*spathan@networking.khu.ac.kr, cshong@khu.ac.kr*

## Abstract

*In this paper we present a detailed review of the works on public key cryptography (PKC) in wireless sensor networks (WSNs). In the early days of sensor networks, public key cryptography was thought to be completely unfeasible considering its computational complexity and energy requirements. By this time, several works have proved that the lightweight versions of many well-known public key algorithms can be utilized in WSN environment. With the expense of a little energy, public key based schemes could in fact be the best choice for ensuring data security in high-security demanding WSN applications. Here, we talk about the notion of public key cryptography in WSN, its applicability, challenges in its implementation, and present a detailed study of the significant works on PKC in WSN.*

## 1. Introduction

Security in wireless sensor network (WSN) has a great number of challenges, ranging from the nature of wireless communications, constrained resources of the sensors, unknown topologies of the deployed networks, unattended environment where sensors might be susceptible to physical attacks, dense and large networks, etc. Each of these issues in fact leads to different research direction. Nonetheless, considering all the different topics of research in WSN security, Public Key Cryptography (PKC) most probably is the hottest topic in recent times.

This is generally perceived that PKC is complex, slow, resource hungry, and as thus not at all suitable for use in ultra-low power network environments like wireless sensor networks. It is therefore a common practice to emulate the asymmetry of traditional PKC services by using a set of protocols based on

symmetric key cryptography (SKC). The main reason for using SKC is that it is comparatively less resource-hungry than PKC. However, with respect to key management, SKC could be inflexible and more difficult to handle than PKC. This is because often the keys for the sensors need to be generated in huge numbers and in many cases they need to be stored in the memories of the sensors prior to their deployment. Such type of pre-distribution of keys could really be cumbersome when a large sensor network with large number of sensors is to be deployed.

In contrast to the commonly held belief of the inefficiency and inapplicability of PKC in WSN, some researchers have come forward and proved that the lightweight versions of many well-known PKC schemes can be applied for tiny low powered sensors.

For implementing PKC-based schemes in WSN, there are mainly two approaches:

(a) Design and development of customized hardware so that they could be used in the sensor boards for supporting public key based operations and computations.

(b) Writing customized software or program that could support PKC operation or using lightweight versions of the public key based schemes.

A combination of hardware support and software optimization also works effectively for such types of implementations. So far, both from hardware and software perspectives, some works have already been done [1]. This paper presents a detailed survey on the applicable PKC-based schemes in WSN. The main objective of this work is to get a broad picture of the current research trend in applying high-standard cryptographic schemes in wireless sensor networks.

The rest of the paper is organized as follows: Section 2 briefly mentions the pros and cons of implementing public key cryptography in WSN, Section 3 presents the detailed survey on the feasible PKC-based schemes for wireless sensor networks, Section 4 discusses the advancements of sensor network technologies and future expectations, and Section 5 concludes the paper.

---

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-521-D00394). Dr. CS Hong is corresponding author.

## 2. PKC in WSN: Pros and Cons

The sensors that build up the network are usually of inadequate memory, processing, and communication capabilities and their energy sources are also very limited. As an example, Crossbow MICA2 mote [2] is a well-known sensor node with an ATmega128L 8-bit processor at 8 MHz, 128KB program memory (flash), 512KB additional data flash memory, 433, 868/916, or 310 MHz multi-channel radio transceiver, 38.4 kbps radio, 500-1000 feet outdoor range (depending on versions) with a size of only 58 x 32 x 7 (mm). Usually it is run by TinyOS operating system and powered by 2 AA sized batteries. It could be realized that a device with this configuration cannot support security mechanisms that require executing a large amount of instructions. In addition, a sensor network usually contains a large number of sensors. The number of sensors in the network might directly affect the use of memory space of nodes, because often they store pre-distributed secret keys, keying information, or the codes to calculate pairwise secret keys between any pair of nodes in the network. Node failure due to high involvement in computation related activities is another problem that could also affect the network severely. If a node is busy relatively longer than other nodes (e.g., performing huge calculations related to security), it might lose its energy rapidly and can fail much sooner than other less active nodes.

Most of the sensor network applications need at least a certain level of security for the communications among the sensor nodes and the base station (BS). The minimum requirements for any type of secure communication are: data privacy, integrity, and authenticity. All of these could be provided by using efficient cryptographic mechanisms. SKC is usually viable on sensor nodes but the size and scalability issue sometimes makes its use inefficient. Again, because of the unattended feature of WSNs, often these networks are vulnerable to physical capture attack. If the nodes carry sensitive key information within themselves, the attackers could get that after physically capturing the sensors (if any other preventive mechanisms are not used). Considering this particular issue, it is much more efficient to use PKC as in this case the nodes only carry public key materials of the BS instead of private secret keys.

## 3. Review of PKC-Based Schemes in WSN

Most of the works related to PKC in WSN are conducted to fit the low-power characteristic of the sensor nodes. Both from software and hardware

perspectives, PKC-based schemes have shown reasonable performances. In this section, we present a detailed study on these exclusive research works.

### Gura et al. [3]

Gura et al., in their work [3], present a comparative analysis of Elliptic Curve Cryptography (ECC) and RSA on 8-bit CPUs. The authors note that they implemented ECC point multiplication and modular exponentiation on two exemplary 8-bit platforms in assembly code. As the first processor, they chose a Chipcon CC1010 8-bit microcontroller which implements the Intel 8051 instruction set. The CC1010 contains 32KB of FLASH program memory, 2KB of external data memory, and 128 bytes of internal data memory. As the second processor, they took an Atmel ATmega128 processor which is frequently used for sensor network research (for example in Crossbow motes). The ATmega128 is an 8-bit microcontroller based on the AVR architecture and contains 128KB of FLASH program memory and 4KB of data memory. For ECC, they implemented point multiplication for three SECG-standardized elliptic curves with some optimizations. For RSA, they implemented RSA-1024 on both of the chosen processors and RSA-2048 on the ATmega128 with some optimizations. Their experimental findings show that on both platforms, ECC-160 point multiplication outperforms the RSA-1024 private-key operation by an order of magnitude and is within a factor of 2 of the RSA-1024 public-key operation. In summary, Gura et al.'s work for the first time gave the practical idea about the applicability of PKC algorithms in small resource-constrained devices.

### Malan et al. [4]

[4] presents the first known implementation of elliptic curve cryptography over  $F_{2p}$  for sensor networks based on 8-bit, 7.3828 MHz MICA2 mote [2]. In this work the authors first demonstrate that the secret-key cryptography is tractable on MICA2 by instrumentation of TinyOS operating system. Then with their method of implementation of multiplication of points on elliptic curves, they argue that PKI (public key infrastructure) for distribution of secret keys is also tractable. The results show that public key based scheme is viable for the modern-era sensors. Breaking the myth of the inapplicability of public key cryptography in sensor networks, this work shows that public keys can be generated within 34 seconds, and that shared secrets can be distributed among nodes in a sensor network within the same, using just over 1

kilobyte of SRAM and 34 kilobytes of ROM, which could easily be provided by the state-of-the-art sensor nodes.

#### **Gaubatz et al. [5]**

In [5], the authors propose a custom hardware assisted approach which makes PKC feasible in WSN. In order to validate their claim, they present proof of concept implementations of two different algorithms; Rabin's Scheme and NtruEncrypt.

Rabin's Scheme was proposed in 1979 in [6]. It is based on the factorization problem of large numbers and is therefore similar to the security of RSA with the same sized modulus. The size of the modulus determines the security of the cipher. The disadvantage of RSA is that the algorithm necessary to implement both encryption and decryption is computation intensive. Rabin's scheme has asymmetric computational costs. It is only necessary to perform a simple modular squaring operation to encrypt a message. Therefore encryption can be preformed relatively efficiently, while the computational cost of the decryption algorithm is comparable to RSA. This asymmetrical property of Rabin's scheme is of significant practical importance in applications such as wireless sensor networks where encryption must be done on very low powered devices. On the other hand, NtruEncrypt [7] claims to be highly efficient and particularly suitable for embedded applications such as smart cards or RFID tags, while providing a level of security comparable to that of other established schemes, in particular RSA.

Taking these asymmetric encryption techniques, the authors analyze the viabilities based on chip area, security level, delay, average power, energy per bit, and throughput. Overall, this work shows that it is possible to design PK encryption architectures with power consumption of less than 20  $\mu$ W using the right selection of algorithms and associated parameters, optimization, and low-power techniques. This work also notes that it might be possible to achieve even better performance for self-powered sensors as this work was based on a regular ASIC (Application-Specific Integrated Circuit) standard cell library that was not specifically optimized for low-power.

#### **Watro et al. [8]**

TinyPK system demonstrated in [8] shows that a PK-based protocol is feasible even for an extremely

lightweight sensor network. TinyPK is a software-based implementation of public key system tested on UC Berkeley MICA2 motes [2]. Incorporating the use of TinySec [9] or any other symmetric encryption service for sensor networks, TinyPK provides the functionality needed for a sensor and a third-party to mutually authenticate each other and to communicate securely. The mote authentication technique, proposed in this paper can be used either to collect evidence that a mote field is valid or to validate a specific mote and to link future traffic to that mote. The notion of a mote credential makes the task of a spoofing a mote network much more difficult than in an environment without credentials. In a TinyPK protected environment, a single stolen and reverse engineered mote cannot be used to impersonate other motes with different credentials. This level of protection is achieved with very little overhead and has been shown to operate on the most limiting of sensor network platforms.

#### **Du et al. [10]**

[10] investigates how to replace the public key authentication with symmetric key operations that are much more efficient. The authors show that due to a unique property of sensor networks, public keys do not need to be authenticated in the same way as it is done in the Internet environment (i.e., using certificates); instead, public keys can be authenticated using one-way hash functions, which are much more efficient than signature verification on certificates. Their scheme uses all sensors' public keys to construct a forest of Merkle trees of different heights. By optimally selecting the height of each tree, they can minimize the computation and communication costs. They also introduce a trimming scheme based on sensor deployment knowledge. The results show that their scheme can save up to 86% of the energy for the public key authentication operation in WSN.

#### **Wander et al. [11]**

In [11], the authors present an analysis of energy consumptions of PKC schemes in wireless sensor networks. They also consider the impact of public key cryptography on battery life and compare PKC to other factors influencing energy consumption, such as idle listening, data reception and transmission, symmetric cryptography, etc. They quantify the energy costs in an 8-bit microcontroller platform with RSA and Elliptic Curve Cryptography (ECC) (energy costs of digital

signature and key exchange computations for RSA-1024, ECDSA-160, RSA-2048, and ECDSA-224). This work shows that: with a given amount of energy, the authors were able to perform 4.2 times the number of key exchange operations (including mutual authentication) with ECC-160 compared to RSA-1024. Overall, this work presents a detailed picture of energy requirements for PKC operations in resource constrained sensors which could be used as a good reference paper for other similar works.

#### **Gaubatz et al. [12]**

[12] shows that special purpose ultra-low power hardware implementations of public key algorithms can be used on sensor nodes. The authors in this work selected three low-complexity PKC schemes (Rabin's scheme, NtruEncrypt, and Elliptic Curve) and for each of the schemes they developed three basic encryption architectures in TSMC 0.13 $\mu$  CMOS standard cell technology. For comparing the inherently different algorithms and their feasibility for ultra-low power implementations, they chose algorithm specific parameter sets to provide approximately the same level of security. For Rabin's scheme, a modulus of 512 bits was selected, which generally provides a security level of around 60 bits. In ECC architecture, for arithmetic operations, a 100 bit prime field was chosen which could provide a security level between 56 to 60 bits depending on the confidence level one puts into the assumption that no significant cryptanalytic progress has been made. Finally, for NtruEncrypt, system parameters  $(N, p, q) = (167, 3, 128)$  were chosen to get a security level of around 57 bits. After analyzing the architectures in conjunction with the full algorithm descriptions, they estimated the overall power and bandwidth requirements of encryption and signature primitives.

Overall, in this work, the authors show that PKC tremendously simplifies the implementation of many typical security services and additionally reduces transmission power due to less protocol overhead.

#### **Blaß and Zitterbart [13]**

Blaß and Zitterbart [13], in their work present efficient and lightweight implementations of PKC algorithms relying on elliptic curves. They checked their codes by running on popular 8-Bit ATMEGA128 microcontroller which is used for MICA2 platform.

For the lightweight implementation of ECC, one of

the most important factors is the key size. In case of ECC, key size means the size of underlying finite field that is; if 53 bit keys are to be used, the elliptic curve must be over  $F_{2^{53}}$ . If the key size is smaller, it provides less security but faster computation. After a detailed primary analysis, the authors chose 113 bit key that means a curve over  $F_{2^{113}}$ . The argument for choosing this curve was that it could offer about 16 times more security than 109 bit keys which could be considered as enough security for today's hardware. For making the entire implementation easier and lightweight, they also considered some other optimizations for memory savings, point multiplications, handcrafting a source to the target platform, sophisticated loop-unrolling, etc.

This work presents the results of the implementations of various ECC-based algorithms; ECDH (Elliptic Curve Diffie-Hellman), El-Gamal (which relies on traditional discrete logarithm problem and that has been adopted to elliptic curves and ECDLP), and ECDSA (Elliptic Curve Digital Signature Algorithm).

#### **Gupta et al. [14]**

In [14], the authors show that ECC not only makes public-key cryptography feasible on "mote"-like, embedded devices, but also it allows one to create a complete secure web server stack that runs efficiently within very tight resource constraints. They present their HTTPS stack, named Sizzle which has been implemented on multiple generations of the Berkeley/Crossbow motes where it runs in less than 4KB of RAM, completes a full SSL handshake in 1 second (session reuse takes 0.5 seconds), and transfers 1 KB of application data over SSL in 0.4 seconds.

Sizzle brings the Internet's dominant security protocol (SSL) to devices with significant computational, memory, and energy constraints. It uses highly optimized implementations of PKC to offer scalable key management and end-to-end security without sacrificing efficiency. Sizzle running on the Berkeley/Crossbow Mica2dot mote represents the world's smallest secure web server in terms of both physical dimensions and resource utilizations.

#### **Jing et al. [15]**

C4W presented in [15] is basically an identity-based public key infrastructure specially designed for WSNs.

Usually any Identity-Based Cryptosystem (IBC) requires heavy computations. However, to reduce the processing burdens, the IBC algorithm of C4W is made lightweight based on ECC optimization techniques in [3] and Combined Public Key (CPK) cryptosystem presented in [16]. The authors in this work show that their identity-based scheme consumes less energy as it is certificateless and thus it is efficient in terms of computation and communication costs.

### **Nyang and Mohaisen [17]**

A distributed and cooperative PK authentication is proposed in [17]. This work is mainly a theoretical work without practical implementation. To facilitate the proposed protocol, some modifications in the MAC frame are needed. In this cooperative mechanism, each node stores a limited number of hashed keys for other nodes which help in the authentication procedure during public key operation. According to [17], this scheme is free from any cryptographic operations, which is designed to make it fit for the constrained resources of the sensors. However the major drawback of the proposed method is that it is designed only for one hop authentication, which makes it impractical and inefficient for use in usual multi-hop WSNs. However, this scheme could be used for individual clusters with cluster heads, if clustering is used in the network. Hence, there remains enough scope of further investigation and enhancement of the protocol to make it feasible for traditional wireless sensor networks.

### **Murphy et al. [18]**

Murphy et al. [18] show that it is possible to implement PK algorithms on resource constrained sensor node platforms. Using a hardware/software co-design approach, they successfully map a public key cryptosystem based on Rabin's scheme onto the motes developed by Tyndall National Institute. The implementation was prototyped on 25mm cube modules provided by the Institute [19]. These cubes consist of several different programmable modules that are interchangeable. The cubes were configured so that each node included a low power 8-bit Atmel microcontroller, a Spartan-IIE FPGA, and a RF transceiver. For their implementation, the different modules were synchronized by a shared 4 MHz clock.

Their implementation mainly focused on efficient architectures that execute the public key algorithms using minimal resources. Their finding is that the

hardware implementation of the encryption algorithm is much faster than the software implementation. Software implementations of the algorithm are also realizable and have the benefit of low cost and high flexibility. However the time necessary to perform encryption and decryption is significantly increased by using a software-only approach.

### **Piotrowski et al. [20]**

Piotrowski et al. [20] investigate four types of nodes; MICA2DOT, MICA2, MICAz, and TelosB and estimate the power consumptions for RSA and ECC operations. Their work presents detailed results of the implementations of RSA-1024, ECC-160, RSA-2048, and ECC-224 on the mentioned platforms.

Based on the results, the authors conclude that: transmission power is not an important factor when comparing cryptographic algorithms. Even sending a 2048 bit RSA signature by a transceiver that requires 1.0  $\mu$ Ws/bit, needs not more than 2 mWs for one signature. This is at least one order of magnitude less than the energy consumption required for computation of the cryptographic operations. However, for large multi-hop networks, it might be important. In that case, a large signature increases the overall transmission power consumption in the network.

### **Yang et al. [21]**

In [21], the authors implement minimized multivariate PKC on Low-Resource Embedded Systems. They illustrate most of their minimization techniques on a current variant in the family of multivariate PKC called the Enhanced TTS (enTTS). TTS [22] is a consequence of the public-key cryptosystem TTM (Tame Transformation Method) and shares many of its superior properties, resulting in low signature delays, fast verification, and high complexity. To evaluate the performance of enTTS on modern sensor nodes, they benchmark enTTS on the Tmote Sky mote. Tmote Sky is equipped with an 8 MHz Texas Instrument MSP430 microcontroller and a Chipcon CC2420 2.4 GHz radio that supports IEEE 802.15.4 wireless low-power medium access control standard. A typical sensor node like the Tmote Sky mote has a small working RAM (10 KBytes in this case), a slightly larger read-only program memory (48 KBytes), and a relatively large flash memory (1 MBytes) for storing collected raw data and other auxiliary information for its operations. The results

from this work show that multivariate schemes could be better contenders against the established PKC schemes if they are a lot better customized and optimized for use in wireless sensor networks.

#### **Batina et al. [23]**

In [23], the authors present a low-cost PKC-based solution for security services such as key-distribution and authentication for WSNs. They propose a custom hardware assisted approach to implement ECC with the goal of obtaining stronger cryptography and minimizing the power consumption.

They also present the details of their designed Elliptic Curve Processor (ECP). Their ECP has the operational blocks: a Control Unit (CU), an Arithmetic Unit (AU), and Memory (RAM and ROM). In ROM, the ECC parameters and some constants could be stored. The RAM contains all input and output variables and it communicates with both the ROM and the ALU. The CU controls the scalar multiplication and the point operations. In addition, the controller commands the ALU which performs field multiplication, addition, and squaring.

With this hardware assisted approach, the authors present their analysis of ECC algorithm and some positive results to allow the use of PKC in WSNs.

#### **Bertoni et al. [24]**

A work on hardware implementation of PKC for elliptic curve over binary extension fields is proposed in [24]. The goal of this work is to evaluate the energy costs of different ECC implementations for low-end systems, and to propose a new hardware coprocessor architecture. In their implementation of the dedicated coprocessor, the operands had a size of 163 bits, while the existing data path was of 8 bits only. They designed two different hardware devices in VHDL, and synthesized them using the 0.18  $\mu\text{m}$  CMOS technology library by ST Microelectronics. Using the Synopsys tools Design Compiler and PrimePower, they obtained the area occupancy, the critical path, and the power consumption for each implementation.

By using their coprocessor, they show that achieving the goals of low energy consumption, reduced silicon area requirements, and significant speed-up (compared to software solutions) are possible. This could be a feasible hardware assisted solution to get improved performance of ECC implementations in WSNs without degrading other performance parameters.

#### **Arazi et al. [25]**

[25] presents an efficient ECC-based method for self-certified key generation in resource-constrained sensor nodes. In particular, the authors provide implementation results on the Intel Mote 2 sensor network platform [26] which demonstrate that such key generation can be established in the order of 60 msec while consuming less than 30 mJ.

The methodologies developed by the authors were implemented on Intel Mote 2 platform [26] which has an Intel PXA271 XScale processor running at a clock frequency ranging from 13 MHz to 416 MHz. The core frequency could be dynamically set in software, allowing the designer to carefully adjust the timing/power trade-off to optimize performance of a specific application.

For implementation of the algorithms, some of the functions used in TinyECC [27] were taken. This package targeted the MICAz platform and provided a basic library of ECC-based functions, including scalar multiplication and exponentiation operations. Also some customizations for the XScale processor (including 32-bit operation optimizations) were carried out. In addition, supplementary functions, e.g., Montgomery arithmetic, were added. All codes were written in NesC running on TinyOS operating system.

#### **Haque et al. [28]**

[28] proposes an efficient PKC based security architecture for WSNs with relatively less resource requirements. The proposed security architecture comprises basically of two parts; a key handshaking scheme based on simple linear operations and the derivation of decryption key by a receiver node. The architecture allows both base-station-to-node or node-to-base-station secure communications, and node-to-node secure communications. Analysis and simulation results show that the proposed architecture ensures a good level of security for communications in the network and could effectively be implemented using the limited computation, memory, and energy budgets of the current generation sensor nodes.

#### **Ugus et al. [29]**

In [29], Ugus et al. implement elliptic curve and finite field arithmetic operations on MICAz mote. The experimental results presented in this work show that scalar point multiplication with a random base takes

1.03s, while it takes only 0.57s in the case of fixed point multiplication, when 2 pre-computed points are employed. Moreover, they claim to achieve at least 44% faster operation than the best previous result for fixed point multiplication. The implementation of the elliptic curve EC-ElGamal encryption shows that the encryption operation only takes 1.19s, when in total 4 pre-computed points are utilized.

#### **Großschädl et al. [30]**

In [30], the authors implement two protocols. The first protocol employs a lightweight variant of the Kerberos key transport mechanism with 128-bit AES encryption. The second protocol is based on ECMQV, an authenticated version of the elliptic curve Diffie-Hellman key exchange, and uses a 256-bit prime field  $GF(p)$  as underlying algebraic structure. They evaluate the energy costs of both protocols on a Rockwell WINS node equipped with a 133 MHz StrongARM processor and a 100 kbit/s radio module. The evaluation considers both the processor's energy consumption for calculating cryptographic primitives and the energy cost of radio communication for different transmit power levels.

Their experiments show that the communication energy cost of Kerberos is between 39.5 mJ and 47.5 mJ (which basically depends on the transmit power level). In their architecture, three of the four Kerberos messages are encrypted. When using 128-bit AES, the overall energy required for encryption and decryption of the messages is less than 0.1 mJ. As a whole, the Kerberos protocol is characterized by high communication energy cost, while the energy needed for encryption or decryption is almost negligible.

They implement the ECDH/ECMQV key exchange using a 256-bit prime field as underlying algebraic structure in order to match the security level of Kerberos key transport with 128-bit AES encryption. According to their findings, a point multiplication over a 256-bit prime field takes approximately  $4.25 \cdot 10^6$  clock cycles on a StrongARM processor when implemented as their architecture. The results show that the overall energy cost of Kerberos key establishment is between 39.6 mJ and 47.6 mJ, while ECMQV key exchange requires an energy of between 79.0 mJ and 84.6 mJ. In other words, the energy consumption of ECMQV and Kerberos differs merely by a factor of between 1.78 (for high transmit power) and 1.99 (for low transmit power).

## **4. Other Works and Future Expectations**

Other than the mentioned works, [31] looks at several additive homomorphic public key encryption schemes and their applicability to WSNs. The authors in this work provide recommendations for selecting the most suitable PK schemes based on the topologies and the scenarios of WSNs. Also, in a recent work, Roman and Alcaraz [32] talk about the applicability of public key infrastructures in wireless sensor networks.

Although PKC in WSN is one of the relatively new areas, a lot of researchers have tried to deal with this issue within a short period of time. Considering all the works done so far, it can be realized that the abundant use of PKC in WSN is just a matter of time. Moreover, next generation sensor nodes are expected to have more energies. They can have ultra-low power circuitry with so-called power scavengers such as Heliomote [33], which allows continuous energy supply to the nodes. At least 8-20 $\mu$ W of power can be generated using MEMS-based power scavengers [34]. Also some other solar-based systems have been developed by this time which could even be able to deliver power up to 100mW for MICA Motes [35].

All these positive signs and achieved results indicate that: with the advancements of fast growing technology for sensors, PKC will no longer be impractical for WSNs, though still now it is expensive for the current generation sensor nodes.

## **5. Conclusions**

Public key cryptography is often ruled out considering the limitations of resources of sensors. The reason is that the data security is not the only requirement for wireless sensor networks. There are also many other aspects which require sharing of the same resources in the sensors. Hence, the main focus is to use cryptographic methods that can ensure the best level of security for sensor data with the least costs. Using of the lightweight versions of some PKC-based schemes can be the solution for this.

Various works presented in this paper show good signs of the practicality of public key cryptography. Based on our findings and current research trend, we can conclude this article with the comment that: public key cryptography is feasible for the current generation sensors with a little expense of energy. Advancements of sensor capabilities might allow abundant use of public key cryptography in wireless sensor networks in near future.



## 6. References

- [1] Arazi, B., Elhanany, I., Arazi, O., and Qi, H., "Revisiting Public-Key Cryptography for Wireless Sensor Networks", *IEEE Computer*, Volume 38, Issue 11, 2005, pp. 103-105.
- [2] [http://www.xbow.com/products/Product\\_pdf\\_files/Wireless\\_pdf/MICA2\\_Datasheet.pdf](http://www.xbow.com/products/Product_pdf_files/Wireless_pdf/MICA2_Datasheet.pdf), accessed 13 September, 2008.
- [3] Gura, N., Patel, A., Wander, A., Eberle, H., and Shantz, S. C., "Comparing Elliptic Curve Cryptography and RSA on 8-Bit CPUs", *LNCS 3156*, Springer-Verlag 04, pp. 119-132.
- [4] Malan, D.J., Welsh, M., and Smith, M.D., "A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography", *Proc. of IEEE SECON'04*, pp. 71-80.
- [5] Gaubatz, G., Kaps, J.-P., and Sunar, B., "Public Key Cryptography in Sensor Networks-Revisited", *ESAS 2004*, *LNCS 3313*, Springer-Verlag 2005, pp. 2-18.
- [6] Rabin, M.O., "Digitalized signatures and public key functions as intractable as factorization", *Mit/lcs/tr-212*, Massachusetts Institute of Technology, 1979.
- [7] Hoffstein, J. and Silverman, J.H., "Optimizations for NTRU", *Proc. of Public Key Cryptography and Computational Number Theory*, de Gruyter, Warsaw, 2000.
- [8] Watro, R., Kong, D., Cuti, S.-f., Gardiner, C., Lynn, C. and Kruus, P., "TinyPK: Securing Sensor Networks with Public Key Technology", *Proceedings of ACM SASN'04*, Washington, DC, USA, 2004, pp. 59-64.
- [9] Karlof, C., Sastry, N., and Wagner, D., "TinySec: A Link Layer Security architecture for Wireless Sensor Networks," *Proc. of ACM SenSys'04*, 2004, pp.162-175.
- [10] Du, W., Wang, R., and Ning, P., "An Efficient Scheme for Authenticating Public Keys in Sensor Networks", *Proc. of ACM MobiHoc'05*, Illinois, USA, 2005, pp. 58-67.
- [11] Wander, A.S., Gura, N., Eberle, H., Gupta, V., and Shantz, S.C., "Energy Analysis of Public-Key Cryptography for Wireless Sensor Networks", *PerCom'05*, pp. 324-328.
- [12] Gaubatz, G., Kaps, J.-P., Ozturk, E., and Sunar, B., "State of the Art in Ultra-Low Power Public Key Cryptography for Wireless Sensor Networks", *Proceedings of the Third IEEE PERCOMW 2005*, pp. 146-150.
- [13] Blaß, E. O. and Zitterbart, M., "Towards Acceptable Public-Key Encryption in Sensor Networks", *Proc. of ACM 2nd Int. Workshop on Ubiqu. Computing*, 2005, pp. 88-93.
- [14] Gupta, V., Wurm, M., Zhu, Y., Millard, M., Fung, S., Gura, N., Eberle, H., and Shantz, S. C., "Sizzle: A Standards-based End-to-End Security Architecture for the Embedded Internet", *SMLI TR-2005-145*, Sun Micro., June 2005.
- [15] Jing, Q., Hu, J., and Chen, Z., "C4W: An Energy Efficient Public Key Cryptosystem for Large-Scale Wireless Sensor Networks", *Proc. of IEEE MASS 2006*, pp. 827-832.
- [16] Tang, W., Xiang-hao, N. and Zhong, C., "Combined Public Key Systems", *Proc. of SoftCOM 2004*.
- [17] Nyang D. and Mohaisen A., "Cooperative Public Key Authentication Protocol in Wireless Sensor Network", *UIC 2006*, *LNCS 4159*, Springer-Verlag 2006, pp. 864-873.
- [18] Murphy, G., Keeshan, A., Agarwal, R., and Popovici, E., "Hardware-Software Implementation of Public-Key Cryptography for Wireless Sensor Networks", *Proc. of Irish Signals and Systems Conference*, 2006, pp. 463-468.
- [19] O'Flynn, B., Barroso, A., Bellis, S., Benson, J., Roedig, U., Delaney, K., Barton, J., Sreenan, C., and O'Mathuna, C., "The Development of A Novel Miniaturized Modular Platform for Wireless Sensor Networks", *Proc. of the fourth IPSN'05*, LA, CA, USA, April 24-27, 2005, pp 370-375.
- [20] Piotrowski, K., Langendoerfer, P., and Peter, S., "How Public Key Cryptography Influences Wireless Sensor Node Lifetime", *Proc. of ACM SASN'06*, VA, USA, pp. 169-176.
- [21] Yang, B.-Y., Cheng, C.-M., Chen, B.-R., and Chen, J.-M., "Implementing Minimized Multivariate PKC on Low-Resource Embedded Systems", *LNCS 3934*, pp. 73-88.
- [22] Chen, J.-M., Yang, B.-Y., and Peng, B.-Y., "Tame Transformation Signatures With Topy-Turvy Hashes", *Proc. of the 2nd IWAP'02*, Oct. 30-Nov. 1, 2002.
- [23] Batina, L., Mentens, N., Sakiyama, K., Preneel, B., and Verbauwhede, I., "Low-Cost Elliptic Curve Cryptography for Wireless Sensor Networks", *LNCS 4357*, pp. 6-17.
- [24] Bertoni, G., Breveglieri, L., and Venturi, M., "Power Aware Design of an Elliptic Curve Coprocessor for 8 bit Platforms", *Proc. of PERCOMW'06*, 2006, p. 337.
- [25] Arazi, O., Elhanany, I., Rose, D., Qi, H., and Arazi, B., "Self-Certified Public Key Generation on the Intel Mote 2 Sensor Network Platform", *2nd IEEE Workshop on Wireless Mesh Networks 2006 (WiMesh 2006)*, pp. 118-120.
- [26] Adler, R., Flanigan, M., Huang, J., Kling, R., Kushalnagar, N., Nachman, L., Wan, C.-Y., and Yarvis, M., "Intel Mote 2: An Advanced Platform for Demanding Sensor Network Applications", *Proc. of ACM SenSys*, 2005, p. 298.
- [27] Liu, A. and Ning, P., TinyECC: Elliptic Curve Cryptography for Sensor Networks (Version 0.1), Technical Report, September 2005, available at: <http://discovery.csc.ncsu.edu/software/TinyECC>
- [28] Haque, M. M., Pathan, A.-S. K., Choi, B. G., and Hong, C. S., "An Efficient PKC-Based Security Architecture for Wireless Sensor Networks", *Proc. of IEEE MILCOM 2007*, October 29-31, Orlando, Florida, USA, 2007.
- [29] Ugus, O., Hessler, A., and Westhoff, D., "Performance of Additive Homomorphic EC-ElGamal Encryption for TinyPEDS", Tech. Rep., 2007, available at: <http://www.ist-ubisecsen.org/publications/EcElgamal-UgHesWest.pdf>
- [30] Großschädl, J., Szekely, A., and Tillich, S., "The Energy Cost of Cryptographic Key Establishment in Wireless Sensor Networks", *Proc. of ASIACCS 2007*, pp. 380-382.
- [31] Mykletun, E., Girao, J., and Westhoff, D., "Public Key Based Cryptoschemes for Data Concealment in Wireless Sensor Networks", *Proc. of IEEE ICC*, 2006, pp. 2288-2295.
- [32] Roman, R. and Alcaraz, C., "Applicability of Public Key Infrastructures in Wireless Sensor Networks", *EuroPKI 2007*, *LNCS 4582*, Springer-Verlag 2007, pp. 313-320.
- [33] Kansal, A., Potter, D., and Srivastava, M., "Performance Aware Tasking for Environmentally Powered Sensor Networks", *In Proc. of ACM SIGMETRICS'04*, 2004.
- [34] Amirtharajah, R. and Chandrakasan, A., "Self-powered signal processing using vibration-based power generation", *IEEE Jnl. of Solid-State Circuits*, Vol. 33, pp. 687-695, 1998.
- [35] Kansal, A. and Srivastava, M., "An Environmental Energy Harvesting Framework for Sensor Networks", *Proc. of ACM/IEEE ISLPED*, 2003.

# A Calculus for Composite Authorities' Policy Derivation in Shared Domains of Pervasive Computing Environments

Morteza Amini

Department of Computer Engineering  
Sharif University of Technology, Tehran, Iran  
m.amini@ce.sharif.edu

Rasool Jalili

Department of Computer Engineering  
Sharif University of Technology, Tehran, Iran  
jalili@sharif.edu

## Abstract

*The decentralized security management in a pervasive computing environment, requires apportioning the environment into several security domains. In each security domain, an administrator (we call it authority) is responsible for specifying the security policies of the domain. Overlapping of security domains results in the requirement of cooperative security management in the shared/ overlapping domains. To satisfy this requirement, we propose an abstract security model, as well as its supplementary calculus of composite authorities. The security model is based on deontic logic and is independent of the domains' heterogeneity. The model's policy language (we call it MASL) enables multiple authorities to specify their domain policies, including obligations and authorizations. Our proposed calculus of composite authorities, enables the security system to infer policy statements of composite authorities from the cooperating primitive authorities. The calculus offers three styles of cooperative administration including collaborative, disjunctive, and delegative administration. Abstraction and automated composite authorities' policy derivation are the main advantages of the proposed logical model.*

## 1 Introduction

In pervasive computing environments, resources like information and services are accessible anywhere and anytime via any devices. Thus, users may access resources and services remotely. There are different sorts of users and services and some of them may be unknown or not predefined [13]. The distribution of resources in these environments, forces us to employ decentralized security management. In this approach, the environment is divided into the number of domains based on different factors like geographical situation. For each domain, there is a security agent with an administrator (we call it *authority*) and each resource, like

a service, can register itself in one or more domains. Thus, the authorities are responsible for preserving the security of resources that are under their protection.

The requirement of policy specification in different domains by different authorities motivated us to propose a multi-authority version of deontic logic to specify security policies including obligation policies as well as authorization ones. The unresolved problem in multiple security domains is the overlapping of domains (which results in shared domains or subdomains) and their administration issues in such a situation. The main contribution of this paper is proposing a logical system to enable cooperative administration in the overlapping or shared domains. To this aim, three styles of cooperative administration are introduced; collaborative, disjunctive, and delegative administration. The axioms that enable us to infer security policies specified by composite authorities (which are obtained based on the cooperative administration styles), as well as the semantics of the proposed syntax, and a soundness proof of the logical system are described in the rest of this paper.

The rest of this paper is organized as follows. After introducing the relative researches in the next section, we introduce our proposed security model and its logical security policy language in section 3. In section 4, the calculus of composite authorities accompanying with three styles of administration is presented. Section 5 demonstrates our proposed process for enforcing obligation policies and access control. In section 6, we present a case study. Finally, in the last section, we conclude the paper.

## 2 Related Work

Using logics to tackle the problems of specifying and proving the security of distributed systems started from 1988 by Glasgow [8]. Since then, various sorts of logic have been used to model user beliefs and inference abilities, specification of security policies, context and temporal constraints, and historical interactions in distributed envi-

ronments like pervasive computing or ubiquitous environments.

The access control logic proposed by Abadi, Lampson, and others in [2], provides a logical system (based on the modal logic) for specifying composite principals, access control lists, and access delegation in distributed systems. The first attempt in providing a general framework for authorization made by Woo and Lam in [22]. They proposed the use of Default logic, which is a kind of non-monotonic logic, to specify authorization policies. Undecidability of the logic proposed by Woo and Lam, encouraged Jajodia, *et al.*, [10, 9] to define an authorization specification language (ASL) based on the stratified first-order logic that is not only decidable but also linear. Barker, *et al.*, [5] took a similar approach in specification of multiple types of policies (with emphasis on RBAC policies) using stratified Horn-clauses logic. Freedom in using constrained negation in this language in comparison with ASL [9], encouraged the authors to leverage a partial-deduction approach for specializing access control on deductive databases in [4]. Kushik, *et al.*, [15] introduced a constraint logic programming (CLP) based framework to determine access to partial or full ontologies in order to preserve confidentiality in open World Wide Web. In this framework, policies are CLP programs (which are stratified Horn clauses with constructive negation) that prevent disclosing sensitive portions of an ontology by renaming or hiding some concepts or relationships.

We propose and use a multi-authority version of deontic logic in this paper. Using deontic logic in access control was considered by Cuppens, *et al.* [6] for information flow control. Kagal, Finin, and others proposed a policy language called Rei [14] based on Semantic Web languages (RDF and DAML+OIL) for pervasive computing environments. In Rei, concepts of deontic logic like permission, obligation, and prohibition were used. However, this policy language does not support multiple administrators' policies resolution and also cooperative administration.

In this paper we propose a kind of modal logic (multiple-authority version of deontic logic) to provide a platform for specifying security policies by multiple administrators in pervasive and distributed environments abstract from implementation details. This enables us to infer composite authorities' policies from the policies specified by primitive administrators. Multiple security domains were introduced in some papers to split the environment into several administration domains. This concept is used in [13] as a security framework in a pervasive computing environment and it is used in [3] for mobile computing environments in controlling mobile users' accesses to their home and visiting (foreign) domains. Furthermore, a policy language named X-RABAC is proposed by Joshi, *et al.* [12, 18], for multi-domain environments. In these researches the security control on shared resources is realized through the specification

of mediation policies [19] and enforcing them for inter domains access control [21]. The approach proposed in this paper, tries to get rid of specifying inter domain security policies through automatic inference of composite authorities' policies. A composite authority is a representative of an administrator that states security policies of the shared domains based on the policies of the participating individual domains.

### 3 Logical Security Model

In distributed and pervasive computing environments, the environment is usually separated into multiple security domains [13, 20]. Each security domain has a security manager, henceforth called *authority*. In each domain a set of resources (objects) are protected based on the domain's security policies. The duty of an authority in a domain is composing the security policies of the domain. It is worthwhile to note that domains may have got overlapped with each other. In this way, the owner or creator of each resource (object) can register with one or more domains and rely on their authorities' policies to secure his/her resource.

To state security policies by each domain's authority, we need to establish a policy language. MASL (multi-authority security policy language), which is presented in this paper, is a logical language that enables authorities to compose their security policy rules, and to infer implicit policy rules from the explicit ones. MASL is a kind of modal logic languages that is founded on the multi authorities version of deontic logic to specify *authorization* as well as *obligation* policies. Statements like it is obligatory that (denoted by OB), permissible that (denoted by PE), and impermissible that (denoted by IM) can be specified by this logic.

#### 3.1 Multi-Authority Security Policy Language (MASL)

The alphabet of MASL includes a set of context and conditional propositions, a set of names (terms), a set of context and conditional predicates in which each predicate  $p_k^{i_k}$  is an  $i_k$ -ary relation on terms like  $p_k^{i_k}(t_1, t_2, \dots, t_{i_k})$ , a definite set  $AU$  of primitive authorities and composite authorities resulted from composing the primitive ones (using the calculus of authorities introduced later), deontic statuses symbols including OB, PE, and IM, the propositional primitive relators  $\wedge, \neg$  (we may use other abbreviations like  $\vee, \rightarrow$ , and  $\leftrightarrow$ ), and auxiliary symbols (, and ).

**Definition 1 (Formula)** A formula is defined as follows:

- every proposition like  $x_i$  is an atomic formula.
- if  $t_1, t_2, \dots, t_k$  are terms and  $p$  is a  $k$ -ary predicate, then  $p(t_1, t_2, \dots, t_k)$  is an atomic formula.

- if  $\alpha_i$  and  $\alpha_j$  are formulas then so are  $\alpha_i \wedge \alpha_j$  and  $\neg\alpha_i$  (and analogously  $\alpha_i \vee \alpha_j$ ,  $\alpha_i \rightarrow \alpha_j$ , and  $\alpha_i \leftrightarrow \alpha_j$ )
- if  $ds$  is a deontic status,  $u$  is an authority and  $\alpha$  is a formula then  $ds_u \alpha$  is a formula.

In this definition,  $ds_u \alpha$  intuitively means an authority  $u$  declares that the status  $ds$  is established for  $\alpha$ . For example,  $OB_u \alpha$  means  $u$  states that it is obligatory (necessary) to be the case  $\alpha$ .

Following the above language, if we suppose to have an access predicate,  $\mathbf{do}(s, o, a)$  on terms of subjects (or access requesters denoted by  $S$ ), objects (or resources denoted by  $O$ ), and actions (denoted by  $A$ ), security policy rules are defined as follows.

**Definition 2 (Security Policy Rule)** A security policy rule (or in short policy rule) stated by an authority  $u$  is a formula of the form:  $\alpha \rightarrow ds_u \beta$  where  $\beta ::= \mathbf{do}(s, o, a) | ds_v \beta$ ,  $\alpha$  is a formula,  $ds$  is a deontic status,  $u$ , and  $v$  are authorities,  $a \in A$  is an action term,  $s \in S$  is a subject term, and  $o \in O$  is an object term.

Note that a security policy of the form  $ds_u \mathbf{do}(s, o, a)$  can be defined on the object term  $o$  by authority  $u$ , if objects described by the term  $o$  are registered in the security domain of authority  $u$ . In fact, MASL does not control administrative rights on specifying security policies and we should do the above control at the implementation level. Furthermore, as we describe later in this paper, a security policy like  $ds_u(ds_v \mathbf{do}(s, o, a))$  is valid whenever an authority  $u$  has administration delegation right of authority  $v$ .

In specifying security policy rules by authorities using the above language, we have this assumption that each authority is not allowed to have inconsistent policy rules. The set of policy rules specified by authorities and the statements that are describing the context, construct the knowledge base (denoted by KB) in our model.

### 3.2 Proof Theory of MASL

The axioms (except the axioms of the composition of authorities) and inference rules of MASL are as follows.

- [A1.] If  $p$  is a tautology of propositional logic,  
then  $\vdash p$  (TAUT)
- [A2.]  $\vdash OB_u(p \rightarrow q) \rightarrow (OB_u p \rightarrow OB_u q)$  (OB-MK)
- [A3.]  $\vdash OB_u p \rightarrow \neg OB_u \neg p$  (OB-MD)
- [A4.]  $\vdash PE_u p \leftrightarrow \neg OB_u \neg p$  (PE-Def)
- [A5.]  $\vdash IM_u p \leftrightarrow OB_u \neg p$  (IM-Def)
- [R1.] If  $\vdash p$  and  $\vdash p \rightarrow q$ , then  $\vdash q$  (MP)
- [R2.] If  $\vdash p$  then  $\vdash OB_u p$  (OB-MO)

Axiom A1 (TAUT) subsumes propositional logic with in our logic. Axiom A2 (OB-MK) presents that if an authority like  $u$  states that a condition is obligatory, and its antecedent is obligatory from  $u$ 's view point, then the consequence is obligatory from  $u$ 's view point as well. Axiom A3 (OB-MD) tells if  $u$  states that  $p$  is obligatory, then its negation is not obligatory from  $u$ 's view point. The axioms A4 to A5 define the deontic statuses PE, and IM based on the OB (obligation) status. We can easily take one of these four statuses as a primitive and define others. Inference rule R1 (MP) is Modus-Ponens rule. Inference rule R2 (OB-MO) states that if something is a theorem, then its obligation from any other authority's point of view is also a theorem.

### 3.3 Semantics of MASL

To present the semantics of the proposed logic, we use a Kripke structure in which domain of objects and interpretation of names and predicates are included. Note that in this paper, quantifiers (like universal and existential ones) are from meta language and they do not belong to the MASL language. A Kripke-style structure of the presented logic is a 5-tuple like  $\mathcal{M} = \langle W, R, \Phi, \Delta, I \rangle$ , where:

- $W$  is a non-empty set of possible worlds. Each world is a global authorized state in the environment.
- $R$  is an interpretation function that maps each authority to a binary relation on  $W$ . This mapping function for primitive authorities is defined as  $r = AU \rightarrow \mathcal{P}(W \times W)$  and for composite authorities is defined as an extension on  $r$  which is presented in section 4. Each relation  $r(u)$  must be reflexive and serial, i.e., for all  $u \in AU : \forall w \in W : w \in r(u)(w)$  [reflexiveness] and  $\exists w' \in W, w' \in r(u)(w)$  [seriality]. However, reflexivity is enough, because, it redounds to seriality. Extension  $R$  of function  $r$  preserves the two properties.
- $\Phi$  is an interpretation function that maps each formula to a subset of possible worlds in which the formula is correct. This function for atomic formulas is defined as  $\phi = \text{Atomic Formulas} \rightarrow \mathcal{P}(W)$  and for complicated formulas is defined inductively as follows:

$$\begin{aligned}\Phi(p) &= \phi(p), \text{ if } p \text{ is an atomic formula} \\ \Phi(\neg\alpha) &= W - \Phi(\alpha) \\ \Phi(\alpha \wedge \alpha') &= \Phi(\alpha) \cap \Phi(\alpha') \\ \Phi(OB_u \alpha) &= \{w | R(u)(w) \subseteq \Phi(\alpha)\}\end{aligned}$$

- $\Delta$  is a non-empty set of objects in all worlds of  $W$ . We assume that all possible worlds have a shared domain of objects.
- $I$  is an interpretation function that in each world  $w$  assigns to every name (term)  $t$ , a set of objects in  $\Delta$ , i.e.,

$I(w)(t) = t_w^I \subseteq \Delta$ , to every predicate  $p(t_1, \dots, t_n)$ , an  $n$ -ary relation on  $\Delta^n$ , i.e.,  $I(w)(p) = p_w^I \subseteq \Delta^n$  which is a set of  $n$ -tuples  $t_{1w}^I \times \dots \times t_{nw}^I$ .

This function must satisfy this limitation that the interpretation of a name (term) must be identical in all worlds, i.e., for all  $w, w' \in W : I(w)(t) = I(w')(t)$ .

Function  $\Phi$  must satisfy this constraint that if  $p$  is a predicate,  $w \in \Phi(p)$  if and only if there exist a set of  $n$ -tuples in the world  $w$  that the function  $I$  assigns them to  $p$ .

**Definition 3 (Truth)** A formula  $\alpha$  in a model  $\mathcal{M} = \langle W, R, \Phi, \Delta, I \rangle$  at a world  $w \in W$  is true, denoted  $\models_w^M \alpha$ , if and only if  $w \in \Phi(\alpha)$ . Analogously,  $\alpha$  at a world  $w$  is not true, denoted  $\not\models_w^M \alpha$  iff  $w \notin \Phi(\alpha)$ .

Regarding to the above semantics and definition of truth, we can prove that the presented axiomatic system is sound, i.e., if  $\vdash \alpha$  then  $\models \alpha$ . The soundness proof can be obtained by proving that the axioms are sound and inference rules preserve soundness. Due to space limitation we do not present the soundness proof of the proposed logic in this paper.

## 4 Calculus of Cooperative Administration

To enable authorities to enact, manage and enforce security policies cooperatively in their shared security domain, three styles of cooperative administration are introduced in this section including collaborative, disjunctive, and delegative administration. Before getting through the details of these three cooperative administration styles, we introduce a virtual authority, i.e., *composite authority*, as an authority of a shared domain that is managed by cooperative approach.

**Definition 4 (Composite Authority)** By having a set of primitive authorities, and  $\&, \triangleright, |$  notions, and parenthesis as alphabets, a composite authority is defined as follows:

- each primitive authority is a composite authority.
- if  $u_i$  and  $u_j$  are composite authorities, then so are  $(u_i \& u_j)$ ,  $(u_i \triangleright u_j)$ , and  $(u_i | u_j)$ .

Adding composite authorities to the proposed security policy language, i.e., MASL, and promoting it with new axioms resulted from adding the concept of composite authorities, enables us to infer security policies specified by composite authorities on shared domains. Automation of policy derivation for shared domains based on the security policies of the participating individual domains is the main advantage of such a logic-based cooperative administration framework. In the rest of this section axioms, semantics and sample soundness proof of proposed axiomatic system are presented separately for each cooperative administration style.

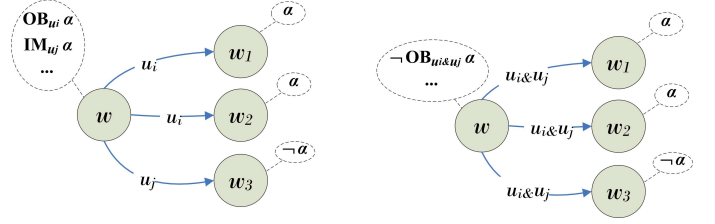


Figure 1. Semantics of collaborative composition of authorities

### 4.1 Collaborative Administration

There exist some situations in which two authorities require to collaborate in making an access decision. For example whenever two authorities have overlapped domains may need to establish security policy on the shared under protection objects in concert.

In order to support this kind of composite authority we add  $u_i \& u_j$  notion to our calculus of authorities.

**Axioms:**

- $\vdash OB_{u_i \& u_j} \alpha \leftrightarrow OB_{u_i} \alpha \wedge OB_{u_j} \alpha$  (CAU)
- $\vdash OB_{u_i \& u_i} \alpha \leftrightarrow OB_{u_i} \alpha$  (CAI)
- $\vdash OB_{u_i \& u_j} \alpha \leftrightarrow OB_{u_j \& u_i} \alpha$  (CAC)
- $\vdash OB_{u_i \& (u_j \& u_k)} \alpha \leftrightarrow OB_{(u_i \& u_j) \& u_k} \alpha$  (CAA)

Axiom CAU demonstrates that whenever two authorities make a statement collaboratively, means both of them have the same statement separately and they agree upon. The CAI, CAC, and CAA axioms show that collaborative operator ( $\&$ ) is idempotent, commutative, and associative over authorities.

**Semantics:** The extension of interpretation function  $R$  in the proposed Kripke model for  $\&$  notion is as follows:  $R(u_i \& u_j) = R(u_i) \cup R(u_j)$

We may expect intersection instead of union in the interpretation of collaboration notion ( $\&$ ). Whereas, union of relations associated to the participating authorities results in decreasing the norms which are available in possible worlds from the resulted composite authority. In other words, if  $u_i$  states in the world  $w$  that  $\alpha$  is obligatory, and  $u_j$  states the inverse one (like  $IM_{u_j} \alpha$ ), in collaborative combination of  $u_i$  and  $u_j$  we cannot infer the obligation of  $\alpha$  in the world  $w$  anymore. It is clear that in all the worlds resulted from the union of the possible worlds of  $w$  from  $u_i \& u_j$  point of view,  $\alpha$  is not true, and hence  $OB_{u_i \& u_j} \alpha$  is not true. Figure 1 shows more precisely.

**Corollary 5** Assume,  $u_i$  and  $u_j$  are authorities. Then:

- $\vdash PE_{u_i \& u_j} \alpha \leftrightarrow PE_{u_i} \alpha \vee PE_{u_j} \alpha$
- $\vdash IM_{u_i \& u_j} \alpha \leftrightarrow IM_{u_i} \alpha \wedge IM_{u_j} \alpha$

The above corollary shows that for obtaining the collaborative obligation statement on something (e.g., doing an action on a object) we require to have each of authorities' obligation statement on it. However, for permission statement, having each of the participating authorities' permission statement is enough.

## 4.2 Delegative Administration

In this paper, by *delegation* we mean *administration delegation*.  $u_i \triangleright u_j$  denotes  $u_i$  on behalf of  $u_j$  privileges can enact statements. For example,  $OB_{u_i \triangleright u_j} \alpha$  means  $u_i$  states on behalf of  $u_j$  that  $\alpha$  is ought to be the case.

It is worthwhile to note that an authority for specifying statements on behalf of another authority needs to have a privilege. Using logic, abstracts us from these implementation details. However, we suppose that trust infrastructures like PKI or delegation network which is proposed in [11] handles delegation details.

### Axioms:

$$\vdash OB_{u_i \triangleright u_j} \alpha \leftrightarrow OB_{u_i} (OB_{u_j} \alpha) \quad (DAU)$$

$$\vdash OB_{u_i \triangleright u_i} \alpha \leftrightarrow OB_{u_i} \alpha \quad (DAI)$$

$$\vdash OB_{u_i \triangleright (u_j \triangleright u_k)} \alpha \leftrightarrow OB_{(u_i \triangleright u_j) \triangleright u_k} \alpha \quad (DAA)$$

Axiom DAU represents that in delegative administration of authority  $u_i$  on behalf of  $u_j$  rights, a case like  $\alpha$  is obligatory from  $u_i \triangleright u_j$  point of view if and only if in any state,  $u_i$  accepts that  $\alpha$  is obligatory from  $u_j$ 's point of view. Axioms DAI, and DAA impose that delegation operator ( $\triangleright$ ) is idempotent, and associative. Idempotency of delegation compels the relations  $R(u_i)$  in the proposed Kripke semantics to be reflexive.

**Semantics:** The extension of  $R$  to include delegative composition of authorities is defined as  $R(u_i \triangleright u_j) = R(u_i) \circ R(u_j)$ .

The relation composition reflects this fact that from  $u_i \triangleright u_j$  point of view, a world  $w'$  is (normatively) possible from a world  $w$  if and only if  $w'$  is (normatively) possible from other world like  $w''$  in  $u_j$ 's point of view and  $w''$  is (normatively) possible from  $w$  in  $u_i$ 's point of view (figure 2). Hence,  $OB_{u_i \triangleright u_j} \alpha$  is true at a world  $w$  when  $\alpha$  is true at every world possible from  $w$  by the composite relation.

**Corollary 6** Assume,  $u_i$  and  $u_j$  are authorities. Then:

$$\vdash PE_{u_i \triangleright u_j} \alpha \leftrightarrow PE_{u_i} (PE_{u_j} \alpha)$$

$$\vdash IM_{u_i \triangleright u_j} \alpha \leftrightarrow OB_{u_i} (IM_{u_j} \alpha)$$

The above results show that we can infer that a case  $\alpha$  is permissible or gratuitous from the view point of  $u_i$  on behalf of  $u_j$ , if permissible or gratuitous statement of  $u_j$  is permissible from  $u_i$ 's point of view. However, for obligatory or impermissible cases, we do more strict and obligatory statement of  $u_i$  over the obligation or impermissible statement of  $u_j$  is required.

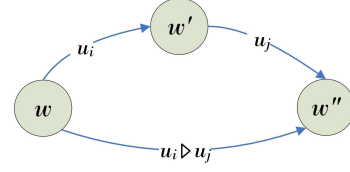


Figure 2. Semantics of delegative authority with possibility relation composition.

## 4.3 Disjunctive Administration

Disjunctive administration can be employed when two authorities like to manage the underlying shared domain objects based on the administration opinion of each of them. In this way, each authority can make a statement instead of the whole authorities who are participated in the disjunctive administration of a domain (or subdomain).

In the proposed calculus of authorities,  $u_i | u_j$  denotes  $u_i$  and  $u_j$  enact disjunctively for their underlying domains.

### Axioms:

$$\vdash OB_{u_i} \alpha \vee OB_{u_j} \alpha \rightarrow OB_{u_i | u_j} \alpha \quad (JAU)$$

$$\vdash OB_{u_i | u_i} \alpha \leftrightarrow OB_{u_i} \alpha \quad (JAI)$$

$$\vdash OB_{u_i | u_j} \alpha \leftrightarrow OB_{u_j | u_i} \alpha \quad (JAC)$$

$$\vdash OB_{u_i | (u_j | u_k)} \alpha \leftrightarrow OB_{(u_i | u_j) | u_k} \alpha \quad (JAA)$$

In disjunctive administration, an obligation statement is derived when at least one of the participating authorities enact the obligation statement. This principle is appeared as axiom JAU in the proposed calculus. Idempotency, commutativity, and associativity of disjunction operation ( $|$ ) is added by the JAI, JAC, and JAA axioms respectively.

**Semantics:** The disjunction operation is interpreted as an intersection operation in possibility relation  $R(u)$  in the proposed Kripke structure as  $R(u_i | u_j) = R(u_i) \cap R(u_j)$ .

There is a justification similar to collaborative composition for having intersection of relations  $R(u_i)$  and  $R(u_j)$  instead of union for the semantics of disjunctive composition of authorities  $u_i$  and  $u_j$ .

**Corollary 7** Assume,  $u_i$  and  $u_j$  are authorities. Then:

$$\vdash PE_{u_i | u_j} \alpha \rightarrow PE_{u_i} \alpha \wedge PE_{u_j} \alpha$$

$$\vdash IM_{u_i} \alpha \vee IM_{u_j} \alpha \rightarrow IM_{u_i | u_j} \alpha$$

$$\vdash OB_{u_i | u_j} \alpha \rightarrow PE_{u_i} \alpha \wedge PE_{u_j} \alpha$$

## 4.4 Hybrid Administration

We can have the different combination of the three aforementioned administration styles together. Complicated



composition of authorities results in some axioms which are presented in the following.

**Axioms:**

$$\vdash \text{OB}_{u_i \triangleright (u_j \& u_k)} \alpha \leftrightarrow \text{OB}_{(u_i \triangleright u_j) \& (u_i \triangleright u_k)} \alpha \quad (\text{DDC})$$

$$\vdash \text{OB}_{u_i \triangleright (u_j | u_k)} \alpha \leftrightarrow \text{OB}_{(u_i \triangleright u_j) | (u_i \triangleright u_k)} \alpha \quad (\text{DDJ})$$

Axioms DDC and DDJ show the distribution of delegation operation ( $\triangleright$ ) on collaboration ( $\&$ ) and disjunction ( $|$ ) operations respectively.

#### 4.5 Soundness

All the axioms represented for cooperative policy derivation in the above are sound. Due to the space limitation, we just show the soundness proof of axiom DDC as a sample.

**Proposition 8 (Soundness of DDC)** *Axiom DDC is sound.*

**Proof** Suppose for some model  $\mathcal{M} = \langle W, R, \Phi, \Delta, I \rangle$  and some  $w \in W$ , we have  $\models_w^{\mathcal{M}} \text{OB}_{u_i \triangleright (u_j \& u_k)} \alpha$ . Then:

$$\begin{aligned} w &\in \Phi(\text{OB}_{u_i \triangleright (u_j \& u_k)}) \\ \text{iff } \forall w', [(w, w') \in R(u_i) \circ (R(u_j) \cup R(u_k))] &\rightarrow w' \in \Phi(\alpha) \\ \text{iff } \forall w', [\exists w'', (w, w'') \in R(u_i) \wedge ((w'', w') \in R(u_j) \vee \\ &(w'', w') \in R(u_k))] \rightarrow w' \in \Phi(\alpha) \\ \text{iff } \forall w', [(\exists w'', (w, w'') \in R(u_i) \wedge (w'', w') \in R(u_j)) \vee \\ &(\exists w'', (w, w'') \in R(u_i) \wedge (w'', w') \in R(u_k))] \rightarrow w' \in \Phi(\alpha) \\ \text{iff } \forall w', [(w, w') \in R(u_i) \circ R(u_j) \vee \\ &(w, w') \in R(u_i) \circ R(u_k)] \rightarrow w' \in \Phi(\alpha) \\ \text{iff } \forall w', (w, w') \in R(u_i \triangleright u_j) \cup R(u_i \triangleright u_k) &\rightarrow w' \in \Phi(\alpha) \\ \text{iff } \forall w', (w, w') \in R((u_i \triangleright u_j) \& (u_i \triangleright u_k)) &\rightarrow w' \in \Phi(\alpha) \\ \text{iff } w \in \Phi(\text{OB}_{(u_i \triangleright u_j) \& (u_i \triangleright u_k)} \alpha) \end{aligned}$$

Thus,  $\models_w^{\mathcal{M}} \text{OB}_{(u_i \triangleright u_j) \& (u_i \triangleright u_k)} \alpha$ . Hence, the axiom is sound.

### 5 Policy Enforcement

Enforcement of policies in the security domains are done by two subsystems. *Obligation enforcement* subsystem which is used to enforce the actions which must be done in a specific situation. *Access control* subsystem which is used for controlling subjects' access requests.

#### 5.1 Access Control

Access control in individual domains is easily done based on the security policy rules specified by their authorities. However, for the shared domains, policy enforcement

strategy might be different based on the agreement made between the rightful authorities. In previous sections, cooperative administration is modeled by the concept of composite authority. Hence, for access control, we just need to infer implicit policy rules composed by composite authorities in the shared domains. For example, if an authority  $u_i$  makes agreement with  $u_j$  on having collaborative administration strategy in their shared domain, we must infer implicit policy rules composed by collaborative authority  $u_i \& u_j$ .

The access decision making functions are as follow. In these functions,  $u \in AU$ ,  $s \in S$ ,  $o \in O$ , and  $a \in A$ .

$$\begin{aligned} \text{BADF}(u, s, o, a) &= \begin{cases} \text{Deny} & , \text{ if } \text{KB} \vdash \text{IM}_u \text{do}(s, o, a) \\ \text{Grant} & , \text{ if } \text{KB} \vdash \text{PE}_u \text{do}(s, o, a) \\ \text{Don'tCare} & , \text{ otherwise} \end{cases} \\ \text{FADF}(u, s, o, a) &= \begin{cases} \text{Deny} & , \text{ if } [\text{BADF}(u, s, o, a) = \text{Deny}] \\ & \text{ or } [\text{BADF}(u, s, o, a) = \text{Don'tCare}] \\ & \wedge \text{DefSt} = \text{IM}] \\ \text{Grant} & , \text{ otherwise} \end{cases} \end{aligned}$$

BADF (Basic Access Decision Function) goes through the existing policy rules for permissible or impermissible ones and FADF (Final Access Decision Function) makes the final decision. In the definition of FADF, DefSt returns to the default access right which might be set to one of IM, or PE deontic statuses. Thus, if there is no policy rule for a request, the default strategy determines the final decision for the requested access. It is worthwhile to note that in case of existing contradictory policies (from different authorities on the same resource), we may not be able to infer an access decision. In this case, DefSt solves our problem and resolves the conflicts by enforcing the default decision.

#### 5.2 Obligation Enforcement

The issues related to the inference of obligation policies are similar to the inference of authorization ones. We follow the same approach for having cooperative administration of security obligations here. However, enforcing obligations is different from the authorizations and requires its enforcement mechanisms and services apart from the access control services.

Monitoring and enforcing obligations (or its restricted type, provisions) is one of the open problems in security management. Since, we concentrated on the specification and inference of security policies including obligation policies in this paper, we do not scrutiny the details of obligation enforcement in this paper. Some approaches for monitoring and enforcement of obligations and provisions are addressed in [7, 16, 17].

## 6 Case Study

In the MIT Oxygen project [1], a *collaborative meeting region* is defined as an area with a set of devices for meeting. This region has a set of trust and authorization rules that specify what happens during a meeting. For security management of collaborative region in Oxygen, we suggest applying our proposed model.

In a collaborative meeting room, all meeting members (who are participating in a meeting) are authorities of the meeting room (with all resources available there). Each meeting member has his/her own security policy rules. In our case, Bob and Alice are meeting members with the following security policy rules. We suppose that the meeting room has a security agent that we call it MSA in short. MSA is another authority of the room. We will see the role of MSA in this case later.

**Alice's Security Policy Rules:** Alice likes to delegate the administration to the meeting room's security agent (MSA) by transferring her security policy rules to it. She does not allow printing the confidential documents (in her domain) in any case. She also gives the read-only permission on customers' information to a meeting member when the location of meeting is outside their company. The logical representation of Alice's security policy rules are as follow. Note that in the following policy rules, each statement placed in the parentheses in the left side of a rule, is a proposition.

[AP1] (Alice's location is meeting room)  $\wedge$  (time is meeting time)  $\rightarrow$   $\text{OB}_{\text{Alice}}$  **do**(AliceSecAgent, MeetSecPolicies, transferTo(MSA))

[AP2]  $\text{True} \rightarrow \text{IM}_{\text{Alice}}$  **do**(Any, ConfDocs, print)

[AP3]  $\neg$  (requester's location is company)  $\wedge$  (location is meeting room)  $\rightarrow$   $\text{PE}_{\text{Alice}}$  **do**(MeetMember, CustInfo, read)

**Bob's Security Policy Rules:** Bob similar to Alice likes to delegate the administration to the meeting room's security agent (MSA). He does not allow anybody who is outside the company to write or update customers' information.

[BP1] (Bob's location is meeting room)  $\wedge$  (time is meeting time)  $\rightarrow$   $\text{OB}_{\text{Alice}}$  **do**(BobSecAgent, MeetSecPolicies, transferTo(MSA))

[BP2]  $\text{True} \rightarrow \text{PE}_{\text{Bob}}$  **do**(MeetMember, CustInfo, read)

[BP3]  $\neg$  (requester's location is company)  $\rightarrow$   $\text{IM}_{\text{Bob}}$  **do**(Any, CustInfo, write)

By entering Alice and Bob into the meeting room, following their first obligation policy, their agent (which is installed on their PDAs) must send all the security policies to the meeting security agent (MSA). MSA establishes its

security policies based on the received security policies as follows. In this way, MSA obligates every obligation policy specified by Alice or Bob and makes permissible every authorization policy specified by them.

[MP1]  $\text{True} \rightarrow \text{OB}_{\text{MSA}}(\text{IM}_{\text{Alice}} \text{ do(Any, ConfDocs, print)})$

[MP2]  $\neg$  (requester's location is company)  $\wedge$  (location is meeting room)  $\rightarrow$   $\text{PE}_{\text{MSA}}(\text{PE}_{\text{Alice}} \text{ do(MeetMember, CustInfo, read)})$

[MP3]  $\text{True} \rightarrow \text{PE}_{\text{MSA}}(\text{PE}_{\text{Bob}} \text{ do(MeetMember, CustInfo, read)})$

[MP4]  $\neg$  (requester's location is company)  $\rightarrow$   $\text{OB}_{\text{MSA}}(\text{IM}_{\text{Bob}} \text{ do(Any, CustInfo, write)})$

We set the security enforcement system for the collaborative meeting room to enforce the security policies based on collaborative administration of MSA on behalf of each meeting member. Formally, the desired composite authority for our case is (MSA  $\triangleright$  Alice)  $\&$  (MSA  $\triangleright$  Bob).

Suppose that in the meeting of Alice and Bob, Alice request to update a customer information through her PDA device. By her request, she sends a certificate that she is a meeting member. The security enforcement system calls FADF for checking the read and write access of Alice (as a meeting member) to customers information (CustInfo) as follows. Note that update operation is interpreted to read and write operations.

FADF((MSA  $\triangleright$  Alice)  $\&$  (MSA  $\triangleright$  Bob), MeetMember, CustInfo, read)  
FADF((MSA  $\triangleright$  Alice)  $\&$  (MSA  $\triangleright$  Bob), MeetMember, CustInfo, write)

Following MP2, and MP3 rules, BADF infers the following statement for read operation, and so returns *Grant*.

$\text{PE}_{(\text{MSA} \triangleright \text{Alice}) \& (\text{MSA} \triangleright \text{Bob})} \text{ do(MeetMember, CustInfo, read)}$

However, for write operation, since, it cannot infer neither PE nor IM, returns *Don'tCare*. Therefore, in this situation, FADF grants read operation and denies write operation following the default strategy for access control (suppose DefSt=IM).

## 7 Conclusions

In this paper, we proposed a logical security model for pervasive computing environments based on MASL logical security policy language. A pervasive computing was taken as a multi-security-domain environment in the proposed model, and three styles of cooperative administration were proposed for shared domains; namely collaborative, disjunctive, and delegative administration. Following each cooperative administration style, a set of logical axiom were presented for cooperative policy derivation.



In future steps of our research we are about to take the logical model of domains into account. In this paper we left this concept informally and just supposed that each authority is responsible of one or more domains.

## References

- [1] MIT project Oxygenom, 2004. <http://oxygen.csail.mit.edu/>, Accessed in Nov 2007.
- [2] M. Abadi, M. Burrows, B. Lampson, and G. Plotkin. A calculus for access control in distributed systems. *ACM Transactions on Programming Languages and Systems*, 15(4):706–734, September 1993.
- [3] R. Au, M. Looi, P. Ashley, and L. T. Seet. Secure authorisation agent for cross-domain access control in a mobile computing environment, 2001.
- [4] S. Barker, M. Leuschel, and M. Varea. Efficient and flexible access control via logic program specialisation. In *ACM SIGPLAN symposium on Partial evaluation and semantics-based program manipulation (PEPM'04)*, pages 190–199, Verona, Italy, 2004. ACM Press.
- [5] S. Barker and P. J. Stuckey. Flexible access control policy specification with constraint logic programming. *ACM Transactions on Information and System Security (TISSEC)*, 6(4):501–546, 2003.
- [6] F. Cuppens and R. Demolombe. A deontic logic for reasoning about confidentiality. In *3rd International Workshop on Deontic Logic in Computer Science*, pages 66–79, Sesimbra, Portugal, 1996.
- [7] P. Gama and P. Ferreira. Obligation policies: An enforcement platform. In *The 6th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY'05)*, pages 203–212, Stockholm, Sweden, 2005.
- [8] J. I. Glasgow, G. H. MacEwen, and P. Panangaden. Reasoning about knowledge and permission in secure distributed systems. In *First IEEE Computer Security Foundations Workshop (CSFW'88)*, pages 139–146, Franconia, New Hampshire, USA, 1988. MITRE Corporation Press.
- [9] S. Jajodia, P. Samarati, M. L. Sapino, and V. S. Subrahmanian. Flexible support for multiple access control policies. *ACM Transaction on Database Systems*, 26(2):214–260, 2001.
- [10] S. Jajodia, P. Samarati, and V. S. Subrahmanian. A logical language for expressing authorizations. In *IEEE Symposium on Security and Privacy*, pages 31–42, Oakland, CA, USA, 1997.
- [11] A. Josang, D. Gollmann, and R. Au. A method for access authorisation through delegation networks. In *The 2006 Australasian Workshops on Grid Computing and E-Research (ACSW Frontiers '06)*, pages 165–174, Hobart, Tasmania, Australia, 2006. Australian Computer Society, Inc.
- [12] J. B. Joshi, R. Bhatti, E. Bertino, and A. Ghafoor. Access-control language for multidomain environments. *IEEE Internet Computing*, 8(6):40–50, 2004.
- [13] L. Kagal, T. Finin, and A. Joshi. Trust-based security in pervasive computing environments. *IEEE Computer*, 34(12):154–157, 2001.
- [14] L. Kagal, T. Finin, and A. Joshi. A policy-based approach to security for the semantic web. In *2nd International Semantic Web Conference (ISWC03)*, Sanibel Island, Florida, USA, Oct 2003.
- [15] S. Kaushik, D. Wijesekera, and P. Ammann. Policy-based dissemination of partial web-ontologies. In *the 2005 Workshop on Secure Web Services (SWS '05)*, pages 43–52, Fairfax, VA, USA, 2005. ACM Press.
- [16] Z. Liu, A. Ranganathan, and A. Riabov. Specifying and enforcing high-level semantic obligation policies. In *The 8th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY'07)*, pages 119–128, Bologna, Italy, 2007.
- [17] M. C. Mont and R. Thyne. A systemic approach to automate privacy policy enforcement in enterprises. In *6th Workshop on Privacy Enhancing Technologies*, volume 4258 of *Lecture Notes in Computer Science (LNCS)*, pages 118–134, Cambridge, UK, 2006. Springer.
- [18] S. Piromruen and J. B. D. Joshi. An RBAC framework for time constrained secure interoperation in multidomain environments. In *the 10th IEEE International Workshop on Object-Oriented Real-Time Dependable Systems (WORDS05)*, pages 36–48, Sedona, Arizona, USA, 2005. IEEE Computer Society.
- [19] B. Shafiq, J. B. D. Joshi, E. Bertino, and A. Ghafoor. Secure interoperation in a multidomain environment employing rbac policies. *IEEE Transactions on Knowledge and Data Engineering*, 17(11):1557–1577, 2005.
- [20] Y. Tang, S. Zhang, and L. Li. A mobile access control architecture for multiple security domains environment. In *the International Conference on High Performance Computing and Applications*, pages 457–461, Shanghai, P.R. China, 2004. Springer.
- [21] Z. Tang, R. Li, and Z. Lu. A request-driven role mapping for secure interoperation in multi-domain environment. In *The IFIP International Conference on Network and Parallel Computing Workshops (NPC 2007)*, pages 83–90, Dalian, China, 2007.
- [22] T. Y. C. Woo and S. S. Lam. Authorizations in distributed systems: A new approach. *Journal of Computer Security*, 2(2-3):107–136, 1993.

# A Grid-Enabled Framework of Expertise Search Engine Using Web-Based Online Communities

Mohammad Mehedi Hassan<sup>1</sup>, Pil-Woo Lee<sup>2</sup> and Eui-Nam Huh<sup>3</sup>

<sup>1,3</sup> Internet Computing and Network Security Lab

Department of Computer Engineering, Kyung Hee University, South Korea

<sup>2</sup> KISTI, Daejeon, South Korea

Email: {hassan, johnhuh}@khu.ac.kr, pwlee@kisti.re.kr

## Abstract

*Existing search engines like Google, Yahoo!, Live etc. are not yet capable to answer queries that require deep semantic understanding of the query or the document. Instead, it is preferable to find and ask someone who has related expertise or experience on a topic and thus Web-based online communities have become important places for people to seek and share expertise. We need to gather the data that describes these online communities. But extracting, aggregating and analyzing data from these communities for finding experts on a single framework is a challenging task. Also there is no universal standard data structure for the outline of user participation in these communities. In this paper, we present a Grid-enabled framework of expertise search (GREFES) engine which utilizes online communities as sources for experts on various topics. This is a work in progress. We also propose an open data structure called SNML (Social Network Markup Language) to outline user participation in online communities. The architecture addresses major challenges in crawling of community data and query processing by utilizing Grid. An expert ranking algorithm is also discussed and evaluated.*

## 1 Introduction

Locating the expertise necessary to solve difficult problems is a social and collaborative problem. Recent advances in Web 2.0 [1] have enabled the proliferation of many online communities and interactive collaboration spaces like wikis, forums, blogs, scientific community and other social networking services which are gaining popularity and have become places for people to seek and share expertise [2].

Existing search engines like Google, yahoo etc. are not yet capable to answer queries that require deep semantic understanding of the query or the document [3]. Instead, it is preferable to find and ask someone who has related expertise or experience on a topic and these online communities are the places for people to seek advice or help. These online communities usually bound by shared professions, interest or products among their participants. The topics range from advice on medical treatment, programming, software, building a computer from scratch to repairing the kitchen sink. Our work seeks to create a next generation of expert finding searching framework based on these online communities.

But before such analysis can be done, we need to gather the data (questions and answers, social support or

discussion, comments or advice, content rating, social relations etc) that describes online communities. But their data representations are very different from general web pages. In online communities different users make use of different tools and social information is scattered. There are no standards for the outline of user participation. Also as these communities rarely interoperate, each is typically only aware of its own social data and cannot benefit from other communities data.

On the other hand, there are many challenges in crawling of online community data and query processing. The efficiency problem in query processing is due to the need to quickly evaluate a query over a rather large index in the presence of many user queries being submitted concurrently [4]. In case of Web crawling, efficiency problem is due to the large scale of the Web or online communities as well as the Web's constantly evolving nature, which require pages to be downloaded and indexed frequently [5]. Another problem in case of Web crawling is the freshness of the collection. Moreover, Web crawling needs a large amount of computational resources, high network bandwidth and a large amount of volatile memory to store and manage the data structures that grow quickly and continuously during the crawl [6]. Thus extracting, aggregating and analyzing data from these communities on a single framework is a challenging task.

In recent years, Grid computing has evolved to co-ordinate sharing of distributed and heterogeneous resources and to support a very powerful way of managing, processing and storing the huge amount of distributed data [7].

In this paper, we present a Grid-enabled framework of expertise search (GREFES) engine which is different from traditional ones in both design philosophy and functionality. This work is currently in progress. The framework collects, analyzes and aggregates data from different online communities or social networks available on the Web to find people with suitable expertise on various topics. A relevance ranked list of expertise results are returned by interpreting a search string and mapping it onto related keywords. The main contributions of this paper are as follows:

- A novel framework of expertise information retrieval that utilizes online virtual communities as sources of experts. It also addresses major challenges in crawling

and query processing in online communities by utilizing the computational power and high bandwidth inherently available in the Grid.

- An open data structure called SNML (Social Network Markup Language) to outline user participation in online communities.
- An expert ranking algorithm for finding experts from communities

The rest of the document is structured as follows: Section 2 briefly reviews related works. Section 3 presents the proposed Grid-enabled expert finding search engine architecture and its features. Section 4 depicts SNML and open APIs. Section 5 presents an expert ranking algorithm. Section 6 presents the evaluation of our ranking algorithm and Section 7 concludes the paper.

## 2 Related Work

### 2.1 Online Communities

Online communities are used for a variety of groups interacting via the Internet for social, professional, educational or other purposes. Examples are Flickr, Facebook, Del.icio.us, Myspace, Twitter and various forums, wikis etc. These communities do not always focused on social relationships. Instead it also reflects community member's shared interest.

In online communities, a user posts a topic or question and then some other user post replies to either participate in the discussion or to answer a question posed in the original post. The reason that users replies to a topic is usually because of an interest in the content of the topic rather than who started the thread. This also indicates that the replier has superior expertise on the subject than the asker. So there is a possibility to form an expertise network [8]

### 2.2 Expert Search

Expert finding systems have been explored in a series of studies including Krulwich and Burkey [26], and McDonald and Ackerman [27] as well as the studies in Ackerman et al. [9]. Newer systems, which use a social network to help find people, have also been explored, most notably in Yenta [28] and Referral Web [10]. These systems attempt to leverage the social network within an organization or community to help find the appropriate others.

In most recent work [11], authors propose to find expert in an organization using social network. The social networks are built from two sources, from email and extracting co-occurrences of people from web pages. An expertise propagation algorithm is proposed based on social network.

Also in [12], authors proposes an expertise search system called Arnetminer which constructs a social network among researchers through their co-authorship,

and utilizes this network information as well as the individual profiles to facilitate expertise oriented search tasks. Our approach is different from above approaches as we are utilizing already existing online communities or social networks as sources of experts.

Online social communities are often huge and there is no standard data structure to outline user participation in these communities. Therefore, crawling these networks could be both challenging and interesting. There are few researches on the crawling of online social community data. Heer and Boyd [13] mentioned their use of a crawler to gather Friendster data for their Vizster social network visualization system, but they did not go into the details of the design and implementation of the crawler. Chau, Pandit et al [14] describes parallel crawlers for online social network utilizing a centralized queue. But all these are mostly centralized systems, not suitable for geographically distributed social network services. In [30] authors define a universal data interchange format called DyNetML to enable exchange of rich social network data and improve compatibility of analysis and visualization tools. Arikan and Erdogan [25] propose an open data structure called ULML (User Labor Markup Language) for retrieving the statistics of content (how many photos, friends, comments etc.) but not for actual content. We propose a universal data structure called SNML to outline user participation in online communities.

### 2.3 Grid Enabled Search Engines

The use of the grid for information retrieval is relatively new. To the best of our knowledge, GRACE (Grid Search and Categorization) [15] and SE4SEE (Search Engine for South-East Europe) [16] are the two attempts to develop grid-enabled search engines. But our work deals with finding people with expertise rather than documents from existing online communities using Grid technology.

## 3 Proposed Framework

### 3.1 Overview of the Architecture

The GREFES infrastructure is framed by the Open Grid Services Architecture (OGSA) framework and thus relies on Service Oriented Architecture (SOA) principles and second generation Web Service standards. As the underlying grid middleware is able to distribute the work evenly, load balancing is not an issue for the current system. The framework is shown in the figure 1. The terminology used to describe the system architecture is listed in table 1. A user can access our search web portal using a web browser from his/her desktop or mobile devices like laptop, mobile phone or PDA. In order to prevent the misuse of grid resources, the user needs to register their profiles to have a valid GREFES account, which is verified by the Grid security interface in the portal. The Web portal acts as a mediator between the user

and the Grid. It converts the user query into a Grid job and submits it through a Grid user interface node (GUI) to a Grid worker node (GWN). GWNs are responsible for executing the crawler, expert ranking and indexing tasks. The generated expert ranking results are stored in the distributed GWNs databases. Search and resource broker (SRB) is not only coordinates the jobs and handles their assignments, but also responsible for submits the search query to OGSA-DAI (Open Grid Service Architecture-Data Access and Integration) after getting appropriate content source from CSS (Content Source Selection) service. OGSA-DAI deliver the ranking results to web portal to satisfy user query.

We can see that our architecture has two parts: processing part (social network data extraction, expert ranking and indexing) and expert searching part. We will present them in detail in section 3.2 and section 3.3.

### 3.2 Processing Framework

The major components of processing parts are explained briefly as follows:

*Crawler* – Since GREFES framework deals with expertise search, which serves a large number of users each with specific, personal expert needs, here a crawler is used to collect SNML files (described detail in section 4) which contain user activities (log data – comments or advice) data, profile data (may include expertise areas) and relation data from online communities. Furthermore, in order to be able to adapt to the heterogeneous nature of the Grid infrastructure, a platform independent crawler should be preferred which is capable of executing on different architectures, thus preventing the recompilation overhead and compatibility issues. Usually many crawler threads execute concurrently on GWNs in order to overlap network operations with CPU processing, thus increasing the throughput. The crawlers in Grid worker nodes retrieve community data in a breadth-first fashion.

The crawlers in Grid worker nodes retrieve community data in a breadth-first fashion. A queue data structure is used to store the list of pending users which had been seen but not crawled. Initially, a seed set of users was inserted into the queue. Then at each step, the first entry of the queue is popped, all information for that user was crawled, and every user who had left (but was not yet seen) was enqueued. Once all of the user's information was crawled, the user was marked as visited, and stored in a separate queue.

*Expert Ranking Service* – The main goal of our search engine is to find suitable experts on a given topic. So after collecting SNML files of users through crawlers from different communities, ERS service is used to find users knowledge or expert areas (If not mentioned by users) and also used to calculate their expert ranking values on these areas using a ranking algorithm. Finally, a relevance

ranked list of experts is presented to the user from different communities with similar topic. If ranking values are same for experts on a topic from different communities, community rank value is used for final ranking of experts. The community which has more members will be ranked high. We will explain more detail about the expert ranking algorithms in section 5.

*Search and Resource Broker (SRB)* – It coordinates and schedules the jobs on worker nodes and also responsible to send search query to OGSA-DAI by contacting content source selection (CSS) service.

*Indexer* – Expert ranking results are indexed by indexers and stored in the distributed worker nodes databases. It helps to quickly find suitable expertise on various topics and thus reduce query response time.

*Monitoring and Event Notification (MTEN) Service* – It monitors distributed worker nodes as well as notify them if any profile update is available in the online communities. Then resource broker schedules worker nodes to execute crawlers.

The processing work flow as shown in figure 1 is explained step by step as follows:

*Step 1 – Resource broker submits and coordinates the jobs of distributed worker nodes. A worker node runs multiple crawler threads concurrently for collecting social network data from online communities.*

**Table 1: List of commonly used terms**

Terminology	Description
Crawlers	Locating, fetching, and storing the content residing within the online communities
Indexers	The expert data is transformed to expert indexes by indexers
Grid Worker Node (GWN)	Distributed computing resource which executes crawling, expert ranking and indexing tasks
SNML	An open data structure to outline user participation in online communities
Expert Ranking Service (ERS)	Ranking experts on various topics based on a ranking algorithm
Monitoring and Event Notification	MTEN monitors Grid worker nodes and notify them if any profile changes in online communities
Web Portal	Web interface for searching expert
GSI security	Grid security interface to validate registered user
OGSA-DAI	Connect to distributed databases and deliver results to the web portal
Grid User Interface	Through this interface, Grid jobs are submitted to search and resource broker
Search and Resource Broker (SRB)	Coordinates the jobs and handles their assignments to worker nodes. Also sends search query to CSS service
Content Source Selection (CSS)	Generates and maintains content source descriptions and optimizes the distribution of queries by selecting suitable sources
Mashup Service	Integrates services like Google map to display

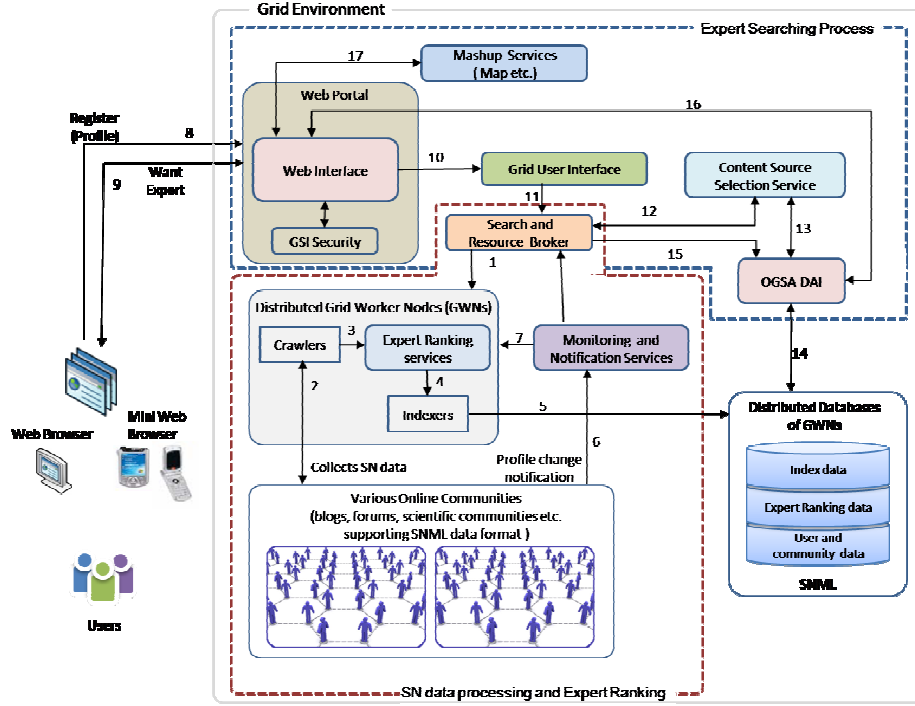


Fig. 1: The Next Generation Expert Finding Search Engine Framework on Grid

Step 2 – After collecting and processing SN data, expert ranking tasks are performed in the worker nodes using our ranking algorithm to find experts on various topics in different online communities.

Step 3 – Then the indexers in the worker nodes create expert indexes for efficient query and store them in worker node databases.

Step 4 – If any profile changes in online communities, MTEN service notifies worker nodes which then updates the expert ranking results.

### 3.3 Searching Framework

This framework provides a relevance-ranked list of experts to user on a topic. The major components of this framework are explained as follows:

**Web Portal** – It is the only interaction point between the user and the GREFES back-end and thus a major component of our architecture. Users register their profiles as well as submit their queries through the portal. Queries are submitted as a Grid job through Grid user interface (GUI). The portal has to be user-friendly, even though it requires a more complex interface than classic search engines due to the application's added capabilities like mashup services (i.e. Google map etc.).

**Content Source Selection (CSS) Service** – It generates and maintains expert content source descriptions which include partial content indices, summative content indices, or result traces from training or past queries. It uses CORI [17] selection algorithm (the standard collection selection

mechanism) and Language Modeling techniques [18] to find appropriate expert index collections. Thus it optimizes the distribution of queries by selecting target expert index sources which are relevant to a given query.

**OGSA-DAI** – Open Grid Service Architecture-Data Access and Integration [19] is an extensible framework for data access and integration. It provides three basic activities- querying data, transforming data, and delivering the results using ftp, e-mail, and push services. In our architecture, OGSA-DAI is used to query expert indexes from distributed worker nodes databases and deliver the query results to the web portal.

The workflow of the expert search process is shown in figure 2 and is explained step by step as follows;

The first stage of the process is the warm up sequences denoted in the figure 2 as steps (1)-(6) and steps (7) – (13) denote the querying sequence

Step 1: The SRB contacts the CSS to initialize the service of collecting expert index descriptions

Step 2: The CSS contacts the OGSA-DAI to acquire descriptions of each expert index.

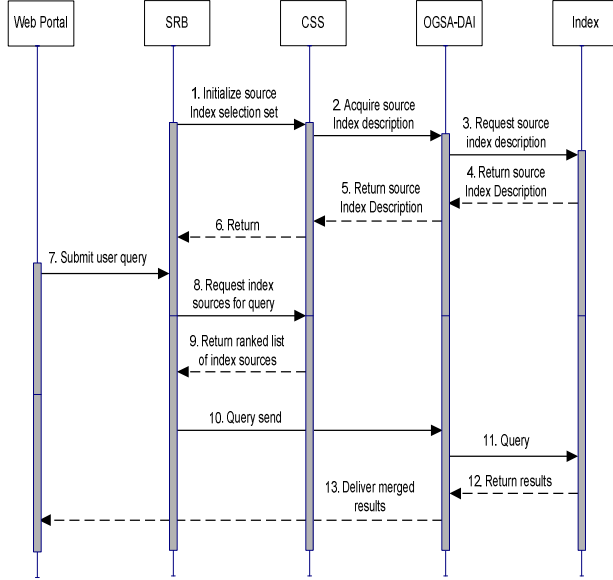
Step 3: The OGSA-DAI requests the description from the expert index

Step 4 and 5: The expert index provides the collection statistics in a pre-defined format to the CSS through OGSA-DAI.

Step 6: Steps (2)-(5) are repeated for each index, until all the expert indexes are described, and the warm up of the CSS is completed.

Step 7: A query is submitted to the SRB through the web portal.

Step 8: The query is sent to the CSS for selection of the set of expert indexes.



**Fig. 2: Sequence diagram of the search process**

Step 9: The CSS returns a ranked list of expert indexes. This may be the top  $n$  indexes which are most likely to satisfy the query, or all those indexes above a certain threshold. These criteria can be defined when performing the selection.

Step 10 and 11: The selected indexes are then contacted through OGSA-DAI and queried for finding relevant experts on a topic. This process is performed concurrently for efficiency (via threading).

Step 12 and 13: The expert ranking result sets, produced by each index from distributed worker node's databases, are merged and final ranking results of experts are delivered to the web portal through OGSA-DAI.

### 3.4 Architectural Features

The distinguishing features of proposed framework are as follows:

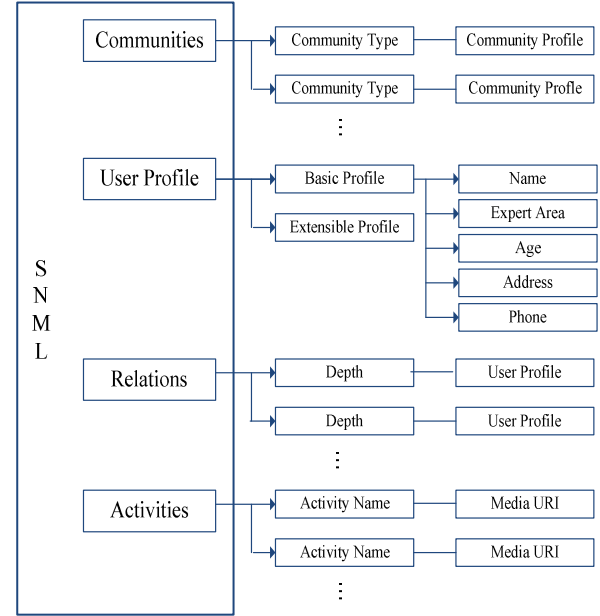
- 1) Utilizing online Web communities as sources of expert
- 2) The computational burden of Web crawling and indexing is alleviated by the utilization of resources available in the Grid.
- 3) Queries are run against multiple distributed collections of data sources in parallel with a likelihood of possessing relevant sources, thus eliminating the need of searching other unnecessary collections.
- 4) Rather than seeking sub-second response times over billions of documents, as monolithic search engines

do, our system seeks such performance over far smaller collections. This will enable more complex query processing.

- 5) Because of their smaller size and locality to a particular collection source, the distributed collections can be updated more quickly, thus eliminating the delay among harvest runs exhibited by search engines.

### 4 SNML Data Format

SNML is an open data structure protocol based on XML that can be generated by the service provider and is followed by our framework. Crawlers collect SNML documents which describes community data in a universal format. SNML document contains community information, user's participation information like his/her experiences, expertise areas, various activities, relations etc. SNML consist of four sections: as shown in figure 3 and describes briefly as follows:



**Fig. 3: SNML Data structure with four sections**

**<Community>**: The community a user is belongs to. It consists of community type and community profile. Community type describes type of media example blog, wikis, forum etc. and community profile describes the community activities or services.

**<Profile>**: Consists of basic profile and extensible profile. Basic profile includes personal information like name, age, expertise area, phone, address etc. Extensible profile depends on different social network or communities. For example Facebook's profile data which includes other information than basic.

**<Relation>**: Consists of depth and connected user profiles to the current user. It may include no. of friends connected to the user, their profile etc.

**<Activities>**: Consists of activity name and media URI (Universal Resource Identifier). It includes user comments, posts etc.

The SNML data format is shown in figure 4.

```
<? xml version="1.0"?>
<SNML xmlns="urn:GIR:2008:08-SNML-NS"
 xmlns:SNML="urn:GIR:2008:08-SNML-NS">
 <community mimeType="text/plain">
 <type/>
 <profile/>
 </community>
 <profile mimeType="text/plain">
 <basic>
 <name/>
 <expertise area>
 <item/>
 <descriptor/>
 </expertise area>
 <age/>
 <address/>
 <phone/>
 </basic>
 <extension>
 <!-- extensible profile -->
 </extension>
 </profile>
 <relation mimeType="text/plain">
 <depth/>
 <profile>
 <!-- user profile based on profile in SNML -->
 </profile>
 </relation>
 <activities mimeType="text/plain">
 <name/>
 <uri/>
 </activities>
</SNML>
```

**Fig. 4: SNML Data Format Structure**

## 5 Expert Ranking Algorithm

Most current systems use modern information retrieval techniques to discover expertise from implicit or secondary electronic resources. Campbell et al. [20] and Dom et al. [21] used graph-based ranking algorithms in addition to content analysis to rank users' expertise levels. However, there is a weakness in these studies. The size of their networks is very small and does not reflect the characteristics of realistic social networks.

Yu and Singh's referral system [22] explicitly argues for a specific expertise-finding algorithm. In their experiment, they use the similarity between a query vector and a neighbor's expertise vector, plus some consideration of its historical referring performance, as the criteria for picking the next agent in a referral graph. However, the query vector and expertise vector in their experiment are manually coded; and each is a combination of preset

topics from a taxonomy. This approach is not practical in real world scenarios:

Adamic et al's best connected search (BCS) algorithm [23], which makes use of the skewed degree distribution of many networks, can also be used to find experts. However, Adamic et al. also found that the BCS algorithm is not always efficient in all networks.

Here we need to find experts in online communities. We also use graph-based algorithm. The task of determining expert is naturally decomposed into two stages.

1. Discovering and identifying possible knowledge areas.
2. Measuring the person's competency in these areas compare to others that is finding his/her expertise ranking.

We represent the profile of a person  $pr$  as a vector of  $n$  values, where the  $i$ th element of the vector corresponds to the knowledge area  $ka_i$ , and  $score(pr, ka_i)$  reflects the individual's knowledge in the given area. To find the knowledge areas of a candidate, we use keyword similarity of candidates and knowledge areas as described in [31]. After parsing the questions, answers, comments or activities, we can extract the top 20 keywords using the TF-IDF weighting formula [29]. Let  $KW(d)$  be the set of keywords that are extracted from profile document  $d$ .

We can obtain the set of keywords of a knowledge area  $ka$  by looking at a query-biased subset of documents  $D_{ka}$ , using the top  $n$  documents retrieved using  $ka$  as the query and the W3C corpus as the document collection. Then we use all keywords from these documents to form a list of keywords of the given area. Formally:

$$KW_{ka} = \bigcup_{d \in D_{ka}} KW(d) \dots\dots\dots (1)$$

Similarly, we can obtain keywords of individuals, by taking keywords of profile documents that they are associated with:  $KW_{pr} = \bigcup_{d \in D, A(d, pr)=1} KW(d) \dots\dots\dots (2)$

Having the set of keywords both for knowledge areas and for candidates in hand, we estimate the profile scores with the ratio of co-occurring keywords:

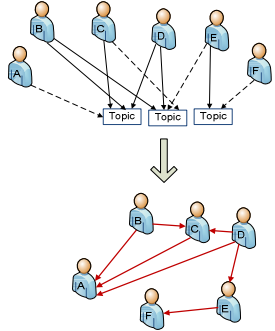
$$score(pr, ka) = \frac{|KW_{ka} \cap KW_{pr}|}{|KW_{ka}|} \dots\dots\dots (3)$$

After determining his knowledge areas based on the highest score, we now need to calculate his expertise ranking on this area.

Online communities can be shown as a post-reply network as shown in figure 5. In this figure, we map a replying relationship into a directed graph. On the left we have a bipartite graph of users (circles) and the discussion threads (squares) they participated in. This is transformed to a directed graph where an edge is drawn to the user making the initial post (the dashed edge) from everyone who replied to him. From figure 5, a user's expertise may be found by counting the number of questions he/she answered. We call it the "AnswerNum" measure. There is a potential problem in counting the number of answers one posted. A user who answers 100 newbies' questions will be ranked as equally expert as another user who answers



100 advanced users' questions. Obviously the latter usually has greater expertise than the former. We propose to generate a measure that not only considers how many other people one helped, but also whom he/she helped. We call it  $Exp_{rank}$ . From figure 5, we can see that if C is able to answer A's question, and B is able to answer C's question, B's expertise rank should be boosted not just because they were able to answer a question, but because they were able to answer a question of someone who herself had some expertise.



**Fig. 5: Post-reply network**

This is similar to PageRank. In a sense,  $Exp_{rank}$  propagates expertise scores through the question answer network. So the proposed  $Exp_{rank}$  algorithm is described as follows:

Suppose User B has answered 50 questions and helped users  $U_1 \dots U_n$ . Then  $Exp_{rank}(B)$  is calculated as follows:

$$Exp_{rank}(B) = (1-d) * NA(B) + d * \sum_{i=1}^n NA(U_i) / H(U_i) \dots (4)$$

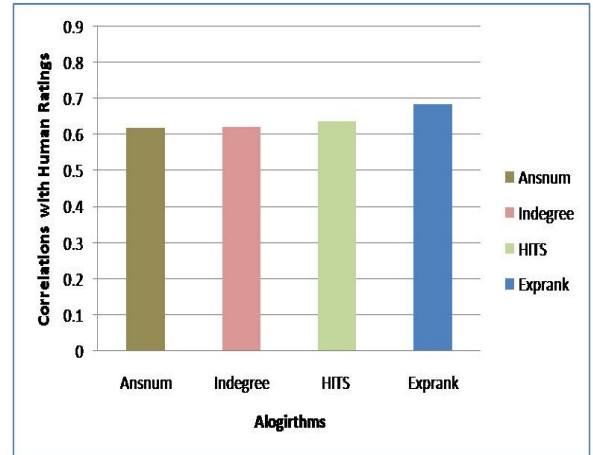
where  $NA(B)$  = no. of replies or answers of user B,  
 $NA(U_i)$  = no. of answers of user  $U_i$ ,  $i=1..n$   
 $H(U_i)$  = no. of users helping  $U_i$  and the parameter  $d$  is a damping factor which can be set between 0 and 1. The damping factor allows the random walker to 'escape' cycles by jumping to a random point in the network rather than following links a fraction  $(1-d)$  of the time.

## 6 Evaluation of the Expert Ranking Algorithm

For the evaluation of the ranking algorithm, we collected the data from a Korean Social network website called "www.rukie.com". This site mainly focuses on different programming issues and user asks questions and some people answer them. Since there was no explicit user-supplied expertise ranking data in this site, we needed to use human raters to generate a "Standard" for comparison. Because it was not possible for us to rate a large number of these users, we randomly selected 130 users from the network for use as a comparison sample. By omitting those users posting fewer than 10 times, we ensured that the sampled users had generated enough Forum content for a reviewer to evaluate their expertise levels. Raters

must read from ten to hundreds of messages posted by a user to evaluate his/her expertise level. It is also difficult to compare two users when they both have posted many messages but have not replied to each other. Based on our observation of the forum and the results of a pilot rating set, we decided to categorize the users into 3 expertise levels like professional programmers, Learners (who knows some basics) and newbies.

We found two raters who are programming experts to rate the 130 users' expertise. We use Kendall's Tau correlation measure in our evaluation [32]. Kendall's Tau gives equal weight to any interchange of equal distance, no matter where it occurs. For instance, an interchange between rank 1 and 2 will be just as bad as interchange between rank 100 and rank 101. The Kendall's Tau distance between the two human raters was 0.752 for the 130 users. Therefore, we may expect that any automated algorithm would at best achieve around a 0.752 correlation with the human raters. For each of these users, in the data analysis below, we summed the ratings from the two raters together as the standard human rating (HR). Figure 6 shows the Kendall's Tau statistical correlations between various algorithms and the human ratings of the 130 users. From the figure, one can see that our  $Exp_{rank}$  algorithm performs well than other existing ones.



**Fig. 6: The performance of various algorithms in Kendall's Tau statistical metric**

## 7 Conclusion and Future Work

No prior work done in the expertise search domain has considered Grid-based search engine to utilize web communities as sources of experts. Many of them find experts in an organization or in a social community or build social networks from emails or web pages to find relevant experts on a topic. So in this paper, we have presented a novel framework of expertise information retrieval on Grid utilizing online web communities. The system returns a relevance ranked list of experts by



interpreting a search string (topic) and mapping it onto related keywords. We have proposed an open data structure called SNML for sharing community data efficiently and effectively. Also we have presented an expert ranking algorithm and evaluated its performance. Currently this work is in progress. In future, we will analytically evaluate the performance of our architecture with other distributed information retrieval system.

## Acknowledgement

This research was partially supported by KISTI (Korea Institute of Science and Technology Information) and the MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement) (IITA-2008-C1090-0801-0002)

## References

- [1] T. O'Reilly. *What Is Web 2.0, Design Patterns and Business Models for the Next Generation of Software*, 2005
- [2] J. Zhang and M. S. Ackerman. *Searching For Expertise in Social Networks: A simulation of Potential Strategies*. ACM GROUP'05, Nov, 2005, USA
- [3] J. Bian, Y. Liu, E. Agichtein and H. Zha. *Finding the Right Facts in the Crowd: Factoid Question Answering over Social Media*. ACM WWW 2008, April 21-25, Beijing, China
- [4] B. B. Cambazoglu and C. Aykanat. *Performance of query processing implementations in ranking-based text retrieval systems using inverted indices*. Information Processing & Management, 42(4), 875-898, 2006
- [5] B. B. Cambazoglu, Turk and C. Aykanat. *Data-parallel Web crawling models*. Lecture Notes in Computer Science, 3280, 801-809, 2004
- [6] J. Cho and H. Garcia-Molina. *The evolution of the Web and implications for an incremental crawler*. In Proceedings of the 26<sup>th</sup> international conference on very large data bases (pp. 200-209). 2000, Cairo, Egypt
- [7] I. Foster and C. Kesselman. *The grid 2: Blueprint for a new computing infrastructure*. San Francisco: Morgan Kaufman, 2003
- [8] J. Zhang, M. S. Ackerman and L. Adamic. *Expertise Networks in online Communities: Structure and Algorithms*. WWW 2007, May 8-12, Banff, Alberta, Canada
- [9] M. S. Ackerman, V. Wulf and V. Pipek, (eds.). *Sharing Expertise: Beyond Knowledge Management*. MIT Press, 2002.
- [10] H. Kautz, B. Selman. and M. Shah. *Referral Web: combining social networks and collaborative filtering*. Commun. ACM, 40 (3). 63-65
- [11] Y. Fu, R. Xiang, Y. Liu, M. Zhang and S. Ma. *Finding Experts Using Social Network Analysis*. IEEE/WIC/ACM Intl. Conf. on Web Intelligence, 2007
- [12] J. Li, J. Tang, J. Zhang, Q. Luo, Y. Liu, (eds.) *EOS: Expertise Oriented Search Using Social Networks*. WWW 2007, May 8-12, 2007, Banff, Alberta, Canada
- [13] J. Heer, and D. Boyd. *Vizster: Visualizing Online Social Networks*. IEEE Symposium on Information Visualization (InfoVis), 2005
- [14] D. H. Chau, S. Pandit et al. *Parallel Crawling for Online Social Networks*. ACM WWW 2007, May 8-12, Banff, Alberta, Canada
- [15] F. Scholze, G. Haya, J. Vigen, and P. Prazak. *Project GRACE: a grid based search tool for the global digital library*. In 7<sup>th</sup> international conference on electronic theses and dissertations. Lexington, KY, 2004
- [16] B. B. Cambazoglu, E. Karaca, T. Kucukyilmaz, A. Turk and C. Aykanat. *Architecture of a Grid Enabled Web Search Engine*. Journal of Information Processing and Management 43 (2007) 609-623, Elsevier
- [17] J. P. Callan, Z. Lu, and W. B. Croft. *Searching distributed collections with inference networks*. In SIGIR '95: Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, pages 21-28, 1995
- [18] L. Si, R. Jin, J. Callan, and P. Ogilvie. *A language modeling framework for resource selection and results merging*. In CIKM '02: Proceedings of the eleventh international conference on Information and knowledge management, pages 391- 397, 2002.
- [19] www.ogsadai.org
- [20] C.S. Campbell, P. P. Maglio, A. Cozzi and B. Dom. *Expertise identification using email communications*. In the twelfth international conference on Information and knowledge management, New Orleans, LA, 2003, 528-231
- [21] B. Dom, I. Eiron, A. Cozzi and Y. Zhang. *Graph-based ranking algorithms for e-mail expertise analysis*. In DMKD, New York, NY, 2003, ACM Press, 42-48.
- [22] B. Yu, and M. P. Singh. *Searching Social Networks*, Proceedings of Second International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2003, 65-72.
- [23] I. A. Adamic, and E. Adar. *How to search a social network*. Social Networks, 27(3), 2005, 187-203.
- [24] I. Guy, M. Jacovi, E. Shahar S. Farrell et al. *Harvesting with SONAR- The Value of Aggregating Social Network Information*. ACM CHI 2008, April 5-10, 2008, Florence, Italy
- [25] B. Arikan and E. Erdogan. *A framework for sustaining user labor across the web*. 2008. www.userlabor.org
- [26] B. Krulwich. and C. Burkey. *ContactFinder agent: answering bulletin board questions with referrals*. In the 13<sup>th</sup> National Conference on Artificial Intelligence, Portland, OR, 1996, 10-15
- [27] D. W. McDonald and M. S. Ackerman. *Expertise Recommender: A Flexible Recommendation Architecture*. Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '00), 2000, 231-240.
- [28] L. N. Foner, Yenta. *A multi-agent, referral-based matchmaking system*. In Proceedings of Agents '97, ACM Press, Marina del Rey, CA, 1997, 301-307
- [29] B. Yates and R. Neto *Modern Information Retrieval*. ACM Press, 1999
- [30] M. Tsvetov, J. Reminga and Kathleen Carley. *DyNetML: A Robust Interchange Language for Rich Social Network Data*. Institute for Software Research, International Carnegie Mellon University, 2005
- [31] K. Balog and M. Rijke. *Determining Expert Profiles*. IJCAI, 2007
- [32] Herlocker, J.L., Konstan, J.A., Terveen, L.G. and Riedl, J.T. *Evaluating collaborative filtering recommender systems*. ACM Transactions on Information Systems, 22 (1). 5-53

# Disassembling SLAs for Follow-Up Processes in an SOA System

Chih-Hao Hsu<sup>1</sup> Yun-Wei Liao<sup>2</sup> Chien-Pang Kuo<sup>3</sup>

Institute for Information Industry  
{ eden ,lyw, jacob }@iii.org.tw

## Abstract

This work presents a novel scheme for Service Level Agreements (SLAs) in a service-oriented architecture (SOA). An SLA is a contract that guarantees quality of service (QoS) between service providers and consumers. An SLA can be split into several parts for an SOA system which is consisting of different services. A split SLA is used to guarantee quality of services for one of services in an SOA system. However, these contract rules do not regulate follow-up process in services. SLA violations may occur consequently. For instance, interactions between services are not encrypted because the encryption parameter isn't included in the split SLA contract. Confidential data would be transmitted through unsafe channels. There would be a security issue of SLA Management for follow-up processes in an SOA system. The novel scheme in this work addresses these problems. Using SLAs management proposed in this work can identify the services that consumer needs and split the contracts for follow-up services. SLAs are not only guaranteed between service providers and consumers but also assured in follow-up processes. The proposed method improves the security and completeness of SLA using in an SOA system.

**Keywords** : Service Level Agreement, Service-oriented architecture

## 1. Introduction

An increasing number of business applications are currently deployed with network topology. To fulfill the requirements of customers all over the world, companies need a

mechanism with increased efficiency for software development. Therefore, Service Oriented Architecture (SOA) was proposed. SOA could easily integrate different business processes for one specific purpose. Services programmed in different languages are deployed on the web and served for different user requests. Additionally, every web service is specific to one purpose and has no common interface for communication. Entire system must be redesigned to fit new goals when the purpose differs. SOA, which is a low-cost and time-saving method for building services, is loosely coupled between web services. The services in an SOA system are already tested and verified. They use the same interface to interact with each other. Even when some services are replaced, the system will be online within a short time.

When a service is fully developed, the relationship between service consumers and providers must be stipulated. SLAs as contracts between service providers and consumers are proposed. An SLA uses formal terms to define the responsibilities of service providers and service consumers. Such contracts also outline penalties for service providers who do not reach the level of quality defined in the contract. Service quality must be measured and monitored to ensure it meets parameters defined in an SLA. Once violations occur, service providers have to adjust system resources to meet parameters in the contract.

An SLA in an SOA system only exists between service consumers and providers. However, there is not any limitations existing in follow-up processes and potential agreement violations may occur.

---

\* This study is conducted under the Intelligent Web - enabled Service Research and Development Project of the Institute for Information Industry which is subsidized by the Ministry of Economy Affairs of the Republic of China

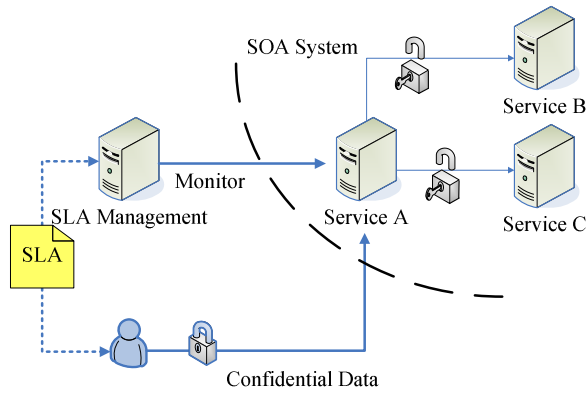


Figure1. Potential violation in SOA system

The following example is utilized to describe the insufficiencies of an SLA in an SOA system. The service consumer requests information from service A for a financial process (Fig. 1). First, SLA management component negotiates with the consumer. Results of negotiation are then stored in databases. Second, the service consumer starts transferring confidential data which are already encrypted to Service A and SLA management component monitors the connection to prevent violations. Financial process procedures include multiple services such as stocks or fund applications in banks or companies. Additionally, user information must be transmitted to other organizations. Service A must send requests to servers B and C for other financial services. These requests include personal information such as credit card ID, which are not encrypted and violate SLA. SLA manager will not detect violations between service A and services B or C. Thus, this work presents a novel split SLA for services for follow-up processes in an SOA system. Figure 2 shows this scheme. SLA was split for different services relative to request and these services must meet the requirements of split SLAs. Split contracts inherit encryption from upstream contracts. Transmission of data between services is safer than traditional systems.

The reminder of this paper is organized as follows. Section 2 presents literature reviews. Section 3 describes the proposed system architecture. Section 4 characterizes the message flow in the proposed scheme. Section 5 provides a scenario that corresponds to a real case. Section 6 is the conclusion.

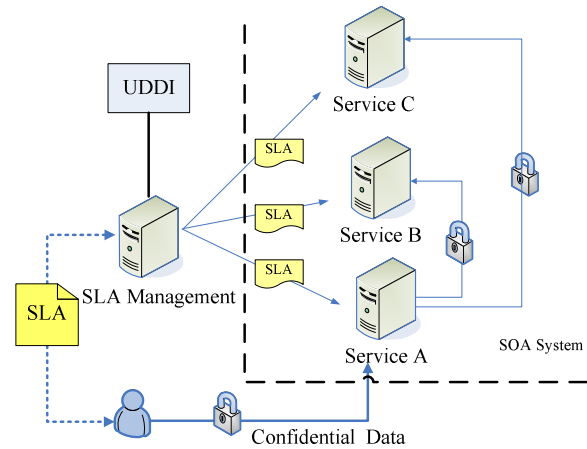


Figure2. Split SLA for preventing violation

## 2. Relative Work

Business processes can be quantified and described using XML-based language [1, 5]. The SLA language for dynamic electrical businesses [2] is collected and analyzed the requirement of SLAs. The SLA language must support a variety of parameters and obligations. Parameters can not be ambiguous for monitors and all parameters can be split. It also addressed the structure and policies for SLA language. In the proposed scheme, parameters in an SLA between consumers and providers are formatted to prevent ambiguous explanations arising from different platforms. Splitting an SLA is what primarily discussed in this paper. An SLA is a contract that defines resource allocations between users and service providers. Violations are detected when resources are insufficient. The Web Service Level Agreement (WSLA) framework [1] was proposed for monitoring SLAs. A WSLA provides a flexible language that supports various SLAs and monitored by the third party. The WSLA was developed for contracts between service consumers and providers. The WSLA has two parts: a script language that describes the SLA which is based on XML schema; and, an execution architecture that includes service monitoring. When the WSLA is applied to an SOA system [6, 7, 8, 9, 10], communication between services can not fulfill contractual agreements between users and service providers. Service management always uses related services to complete a complex job. Third-party monitoring only supervises the interactions between users and service providers. Follow-up processes of services could transmit

data without constraints and cause security problems.

Service oriented architectures have been recently applied to business systems. Stakeholder-driven service [3] was proposed to develop dynamic software system. Stakeholder-driven service provides a life-cycle for developing software and more flexible schema that replaces services dynamically. This work provides a better mechanism to develop software than existing structure. However, managing Stakeholder-driven model can also result in violation when consumers can be replaced. The proposed scheme also solves these problems. Integrating an SLA with an SOA system [4] generates some problems. The system utilizes different software to provide integrated service for users. Services are working concurrently and may change command value in database. A contract can be utilized to limit communication between services.

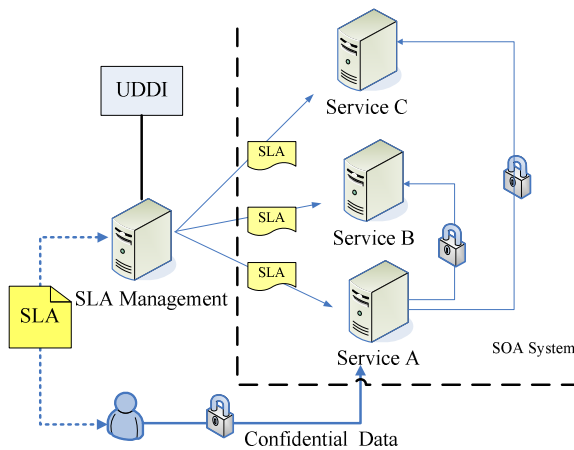


Figure 2. Split SLA for preventing violation

### 3. System Architecture

This section describes all components in the proposed system. Figure 3 presents the system architecture. A conventional SLA can be split and parts of contract can be sent to other services; however, split contract can not guarantee follow-up services. The proposed architecture guarantees that the SLA will be followed. The capabilities of components are discussed in detail as follows.

#### 3.1 Service consumer

Users or services who request service from service provider are so called service consumers.

#### 3.2 Service controller

The service controller is responsible for searching, monitoring and managing resources or services for service consumers. Service controller can be divided to two parts.

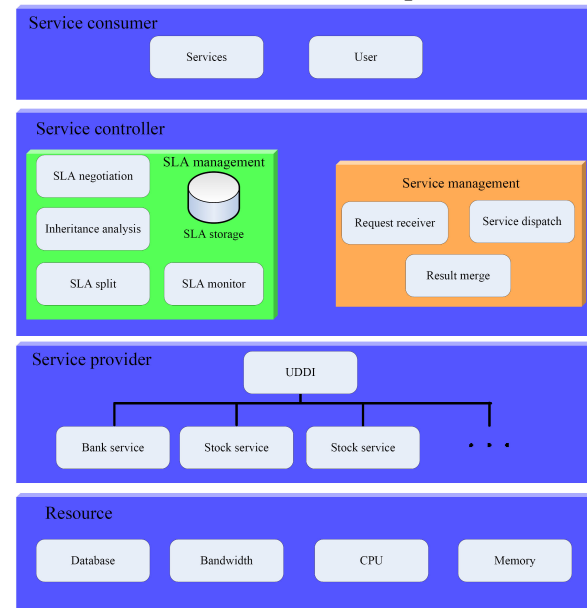


Figure 3. System architecture

#### 3.2.1 SLA management

The SLA manager is in charge of negotiating with consumers, analyzing SLAs, splitting contracts and monitoring requested services.

##### 3.2.1.1 Negotiation

The component is responsible for providing contract content to service consumers. Various service consumers need the same service, but they offer different definitions for service-level objects. Parameters in an SLA contract typically vary and may be ambiguous. Negotiation is used to define appropriate parameters for follow-up services.

When receiving requests, the SLA manager must negotiate with consumers. The contract is in XML-based language. The contract defines communication protocol and other quality values for service consumer.

##### 3.2.1.2 Inheritance analysis engine

This component computes the relations among parameters in a contract. A contract may be split for requesting other services in an SOA system. Every service provider only receives parts of a contract and guarantees partial obligations. However, some characteristics of the original contract have to be extended to split contracts. Otherwise, violations will occur. This

component is used to calculate relationships among parameters and determine which arguments must be inherited. For example, encrypted data from consumers are decrypted according to the key stored in an SLA. The data are then split and transmitted to different services providers in plain code as the key was not inherited. The proposed scheme provides for inheritance of original and split SLAs. This component prevents conditions that violate the SLA.

### 3.2.1.3 Disassembler

When a service consumer signs the SLA, the contract will be analyzed. If only one service is needed, a complete contract will be transmitted; otherwise, the disassembler analyzes and splits requests. Parameters in an SLA are functional and can be split. When splitting a contract, inheritance relationships are computed by inheritance analysis engine. According to computational results, the disassemblers use independent parameters and inherited parameters to build new service SLAs.

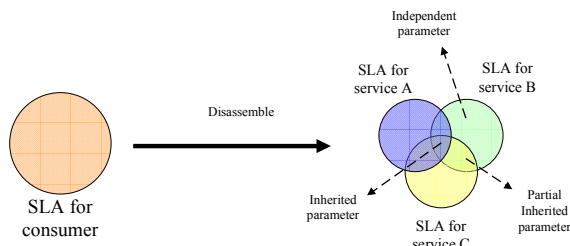


Figure 4. relations of disassemble SLA

For example, a consumer request could involve numerous services. Disassembling and inheritance analysis compute a request that includes three services and relationship among parameters. Figure 4 shows the relationships between split SLAs and parameters. Independent parameters are only restricted for one specific service, not for any services. All of the SLAs must include inherited parameters to prevent violations. Partially inherited parameters, which exist in some services, use the same data from consumers and must follow the related rules.

### 3.2.1.4 Monitor

Monitor maps service level objects to resources correctly for detecting violations. All services are measured by an SLA monitor. Resource information is collected periodically to prevent violations. When violations are detected, the monitor notifies SLA manager and

service consumers. The SLA manager then adjusts resources allocations in compliance with contract rules. A service consumer can charge compensation from service provider according to the record of violations.

### 3.2.1.5 Storage

The SLA storage stores all agreements, including complete and split contracts, in the database. A complete SLA is referred to users and service providers for the purpose of maintaining the overall agreement. A Split SLA is referred to service manager and service providers for the purpose of handling partial rules.

## 3.3 Service management

This module manages all of the services. As a bridge to other components, Service management could coordinate negotiations between consumers and SLA managers and allocate resources to services.

### 3.3.1 Service dispatch

According to different services, requests from service consumers are analyzed and split. Based on analytical results, a dispatcher looks up a service provider and locates corresponding services. The service dispatcher forwards split SLAs and requests the required services. When violations occur, the SLA monitor will allocate resource using this component.

### 3.3.2 Service merger

The service provider returns computed results which are responded from different services to consumers. The service manager must merge these results and forward this data to consumers. The merging component assesses the completeness of results since consumers request the whole one service but not three-quarters or half service. Merging component would guarantee the integrity of service responses.

### 3.3.3 Request receiver

All users and service provider requests are received by this component. When service response results sent to consumers, request receiver component would work as an interface for consumers.

## 3.4 Service provider

The Universal Description, Discovery, and Integration (UDDI) provides an index of services for service management. All service providers must register in the UDDI. Service managers will search for the UDDI when a

service request is received and split.

### 3.5 Resource

System resources include CPU, memory, bandwidth and encryption key. These resources require flexible definitions for different requests form numerous platforms. All resources and related parameters are used to provide services.

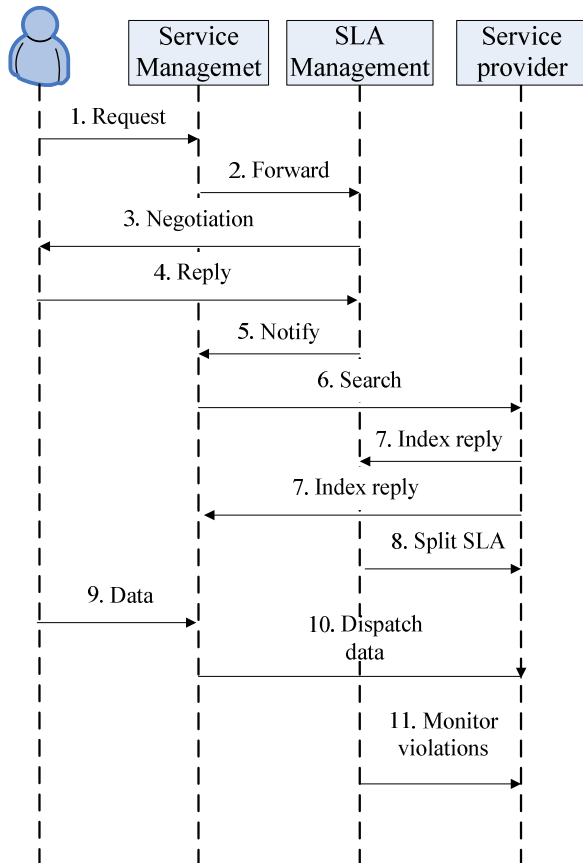


Figure 5. Message Sequence

#### 4. Diagram of message flow

This section presents an overview of the message sequence when a consumer sends a request to a service provider. Figure 5 shows this procedure.

1. Request: Initially, service consumers connect to the service management and send a request to request receiver. The request receiver provides an interface for receiving different requests from consumers and sending responses.
2. Forward: The request receiver send a request to SLA manager who starts negotiating with consumer.
3. Negotiation: The SLA manager provides a contract to service consumer. According to requests, all relative parameters are used as contract content

in SLA negotiation. This stage will be pending until the consumer responds.

4. Reply: If a service consumer disagrees with the contract, the process is interrupted; otherwise, the contract is approved and stored in a database.
5. Notify: The SLA manager notifies the service manager that the agreement has been approved. According to the decomposed SLA received from the SLA manager, the service manager finds the required services.
6. Searching: The service manager finds the required services using the UDDI component and starts communicating with them.
7. Index reply: The UDDI provides an index of services for the service manager. The service manager dispatches split requests to the corresponding services.
8. Split SLA: The SLA manager computes the inheritance of services and splits the contract for specific services. These split contracts and inherited parameters are combined into different SLAs and stored in a database. The SLA manager then sends these new contracts to the corresponding services.
9. Data: The data from consumer in runtime are transmitted to the service manager. The request receiver remains the interface for consumers.
10. Dispatch data: According to the initial request, data are divided and dispatched to the corresponding services listed in the SOA system.
11. Monitor violation: The SLA manager monitors connections between services. Once the violation occurs, the monitor will notify consumer and the SLA manager to revise resources allocations.

#### 5. Scenario

This scenario applies the scheme to financial services. The number of electronic business services on Internet today is increasing. Service consumers, like the users, want to trade via the Internet trading and perform account transfers. First, users must transmit bank account and stock information to service controller. The request is then analyzed to determine which services are needed. At the same time, the service controller negotiates an agreement with



users. This agreement includes authentication, encryption, bandwidth usage, transfer limits, and stock trading limits. In this case, the service provider finds the Internet trading service offered by the investment company and Internet banking service which needs the account information and a password. The service controller obtains the information of services and split the SLA for these services. The decomposed SLA also inherits the necessary limitations from the original contract. Individual contracts are then transmitted to different services and connections are monitored. The user sends account numbers and passwords to the service controller and divided information into two parts. One part of information is stock investment service, and the other is for banking company. However, stock investment company also requires banking service to charge, it becomes a service consumer in the system. The connection between the stock investment company and bank must follow the split SLA stored at stock investment company. Under this scheme, communication in follow-up process totally complies with the contract.

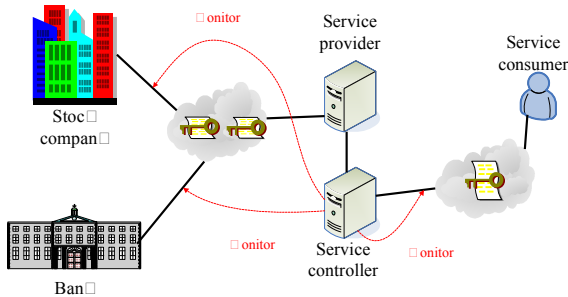


Figure 6 Process of financial service

## 6. Conclusion

The proposed scheme for SLAs uses more resources than the traditional structure. This is unavoidable since additional processes are necessary to monitor the same parameter in numerous services. Furthermore, the additional processes are needed to monitor communications between services in the SOA system. Contract rules are extended in the follow-up processes to prevent potential violation which isn't completely guaranteed in traditional SLAs architecture.

This work presented a novel scheme for SLAs in an SOA system. The security issue of services in an SOA system and follow-up processes is ensured. Services involved will comply with relevant contract rules and improve the accuracy of an SLA.

## Reference

- [1] A. Keller, H. Ludwig, "The WSLA Framework: Specifying and Monitoring Service Level Agreements for Web Services", *IBM Research Report*, 2002.
- [2] H. Ludwig, A. Keller, A. Dan, R. King, "A Service Level Agreement Language for Dynamic Electronic Services", *Electronic Commerce Research*, 3: 43–59 (2003).
- [3] Q. Gu, P. Lago, "A stakeholder-driven Service Life Cycle Model for SOA", *International Workshop on Service Oriented Software Engineering (IW-SOSWE)*, Sep. 3, 2007.
- [4] D. Ameller, X. France, "Service Level Agreement Monitor (SALMon)", *IEEE Seventh International Conference on Composition-Based Software Systems* 2008.
- [5] H. Kreger, "Fulfilling the Web services promise", *Communications of the ACM*, pp. 29-34, 2003.
- [6] P. Vittorini, S. Necozone, F. D. Orio, "Towards a SOA Infrastructure for Statistically Analysing", *ACM Workshop on CyberInfrastructure: Information Management in eScience (CIMS)*, 2007.
- [7] A. Michlmayr, F. Rosenberg, C. Platzer, M. Treiber, S. Dustdar, "Towards Recovering the Broken SOA Triangle – A Software Engineering Perspective", *ACM international workshop on service oriented software engineering (IW-SOSWE)*, 2007.
- [8] H. Yoon, E. Ji, B. Choi, "Building Test Steps for SOA Service Orchestration in Web Service Testing Tools", *In Proceedings of the 2nd international conference on Ubiquitous information management and communication*, 2008.
- [9] M. Tanuan, "Using Sybase WorkSpace to Build Service Oriented Architecture (SOA) Applications Quickly", *ACM SIGPLAN International Conference on Object-Oriented Programming, Systems, Languages, and Applications*, 2007.
- [10] L. Resende, R. Feng, "Handling Heterogeneous Data Sources in a SOA Environment with Service Data Objects (SDO)", *ACM Special interest group on management of data (SIGMOD)*, 2007.

## A History Based Semantic Aware Access Control Model Using Logical Time

Ali Noorollahi Ravari, Morteza Amini, Rasool Jalili, and Jafar Haadi Jafarian

*Network Security Center, Computer Engineering Department,*

*Sharif University Of Technology, Tehran, Iran*

*noorollahi@ce.sharif.edu, jafarian@ce.sharif.edu, m\_amin@ce.sharif.edu, jalili@sharif.edu*

### Abstract

With the advent of semantic technology, access control cannot be done in a safe way unless the access decision takes into account the semantic relationships among the entities in a semantic-aware environment. The SBAC model (Semantic Based Access Control model) considers this issue in its decision making process. However, time plays a crucial role in new computing environments, which is not supported in SBAC. In this paper, we propose the Temporal Semantic Based Access Control (TSBAC) model, as an extension of SBAC, which enhances the specification of user-defined authorization rules by constraining time interval and temporal expression over users' history of accesses. TSBAC uses logical time, rather than to real time, in its authorization rules. A formal semantics for temporal authorizations is provided and conflicting situations (due to the semantic relations of the SBAC model and a sub-interval relation between authorizations) are investigated and resolved in our proposed model. An architecture for the access control system based on TSBAC is presented.

### 1. Introduction

An important requirement of any information management system is to protect data and resources against unauthorized disclosure (confidentiality) and unauthorized or improper modifications (integrity), while at the same time ensuring their availability to legitimate users (no denials-of-service). Therefore, enforcing protection requires that every access to a system and its resources be controlled and all and only authorized accesses can take place [1]. The development of an access control system requires the definition of the regulations according to which access is to be controlled and their implementation as functions executable by a computer system.

An important requirement, common to many applications, is related to the temporal dimension of access permissions. In these systems, permissions are granted based on previous authorizations given to the users of the system or denied in specific time points (in the past). Another critical requirement is the possibility of expressing the semantic relationships that usually exist

among different authorization elements, i.e. subjects, objects, and actions. To overcome this challenge, our model is constructed based on the SBAC model (Semantic Based Access Control model) [2, 3] which is a semantic-based access control model. SBAC authorizes users based on the credentials they offer when requesting an access right. Ontologies are used for modeling entities along with their semantic interrelations in three domains of access control, namely subjects domain, objects domain, and actions domain. To facilitate the propagation of policies in these three domains, different semantic interrelations can be reduced to the subsumption relation.

In this paper we unify the two concepts mentioned above; that is, we use the SBAC model (as the base model), and associate a temporal expression with each authorization. Thus, in this paper we employ logical time operators to specify historical constraints over the elements of the history base in authorization rules. Furthermore, a time interval bounds the scope of applicability of the temporal authorization rules, and usage of elements of history (e.g., [1, 2, 20] shows that the authorization is valid for time interval starting at '1' and ending at '20', and history elements with the time stamp greater than 2 are taken into account). Thus, the main feature provided by our model is the possibility of specifying authorization rules which express temporal dependencies among access events in the past (stored in a *history base* in the form of  $done(t, s, o, a)$  and  $denied(t, s, o, a)$ )

The rest of this paper is structured as follows: Section 2 discusses the related work in this topic. Section 3 gives a brief introduction of the SBAC model and describes the model of time used throughout our work. In section 4, we represent our authorization rules in detail, offer their formal semantics, and conflict resolution in access decision point. Section 5 gives an architecture for the access control system based on the proposed model. In Section 6 an implementation of the basic components of the model using CLIPS is presented. In section 7, a case study in a banking environment is described. A qualitative and quantitative evaluation of the model is discussed in section 8. Finally, Section 9 concludes the paper and gives some future works on the topic.



## 2. Related Work

Access control systems for protecting web resources along with credential based approaches for authenticating users have been studied in recent years. With the advent of Semantic Web, new security challenges were imposed to security systems. Bonatti *et al.* [4] have discussed open issues in the area of policy specification for Semantic Web community such as important requirements for access control policies. Developing security annotations to describe security requirements and capabilities of web service providers and requesting agents have been addressed by Rabitti *et al.* [5]. Qin and Atluri [6] proposed a concept level access control model that considers some semantic relationships in the level of concepts in objects domain. SBAC which is the basis of our model, is proposed by Javanmardi *et al.* [2, 3], and is based on the OWL ontology language and considers the semantic relationships in the domains of subjects, objects, and actions to make decision about an access request.

The first security policy based on the history of events was introduced as Chinese Wall Security Policy (CWSP) [7]. The objective of CWSP is to prevent information flows which cause conflict of interest for individual consultants. Execution history also plays a role in Schneider's security automata [8] and in the Deeds system of Edjlali [9]. However, such works focus on collecting a selective history of sensitive access requests and use this information to constrain further access requests; for instance, network access may be explicitly forbidden after reading certain files. Another approach that considers the history of control transfers, rather than a history of sensitive requests, is presented by Abadi and Fournet [10].

In a basic authorization model, an authorization is modeled by a triple  $(s, o, \pm a)$ , interpreted as "subject  $s$  is (not) authorized to exercise access right  $a$  on object  $o$ ". Recently, several extensions to this basic authorization model have been suggested. One of them is the temporal extension, which increases the expressive power of the basic authorization model [11-15]. In the model proposed by Bertino *et al.* [11], an authorization is specified as  $(time, auth)$ , where  $time = (t_b, t_e)$  is the time interval in which the authorization  $auth = (s, o, m, pn, g)$  is valid. In  $auth$ ,  $s$  represents the subject,  $o$  the object, and  $m$  the privilege,  $pn$  is a binary parameter indicating whether an authorization is negative or positive, and  $g$  represents the grantor of the authorization. This model also allows operations *WHENEVER*, *ASLONGAS*, *WHENEVERNOT*, and *UNLESS* on authorizations. For example, *WHENEVER* can be used to express that a subject  $s_i$  can gain privilege on object  $o$  whenever another subject  $s_j$  has the same privilege on  $o$ . Later Bertino *et al.* [14] extended the temporal authorization model to support periodic authorizations. They completed their research in [16] by presenting a powerful authorization mechanism that

provides support for: (1) periodic authorizations (both positive and negative), that is, authorizations that hold only in specific periods of time; (2) user-defined deductive temporal rules, by which new authorizations can be derived from those explicitly specified; (3) a hierarchical organization of subjects and objects, supporting a more adequate representation of their semantics. From the authorizations explicitly specified, additional authorizations are automatically derived by the system based on the defined hierarchies.

## 3. Preliminaries

In this section we give a brief introduction to the SBAC model, proposed by Javanmardi *et al.* [2, 3], and introduce the model of time used throughout our work.

### 3.1 Introduction to SBAC

Fundamentally, SBAC consists of three basic components: Ontology Base, Authorization Base, and Operations. Ontology Base is a set of ontologies: Subjects–Ontology (SO), Objects–Ontology (OO), and Actions–Ontology (AO). By modeling the access control domains using ontologies, SBAC aims at considering semantic relationships in different levels of ontology to perform inferences to make decision about an access request. Authorization Base is a set of authorization rules in the form of  $(s, o, \pm a)$  in which  $s$  is an entity in SO,  $o$  is an entity defined in OO, and  $a$  is an action defined in AO. In the other words, a rule determines whether a subject which presents a credential  $s$  can have the access right  $a$  on object  $o$  or not.

The main feature of the model is reduction of semantic relationships in ontologies to subsumption relation. Given two concepts  $C$  and  $D$  and a knowledge base  $\Sigma$ ,  $C < D$  denotes that  $D$  subsumes  $C$  in  $\Sigma$ . This reasoning based on subsumption represents that  $D$  (the subsumer) is more general than  $C$  (the subsumee).

### 3.2 Modeling of Time

We assume that the system is composed of a single process, and we timestamp each event with a counter based clock. The clock ticks every time an event occurs.

## 4. SBAC with Logical Time Constraints

In some applications, only the logical sequence of events is of considerable importance, or due to the system specifications, time stamping of events with real time is impossible. Thus, in these situations, we use logical time instead of real time and timestamp events using the logical time scheme introduced in section 3.2.

**Definition (Temporal Authorization Rule)** A temporal authorization rule is a triple  $([t_{sv}, t_{sh}, t_f], (s, o, \pm a), F)$ , where  $t_{sv}, t_{sh}, t_f \in LogicalTime$ , and  $t_{sv}, t_{sh} \leq t_f$ . In this notation,  $[t_{sv}, t_{sh}, t_f]$  represents the authorization  $(s, o, \pm a)$  validation time interval, and formula  $F$  is a temporal constraint which is formally defined in Table 1.

**Table 1. Definition of temporal predicate  $F$**

$A ::= done(s, o, a)   denied(s, o, a)  $ $\neg done(s, o, a)   \neg denied(s, o, a)$
$E ::= prev(A)   past\#(A)   H(A)   sb\#(A, A)  $ $ab(A, A)   ss(A)   during(A, A)$
$F ::= true   false   E   \neg E   E \wedge E   E \vee E   E \rightarrow E   E \leftrightarrow E$

The temporal authorization rule  $([t_{sv}, t_{sh}, t_f], (s, o, \pm a), F)$  states that subject presenting credential  $s$  is allowed (or not allowed) to exercise access  $a$  on object  $o$ , in the time interval  $[t_{sv}, t_f]$  if formula  $F$  is evaluated to true by the access events occurred in the time interval  $[t_{sh}, t_f]$ .

**Definition (Temporal Authorization Base)** A temporal authorization base (TAB) is a set of temporal authorization rules in the form of  $([t_{sv}, t_{sh}, t_f], (s, o, \pm a), F)$ .

**Definition (History Base)** A History Base is a set of authorizations with timestamp, in the form of  $done(t, s, o, a)$  that means access  $a$  has been granted to subject presenting credential  $s$  on object  $o$  at logical time  $t$ , and  $denied(t, s, o, a)$  that means the system has denied access  $a$  on object  $o$  at logical time  $t$  requested by subject presenting credential  $s$ .

#### 4.1 Informal Meaning of Temporal Authorization Rules

The intuitive meaning of (logical time) temporal authorization rules is as follows (in these statements *auth* is used instead of  $(s, o, \pm a)$ ):

- $([t_{sv}, t_{sh}, t_f], auth, prev(A))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A$  is evaluated to true at the time  $(t - 1)$ , and  $t - 1 \geq t_{sh}$ .
- $([t_{sv}, t_{sh}, t_f], auth, past\#(A))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A$  is evaluated to true  $\#$  of times in the time interval  $[t_{sh}, t]$  in the history base.
- $([t_{sv}, t_{sh}, t_f], auth, H(A))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A$  is evaluated to true at all time-points in the time interval  $[t_{sh}, t]$  in the history base.
- $([t_{sv}, t_{sh}, t_f], auth, sb\#(A_1, A_2))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A_1$  is evaluated to true  $\#$  of times before the last occurrence of  $A_2$  in the time interval  $[t_{sh}, t]$ .
- $([t_{sv}, t_{sh}, t_f], auth, ab(A_1, A_2))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A_1$  is evaluated to true at  $t'$  ( $t_{sv} \leq t' \leq t$ ), and there exist a time-point  $t''$  ( $t' \leq t'' \leq t$ ), in which  $A_2$  is evaluated to true.

- $([t_{sv}, t_{sh}, t_f], auth, ss(A_1, A_2))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A_1$  is always evaluated to true, from the first occurrence of  $A_2$  in the time interval  $[t_{sh}, t]$  till  $t$ .
- $([t_{sv}, t_{sh}, t_f], auth, during(A_1, A_2))$ : Authorization *auth* is valid at the time of request  $(t)$ ,  $t_{sv} \leq t \leq t_f$ , if  $A_1$  is not true before the first, or after the last time-point in which  $A_2$  is true in the time interval  $[t_{sh}, t]$ .

#### 4.2 Comparison with Linear Time Temporal Logic Operators

As mentioned in previous sections, TSBAC is an access control model that makes its access control decision based on the temporal relation between users' access events in the history. Due to this requirement, we considered a modified set of operators of the Propositional Linear Temporal Logic (Future and Past version) [17]. In addition, some of the operators were added due to the requirements of real environments where the access control system based on the model is applicable. In the following, an overall comparison of the operators of TSBAC and LTL is presented. The newly added operators are described next.

The basic temporal operators of this system are '*Fp*' (sometime *p* or eventually *p*), '*Gp*' (always *p* or henceforth *p*), '*Xp*' (next time *p*), '*p U q*' (*p* until *q*), and the modality '*p B q*' (*p* precedes *q*).

'*Prev*' in TSBAC is the past time equivalent of '*X*' in LTL (or *X<sup>-</sup>*), '*H*' is the past time equivalent of '*G*' (or *G<sup>-</sup>*), '*sb#*' is the past time equivalent of '*B*' which takes into account the number of instants in which the first operand ( $A_1$ ) is evaluated to true. In some situations we need to identify the situation in which an event ( $A_1$ ) has been repeated in all timepoints in the past. Thus, the operator '*ab*' seems necessary. '*ss*' is the '*Since*' operator but with a minor modification; in addition to being a past time equivalent, it is required that the first operand ( $A_1$ ) holds from the moment that  $A_2$  is evaluated to true.

In order to describe a situation in which an event occurred only between the first and last occurrence of another event, during operator seems necessary. '*past#*' is the modified version of '*prev*' which makes its decision based on the desired number of its operand evaluated to true.

For the purpose of increasing the expressiveness of the model, the propositional logic combination of the predicates generated so far was taken into account to create a temporal expression. In this manner, we have a fully functional expressive set of operators that could be used to express different temporal relations between the elements of a history of accesses.

### 4.3 Formal Semantics of Logical Time Authorization Rules

To formalize the semantics of temporal authorization rules, we first define an evaluation function  $\Phi$ . This function evaluates the predicate  $F$  of temporal authorization rules at a logical time point  $t$ , and based on the elements stored in *History Base*. The semantics of such an evaluation is given in first order logic and is reported in Table 2.

**Table 2. Formal semantics of the  $\Phi$  evaluation function**

$\Phi_t(t_{sh}, done(s, o, a))$ $= \begin{cases} true, & t \geq t_{sh} \wedge done(t, s, o, a) \in HB \\ false, & t \leq t_{sh} \vee done(t, s, o, a) \notin HB \end{cases}$ $\Phi_t(t_{sh}, denied(s, o, a))$ $= \begin{cases} true, & t \geq t_{sh} \wedge denied(t, s, o, a) \in HB \\ false, & t \leq t_{sh} \vee denied(t, s, o, a) \notin HB \end{cases}$ $\Phi_t(t_{sh}, \neg done(s, o, a))$ $= \begin{cases} true, & \nexists t, t \geq t_{sh}, done(t, s, o, a) \in HB \\ false, & \exists t, t \geq t_{sh}, done(t, s, o, a) \in HB \end{cases}$ $\Phi_t(t_{sh}, \neg denied(s, o, a))$ $= \begin{cases} true, & \nexists t, t \geq t_{sh}, denied(t, s, o, a) \in HB \\ false, & \exists t, t \geq t_{sh}, denied(t, s, o, a) \in HB \end{cases}$ $\Phi_t(t_{sh}, prev(A)) = \Phi_{t-1}(t_{sh}, A)$ $\Phi_t(t_{sh}, past\#(A)) = \exists t_1, \dots, t_{\#} \cdot \bigwedge_{k=1}^{\#} \Phi_{t_k}(t_{sh}, A)$ $\Phi_t(t_{sh}, H(A)) = \forall t_i, t_{sh} \leq t_i \leq t, \Phi_{t_i}(t_{sh}, A)$ $\Phi_t(t_{sh}, sb\#(A_1, A_2)) = \exists t_{22}, t_{sh} \leq t_{22} \leq t, \exists t_1, \dots, t_{\#}$ $\leq t_{22} \cdot \Phi_{t_{22}}(t_{sh}, A_2) \wedge \bigwedge_{k=1}^{\#} \Phi_{t_k}(t_{sh}, A_1)$ $\Phi_t(t_{sh}, ab(A_1, A_2)) = (\exists t_1, t_{sh} \leq t_1 \leq t \cdot \Phi_{t_1}(t_{sh}, A_1))$ $\rightarrow (\exists t_2, t_1 \leq t_2 \leq t \cdot \Phi_{t_2}(t_{sh}, A_2))$ $\Phi_t(t_{sh}, ss(A_1, A_2)) = \forall t_2, t_{sh} \leq t_2$ $\leq t, (\Phi_{t_2}(t_{sh}, A_2)$ $\rightarrow (\forall t_i, t_2 \leq t_i \leq t \rightarrow \Phi_{t_i}(t_{sh}, A_1)))$
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

$\Phi_t(t_{sh}, during(A_1, A_2))$ $= \left( \left( \left( \exists t_{min}, t_{sh} \leq t_{min} \right. \right. \right.$ $\leq t \wedge \Phi_{t_{min}}(t_{sh}, A_2)$ $\wedge \left( \nexists t_x, t_{sh} \leq t_x \right.$ $\leq t_{min} \wedge \Phi_{t_x}(t_{sh}, A_2) \left. \left. \right) \right)$ $\wedge \left( \left( \exists t_{max}, t_{sh} \leq t_{max} \right. \right.$ $\leq t \wedge \Phi_{t_{max}}(t_{sh}, A_2)$ $\wedge \left( \nexists t_y, t_{max} \leq t_y \right.$ $\leq t \wedge \Phi_{t_y}(t_{sh}, A_2) \left. \left. \right) \right)$ $\rightarrow (\forall t_1, t_1 \leq t$ $\rightarrow (\Phi_{t_1}(t_{sh}, A_1) \rightarrow t_{min} \leq t_1 \leq t_{max}))$ $\Phi_t(t_{sh}, \neg E) = \Phi_t(t_{sh}, E)$ $\Phi_t(t_{sh}, E_1 \wedge E_2) = \Phi_t(t_{sh}, E_1) \wedge \Phi_t(t_{sh}, E_2)$ $\Phi_t(t_{sh}, E_1 \vee E_2) = \Phi_t(t_{sh}, E_1) \vee \Phi_t(t_{sh}, E_2)$ $\Phi_t(t_{sh}, \neg E) = \neg \Phi_t(t_{sh}, E)$ $\Phi_t(t_{sh}, E_1 \wedge E_2) = \Phi_t(t_{sh}, E_1) \wedge \Phi_t(t_{sh}, E_2)$ $\Phi_t(t_{sh}, E_1 \vee E_2) = \Phi_t(t_{sh}, E_1) \vee \Phi_t(t_{sh}, E_2)$ $\Phi_t(t_{sh}, E_1 \rightarrow E_2) = \Phi_t(t_{sh}, E_1) \wedge \neg \Phi_t(t_{sh}, E_2)$ $\Phi_t(t_{sh}, E_1 \leftrightarrow E_2) = (\Phi_t(t_{sh}, E_1) \rightarrow \Phi_t(t_{sh}, E_2))$ $\wedge (\Phi_t(t_{sh}, E_2) \rightarrow \Phi_t(t_{sh}, E_1))$
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

By the definition of the evaluation function  $\Phi$  and by the assumption described above, the semantics of authorization rules are presented in Table 3. In the following,  $grant(t, (s, o, a))$  denotes subject presenting credential  $s$  is granted to exercise action  $a$  on object  $o$ , and analogously,  $deny(t, (s, o, a))$  denotes the access request of subject having credential  $s$  for exercising an access  $a$  on object  $o$  is denied.

**Table 3. Semantics of logical time authorization rules**

$([t_s, t_f], (s, o, +a), F)$ $\Leftrightarrow \forall t (t_{sv}, t_{sh} \leq t \leq t_f \wedge \Phi_t(t_{sh}, F))$ $\rightarrow grant(t, (s, o, a))$ $([t_s, t_f], (s, o, -a), F)$ $\Leftrightarrow \forall t (t_{sv}, t_{sh} \leq t \leq t_f \wedge \Phi_t(t_{sh}, F))$ $\rightarrow deny(t, (s, o, a))$
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### 4.4 Access Control

The centric security mechanism in each system is an access control system. Upon receiving an access request in such a system, we need to make a decision whether to grant the requested access or deny it. Following the proposed model of temporal authorization in the previous sections, upon receiving an access request  $(s_r, o_r, a_r)$  at

time  $t$ , the access control system performs the following steps:

1. Determine the explicit and implicit valid authorization rules in TAB at time  $t$  (following the definition of valid authorization rules), satisfying the following conditions:
  - $t_s \leq t \leq \min(t_f, t_d)$ ,  $t_d$  is the time of deletion of a specific authorization rule.
  - Temporal predicate  $F$  is evaluated to true at time  $t$  (based on  $\varphi_t$  evaluation function).
2. Extract the set of valid authorization rules such as  $([t_{sv}, t_{sh}, t_f], (s, o, \pm a), F)$  which match the access request. These authorization rules must satisfy, at least, one of the following conditions:
  - $s = s_r, o = o_r, a = a_r$
  - Following the propagation rules of the SBAC model, in the case of a positive action ( $+a$ ), we have  $s_r < s, o_r < o, a_r < a$ , and in the case of a negative action ( $-a$ ), we have  $s_r < s, o_r < o, a < a_r$ .
3. If there exist just positive valid authorization rule(s) such as  $([t_{sv}, t_{sh}, t_f], (s, o, +a), F)$  in MVA, grant the requested access.
4. If there exist just negative valid authorization rule(s) such as  $([t_{sv}, t_{sh}, t_f], (s, o, -a), F)$  in MVA, deny the access request.
5. If there exist both positive and negative authorization rules in MVA, resolve the conflict and follow the result.
6. If there exists no valid authorization rule, which matches the requested access, follow the default access policy.
7. Store  $done(t, s_r, o_r, a_r)$  in case of the requested access is granted and  $denied(t, s_r, o_r, a_r)$  in case of the access request is denied.

#### 4.5 Conflict Detection and Resolution

A conflict occurs when two or more access policies cannot be applied in the same time. In access control, due to modal conflict between *matched valid authorizations*, we need a conflict resolution strategy.

##### 4.5.1 Conflict Occurrence

In TSBAC, conflict occurs due to the semantic relations between the entities (in the domains of subjects, objects, or actions) and applying the inference rules of SBAC, or due to the sub-interval relationship between temporal authorization rules of TSBAC.

- *Conflict due to the semantic relations between the entities*: as mentioned before, in the domains of subjects and objects, subsumee has all the privileges (positive and negative) of subsumer, but, in the domain of actions, positive access rights is propagated from subsumer to subsumee, while negative access rights is propagated in the opposite direction (that is from subsumee to subsumer). These semantic relationships and propagation of negative and positive authorizations between entities may result in conflicting situations.

##### 4.5.2 Conflict Resolution

In access control, due to modal conflicts between valid matched authorization rules (in set MVA), it is required to have a conflict resolution strategy to resolve conflicts. The conflict might be a result of semantic relationships between the entities (i.e. subjects, objects, and actions) and applying the inference rules of SBAC model, or the sub-interval relation between authorizations (i.e.  $[t_{sv2}, t_{f2}]$  is a sub-interval of  $[t_{sv1}, t_{f1}]$ ).

TSBAC supports three predefined strategies for conflict resolution; negative authorization rule takes precedence (NTP) strategy, positive authorization rule takes precedence (PTP) strategy, and most specific rule takes precedence. Similar to the default access policy, the conflict resolution strategy is determined by the administrator.

## 5. Architecture

In order to guarantee the applicability of the model and usefulness in semantic based and temporal environments, an architecture for the access control system based on the proposed model is presented.

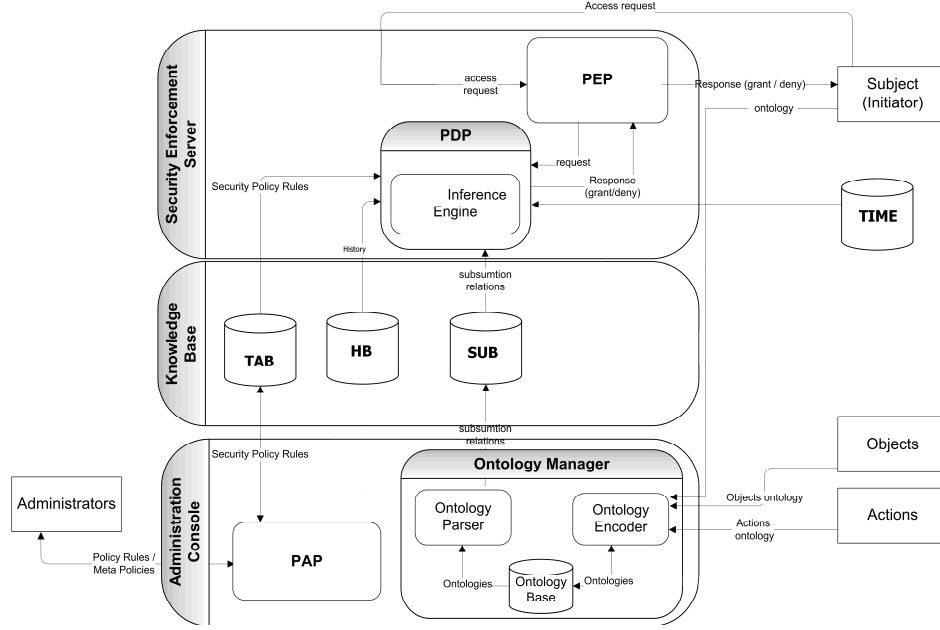


Figure 1. An architecture for the access control system based on TSBAC model

## 6. Implementing TSBAC using CLIPS

In order to make inference by CLIPS engine, we must introduce TSBAC to the CLIPS inference engine. This procedure can be summarized into these steps:

- Expressing description logics axioms.
- Expressing subjects ontology, objects ontology, and actions ontology.
- Expressing fix facts of the model.
- Expressing inference rules of SBAC.
- Feeding history base to the inference engine.
- Expressing temporal authorization rules.
- Applying access control and conflict resolution.

The description and details of the implementation could be found in the following URL:

<http://ce.sharif.edu/~noorollahi/Downloads.html>

## 7. Case Study

The required information for applying access control is as follows: two subjects, that is,  $s_1$  and  $s_2$  exist in the environment, and objects ontology and actions ontology can be seen in Figure 2.

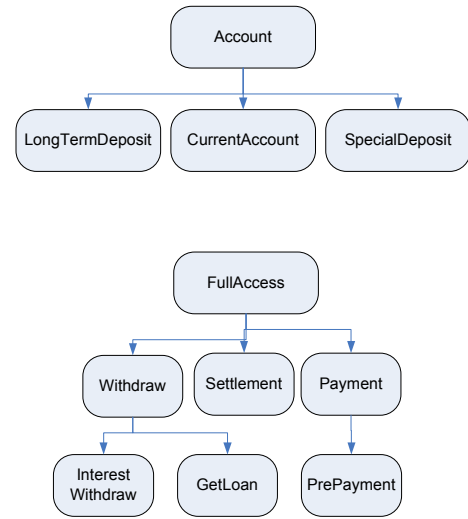


Figure 2. Objects and actions ontology in the sample banking environment

Sample temporal authorization rules are as follows:

$$\begin{aligned}
 R_1: & \left( [10, \infty], (s_1, LongTermDeposit_1, +InterestWithdraw), \right. \\
 & \left. H(\neg done(s_1, LongTermDeposit_1, Withdraw)) \right) \\
 R_2: & \left( [0, \infty], (s_1, CurrentAccount_1, -Withdraw), \right. \\
 & \left. sb0(done(all, CurrentAccount_1, Settlement), true) \right) \\
 R_3: & \left( [10, \infty], (s_1, CurrentAccount_1, +GetCheque), \right. \\
 & \left. ss(done(s_1, CurrentAccount_1, GetCheque), \neg denied(all, CurrentAccount_1, pass), 1) \right)
 \end{aligned}$$

$$R_4: \left( \begin{array}{l} [120, \infty], (s_2, \text{SpecialDeposit}_2, +\text{GetLoan}), \\ ss \left( \begin{array}{l} \text{done}(s_2, \text{SpecialDeposit}_2, \text{PrePayment}), \\ \text{done}(s_2, \text{SpecialDeposit}_2, \text{Payment}), 30 \end{array} \right) \wedge \\ \text{past120}(\text{done}(s_2, \text{SpecialDeposit}_2, \text{Payment})) \end{array} \right)$$

The elements of history base are as follows:

- 1:  $\text{done}(20, s_1, \text{LongTermDeposit}_1, \text{Withdraw})$
  - 2:  $\text{done}(30, s_1, \text{LongTermDeposit}_1, \text{Withdraw})$
  - 3:  $\text{done}(25, s_1, \text{CurrentAccount}_1, \text{Settlement})$
  - 4:  $\text{done}(35, s_1, \text{CurrentAccount}_1, \text{Settlement})$
  - 5:  $\text{denied}(50, s_2, \text{CurrentAccount}_1, \text{Pass})$
  - 6:  $\text{done}(51, s_2, \text{CurrentAccount}_1, \text{Pass})$
  - 7:  $\text{done}(130, s_2, \text{SpecialDeposit}_2, \text{PrePayment})$
  - 8:  $\text{done}(140, s_2, \text{SpecialDeposit}_2, \text{Payment})$
  - 9:  $\text{done}(150, s_2, \text{SpecialDeposit}_2, \text{Payment})$
  - 10:  $\text{done}(160, s_2, \text{SpecialDeposit}_2, \text{Payment})$
  - 11:  $\text{done}(170, \text{SpecialDeposit}_2, \text{Payment})$
- **First Access Scenario:** Suppose that request  $\text{req}_1 = (s_1, \text{LongTermDeposit}_1, \text{InterestWithdraw})$  is initiated at the time  $t_{\text{req}} = 40$ . Regarding  $R_1$  and elements #1 and #2 of History Base, the temporal predicate is evaluated to false, and this access request ( $\text{req}_1$ ) is denied (in the case of a closed access policy, or granted in the case of an open access policy). Finally,  $\text{denied}(40, s_1, \text{LongTermDeposit}_1, \text{InterestWithdraw})$  is added to History Base.
  - **Second Access Scenario:** Suppose that request  $\text{req}_2 = (s_1, \text{CurrentAccount}_1, \text{Withdraw})$  is initiated at the time  $t_{\text{req}} = 75$ . Regarding  $R_2$  and elements #3 and #4 of History Base, no negative authorization is issued for this request. Therefore, in case of an open access policy, the authorization is granted and  $\text{done}(75, s_1, \text{CurrentAccount}_1, \text{Withdraw})$  is added to History Base, but in case of a negative access policy, the authorization is denied and  $\text{denied}(75, s_1, \text{CurrentAccount}_1, \text{Withdraw})$  is added to History Base.
  - **Third Access Scenario:** Suppose that request  $\text{req}_3 = (s_1, \text{CurrentAccount}_1, \text{GetCheque})$  is initiated at the time of  $t_{\text{req}} = 100$ . Regarding  $R_3$  and element #5 that exists in History Base, this access request is denied and  $\text{denied}(100, s_1, \text{CurrentAccount}_1, \text{GetCheque})$  is added to History Base.
  - **Fourth Access Scenario:** Suppose that request  $\text{req}_4 = (s_2, \text{SpecialDeposit}_2, \text{GetLoan})$  is initiated at the time  $t_{\text{req}} = 200$ . Regarding  $R_4$  and elements #7...#11 exist in History Base, in spite of existence of element #7 (prepayment) and elements #8 through #11 (monthly settlement),

this access request is denied (because the number of monthly payments is less than 120) ,and  $\text{denied}(200, s_2, \text{SpecialDeposit}_2, \text{GetLoan})$  is added to History Base.

## 8. Discussion and Evaluation

The best way to evaluate an access control model is to qualitatively evaluate it against the security requirements of the environment. Moreover, we can take some quantitative criteria into account, but this consideration is only possible if an implementation of an access control system based on the proposed model exists.

### 8.1 Qualitative Evaluation of TSBAC Model

In this section we evaluate TSBAC regarding requirements of semantic-aware environments.

- **Conditional Authorization:** With the existence of temporal operators, TSBAC supports this type of authorization. In this model, due to wide spectrum of temporal operators, and using first order logic operators for combining temporal expressions, conditional authorization is provided, on the basis of existence or non-existence of specific authorizations in the past.
- **Conflict Detection and Resolution:** Conflict occurrence may be a result of semantic relationships between authorizations, or, sub-interval relations between validity constraint intervals. TSBAC detects these conflicts, and resolves them. Different conflict resolution policies include: denials take precedence, positives take precedence, and most specific takes precedence.
- **Supporting History-based Information:** The main feature of TSBAC is that authorizing an access request is done based on granted or denied access requests (done and denied access requests, which are stored in History Base), or, access requests that have not been done or have not been denied in the system (which can be inferred from History Base). These elements could be combined with temporal operators, or first order logics operators to compose temporal expressions.

### 8.2 Quantitative Evaluation of TSBAC

Due to the existence of an implementation for TSBAC, quantitative evaluation of the model is possible. The time and space complexity of the system based on the logical time operators of the model is as follows.

- **Time Complexity:** Since every access request is validated at the time of the request, and the process of authorization is based upon searching History Base and evaluating the temporal predicate, due to vast amount of elements of History Base and temporal predicate complexity,

access control in TSBAC is time consuming. In order to clarify the subjects mentioned above, we give a brief complexity analysis on *logical time* operators (in case of existence of  $n$  elements in History Base) of the proposed model.

$$\begin{aligned} \text{complexity}(\text{prev}) &= n \\ \text{complexity}(H) &= (t - t_s) \times n \\ \text{complexity}(\text{past}\#) &= \# \times n \\ \text{complexity}(\text{sb}\#) &= (\# + 1) n \\ \text{complexity}(\text{ab}) &= n \\ \text{complexity}(\text{ss}) &\leq (t - t_s) \times n \\ \text{complexity}(\text{during}) &= n \end{aligned}$$

- *Space Complexity:* All of the access requests (granted or denied) are stored in History Base. Storing all the requested accesses in the system, gradually, requires a huge amount of storage space. In case of a vast amount of history elements, and thus, incapability of keeping all these elements on volatile storage, time complexity of access control process is amplified too.

## 9. Conclusion

In this paper, an extension of the Semantic Based Access Control model (SBAC), in order to enhance its capabilities by taking the history of accesses of the system into account was proposed. The proposed model (named TSBAC) uses the same semantic relationships of SBAC, and moreover, it is capable of using temporal relations between the access events occurred in the past (composed as a temporal expression) in specifying authorization rules. In this model, a dynamic aspect are also given to the authorizations by attaching a time interval to each of the authorizations that restricts the validity period of them. These authorization rules (which are composed of base authorization of SBAC, validity time interval, and temporal expression over the history of users' accesses in the system), provides the ability to derive new authorizations based on the occurrence (or not occurrence) of other accesses in the past.

The formal semantics of our authorization rules, plus the access control and conflict resolution procedures were proposed. An architecture for an access control system based on TSBAC was presented. Finally, a banking environment was modeled using TSBAC as a case study.

A generalized history-based access control model that could be applied to other access control policies (such as RBAC) is one of our important future works that will be done.

## References

1. Samarati, P. and S. Vimercati. *Access Control: Policies, models, and mechanisms*. in *Foundations*

2. Bertinoro, Italy: Springer-Verlag.
3. S. Javanmardi, A. Amini, and R. Jalili. *An Access Control Model for Protecting Semantic Web Resources*. in *Web Policy Workshop*. 2006. Ahens, GA, USA.
4. Javanmardi, S., et al. *SBAC: "A Semantic-Based Access Control Model"*. in *NORDSEC-2006*. 2006.
5. Bonatti, P.A., et al., *Semantic web policies: a discussion of requirements and research issues*. ESWC, 2006. **2006**: p. 712-724.
6. Rabitti, F., et al., *A Model of Authorization for Next-Generation Database Systems*. ACM Transactions on Database Systems, 2001. **16**(1): p. 88-131.
7. Qin, L. and V. Atluri. *Concept-level access control for the Semantic Web*. in *2003 ACM workshop on XML security*. 2003.
8. Brewer, D.F.C. and M.J. Nash. *The Chinese Wall Security Policy*. in *IEEE Symposium on Security and Privacy*. 1989. Oakland, California.
9. Dias, P., C. Ribeiro, and P. Ferreira, *Enforcing History-Based Security Policies in Mobile Agent Systems*. 2003.
10. Edjlali, G., A. Acharya, and V. Chaudhary. *History-based access control for mobile code*. in *5th ACM conference on Computer and communications security*. 1998.
11. Abadi, M. and C. Fournet. *Access control based on execution history*. in *10th Annual Network and Distributed System Security Symposium*. 2003.
12. Bertino, E., et al., *A temporal access control mechanism for Database Systems*. IEEE Trans. Knowl. Data Eng. 1996. **8**(1): p. 67-80.
13. Thomas, R.K. and R.S. Sandhu. *Sixteenth National Computer Security Conference*. 1993. Baltimore, Md.
14. Bertino, E., C. Bettini, and P. Samarati. *A temporal authorization model*. in *Second ACM Conference on Computer and Communications Security*. 1994. Fairfax, Va.
15. Bertino, E., et al., *An access control model supporting periodicity constraints and temporal reasoning*. ACM Trans. Database Systems, 1998. **23**(3): p. 231-285.
16. Ruan, C., *Decentralized Temporal Authorization Administration*. 2003.
17. Bertino, E., et al., *Temporal authorization bases: From specification to integration*. Journal of Computer Security, 2000. **8**: p. 309-353.
18. Emerson, E.A., *Temporal and modal logic*. Handbook of Theoretical Computer Science, 1990. **B: Formal Models and Semantics (B)**: p. 997-1072.

# A new adaptive routing approach based on Ant Colony Optimization (ACO) for Ad hoc Wireless Networks

Niaz Morshed Chowdhury<sup>1</sup>, Syed Murtoza Baker<sup>2</sup>, Ershadul H. Choudhury<sup>3</sup>

University of Dublin, Trinity College<sup>1</sup>, East West University<sup>2,3</sup>  
[chowdhun@cs.tcd.ie](mailto:chowdhun@cs.tcd.ie)<sup>1</sup>, [galib@ewubd.edu](mailto:galib@ewubd.edu)<sup>2</sup>, [ershad@ewubd.edu](mailto:ershad@ewubd.edu)<sup>3</sup>

## Abstract

*The goal of this work is to design a new adaptive routing technique for ad hoc wireless networks. This paper proposed the basic design of the algorithm that works based on the principle of Ant Colony Optimization (ACO). This is a probabilistic adaptive technique that changes its routes with the change of network topology over the period of time by learning its environment. It identifies appropriate paths with the feedback of previously travelled packets and maintains routing table accordingly. A self-made simulator implemented on C++ is used to evaluate performance of this algorithm on the basis of diverse adaptive issues such as Change of Probability, Growth of Pheromone Intensity, Randomness of the Selection and Packet sending rate through different paths.*

**Keywords:** Ant Colony Optimization, Ad hoc Network, Adaptive System, Routing Protocol

## 1. Introduction

Mobile ad hoc network (MANET) started to become popular during mid 90's [1]. In between 1994 and 2004, hundreds of routing protocols came though most of them failed to survive due to their weak proactive nature [2, 3, 4]. During this period four Ad hoc routing protocols became extremely popular. In 1994 Charles E. Perkins proposed his first routing protocol *Dynamic Destination Sequenced Distance Vector (DSDV)* [5]. Two years later in 1996, David B. Johnson proposed his famous Ad hoc routing protocol *Dynamic Source Routing (DSR)* [9,10,11]. In 2002 Niklas Beijar proposed another protocol, *Zone Routing Protocol (ZRP)* [12] though it maintains few restrictions and also lacks in its generalized nature. During this period Perkins was working in Norka Research Centre and developing a new routing protocol and finally in July 2003 his most famous and most commonly used routing protocol *Ad hoc On-demand Distance Vector (AODV)* [6] routing protocol got accepted as Request For Comments (RFC) [7].

Since then DSR and AODV have been dominating ad hoc arena with promising success.

Though DSR and AODV are the most popular routing protocols of current time, they are not perfect yet. Biggest drawback of DSR is its packet overhead as it has to send large control information with the packet. On the other hand, AODV broadcasts huge control data (RREQ packet) to its neighbouring router while establishing a route which increases its bandwidth overhead. Since then a number of approaches has been tried to either improve their performances or implement parallel routing protocols.

## 2. Related study

Due to the NP complex nature of ad hoc routing problem, it is pretty hard to get a perfect solution or even a better solution without increasing the bandwidth or packet overhead. All those problems led the researchers to work with Adaptive protocols for ad hoc routing. In 2001 a routing protocol came named Adaptive Distance Vector Routing [14]. In the same year Adaptive Demand Driven Multicast [15] routing came. There was another algorithm named Fuzzy Sighted Link State Algorithm [15]. In July 2006 M. Roberts Masillamani proposed a GA based routing protocol for conventional distance vector routing and they are currently working to convert this protocol to a GA based adaptive AODV [16].

## 3. Ant Colony Optimization

Ant Colony Optimization (ACO) [17, 18] is an adaptive technique to find partial solutions for the problems where identifying exact solution is either difficult or impossible [19, 20]. ACO is proposed by Dr. Marco Dorigo in 1992 and inspired by the behaviour of ants in finding paths from the colony to food. In 1997, Dr. Marco Dorigo in his work *Ant Colony System: A comparative learning approach to the Travelling Salesman Problem* [21] hinted that network routing can be implemented based on Ant



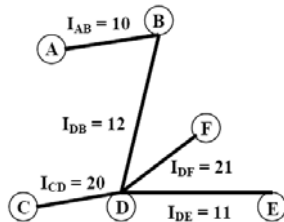
Algorithm. Very next year he published another research work [22] where he discussed that packets and ants are typically identical for an adaptive system and routing algorithm can be formed or improved based on ant algorithm. In 1999, in another paper, *Ant Algorithm for Discrete Optimization* [31] he proposed the same idea. Despite the fact that he always recommended his algorithm to be used in computer networks, neither he nor other researchers showed enough interest in this regard.

#### 4. Design

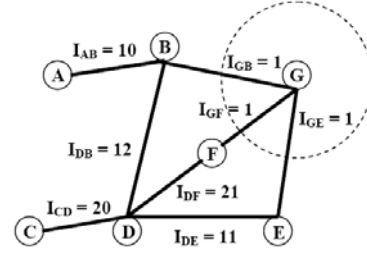
ACO is a generalized adaptive technique [27, 28, 29, 30] like Fuzzy Logic [25, 26], Genetic Algorithm [16] or Simulated Annealing [24]. It has different phases such as Initialization, Probability Calculation, Selection, Pheromone Deposition, Pheromone Evaporation etc. which are problem dependent [31]. Each of those phases needs to be designed based on the problem demand and nature. Rest of the part of this section will talk about the design issue of those phases for the proposed routing technique.

##### 4.1 Initialization

ACO is a meta-heuristic method that requires to be initialized at the beginning of the procedure [23]. Conventionally in ACO, all the ant tracks (here routing paths) are initialized with numerical value *zero* because it is assumed that initially there is no pheromone on the tracks. But proposed algorithm initializes its paths in a different way; with numerical value of *one* instead of zero. The rational behind this modification is quite straightforward – in this algorithm a non-zero value of intensity represents a connection which means pheromone intensity also indicates connectivity of nodes and paths having zero intensity indicate there is no connection between two particular nodes.



**Figure 1:** A network topology with a stable pheromone intensity of the paths.



**Figure 2:** Changed topology when a new router G joined the network.

In Figure 1, a stable network topology has been shown where pheromone intensities of the paths are  $I_{AB}$ ,  $I_{CD}$ ,  $I_{DB}$ ,  $I_{DF}$  and  $I_{DE}$ . Let us assume that a new node G has just joined the network (shown in Figure 2), hence creates three new paths GB, GF and GE. Conventional approach of ACO would initialize these three paths with zero intensity as those new paths have not been travelled by any ant (packet) yet. But, as per our proposed algorithm initialization will be done with one in order to indicate that there is a path available between B and G, F and G and E and G.

##### 4.2 Probability Calculation

Probability calculation is one of the most important issues for any adaptive algorithm and success of the proposed technique is highly influenced by the proper calculation of the path probability as while sending a packet from one node to another node, which path should get selected is determined by this probability. Proposed technique calculates this parameter based on the pheromone intensity of different paths. Probability of a path is defined as,

$$P_i = \frac{I_i}{I_t}$$

**Equation 1:** Path probability

Here,

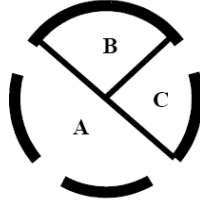
$P_i$  is the probability of a particular path  $i$  connected with the respective node.

$I_i$  is the intensity of a particular path  $i$  connected with the respective node.

$I_t$  is the summation of intensity of all paths connected with the respective node.

##### 4.3 Selection

Selection is done by single spin roulette wheel method. Space allocation for each path on the wheel is proportional to its probability. Say for example, there are three paths A, B, and C connected with a router having probability  $P_A=0.50$ ,  $P_B=0.25$  and  $P_C=0.25$ . Arrangement on the wheel for this particular case is shown in Figure 3.



**Figure 3:** Roulette Wheel arrangement for PA=0.50, PB=0.25 and PC=0.25.

Another important issue in selection is generation of pseudo random number as weakness of random number generation may create initial fluctuation problem [31]. In real life, most of the built-in random number generator functions return close value if they have been called with in a short duration of time (such a 1 millisecond) as they take current time in second as seeds. A method for well distributed random number generator has been proposed with this algorithm. A C++ code segment has been given bellow. This function takes a *value* as a parameter and returns a well distributed random number in between *zero* and *value*.

#### Random Number Generator Function

```
Int random(int highest)
{
 srand((unsigned)time(0));
 int random_integer;
 int lowest=0;
 int range=(highest-lowest)+1;
 random_integer = lowest
 +int(range*rand()/(RAND_MAX
 + 1.0));
 return random_integer;
}
```

#### 4.4 Pheromone Deposition

In this proposed algorithm, pheromone deposition is done by the *backward* packets. After a successful delivery, each packet increases the intensity of each path that it has travelled to deliver the packet based on the following formula.

$$I_{new} = I_{old} + \frac{Q}{Lenth\_Ratio}$$

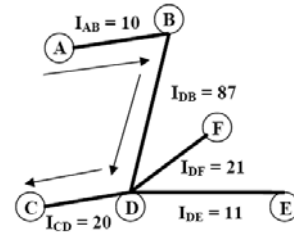
**Equation 2:** Pheromone Deposition Formula

Here,  $Q$  is a unit quantity of pheromone to be deposited. But, all the paths wont get the full portion of  $Q$  rather deposition of  $Q$  will be inversely proportional to the travelled Path.

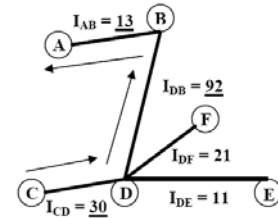
Here **Length Ratio** varies from node to node. For each node on the path it is measured as the path length between the source and the calculated node.

In this technique, backward packet will deposit more pheromone on the shorter path compare to the longer path while coming back from the destination. As the probability calculation is dependent of the pheromone density which ultimately increases the chance of best path to be selected.

A pictorial example is shown here to describe the above literature. Consider the following network topology shown in Figure 4 and 5. In Figure 4, it shows that A has made a successful delivery of a packet to C. During this travel, this packet marked the paths that it has travelled.



**Figure 4:** A is sending a packet to C.



**Figure 5:** Intensity of the paths is being changed by the backward packet.

Let us consider  $Q = 10$ . Based on the Equation 2 pheromone of each track will be updated as follows:

$$\begin{aligned} I_{AB(new)} &= 10 + (10 / 3) = 13 \\ I_{DB(new)} &= 87 + (10 / 2) = 92 \\ I_{CD(new)} &= 20 + (10 / 1) = 30 \end{aligned}$$

Figure 5 shows backward packet sent by C to A. This backward packet will increase the intensity of the paths (underlined) that it has travelled while coming from A to C.

#### 4.5 Pheromone Evaporation

Pheromone evaporation is another issue in ACO. It decreases pheromone from less travel paths. In this algorithm pheromone evaporation is done based on a parameter called *evaporation parameter*,  $\lambda$ . Value of  $\lambda$  depends on the traffic density of the network. If Intensity of any path at a junction crosses the value of  $\lambda$ , intensity for all the paths of that junction will be changed as follows:

$$I_{i(new)} = \frac{I_{i(old)}}{GCD(I_{1(old)}, I_{2(old)}, \dots, I_{k(old)})}$$

**Equation 3:** Pheromone Evaporation Formula

Here,

$I_{i(new)}$  is the Intensity of  $i$  path after the evaporation.

$I_{i(old)}$  is the Intensity of  $i$  path before the evaporation.

$k$  is the number of connected paths with that router.

## 5. Operation

A routing technique has to perform three basic operations, namely *Route Establish*, *Route Discovery* and *Route Maintenance* [13]. Nature and performance of a routing algorithm is determined on how does it deal with these three tasks. Followings are the brief operational summary of the proposed technique.

### 5.1 Route establish

Route establishment is a prime and foremost task for any routing procedure. In our proposed technique, at the beginning of the network or when a new node enters into the network, it finds its neighbouring node and initializes local intensity value of the corresponding paths with unit value (i.e. 1). Initially it also assumes that all the nodes of the network are reachable through any of those paths. It maintains a vector table like distance vector routing where it puts the intensity of different paths. This table is arranged in the conventional form of “*from A to B via C*” [13]. But instead of distance value, it holds intensity of a particular link. As time passes, gradually best outgoing link for a particular node becomes stable in the intensity table.

### 5.2 Route discovery

This routing technique does not send any route request packet like AODV or any service discovery like DSR. Still it maintains a reactive on demand nature. Before sending any packet to a particular destination, it finds the path based on a probabilistic selection mentioned earlier in

section 3. Path having been travelled most to reach a destination will get higher priority in the selection. This route discovery is an adaptive method based on Ant Colony Optimization and it becomes stable eventually by learning the environment.

### 5.3 Route maintenance

Proposed technique maintains its route in an adaptive manner. When a node gets down from the network, no packet can successfully be delivered through the path that goes over that particular node. As a result pheromones of that path will be evaporated and other paths will start to get the priority in selection. If the down node gets back to the network then this path will be established again otherwise due to periodical pheromone evaporation this path will become a dead link.

## 6. Result and Evaluation

A self made simulator implemented on C++ is used to evaluate performance and behaviour of the proposed technique. This simulation mainly concentrates in measuring and evaluating adaptive natures of the algorithm. A 50 node scenario is considered taking two nodes as source and destination where three different and complex paths are available to reach the destination from the source node. Among the paths, *Path 2* is the best one while *Path 3* and *Path 1* are the second and third best respectively. We measured following four issues to evaluate our proposed technique:

1. Change of Probability
2. Growth of Pheromone Intensity
3. Randomness of the Selection
4. Packet Sending Performance

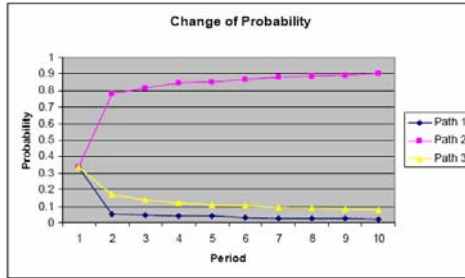
### 6.1 Change of Probability

Change of probability is measured over the period of 900 packet send time from the source node to the destination. Initially when the very first packet was sent, the probability of each path was equal, i.e. 0.33. As time increased, gradually best path was established and the worst two got almost down. Figure 8.1 shows change of probability of different paths within an interval of 100 packet send period.

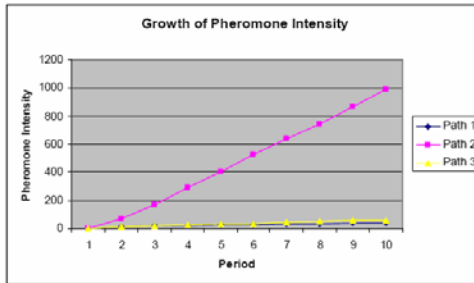
### 6.2 Growth of Pheromone Intensity

In this part of the evaluation once again 900 packets were sent and within an interval of each 100 packet send time measurement was taken. Initially pheromone intensity of all three paths was 1. Figure 8.2 shows change of intensity due to pheromone deposition on the experimental

paths. Result clearly shows that best path got more pheromone deposition and its intensity increased. Intensity of other two paths was increasing with an extremely slow rate.



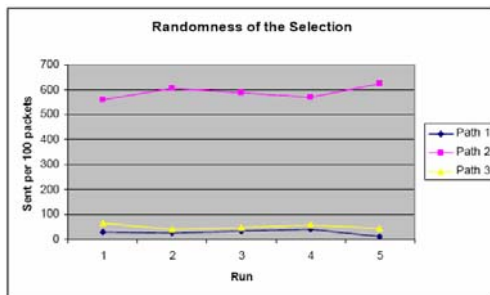
**Figure 6:** Change of Probability in the test topology



**Figure 7:** Growth of Pheromone Intensity

### 6.3 Randomness of the Selection

Randomness of the selection has been measured by this experiment maintaining the same network topology over five different runs. Figure 8.3 shows the result where it is clearly noticeable that the variation is very marginal in different runs.

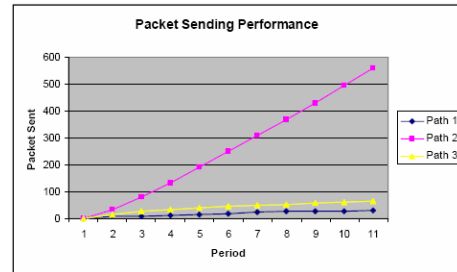


**Figure 8:** Randomness of the Selection

### 6.4 Packet Sending Performance

In this final evaluation total number of packet sent through each path is measured for the same topology over the period of 100 packet sent time. Figure 8.4 shows, as time increased, Path 2 (best

path) got selected more frequently for sending packets from the source node to the destination.



**Figure 9:** Packet Sending Performance

## 7. Conclusion and Future Work

An ant navigation based adaptive routing technique is proposed and evaluated in this paper and we are confident that it will work effectively in the real ad hoc network scenario. This work mainly concentrates on designing different adaptive parameters of the technique rather than concentrating in the network layer issues. Our future objective is to implement this algorithm using OPNET [32] for mobile ad hoc network and compare its performances with AODV and DSR.

### Reference:

- [1] Larry L. Peterson and Bruce S. Devis, "Computer Networks – A Systems Approach". San Francisco, Morgan Kaufmann Publishers Inc. ISBN 1-55860-368-9
- [2] Dimitri Bertsekas and Robert Gallager, "Data Networks – 2nd Ed". Prentice Hall, New Jersey, ISBN 0-13-200916-1
- [3] Scott Corson, S. Papademetriou, Philip Papadopolous, Vincent D. Park and Amir Qayyum, "An Internet MANET Encapsulation Protocol (IMEP) Specification". Internet Draft, draft-ietf-manet-imepspec01.txt, August 1998. Work in progress.
- [4] Santashil Pal Chaudhuri, Sridhar Lavu. A performance Comparison of Ad Hoc Networks Routing Protocols.
- [5] Charles E. Perkins and Pravin Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for mobile Computers". In the proceedings of SIGCOM '94 Conference on Communications Architecture Protocols and Applications, pages 234 – 244, August 1994

- [6] Charles E. Perkins, "Ad hoc On Demand Distance Vector (AODV) Routing", Internet Draft, draft-ietfmanet-aodv-01.txt, August 1998. Work in progress.
- [7] C. Perkins. Ad Hoc On-Demand Distance Vector Routing (RFC - 3561)
- [8] Martha Steenstrup, "Routing in Communication Networks". New Jersey, Prentice Hall. ISBN 0-13-010752-2
- [9] Josh Broch, David A. Maltz, David B. Johnson, Yih-Chun Hu and Jorjeta Jetcheva, "A performance Comparison of Multi-hop Wireless Ad hoc Network Routing Protocols". Mobicom '98, Dallas Texas, 25-30 October, 1998.
- [10] David A. Maltz, David B. Johnson, "Protocols for adaptive wireless and mobile computing". In IEEE Personal Communications, 3(1), February 1996.
- [11] David A. Maltz, David B. Johnson, "Dynamic source routing in ad hoc wireless networks". In mobile computing, edited by Tomasz Imielinsky and Hank Korth, Chapter 5, pages 153-181. Kluwer Academic Publishers.
- [12] Zygmunt J. Haas and Marc R. Pearlman, "The Zone Routing Protocol (ZRP) for Ad Hoc Networks". Internet Draft, draft-ietfmanet-zone-zrp-01.txt, August 1998. Work in progress.
- [13] Tony Larsson, *Routing Protocols in Wireless Network*
- [14] <http://www.comp.brad.ac.uk/~sburuha1/routingprot.htm> (Last accessed 25 July, 2008).
- [15] [http://en.wikipedia.org/wiki/Ad\\_hoc\\_routing\\_protocol\\_list](http://en.wikipedia.org/wiki/Ad_hoc_routing_protocol_list) (Last accessed 25 July, 2008)
- [16] M. Roberts Masillamani, Avinankumar Vellore Suriyakumar, Rajesh Ponnurangam, G.V. Uma, *Genetic Algorithm and Distance Vector Routing Technique*
- [17] Dorigo, M. (1992). *Optimization, learning and natural algorithms* (in Italian). Unpublished doctoral dissertation, Politecnico di Milano, Dipartimento di Elettronica, Italy.
- [18] Dorigo, M., Maniezzo, V., & Colormi, A. (1991). *Positive feedback as a search strategy*. (Tech. Rep. 91-016). Milan, Italy: Politecnico di Milano, Dipartimento di Elettronica
- [19] Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *From natural to artificial swarm intelligence*. New York: Oxford University Press
- [20] Corne, D., Dorigo, M., & Glover, F. (Eds.). (1999). *New ideas in optimization*. Maidenhead, UK: McGraw-Hill
- [21] Marco Dorigo, "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem". Universite Libre Bruxelles, Belgium.
- [22] AntNet: Distributed Stigmergetic Control for Communications Networks.
- [23] Pasteels, J. M., Deneubourg, J.-L., & Goss, S. (1987). Self-organization mechanisms in ant societies (i): Trail recruitment to newly discovered food sources. *Experientia Supplementum*, 54, 155–175
- [24] Hin Leong Tan, Saul B. Gelfand, "A Cost Minimization Approach to Edge Detection Using Simulated Annealing.", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 14, No. 1.
- [25] Raghu Krishnapuram, "Fuzzy and Possibilistic Shell Clustering Algorithms and Their Application to Boundary Detection and Surface Approximation-Part I", IEEE TRANSACTIONS ON FUZZY SYSTEMS, VOL. 3, NO. 1
- [26] Patrick K. Simpson, "Fuzzy Min-Max Neural Networks-Part 2: Clustering", IEEE TRANSACTIONS ON FUZZY SYSTEMS, VOL. 1, NO. 1
- [27] Fogel, D. B. (1995). *Evolutionary computation*. Piscataway, NJ: IEEE Press
- [28] Holland, J. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press
- [29] Rechenberg, R. I. (1973). *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart, Germany: Frommann-Holzboog
- [30] Schwefel, H.-P. (1977). *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*. Basel, Switzerland: Birkhäuser
- [31] Marco Dorigo Gianni Di Caro, "Ant Algorithms for Discrete Optimization". Artificial Life, MIT Press 1999
- [32] [www.opnet.com](http://www.opnet.com) ((Last accessed 25 July, 2008))

# Capacity Based Channel Assignment in Multi-Interface Wireless Mesh Networks

Mahmudur Rahman, Anjali Agarwal, Ayoub Alsarahn  
Department of Electrical & Computer Engineering  
Concordia University  
Montreal, Quebec, Canada  
{mah\_r, aagarwal, a\_alsar}@ece.concordia.ca

## Abstract

*To utilize the available frequency channel space in wireless mesh networks (WMNs), recent work has significantly been focusing on the channel assignment in multi-interface and multi-channel wireless mesh network. This paper presents our design of a distributed channel assignment algorithm that assigns channels based on capacity of the channels. It is shown that the performance of the network is significantly improved when interference and capacity are considered to choose a channel. Simulation results show that our channel assignment algorithm performs better than a similar channel assignment algorithm which is distributed and based only on interference.*

## 1. Introduction

Wireless Mesh Network (WMN) is the next generation wireless internet paradigm that is under progress for commercial deployment. The performance of the entire system of wireless mesh networks is influenced by the radio resource management (routing, transmission power, channel assignment). Power constraints and mobility issues are not the major concern for such networks in current research. Although significant progress has been achieved during the last decade, it still requires more work to achieve higher performance.

Recent research has shown that each node in a wireless mesh network greatly enhances the performance of wireless mesh networks by using multiple radios [1]. A typical wireless mesh network requires a common channel between two nodes to communicate. Deployment of multi radio introduces the availability of having multiple channels to improve the system throughput of WMNs. Use of multiple channels is introduced to avoid the interference in adjacent access points in a network. Proper channel assignment algorithms are therefore required to utilize multiple channels efficiently.

Previous work done in [1] shows that 802.11a offers twelve and 802.11b offers three non overlapping channels. Although multiple channels and multiple interfaces are used to improve the throughput of the network, inappropriate channel assignment will cause the throughput reduction [2]. Thus, a

channel assignment algorithm plays as one of the most important roles for the performance of such networks. In this paper we aim to design a channel assignment algorithm for multi-channel wireless mesh networks that enhances the network performance with better throughput and reduced delay.

The rest of the paper is organized as follows. Section II presents the review of the previous work in channel assignment for wireless mesh network. Section III motivates the benefits of having multiple interfaces and multiple channels. Fundamental issues for channel assignment and different kinds of channel assignment strategies are shown in Section IV. Section V describes our algorithm and evaluates the performance result of our algorithm. Section VI concludes the paper.

## 2. Related Work

Several mechanisms have been proposed in [1-8] for channel assignment in wireless mesh network. These are different in behavior and have focused in separate issues in wireless mesh network. The main purpose of most of these researches is to find optimality for packet transmission and to enable multiple packet transmission with better throughput in the network. These are mainly categorized into two different kinds of algorithms: centralized and distributed.

A centralized algorithm has the ability to process and assign channels to different nodes based on the information stored in a central entity. All nodes send their channel information (interference level, default channels available and channels already assigned) to the central entity which assigns channels for nodes based on this information. As a consequence, more bandwidth is consumed. The main drawback of this approach is the complete failure of the network in a situation where the channel assignment server would fail.

Distributed algorithms have the advantage of having a better ability to cope with failures (fault-tolerance). A distributed algorithm must have the ability to be self stabilizing when used for routing purposes, for example, no matter what the initial configuration is, the algorithm will eventually achieve its intended purpose. In the distributed assignment scheme, each node collects information from his neighbors only to

select the channels and consequently not consume large bandwidth. This kind of approach is scalable but it has the limitation of increased computing power in each node of the network.

In [1] a hybrid channel assignment is proposed where each node of the network is assigned as one fixed and one dynamic channel. It has shown that using this mechanism, nodes can both transmit and receive data simultaneously. The fixed channels in the interface are only used to receive the data and the switchable one chooses a channel according to the demand to transmit the data to other nodes. However, no mechanism is shown in this research to choose the proper channel when it is communicating to other nodes to improve the performance of the network.

The algorithm proposed in [2] is mainly focused on the multicast mechanism in wireless mesh network. Using a centralized approach, this algorithm improves the system throughput by allowing simultaneous close-by transmissions with multi-channel and multi-interference. It assigns all the available channels to interface. The channel assignment and routing is optimized according to the change of topology and node failure. Thus, it does not assure the stability of the network.

The algorithm presented in [3] provides a mechanism to recognize the node failure and to restore the connectivity. It reduces the number of hop count in the network and also decreases the interference between the nodes by considering node levels. It uses a randomized channel assignment in a distributed manner and maintains the network connectivity.

A breadth search based centralized channel assignment was proposed in [4], which mainly focuses on the interference level and better throughput of the network. It uses a default channel for transmission and has a link redirection protocol to improve the stability of the network. However, it does not consider the routing issues.

In [5], the proposed algorithm for channel assignment can minimize the overall interference of a multi-hop wireless mesh network. The result shows that this algorithm preserves connectivity and the interference level decreases with the increase of radios. Routing and stability issues are however not considered in this algorithm.

The channel assignment algorithm described in [6] considers link bandwidth and channel diversity to transmit a packet from source to destination. Although it works better in shortest path, it gives 35% more improvement than any other algorithms for a path that has more than 5 hops.

In [7], the goal of the channel assignment algorithm is to improve the throughput. It does not deal with the interference avoidance. This algorithm is used to assign channels to the network interfaces such that the resulting available bandwidth on these interfaces is at least equal to their expected traffic load. It is measured by calculating the expected bandwidth needed for traffic to each pair of node. Connectivity of the network is not considered in this algorithm.

In [8], a distributed channel assignment algorithm was proposed that has self-stabilizing mechanism which assigns channels to multi-radio nodes in wireless mesh networks

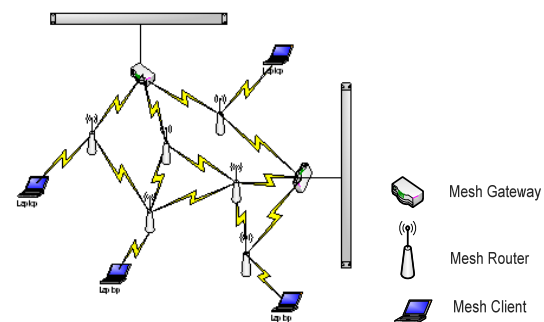
based on the interference of the channels with the neighboring nodes. The scheme is stable to a desirable channel configuration that strikes a good balance between network connectivity and channel diversity, and that routing protocols can exploit to provide better end-to-end system performance. The design considers several constraints present in current 802.11 devices, and its distributed nature ensures it is sufficiently light-weight to be executed on large scale mesh networks. However, issues such as channel capacity and switching delay were not considered here.

We have chosen the algorithm described in [8] to compare the proposed algorithm. In [8], even though the channel assignment mechanism is done in the real test bed, we have chosen almost the similar environment in our simulation. The mechanism shown in [8] focuses to utilize the 802.11 channels within the mesh routers [8]. Unlike the other channel assignment algorithms described previously, it considered the real wireless mesh network environment with 14 nodes. It emphasizes only on the channel assignment issue for wireless mesh network. We are dealing with multi-channel issues for the wireless nodes which were also a key concern for this algorithm. The distributed nature of this channel assignment enhances major improvement by balancing between connectivity and stability of the network and thus assures end to end quality of the network [8].

### 3. Motivation

In this section, we first describe the wireless mesh network architecture. Thereafter we show the motivation to have a proper channel assignment in wireless mesh networks.

#### 3.1 Network Architecture



**Figure 1. Wireless Mesh Network architecture**

Wireless mesh networks provide users with more services when compared to other wireless network types. The network contains two types of nodes: mesh routers and mesh clients [10, 11]. Conventional users can connect to mesh network through mesh routers if their system have a wireless network interface, otherwise they can connect via Ethernet links. Mesh router can have multiple wireless interfaces built on either the same or different wireless access technologies.

As a result, many wireless networks with different technologies can be integrated through mesh routers. WMNs network architecture can be classified into three groups [11, 12]:

**Mesh Gateway:** Mesh gateway act as the backbone of the network. It has necessary functions for mesh networking. They can be connected to other networks through gateways or bridges.

**Mesh Router:** Mesh routers perform routing and configuration functions as the access point. Because they provide extra functions, end user devices need more resources to establish a connection with them. Mesh routers are considered to have less mobility or no mobility at all than clients. It is considered to be immobile or with small mobility.

**Mesh Client:** Mesh clients are the end users of mesh network. They can connect both with the gateway and access point routers. Mesh clients are considered to be mobile.

### 3.2. Benefits of having Multi-Interface

From [1], we found that the previous works in wireless mesh network used single interface. It is because the networks were maintained by single channel for all the nodes. Multi-interface is beneficial when the network is having multi-channel for the nodes. As mentioned in Section I , wireless mesh networks nodes need to be in the same channel to establish a connection between them. If we use a single interface for the nodes then it has to use a channel to send or receive a packet from other node. While it is in communication with another node no other node can communicate with that specific node. As a result, when a single interface is used, packets might have to wait to use that specific node or to find an alternative route which will also have to provide a common channel. This increases end to end delay for the network if different hops are not in the common channel. Another important benefit that multi-interface provides is to transmit and receive data simultaneously. If a node has multi-interfaces working in multi-channels, both interfaces can transmit and receive data simultaneously. This doubles the maximum throughput in a multi-hop route [1].

### 3.3. Necessity of Multi-Channel

The deployment of multi-channel in wireless mesh network makes the whole network very flexible. Multi-channel makes the network robust, since it increases the optimality of the route in the network. The connectivity might not affect the whole network performance if the nodes have multi-channel to use in their interfaces. Also, from the previous section we get the understanding that to run simultaneous communication for multi-interface we need multi-channel to avoid interference. In a single channel environment, if most of the nodes have the same channel then interference will result in reduction of the total throughput. Multi-channel gives the flexibility of selecting routes that uses different channels. Thus, having these benefits really made the difference from a single channel network.

### 3.4. Need for Proper Channel Assignment

Channel assignment determines the connectivity between radios since two radios can communicate with each other only when they are on a common channel, and therefore it determines the network topology. As we know, the routing decisions are made based on the network topology. Thus, channel assignment has a direct impact on routing. Channel assignment should be dynamically adjusted according to the traffic status, which is determined by a routing algorithm. Therefore, channel assignment and routing are tightly coupled with each other.

Channel assignment also puts an impact on the connectivity of the network. The concept of channel assignment is not constant and hence it depends on the network classification. Different channel assignment schemes were proposed for Single Radio networks and Multi-Radio Networks; so, the connectivity changes accordingly. It improves the channel diversity for the nodes in a wireless network which in the long run enhances the quality of the network.

The use of single channel in wireless mesh network cannot fully use the aggregate bandwidth available in the network. Using multi channel avoids this problem and gives a better performance for the network. The purpose of channel assignment in a multi channel wireless mesh network is to assign channels efficiently so that the total interference of each link is reduced, thus resulting in an improvement of the total throughput. As more interfaces within the range are available and use same channel within the interference range, the bandwidth available per interface will surely decrease. Therefore, a proper channel assignment algorithm is needed to balance the network connectivity and improve the aggregate bandwidth. The delay for each link will also depend on the channels used for each node. With a proper channel assignment the delay can also be reduced.

## 4. Channel Assignment in WMNs

In this section we provide the fundamental issues for channel assignment, different channel assignment strategies and our algorithm for channel assignment in wireless mesh networks. Several key performance issues such as: interference, connectivity, stability and routing; are considered in our research to find a proper channel assignment algorithm in wireless mesh networks. Using a proper channel assignment and maintaining these constraints constraints, a good quality of service of a network can be guaranteed.

### 4.1. Fundamental Issues for Channel Assignment

Several channels are employed to establish the communication with the other nodes. Numerous issues are considered to assign the channels. These issues set the performance of the network. Using a proper channel assignment and maintaining these constraints, a good quality



of a network can be guaranteed. The issues that we consider in this paper are described as follows:

**Interference:** The main obstacle to achieve high capacity wireless mesh networks is interference between the mesh links. This is the major factor that degrades the performance of the network. The throughput of the network has a direct relationship with the interference among the links. Due to interference, the quality can change arbitrarily with any change in the routing pattern. As such, these measurements have limited predictive value because they cannot tell whether re-routing existing flows would result in better network throughput or which path is best for a new flow. Mesh networks however operate in the unlicensed bands and have no particular claim to the spectrum they use. They must share frequencies not just with other wireless LANs but with a broad range of products, from cordless phones, microwave ovens to wireless video surveillance cameras. In addition, physical objects like buildings, mountains and bridges, pose major problems for wireless transmissions. Interference can be classified in two categories: co-channel interference and adjacent channel interference.

**Connectivity:** A network is said to be connected, if for every pair of nodes there exists a path between them; otherwise it is disconnected. Connectivity is a fundamental property of wireless mesh networks. Unlike cellular networks, a multi-hop wireless mesh networks has a decentralized approach to establish a connection between links which is more sophisticated. To achieve a connected network, a wireless multi-hop path must exist from each node to each other node. Each single node contributes to the connectivity of the entire network. If a node fails, the connectivity might be changed. The probability for a network to be connected depends on the density of nodes and their transmission ranges. However, the connectivity is not the same for single-channel and multi-channel networks. Based on this, channel assignment can change the topology of the wireless mesh networks. There is a trade-off relationship between connectivity and interference. The more radios are assigned to the same channel, the more connectivity is achieved and more interference is induced.

**Stability:** Stability plays a very important role in the successful deployment of any wireless network. It may be regarded as one of the most important aspects, since a network does not exist without a connection. If a network technology offers a secure routing mechanism along with stable connection among the various participants, the network is anticipated to be a stable one. To maintain stability, a channel assignment is done in such a way that the channels in the interfaces of the nodes do not change back and forth among several choices. When two nodes discover that a channel is under-utilized, they might simultaneously switch to this channel and both begin the transmission on it, and then switch back upon discovering this channel is overloaded.

**Routing:** The topology of the network changes accordingly to the route taken for that network. For wireless mesh networks, the basic factor of making a routing decision is done on channel assignment. If the channels assigned for the nodes in the network are changed, the routing will also be changed,

thus the network topology changes. Even though wireless routers are typically stationary in this kind of network, routes in this network are expected to be unstable. Because of issues like multi-path fading effects, external interference and weather conditions, wireless links vary widely in their qualities. Link interference can lead to variations in the quality of mesh routes, which can result in route change. So, routing is a very important issue to be considered in a channel assignment. For an example, channel assignment schemes in [1] for multi-radio mesh networks assign channels to frequently diversifying routes in the mesh. If routes are expected to change, the mesh radios should also be re-assigning channels in order to ensure optimal network performance.

## 4.2. Channel Assignment Strategies in Interfaces

Channel assignment in multi-interfaces can be classified into three different kinds. These are static or fixed, dynamic, and hybrid channel assignment.

**Static Channel Assignment:** In static channel assignment, each interface of the nodes is assigned a channel permanently or for a long period of time. Such assignments are relatively simple but cannot adapt to the capacity changes. Different channels can be assigned for adjacent links but channel must be assigned to the links with a certain distance. This is the easiest one among all the strategies, but provides the worst channel utilization. It also cannot adapt to different traffic load in different links.

**Dynamic Channel Assignment:** In dynamic assignment, channels are assigned for different nodes for a small period of time. The advantage of this kind of strategy is to have the ability of switching one interface to any channel, thereby having the choice to use many channels in very few interfaces. It has some drawback too. It gives worse performance than static channel assignment under heavy traffic loads.

**Hybrid Channel Assignment:** Hybrid channel assignment is the combination of both static and dynamic channel assignment. It divides the number of channels into two groups, one of which is used for fixed allocation while the other is used to keep the channels to be shared by all the nodes. It needs fewer interfaces than dynamic channel assignment and can also be used with channel reassignment strategy.

Switching the channels between the interfaces has a close relationship with the performance of the network. Multi-interfaces need a proper mechanism to switch the channels properly among the interface to reduce switching delay.

## 5. Capacity based Channel Assignment

In this section, we describe a new distributed channel assignment algorithm. We consider interference and capacity ratio to choose the best channel among all the available channels. To calculate the interference of a channel we have chosen the mechanism shown in [8]. Each node selects a

channel that has less interference with neighbors. Section IV discussed how interference enhances the channel assignment and the performance of the network. We consider the least interference and the highest capacity of a channel to assign channels for a node to improve the performance of the network in terms of throughput and delay.

Capacity of a channel can be defined as the highest upper bound on the amount of information that can be reliably transmitted over a communication channel. In other words, capacity of a channel will decide how fast a node can send or receive information. Every channel has separate capacity and it can vary depending on the frequency spectrum. In our algorithm, each node selects a channel by executing the algorithm shown below.

---

#### *Channel Selection ( node i)*

---

*Input*

*i: node i*

*S<sub>k</sub>: set of channels for node k*

*S<sub>i</sub>: set of candidate channels*

*I: interference for a channel*

*C<sub>i</sub>: the capacity of channel i*

*begin procedure*

*for each i ∈ S<sub>i</sub>*

$$I_i = \sum_{k \in S_k} I(i, k)$$

$$R_i = \frac{C_i}{I_i}$$

*End for*

*Selected \_ Channel = Max(R<sub>i</sub>)*

*end procedure*

---

Each node selects a channel that has less interference and maximum capacity. The cost function to measure the interference,  $f(a, b) \geq 0$ , provides a measure of the spectral overlapping level between channels a and b [8]. The interference cost function in [8] is defined in such a way that  $f(a, b) \geq 0$  and  $f(a, b) = f(b, a)$ , where two channels are justified as non interference channels when the value of the function is 0 [8]. Each node computes the interference between the candidate channel and neighboring channels. From the result observed in [8], which considers only the interference to assign the channel among the nodes, we figured out that considering interference-capacity ratio would

generate noteworthy improvement. The algorithm in [8] chooses a channel based on the least interference. A node selects a channel that has the minimum sum of the interference cost for the set of nodes within the interference range [8]. Considering capacity as a key parameter, we have chosen the maximum value that we can measure after finding the interference capacity ratio for every single channel. Capacity for the channels is taken as a random value since our focus was not to justify the amount of capacity for the channels. Thus, for each node the channels are assigned from the channel list in a way that it gives the best throughput and fast transmission of data from source to destination.

We have already discussed the advantage of multi-interface in wireless mesh networks. The multi channel assigned for all the nodes can be distributed using the algorithm. The interfaces are assigned a channel which is the same channel for the interface used in the destined node. This improves the performance even better if the other interface of the node is used only to transmit and receive the data [7]. Moreover, it assures network connectivity even if any other node enters into the network. That means, this protects the end to end communication from topology changes. In [8], the algorithm likely assigns different channels to the non neighboring nodes. In our proposed algorithm, since the channels are not assigned only based on interference, the information of non neighboring nodes does not make an important impact.

However, network parameters such as link delay, link bandwidth, switching delay and congestions are not the motivation of this paper. These parameters are needed to be considered for an improved routing protocol. The objective of the proposed algorithm was focused to show how an improved channel assignment can enhance the total performance of WMNs. Improving routing protocol was out of our scope for this paper. Several important issues for WMNs environment are considered to evaluate the performance of the channel assignment algorithm. The effects of these issues are discussed in the next section.

## **5.1 Performance Evaluation and Result**

In this section, we show the methodology we used for the simulation and also evaluate the performance of the network. In the result, we have showed the performance of the network in terms of throughput and delay.

## **5.2. Simulation and Methodology**

In our work we used Sinalgo simulation framework that is used for testing and validating network algorithms and abstracts from the underlying layers. Sinalgo focuses on the verification of network algorithms. To justify the performance of the algorithms, we implement both of the algorithms under similar simulation environment. The algorithms are implemented in the same topologies considering a wireless mesh network environment. We have considered a 14 nodes topology. These nodes are considered to have two interfaces and to ensure the performance of the result a random number

of messages are chosen from end to end. The network parameters chosen for evaluating the algorithms and the methodology of the simulation are shown below.

PARAMETER	COMMENT
Number of Nodes	14
Total Number of Channels	30
Number of Channels per Node	3-10
Distribution Model	Random
Mobility	Considered
Connection Type	TCP
Message Type	Data Packet

Table 1. Network Parameters

**Spectrum of interference:** Spectrum of interference for the simulation is implemented as a bi-dimensional array. The values used in the array represent the distances between any two channels. This distance is uniformly distributed. The spectrum is initialized at initialization of simulator and it does not change while the simulator is running. Total thirty channels are considered in the spectrum.

**Channel Assignment:** The nodes are created randomly. This is handled by “Sinalgo” framework. When a node is created the channels are taken randomly from the spectrum of interference. The number of channels for each node is randomly generated as well and each node can have 3 to 10 channels.

**Connectivity:** We implemented our connectivity, based on the commune nodes between channels. If there are commune nodes an edge will be drawn, so the nodes are connected and routing is possible. We have tested the performance of the network when there are any node failure or topology changes. Since the channels are uniformly distributed always two nodes should have commune channels.

**Routing:** Various routing protocols have been already experimented in wireless mesh networks. As per [1], the most common routing protocols such as DSR and AODV selects shortest path to reach the destination. But since the available channels in wireless mesh networks cannot be properly used by the shortest path route, it is not suitable for multi-channel, multi-interference wireless mesh network. It also does not consider the impact of switching channels in a dynamic channel assignment.

Our channel assignment algorithm is not tied to any routing algorithm. Our focus was to improve the throughput by selecting the proper channels for different nodes. The exact set of paths between a communicating node pair is chosen randomly based on the interference and capacity ratio. Our focus was to evaluate the performance of the network for different random routes. For evaluation purposes, we have considered different topologies to measure the performance of the algorithm. Each node maintains a routing table. When a

node (source) wants to send a message to other node (destination), the source node sends a flooding message to the other nodes in neighborhood and waits for an acknowledgement. When the acknowledgement arrives to the source, the destination, the next hop to destination and the number of hops to destination is stored in the routing table.

**Throughput Calculation:** When a node is created, a random number of messages are created and sent to a random number of chosen neighbors. The neighbors are chosen randomly. When the messages arrive to destination, they are stored in a hash table. After all messages arrive the throughput of all messages is summarized and calculated.

**Network Topologies:** The simulation done for analyzing the performance of our algorithm is free from the size, number of hops and change of the topology of the network. We have done the simulation for ten different topologies.

### 5.3. Evaluation Results

The following figures show the result found from the simulation. The solid line (CBCA) in the graph is used to show the result for our algorithm and the dotted line (IBCA) is for the channel assignment algorithm used in [8].

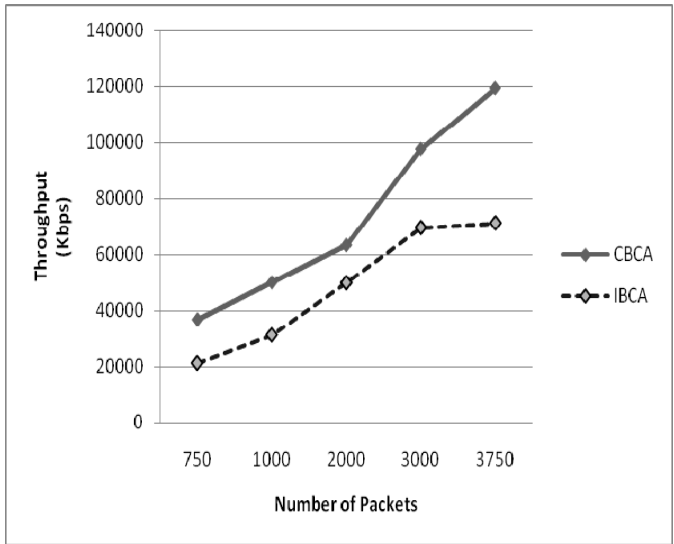
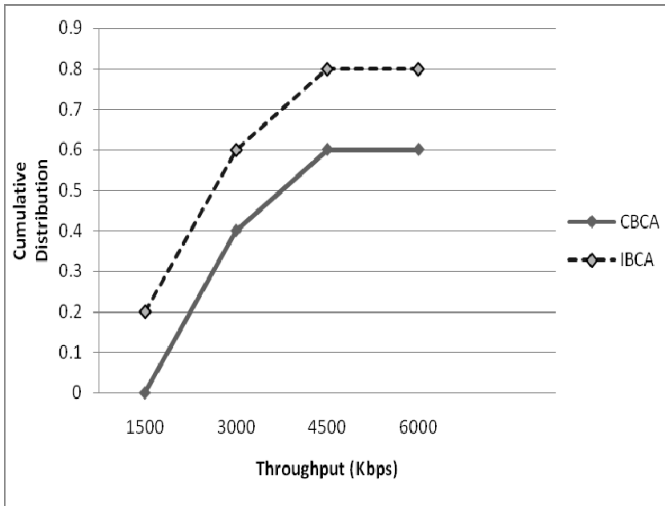


Figure 2. Comparison of throughput

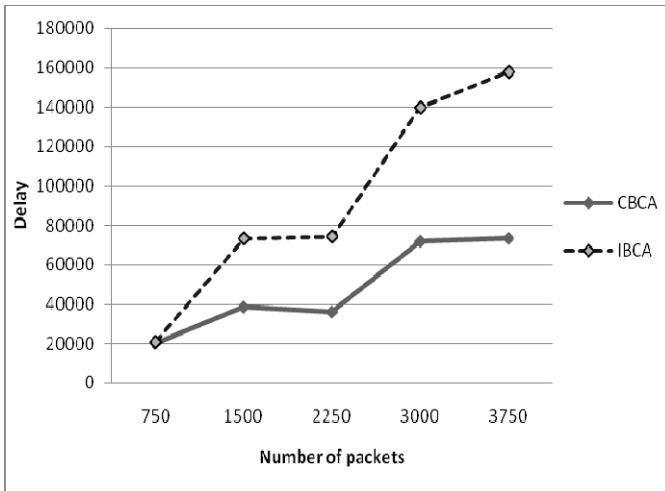
Figure 2 shows the throughput of the two algorithms. The x-axis in the figure represents the throughput in kbps, and the y-axis indicates the number of packets in each run. The performance gain by our mechanism is clear. For instance, at 3750 messages, the throughput of our algorithm reaches 119250.0 kbps, while the best of the [8] achieves no more than 71250.0 kbps.

Figure 3 compares cumulative distribution of the throughput of two algorithms. The results show that our algorithm outperforms the algorithm in [8]. The x-axis in the figure represents the throughput in kbps, and the y-axis indicates the cumulative distribution function of the throughput.



**Figure 3. Comparison of cumulative distribution**

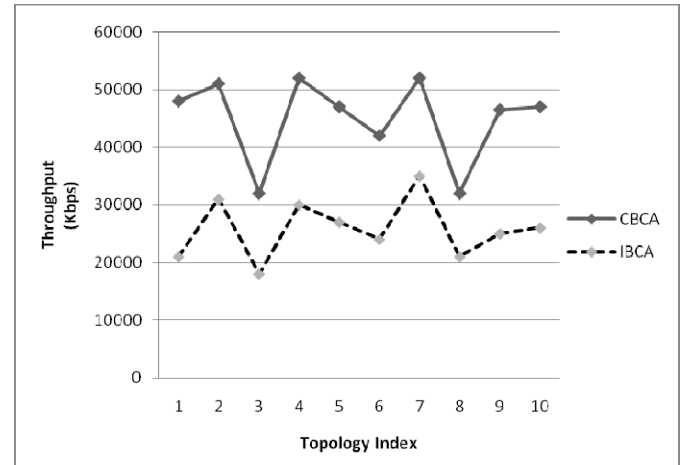
It can be clearly seen that our algorithm has less probability to achieve low throughput. The simulator was run ten times. No run achieved a throughput of less than 1500 Kbps in our scheme whereas the algorithm in [8] shows the cumulative distribution of 0.2 for the same throughput. Moreover, 3 runs of our algorithm achieved throughput around 6000 kbps. Therefore considering channel capacity increases the throughput.



**Figure 4. Comparison of delay**

By reducing the interference and choosing the channel that has the maximum ability to transmit data we compare the delay of the two algorithms in Figure 4. When the load (number of packets in the system) is small, the difference in delay is small because the load is not sufficient to cause congestion in the network. However, by increasing the load (1500, 2250, 3000 and 3750) we see the difference in delay is increasing. Our algorithm reduces the delay because it

chooses the channels that has less interference and has ability to transfer more data, which means less delay.



**Figure 5. Comparison of throughput in different topologies**

The locations of nodes are changed each time the simulator is run, generating different topology for each run. We run the experiments over 10 different topologies. The result in Figure 5 shows that even in this case our algorithm outperforms the one in [8]. The x-axis in the figure represents the topology index, and the y-axis indicates the throughput in kbps.

From the results, it can be clearly seen that the algorithm based on the capacity has better performance for different cases. As a result we can propose that considering capacity along with interference surely improves the performance of the network.

## 6. Conclusion

In this paper, we have proposed an algorithm to assign channels in wireless mesh networks. We have showed by simulation that considering channel capacity in decision function improves the throughput significantly. Other factors can be considered for improving the quality of the network too. These factors include: routing, link delay, switching delay, channel error rate, path loss, etc. The results show that the delay with our algorithm is also less than the algorithm proposed in [8]. It assigns the best channel from interference capacity ratio. The results also show that for different topologies, our algorithm performs better. As a consequence, considering capacity to channel assignment makes a difference. In the future we will focus on issues such as routing and switching delay that can enhance the performance of the network with channel assignment.

## Acknowledgement

The authors would like to thank Dan Ovidiu Fornade for helping us with the simulation results reported in this paper.

## References

- [1] P. Kyasanur and N.N. Vaidya. "Routing and Interface Assignment in Multi-channel Multi Interface Wireless Networks", WCNC, New Orleans, LA, USA, 2005.
- [2] G. Zeng, B. Wang, Y. Ding, L. Xiao and M. Mutka. "Multicast Algorithms for Multi-Channel Wireless Mesh Networks", IEEE ICNP, Beijing, China, 2007.
- [3] M. Shin, S. Lee and Y. Kim. "Distributed Channel Assignment for Multi-radio Wireless Networks", IEEE MASS, Vancouver, Canada, 2006.
- [4] K. N. Ramachandran, E. M. Belding, K. C. Almeroth and M. M. Buddhikot. "Interference Aware Channel Assignment in Multi-Radio Wireless Mesh Networks", IEEE INFOCOM, Barcelona, Catalunya, Spain, 2006.
- [5] A. P. Subramanian, H. Gupta, and S. R. Das, "Minimum-Interference Channel Assignment in Multi-Radio Wireless Mesh Networks", IEEE SECON, San Diego, California, USA, 2007.
- [6] R. Draves, J. Padhye and B. Zill. "Routing in Multi-Radio, Multi-Hop Wireless Mesh Networks", MOBICOM, Philadelphia, PA, USA, 2004.
- [7] A. Raniwala, K. Gopalan and Tzi-cker Chiueh. "Centralized Channel Assignment and Routing Algorithms for Multi-Channel Wireless Mesh Networks", ACM SIGMOBILE Mobile Computing and Communications Review, 2004.
- [8] B-J Ko, V. Misra, J. Padhye and D. Rubenstein, "Distributed Channel Assignment for 802.11 Multi-Radio 802.11 Mesh Networks", WCNC, Hong Kong, 2007.
- [9] A. Das, R. Vijayakumar and S. Roy. "Static channel assignment in multi-radio multi-Channel 802.11 wireless mesh networks: issues, metrics and algorithms", IEEE GLOBECOM, San Francisco, CA, USA, 2006.
- [10] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran and S. Mohanty. "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," Computer Networks vol. 50, 2006, pp. 2127-2159.
- [11] I.F. Akyildiz, X. Wang and W. Wang. "Wireless mesh networks: a survey," Computer Networks Journal 47 (4), 2005, pp. 445-487.
- [12] C.K Toh and B.S Manoj. "Ad hoc mobile wireless networks protocols and systems," Prentice hall, 2002.

## Anycast Routing in Delay Tolerant Networks Using Genetic Algorithms for Route Decision

Ederson Rosa da Silva  
Faculty of Electrical Engineering  
Federal University of Uberlandia  
Uberlandia, MG, Brazil  
ederson@eel.ufu.br

Paulo Roberto Guardieiro  
Faculty of Electrical Engineering  
Federal University of Uberlandia,  
Uberlandia, MG, Brazil  
prguardieiro@ufu.br

### Abstract

*Delay/Disruption Tolerant Networking (DTN) architecture adopts the concept of intermittent networks that may suffer frequent disconnections, with the possibility of never having end-to-end connectivity between the source and destination over a given period of time. Therefore, the design of protocols for those networks becomes a unique challenge. Routing in DTNs is one of the key components and remains open for discussion. In this paper we treat the routing for anycast delivery that is useful in situations where a host, applications or users wish to locate a host which supports a particular service. However, if several servers support the service, any of these servers can be used. Then we implement a routing algorithm using genetic algorithm (GA), because the route and destination decision influence many parameters simultaneously. Our simulation results have shown that the routing using GA performs better than the shortest path algorithm.*

### 1. Introduction

In the last few years, Delay/Disruption Tolerant Networking (DTN) has grown as a research area that is focused on addressing the communication requirements specific to challenged networks. These networks are characterized by frequent disconnection and long duration partition, with the possibility of never having end-to-end connectivity between the source and destination over a given period of time.

RFC 4838 [1] defines an end-to-end message-oriented overlay, called the bundle layer, located between the transport (or other) and application layers of the networks on which it is hosted (see Figure 1).

The protocol's data units passed between entities participating are called bundles (the format of a bundle is described in RFC 5050 [2]).

Routing in DTNs [3,4] is one of the key components and remains open for discussion in the DTN architecture (the architecture for interoperability between and among challenged networks is proposed by Fall [5]). In store-carry-forward routing, a next hop may not be immediately available for the current node to forward the data (bundle). The node will need to buffer the data until the node gets an opportunity to forward the data and must be capable of buffering the data for a considerable duration.

Many of the principles of DTN architecture are reviewed by Fall and Farrel [6], being highlighted design decisions that have persevered through repeated analyses.

In this paper we treat the routing for anycast delivery: a host transmits a datagram to an anycast address and the internetwork is responsible for delivery of the datagram to at least one, and preferably one, of the members in an anycast destination group. There are various applications of anycast in DTN such as resource discovery (people may want to find a determined service without knowing their accurate locations), long distance education (e.g. provides asynchronous Internet access to schools in remote villages with bus and/or motorbike acting as ferries), battle fields (a soldier may want convey the surrounding information to any one of the command centers), and many other applications.

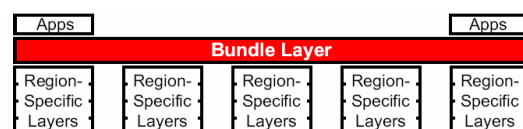


Figure 1. Bundle layer

Due to large transfer delays in DTNs, group membership may change during a message transfer. Under these situations, it is necessary to define the intended receivers of a message. We treat the anycast semantic models that allow users to explicitly specify the intended receivers of a message in section 3.2.

In unicast area in DTNs, many works were found, however, in this case, the destination is determined when the message is generated, while in anycast, the destination can be any one of a node group. Thus, during the routing, both the route for a destination group member and the destination of the anycast message can change dynamically according to the movement of mobile devices and semantic model.

Route and destination decision will influence many parameters simultaneously (like delay, delivery probability). The Genetic Algorithm (GA) has ability to manipulate them. GAs are appropriate when there is no single value to be minimized or maximized, but having multiple objectives.

Moreover, GA can provide adaptive route with optimized delivery probability and delay to deal with directed multi-graph (we talk about DTN graph in section 3.1). In this paper, we applied GA to search routes that optimize the routing objectives.

The rest of the paper is organized as follows: Section 2 describes the related work while Section 3 introduces the anycast in DTNs. We present the routing using GA in the Section 4. Section 5 describes the simulation and performance evaluation and, finally, in Section 6 we conclude the study and present future works.

## 2. Related work

There is a rich literature on anycast routing in the Internet and mobile ad hoc networks (MANET). More specifically, anycast routing using GA in MANETs has been studied quite extensively [7,8], but routing in DTN is different and more challenging due to the frequent partitions and long end-to-end delay.

Abdulla and Simon [9] analyze the multicasting in DTNs, using several multicast routing schemes, incorporating some previous ideas in core-based routing and DTN unicasting, as well as gossiping and epidemic approaches. It is important to note that when the multicast service is used, mobile nodes responsible for assisting in the delivery of messages (when we refer to messages, is the same that bundle), store the messages until it is confirmed that all members of the destination group already have received. In our anycast case, mobile nodes responsible for bringing the message to a member of the anycast group need to

store them until delivery to only one member of the anycast group, which leads to a substantial saving in storage of mobile devices that relayed the messages to a destination group.

Gong et al [10] analyze the anycast semantic for DTN and presented a metric named EMDDA (Expected Multi-Destination Delay for Anycast). The authors assumed that nodes in the network were stationary. The connectivity among the nodes was the mobile devices that act as carrier to deliver messages for the nodes. Also the moving patterns of these mobile carries can be obtained. We used some ideas from [10], but this work is based on estimation, i.e., the intermediate nodes decide whether to store the packet and wait for a better chance, or decide to which nodes to forward. The routing performed by Gong et al [10] is categorized as a stochastic routing protocol. Our routing can be considered deterministic. In a tutorial article [11], is described a simple combinatorial reference model that captures most characteristics of time-varying networks.

In [12], is proposed a new adaptive anycasting approach for mobile nodes where the link availability is opportunistic and the approach is adaptive in nature by communicating with the underlying network for getting the current topological information.

Despite [10] use three anycast semantic models that allow the source explicitly specify the destination of a message through the CM (Current Membership), TIM (Time Interval Membership) and TPM (Temporal Point Membership) models, the network traffic during the selection of routing is not considered. Our anycast routing scheme incorporate both node storage constraint and network traffic dynamics.

## 3. Anycast in DTNs

One of the challenges in designing an anycast routing protocols is to maintain the group membership efficiently. Due to the long delivery delay in DTNs, group membership may already change during the delivery of a message. This leads to the necessity to define new anycast semantic models for DTN.

Moreover, many times the mobile nodes that perform the transport of the messages have limited storage capacity, requiring an efficient anycast routing that does not forward the messages randomly in the networks, avoiding buffer overflows.

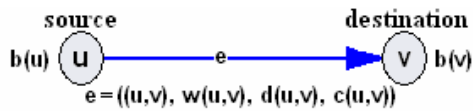
### 3.1. Network model

We consider that the topology of the network may be known ahead of time, categorized as deterministic

case by Zhang [3], i.e., the future movement and connections are completely known.

The DTN graph is a directed multi-graph, in which more than one edge (also called link) may exist between a pair of nodes. Furthermore, the link capacities (storage capacity, propagation delay and departure time) are time-dependent.

An edge in a DTN graph is showed in Figure 2. An edge between node  $u$  and  $v$  means that there exist some mobile devices moving from the initial node  $u$  (source) to the terminal node  $v$  (destination). The storage capacity ( $c(u,v)$ ) on all nodes and the mobile devices is limited. We assumed that every mobile device that moves between the same initial node,  $u$ , and terminal node,  $v$ , has the same moving speed, thus having the same moving delay,  $d(u,v)$ , from the source to the destination. The departure time  $w(u,v)$  of mobile devices on each edge is represented by randomly generated numbers with Poisson distribution.  $b(u)$  and  $b(v)$  are the storage capacity of node  $u$  and  $v$ , respectively.



**Figure 2. An edge in a DTN graph**

We assume nodes in the network are stationary and generate messages. On the other hand, mobile devices do not generate messages themselves. They move from one node to another.

### 3.2. Anycast semantics

In [10] the anycast semantics for DTN are analyzed. It is described that the intended receiver should be clearly defined for a message as group membership changes when nodes join and leave the group. It is showed three anycast semantic models that allow message senders to explicitly specify the intended receivers of a message:

- Current Membership Model (CM): the receiver of the message should be destination group member at the time of message delivery;
- Temporal Interval Membership Model (TIM): a message includes a temporal interval that specifies the period during which the intended receiver must be a member of destination group member;
- Temporal Point Membership Model (TPM): its intended receiver at least should be a member of destination group at some time during membership interval.

In our anycast routing we define the intended receiver when the message is generated. Thus, when the messages arrive to one of these receivers, independently if the receiver is located in the initial group, this will be a valid receptor. This is a particular case of the TIM model, whereas that the temporal interval is the instant of the message generation.

## 4. Routing using genetic algorithms

Genetic algorithms are defined as search algorithms based on the mechanics of natural selection and natural genetics [13]. They combine survival of the fittest among string structures (in our case, routes) with a structured yet randomized information exchange to form a search algorithm. While randomized, genetic algorithms are no random walk. They efficiently exploit historical information to speculate on new search points with expected improved performance.

Before examining the mechanics and the power of a simple genetic algorithm, we must be clearer about our goals when we say we want to optimize a function or a process.

In order to use GA for route decision, in every generation, the GA keeps a population of routes created using randomly selected possible paths (our chromosome) and allows filter routes to combine and produce offspring with new characteristics, which may replace low fitness old routes. Fitness function is a particular type of objective function that quantifies the optimality of a solution.

Diversity in the individual population is highly desirable in order to have a set of alternative routes. The GAs copy routes with some bias toward the best, mate and partially swap (sub) individuals, and mutate occasional possible paths for good measure.

We set the crossover probability to 0.8 and the mutation probability to 0.1. Jong's study [14] suggested the use of a high crossover probability and a low mutation probability (inversely proportional to the population size). We consider values a little higher than the suggested by Jong [14] to increase the diversity in the individual population. The population has 50 individuals. Our GA is controlled by the number of generations.

We applied GA to search the route and consequently the destination that better serve the routing objective (enhance some metric). An example of population representation and crossover is showed in Figure 3. CR1 and CR2 represent the population before the crossover and P1 and P2 the individuals generated. Each number in the each square represent a node, and



each square represent a possible route between the source and destination.

To reduce the time that the GA takes to converge we evaluate and compare the GA when we consider the route obtained by shortest path algorithm as input. Moreover, we test selecting routes with short hops, rejecting the routes with large hops. The results of the two techniques are showed in session 5.

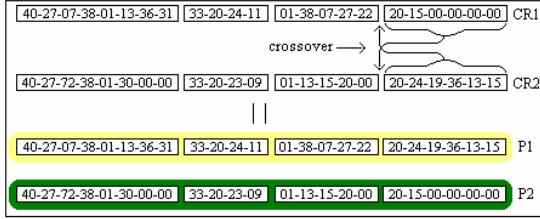


Figure 3. Crossover example

## 5. Simulation and performance evaluation

Routing algorithms in the constructed space-time graph can be developed using shortest path algorithm. For example, Dijkstra's algorithm was adapted to compute the shortest path by Jain, Fall and Patra [4].

We use simulation to compare the performance of the GA-based anycast routing algorithm and the shortest path algorithm modified (the cost function considered was the number of hop) under different environments.

### 5.1. Modeling and simulation

Knowledge oracles are able to answer questions we task of them about the DTN. Accordingly [4], our routing algorithm has a partial knowledge (more practical assumption from an implementation perspective than complete or zero knowledge): our algorithm use information about node and edge queuing, store capacity, moving delay, and departure time of the mobile devices. Those informations are used to compute the route and define the destination.

In our simulation, we employ Waxman Network Topology Generator [15] to generate a random graph of 40 nodes. In the Waxman generator, the nodes follow a Poisson process in the plane and edge density can be controlled. The probability to have an edge between  $u$  and  $v$  is given by

$$P(u,v) = \alpha \cdot e^{-d/(\beta \cdot L)} \quad (1)$$

where  $\alpha > 0$ ,  $\beta \leq 1$ ,  $d$  is the distance from  $u$  to  $v$ , and  $L$  is the maximum distance between any two nodes. We

set  $\lambda$  to 0.6 (intensity of the Poisson process),  $\alpha$  to 0.4 (small values of  $\alpha$  increase the density of short edges relative to longer ones),  $\beta$  to 0.2 (small values of  $\beta$  result in graphs with lower edge densities).

To simulate the behaviors of DTNs, we assume the communication between nodes is carried out by mobile devices. For each edge generated by Waxman generator, we replace it with a mobile device acting as ferry (a set of special mobile nodes that exploits device mobility are utilized to enable or improve communication). We use the statistic toolbox `poissrnd` to generate random numbers from the Poisson distribution to represent the leaving time ( $w(u,v)$ ) of mobile devices on each edge with mean interval time selected randomly from 600 to 6000 seconds. The moving delay ( $d(u,v)$ ) on each edge is a number selected randomly between 60 and 600 seconds, which is multiplied by the distance between the nodes. We assume that the storage capacities ( $c(u,v)$ ) of each mobile device can vary from 300 to 800 messages, and the storage capacities of each node may vary from 600 to 1000 messages (we showed an example of edge between node  $u$  and  $v$  in Figure 2).

We enumerate the nodes in the area 1300 m x 1300 m (see Figure 4). For each anycast session (we consider only anycast traffic in our simulations) we generate a random number between 2 and 5 to represent the number of possible destinations. We randomly pick a node as the anycast source. The first destination member is selected randomly from the possible nodes except the source node. The rest of destination group is the node in sequence if this node is different from source node until completes the number of desired destinations. The messages to send to the destination group for each source can vary between 200 to 500 messages. We generate random numbers from the Poisson distribution between 0.04 and 0.06 to represent the inter-arrival times (messages generated per second). The table 1 has an example of initial traffic for four sessions.

A message is split only at source and different parts (fragments) are routed along same paths. To compare the algorithms we collected statistics about delivery probability (total number of unique anycast bundles received by any anycast group member to the total number of bundles transmitted by the anycast source) and delay (weighted mean of delay, the weighs are the number of bundles relayed).

### 5.2. Simulation experiments

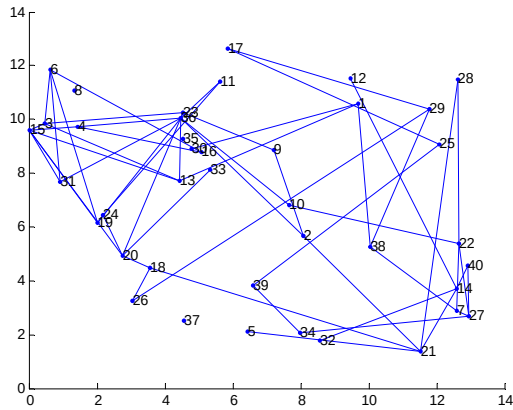
First we simulate a network with only 4 anycast sessions (to facilitate the demonstration of the

algorithms functioning), however the simulation and results with more anycast sessions is presented in the next session. In this case, the simulation is controlled by the number of generations (50). The initial traffic is showed in Table 1. Figure 4 presents the considered network topology.

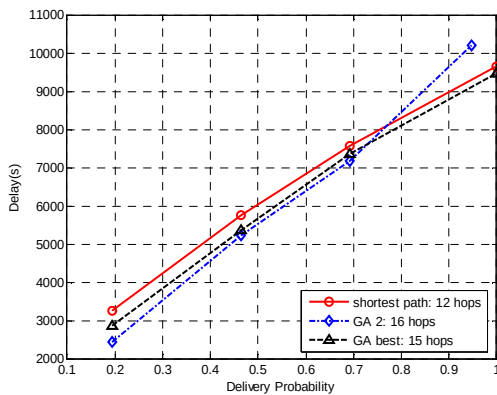
Figure 5 shows that the GA (best individual – GA best) achieve the same delivery probability that the shortest path algorithm (100%), but got a delay a little lesser. The main contribution is this, with GA we can choose a minimum delivery probability (in the example we used 80%) and the algorithm will search the route (optimal or very good) that satisfying the probability and has the lesser delay.

**Table 1. Initial traffic**

	Session 1	Session 2	Session 3	Session 4
Source	40	33	1	20
Destination Group	[30; 31]	[8; 9; 10; 11]	[20; 21; 22]	[15; 16; 17]
Messages to sent	290	400	340	460
Beginning of the session (s)	1400	2700	3900	7000



**Figure 4. Network topology used in example**



**Figure 5. Delay vs. delivery probability**

Analyzing the Tables 2 and 3 we can conclude that the reason for the GA (best route) achieve a delay lesser than the shortest path algorithm is that the route in session 1 and 3, relayed the messages through node 22 (can be a bottleneck). Moreover, the second best route obtained by GA (GA 2), generated a delay larger and a delivery probability lesser than others results showed in the tables, possibly because the route in session 5 performs several hops.

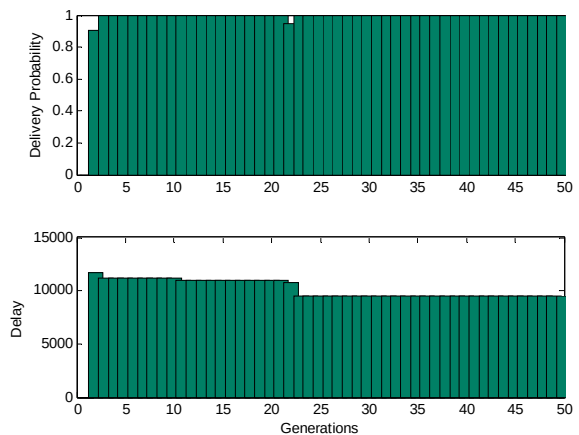
**Table 2. Performance metric comparison**

Results	Total Delay	Route Session 1	Route Session 2	Route Session 3	Route Session 4
Shortest Path	9.651 s	40-27-22-10-36-31	33-20-24-11	1-38-7-22	20-15
GA best route	9.456 s	40-27-7-38-1-13-36-31	33-20-24-11	1-38-7-27-22	20-15
GA 2	10.19 s	40-27-72-38-1-30	33-20-23-9	1-13-15-20	20-24-19-36-13-15

**Table 3. Messages delivered per session**

Results	Deliv. Prob.	Session 1: 290	Session 2: 400	Session 3: 340	Session 4: 460
Shortest Path	1.0000	290	400	340	460
GA best route	1.0000	290	400	340	460
GA 2	0.9463	290	400	340	380

Looking at Figure 5 and Table 2, we can see that if we carried out a crossover between the first (GA best route – CR1) and second (GA 2 – CR2) GA individual, in session 4, we could get a lesser delay (P2) than the results showed in Figure 5 (see Figure 3).



**Figure 6. Best individual evolution**

To see the evolution in every generation of GA, we maintain the best individual and plot in Figure 6. It is

important to note the decrease in the delivery probability in generation 22. This can be easily justified looking at the delay graph, by the fact that our GA, seek the route that has the shortest delay among those that have a probability above the allowed. The best individual is found in generation 23.

### 5.3. Performance evaluation

For performance evaluation, we implemented the proposed algorithms for anycasting under different traffics and number of generations. The GAs take a variable time to produce an optimal or near optimal route depending on the number of sessions and generations used. With 4 anycast sessions the algorithm takes 0.5 seconds per generation. When we have 10 anycast sessions takes 1.5 seconds per generation. If the number of sessions is 20, GA takes 3.2 seconds per generation.

First, we simulated a DTN with 4 anycast sessions. Figure 7 shows that in all cases the GA obtains a delay lesser or equal than the shortest path algorithm and a delivery probability above 80%. The GA converges quickly and has better performance than shortest path because with 4 anycast sessions the diversity of the problem is small.

The difference between Normal GA and SP Input GA is that the second algorithm has the shortest path route as input. We computed the number of average generations to find the best individual. Although this number can change due the randomly of GA, the SP Input GA finds good routes in less generations (18) than Normal GA (27).

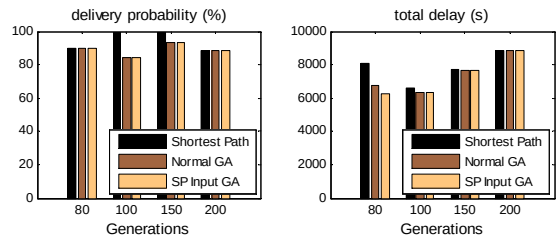


Figure 7. Four sessions result

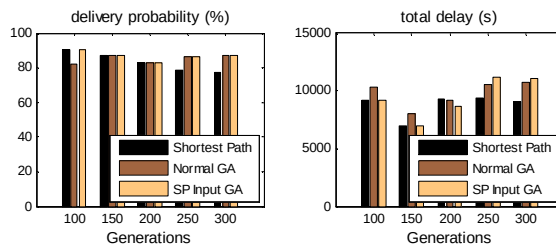


Figure 8. Ten sessions result

Figure 8 shows a simulation with 10 anycast sessions. The algorithm needs a larger number of generations for convergence due to the increase of the complexity. Only with more than 200 generations the GA achieved a better performance than shortest path and again the SP Input GA converge to good routes faster than Normal GA. It is important to note that for 250 and 300 generations the shortest path does not achieve the delivery probability of 80%. However, GA found routes with delivery probability higher than 80%.

Figure 9 shows simulation results with 20 anycast sessions. The GA failed to converge and the shortest path found results better or equal than GA. This can be explaining by the fact that the GA increases the number of generations necessary to converge when the diversity of problem is large. After that, we present a technique to resolve this in the next session.

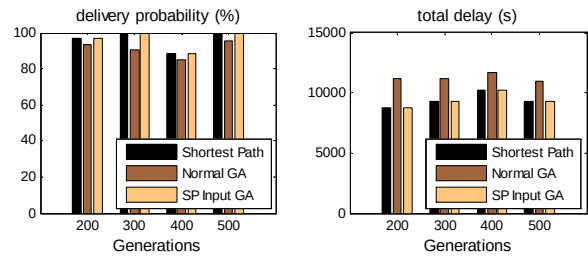


Figure 9. Twenty sessions result

### 5.4. Improving the convergence

To improve the convergence we modified the way that the GA chooses the routes to form the population, because if the GA takes a long time to converge, it could become unfeasible. We viewed that when the numbers of sessions is increased, the algorithm complexity increases too. Then, if the number of session is large, we avoid hold individuals (routes) with a large number of hops, because this routes, normally leads to long delays.

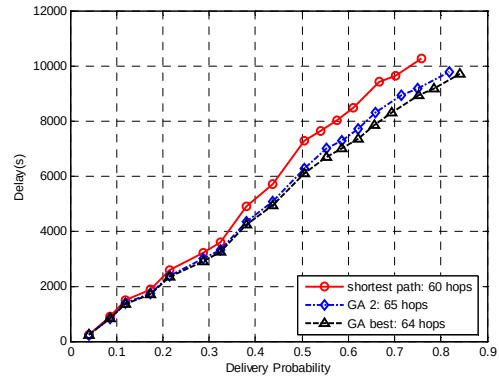


Figure 10. Performance using Normal GA

The improvement can be viewed in Figures 10 and 11. We simulated a DTN with 20 anycast sessions and 300 generations. We can see that the shortest path does not achieve the minimum delivery probability of 80%, although the shortest path has a lesser number of hops than the routes found by the GA. However, the two best individuals found with GA have a delivery probability above the 80%. Again the SP Input GA found the best individual (in generation 59) more quickly than Normal GA (in generation 201).

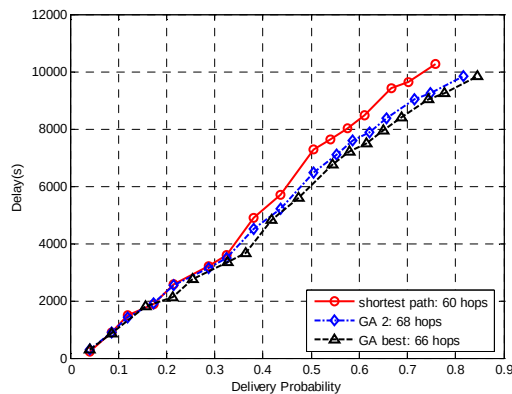


Figure 11. Performance using SP Input GA

## 6. Conclusions and future works

In this paper, we presented a GA-based anycast routing algorithm. Here we are considering a deterministic network with partial knowledge (information) about the network. Moreover, our GA incorporates node storage constraint and permits search a route with the lesser delay between the routes that have a delivery probability above a minimum that can be chosen.

Our simulation results show that GA can reduce the average delay compared to shortest path algorithm. Depending on the traffic, we can employ different strategies to improve the GA performance. For example, when the number of sessions is large, we can efficiently choose the routes to reduce the time to GA convergence.

Otherwise, when we increase a lot the number of sessions and the traffic, the performance of GA is reduced. As future works, we intend to develop different techniques to the GA converge faster and analyze the algorithms performance under different network scenarios.

## References

- [1] V. Cerf et al., "Delay-Tolerant Network Architecture", *IETF RFC 4838*, informational, April 2007. < <http://www.ietf.org/rfc/rfc4838.txt>>
- [2] K. Scott, and S. Burleigh, "Bundle Protocol Specification", *IETF RFC 5050*, experimental, November 2007. < <http://www.ietf.org/rfc/rfc5050.txt>>
- [3] Z. Zhang, "Routing in Intermittently Connected Mobile Ad Hoc Networks and Delay Tolerant Networks: Overview and Challenges", *Communications Surveys & Tutorials*, vol. 8, no 1, pp. 24-37, first quarter 2006.
- [4] S. Jain, K. Fall, and R. Patra, "Routing in a Delay Tolerant Network", *Special Interest Group on Data Communications - SIGCOMM*, August/September 2004.
- [5] K. Fall, "A Delay-tolerant Network Architecture for Challenged Internets", in *Proc. ACM SIGCOMM'03*, pp. 27-34, New York, USA, 2003.
- [6] K. Fall, and S. Farrel, "DTN: An Architectural Retrospective", *IEEE Journal on selected areas in communications*, vol. 26, no 5, pp. 828-836, June 2008.
- [7] K. Kotecha, and S. Popat, "MultiObjective Genetic Algorithm based Adaptive QoS Routing in MANET", *IEEE Congress on Evolutionary Computation (CEC)*, pp. 1423-1428, September 2007.
- [8] M. Al-Ghazal, A. El-Sayed, and H. Kelash, "Routing Optimization using Genetic Algorithm in Ad Hoc Networks", *IEEE International Symposium on Signal Processing and Information Technology*, pp. 497-503, December 2007.
- [9] M. Abdulla, and R. Simon, "A Simulation Analysis of Multicasting in Delay Tolerant Networks", *Proceedings of the Winter Simulation Conference - WSC*, pp. 2234-2241, 2006.
- [10] Y. Gong et al., "Anycast Routing in Delay Tolerant Networks", *Global Telecommunications Conference (GLOBECOM)*, November 2006.
- [11] A. Ferreira, "Building a Reference Combinatorial Model for MANETs", *IEEE Network*, vol. 18, no. 5, pp. 24-29, September-October 2004.
- [12] Fazl-e-Hadi, N. Shah, A. H. Syed, and M. Yasin, "Adaptive Anycast: A New Anycast Protocol for Performance Improvement in Delay Tolerant Networks", *IEEE International Conference on Integration Technology - ICIT*, pp. 185-189, March 2007.
- [13] David E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Massachusetts, Addison-Wesley, 1989.
- [14] K. A. De Jong, "An analysis of the behavior of a class of genetic adaptive systems", doctoral dissertation, University of Michigan, 1975.
- [15] I. Kaj, and R. Gaigalas, "Stochastic Simulation: Waxman random network topology generator", <<http://www.math.uu.se/research/telecom/software/>> accessed in June/2008.

## LCRACO – A New Load and Congestion Controlled Routing Based on Ant Colony Optimization

Rituparna Chaki  
West Bengal University of Technology  
West Bengal, India  
rituchaki@gmail.com

Ditipriya Sinha  
Calcutta Institute of Engineering and Management  
West Bengal, India  
ditipriya\_sinha@yahoo.co.in

### Abstract

*Designing efficient, cost-effective routing solutions for the infrastructure-less MANETs remains a daunting task till date. A proven optimization technique, the Ant colony optimization has been used in this paper to find the optimum in terms of the shortest route in a MANET. Special attention has been given to the load balancing and congestion control in the network. The simulated performance graphs show an improved congestion control for the proposed routing protocol.*

### 1. Introduction

The mobile ad hoc network (MANET) is an autonomous system of mobile routers (and associated hosts) connected by wireless links – the union of which form an arbitrary graph. The routers are free to move randomly and organize themselves arbitrarily; thus, the network topology may change rapidly and unpredictably. Such a network may operate in a standalone fashion, or may be connected to the larger Internet. Routing in adhoc network is very difficult and challenging. ACO algorithms have been employed to solve numerous problems in ad hoc networks. The swarm intelligent mechanisms present in natural ant colonies are used by a powerful class of optimizing algorithms that was derived from the behavior of natural ant colonies. Dorigo *et al* first combined ACO with source routing to produce AntNet, Schoonderwoerd *et al* combined ACO with distance vector routing to produce Ant-Based Control and various other authors produced ACO-based routing algorithms with the focus on different factors such as Quality-of-Service and clustering of large networks. Ants consistently achieve remarkable feats of cooperation, coordination and construction [2]. Of particular interest to this work is how ants can find shortest routes between their nest and food sources. To study this phenomenon, Deneubourg

*et. al.*, set up the binary bridge experiment. In the binary bridge experiment a food source and a nest of ants are separated by a bridge with two branches, one being longer than the other. Ants deposit a substance called pheromone on the ground as they walk. Other ants can smell this pheromone, and their path-taking decisions are influenced by the pheromone. When subsequent ants sense this trail of pheromone, they will follow the trail with a probability that is proportional to the concentration of pheromone – the higher the concentration, the larger the probability that subsequent ants will follow this trail. The pheromone evaporates at an exponential rate, which means that the strength of pheromone encountered by another ant is a function of the original pheromone strength and the time since the trail was laid.

### 2. The State of the Art Scenario

In flat (or non clustered) networks the nodes are not clustered into any form of hierarchical structure, all nodes are homogeneous in terms of network responsibility and packet handling. The first protocols to use ACO used basic source routing or distance vector algorithms where the traditional routing tables were replaced with pheromone tables, while more advanced protocols also incorporated ad hoc characteristics such as mobility and congestion into the pheromone mechanism.

SACO [2] aimed to solve the problem of finding shortest path between two nodes on a MANET. A number of ants are placed on the source node of the network. Each ant incrementally constructs a path to the destination node during each iteration of the algorithm. The next hop node is chosen on the basis of highest pheromone concentration. At each node ant uses the following probability to select the next node. The pheromone deposition in an ACO based system can be influenced by heuristic information. However, SACO

does not take into consideration the effects of load and congestion of the network.

Antnet [3] describes the application of ACO to dynamic routing in packet-switched networks. This is the first application where ACO is used for routing with a source routing mechanism, i.e. the entire route of the packet is listed in the header of the packet. Each node  $s$  in the network periodically generates an artificial ant (called a forward ant) with a randomly selected destination node  $d$  to observe the trip time from source node  $s$  to destination node  $d$ . The forward ant uses the current routing tables to find a path to node  $d$ , and records the route taken. When the forward time, taking the same path as the forward ant in the opposite direction. Each node has its own routing table [4]. Although AntNet forwards data packets using distance-vector principles, route discovery and maintenance are done using source routing. AntNet consequently suffers similar bandwidth and hop-count drawbacks encountered in Dynamic Source Routing (DSR). This translates to very limited scalability, and the results could be checked on networks of only 36 nodes.

ARA [7] is an ACO-based routing scheme using distance vector routing. Route discovery is done by broadcasting Forward Ants (FANT). The FANT sets up a pheromone trail pointing back to the source node as it is broadcast through the network. When a route is found to the destination node, a Backward Ant (BANT) is created. The BANT follows the pheromone trail created by the FANT back to the source node, and sets up a pheromone trail pointing to the destination node. The amount of pheromone deposited by FANT and BANT packets is a function of the length of the route associated with the pheromone. Once the route is established, packet transmission starts. The data packets also reinforce the pheromone trails pointing to the source and destination nodes. In the case of a route failure, the intermediate node attempts to send the packet over an alternative link. If no alternative link is found, the packet is returned to the previous up-stream node, which also tries to send the packet over an alternative link. If the packet is returned in this way to the source node, a new route discovery sequence is initiated using backtracking [7]. ARA implements the packet sequencing mechanism to prevent packet loops. However, in AODV each RREQ packet is also given a RREQ-sequence number to ensure that destinations only respond to a RREQ once. Contradictory to AODV, however, in ARA, destinations respond to multiple FANT packets received from the same node, thus supporting multipath routing [1].

MACO [9] concentrates on load balancing of the network. It uses more than one colony of ants to search for optimal path, and each colony of ants deposits a different type of pheromone of a different color. Although ants in each colony respond to pheromone from its own colony, MACO is augmented with a repulsion mechanism that prevents ants from different colonies to choose same optimal path. The degrees of attraction and repulsion determined by two weighting parameters, combined into a probability function which is used into VWP network. Different sets of mobile agents corresponding to a pheromone table of each colony are used. By adopting the MACO approach, it may be possible to reduce the likelihood that all mobile agents Establish connection using only the optimal path [9]. The advantage of using MACO in circuit switched routing is that it is more likely to establish connections through multiple paths. It helps to balance the load but does not increase routing overhead. MACO can establish connection through multiple paths while traditional ACO is most likely to establish connection through the optimal path.

### 3. The Proposed LCRACO Algorithm

A new routing methodology named Load and Congestion Controlled Routing based on Ant Colony Optimization Algorithm (LCRACO) has been proposed here, giving special attention to load balancing and congestion control of the network. This algorithm uses more than one colony of ants to search for the optimal path. The next hop is chosen based on the highest concentration of pheromone. Connection is established on a multipath basis to balance the load in the network.

#### 3.1 Assumption

We make a couple of important assumptions before we propose the new routing algorithm.

1. There are two colony of ants. Symbol 0 indicates red colony and 1 indicates blue colony. Initially there are two array. One is blue pheromone array and another one is red pheromone array. Two arrays give the information of amount of pheromone between two nodes, which are initialized as follows:

For red pheromone array: [000...0]

For blue pheromone array: [111...1]

2. All the possible routes are needed to reach the destination node are already known by the source node and all the details of next hop node in each route are stored in routing table.

In this algorithm at particular time a source node wants to send m no of messages to a destination node.

### 3.2 Definitions

- Cost Table – Consists of the cost of reaching nodes (next hop). This consist of a set of tuples:  $\{S_{ip}, N_{ip}, C_{sn}\}$ , where,  
 $S_{ip}$  - Source node address,  
 $N_{ip}$  - Neighbor node address  
 $C_{sn}$  - Cost on the path is a derived entity and is a function of time interval of sending “HELLO” message by node  $s$  and receipt of an “ACK” message from node  $n$ .
- The HELLO message:  $S_{ip} D_{ip} N_{ip} S_{t1}$ , where  $D_{ip}$  is the Destination node and  $S_{t1}$  is the start time 1
- The ACK message :  $S_{ip} N_{ip} D_{ip} S_{t1} S_{t2}$ , where  $S_{t2}$  = Start time 2
- Routing Table: It consists of information about the next hop node on the path to particular destination node. Thus the routing table vector is— $S_{ip} N_{ip} D_{ip}$
- Red pheromone array gives the information about the amount of red pheromone deposited by red ants between source and all possible next hop node to reach the destination node. Blue pheromone array gives the information about the amount of blue pheromone deposited by blue ants between source and all possible next hop node to reach the destination node. The structure of the red pheromone array is  $[\delta r_1 \delta r_2 \delta r_3 \dots \delta r_n]$ . The structure of the blue pheromone array is  $[\delta b_1 \delta b_2 \delta b_3 \dots \delta b_n]$ .  
Here,  $\delta r_i$ ,  $\delta b_i$  represent the pheromone value deposited on the way from the source node to  $i$ th node from the next hop list of the routing table. The pheromone value is modified using the eqn. (3), (4) and (5). For evaporation of pheromone, equation (7), (8) is used.

Probability calculation for ant of red colony:

$$P_{kn}^d(t) = \{\delta r_{kn}(t) / \sum_{x \in N_k(t)} \delta r_{kx}(t)\} \quad \dots (1),$$

where  $t$  = Time index

$P_{kn}^d(t)$  = The probability that an ant at node  $k$  at time  $t$  traveling to node  $d$  will select node  $n, 1 \leq p \leq 0$

$\delta r_{kn}(t)$  = The red pheromone level at node  $k$  at time  $t$  associated with using node  $n$  as a next hop node

$N_k(t)$  = The set of feasible node connected to node  $k$  at time  $t$ .

Probability calculation for ant of blue colony:

$$P_{kn}^d(t) = \{\delta b_{kn}(t) / \sum_{x \in N_k(t)} \delta b_{kx}(t)\} \quad \dots (2),$$

where  $t$  = Time index

$P_{kn}^d(t)$  = The probability that an ant at node  $k$  at time  $t$  traveling to node  $d$  will select node  $n, 1 \leq p \leq 0$

$\delta b_{kn}(t)$  = The blue pheromone level at node  $k$  at time  $t$  associated with using node  $n$  as a next hop node.

$N_k(t)$  = The set of feasible node connected to node  $k$  at time  $t$ . Each ant deposits certain amount of pheromone. The amount of pheromone deposited is proportional to the quality of the route selected by the ant.

Updation of red and blue pheromone matrix:

$$\delta r_{kn}(t+1) = \delta r_{kn}(t) + q_r \quad \dots (3)$$

$$\delta b_{kn}(t+1) = \delta b_{kn}(t) + q_b \quad \dots (4)$$

where  $q_{ri}$  is the inverse of the cost function between source and destination node,

$$\text{i.e., } q_r = 1/c_{kn} \quad \dots (5)$$

$C_{kn}$  is the cost function, which gives the average time required to send message between source node  $k$  to next hop node  $n$ .

Evaporation of red and blue pheromone:

$$\delta r_{kn}(t+1) = \delta r_{kn}(t) - \mu * c_{kn} \quad \dots (7)$$

$$\delta b_{kn}(t+1) = \delta b_{kn}(t) - \mu * c_{kn} \quad \dots (8)$$

$\mu$  = evaporation constant. value of it is 0.3

### 3.3. Proposed Algorithm

Step1: Each node in the network sends “HELLO” message to all possible next hop node after certain time interval to reach the destination node and each next hop node sends “ACK” message. According to the total time required to receive “ACK” message, source node calculates the cost on the path of all possible next hop node and stores it into the cost table.

The source node sends “HELLO” message after fixed time interval to check cost table.

- Let start time  $t_1$  (when “HELLO” message is sent) and the neighbor node receives the message at time  $t_2$ . So total time required to send “HELLO” message is  $T_1 = t_2 - t_1$
- If start time  $t_2$  when “ACK” message is sent by the neighbor node and source node receives the message at time  $t_3$ . So total time required to send “ACK” message is  $T_2 = t_3 - t_2$
- Total time  $T = (T_1 + T_2) / 2$
- Update  $C_{sn}$  with  $T_x$   $k$ , where  $k$  is the cost per unit time.

Step2: Step 1 is descriptive-, whereas step 2 is

For each (Sip)

a) {Construction of ant}

for each ant  $k=1$  to  $m$

Generate an alternate sequence of 0 and 1 and stores it in  $v$ .

begin {Probability calculation }

b) If  $v$  is 0 then

i) Each node calculates the concentration of red pheromone for all it's next hop node to reach particular destination node using the equation (1).

ii) The ant chooses next hop node in which concentration of red pheromone is highest.

iii) Red pheromone array is updated.

else

i) Each node calculates concentration of blue pheromone for all it's next hop node to reach particular destination node using equation (2).

ii) The ant chooses next hop node in which concentration of blue pheromone is highest.

iii) Update blue pheromone array.

End {Probability calculation }

Step 3: Evaporation of red pheromone is done by the equation (7) and the same for blue is done by using equation (8).

Step 4: The process continues till all the messages are reached to their destination node.

### 3.4. Illustration of Algorithm

Let's consider a network of four nodes. Suppose node A has to send four messages to node D.

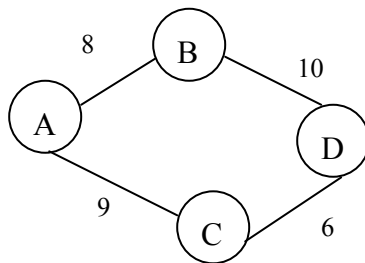


Figure 1: A 4-node network

Thus, A broadcasts a HELLO message. All the neighbours of A, i.e., B and C accepts this message as they have D in their route table. B and C sends ACK message to the source node A. A derives the cost of the

routes from the time interval between broadcasting the HELLO message and receipt of ACK message. The routing table of A is :

	C	B
A	R=4	R=5

Then it finds probability of concentration of red pheromone in which path is highest. It is found that it is highest on the path AB. So A sends it's first packet on this path. Then A updates it's red pheromone table on this path. The red ant deposits red pheromone, the amount of which is equal to  $1/(\text{path length})$ . Now the updated red pheromone table becomes—

	C	B
A	R=4	$R=5+1/8=5.125$

Red Pheromone table

If  $v$  is 1 then it follows blue pheromone table and chooses the path on which concentration of blue pheromone is highest. The blue pheromone table is given below:

	C	B
A	B=6	B=4

Blue Pheromone table

At that moment ant chooses the path AC because concentration of blue pheromone is highest on this path and updates the blue pheromone table in same way as the red pheromone table.

	C	B
A	$B=6+1/9=6.111$	B=4

Blue Pheromone table



After certain time interval (after sending 3 messages) pheromone is evaporated. In this way the algorithm goes on and it controls congestion as well as load in the network.

#### 4. Performance analysis of LCRACO

The theoretical analysis of the algorithm helps us to notice the following points:

**Load Balancing:** Here more than one colony of ants are used to search for optimal path, and each colony of ants deposits a different type of pheromone by a different color and ants in each colony respond to pheromone from it's own colony. Thus the algorithm balances the load in the network load.

**Congestion Control:** Here one colony of ant chooses the path in which deposition of pheromone of that colony is highest and another colony of ant chooses the path in which deposition of pheromone of that colony is highest. The congestion in the network can thus be controlled to a large extent.

##### 4.1 The simulation results

###### 4.1.2 Variation of maximum probability value:

When a particular node wants to send set of packets, it chooses path alternately according to the concentration of red pheromone matrix and blue pheromone matrix. After selecting a particular path, the concentration of pheromone of that particular path is increased.

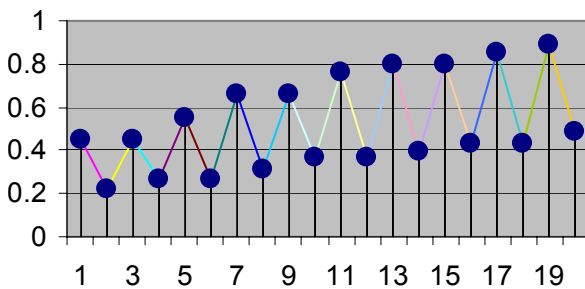


Figure 2: Pheromone level vs. Number of messages

The graph varies as maximum probability value of pheromone is changing. The variation of maximum probability value confirms that alternately the source node chooses the highest concentration of red and blue pheromone for sending each packet.

**4.1.2. Pheromone level variance:** The x axis of the graph in figure 4 shows the number of packets sent by a particular node and Y axis gives the performance of red

ant and blue ant.

The red curve indicates red pheromone level and blue curve shows the blue pheromone level. From this graph it can determine that for each message sent, there is a variance in pheromone level. Thus, the source node does not follow same path for sending all messages. In this way load as well as congestion in the network is controlled.

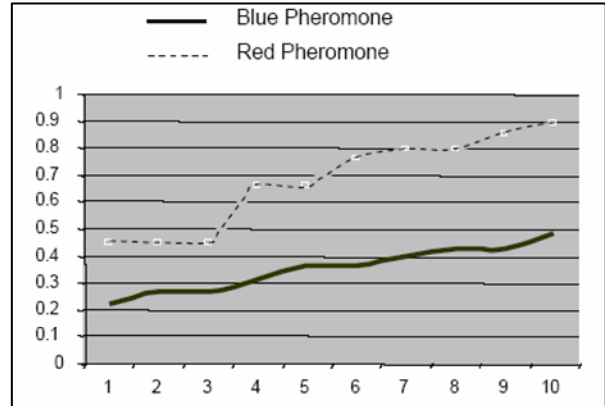


Figure 3: Pheromone Level for Two Ant colonies

Next, we studied the variance of packet delivery rate against the number of nodes. The X axis in figure 4 shows the number of packets sent by source node and the Y axis shows the delivery rate of each packet.

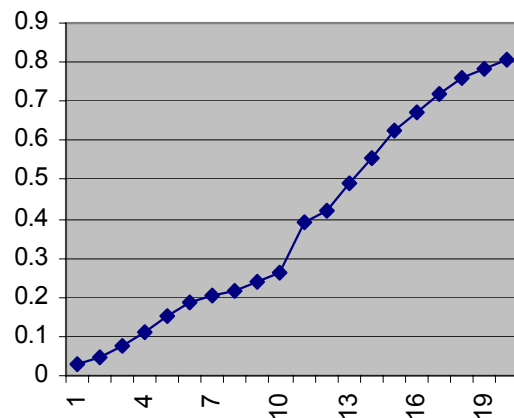


Figure 4: Delivery rate vs. Number of nodes

#### 5. Conclusion

A brief review of the well known methodologies in

this area shows that very little consideration had been given towards load balancing and congestion control of the network. This paper proposes a new routing methodology LCRACO based on the theory of ant colony optimization. Theoretical evaluation shows that this algorithm can take care of load balancing and congestion control. of the network. The algorithm will be useful in cases where the bandwidth of the packets sent by a node exceeds the bandwidth of the channel.

## 6. References

- [1] Siva Kumar. D and Bhuvaneshwaran.R.S, "Proposalon Multi agent Ants based Routing Algorithm for Mobile Ad-hoc Networks", in IJCSNS International Journal of Computer Science and Network Security, Vol-7, No. 6, June,2007.
- [2] M.Dorigo, "Ant Colony Optimization", University of Pretoria Etd, du Plessis J, 2006
- [3] V.Laxmi, Lavina Jain, M.S.Gaur, "Ant Colony Optimization based Routing on NS-2" in the preceedings of International Conference On Wireless Communication and Sensor Networks,WCSN 2006
- [4] Payman Arabshahi, Andrew Gray, Ioannis Kassabalidis, Arindam Das, "Adaptive Routing in Wireless Communication Network using Swarm Intelligence", Jet Propulsion Laboratory and University of Washington.,2004
- [5] Kwang Mong Sim and Weng Hong Sun, Member IEEE, "Ant Colony Optimization or Routing and Load-Balancing: Survey and New Direction, part-A: System and Humans, Vol-33,No. 5 Sep,2003.
- [6] Sundaram Rajagopalam, Chien-Chung Shen, "A Routing Suite for Mobile Ad hoc Networks using Swarm Intelligence", Department of Computer and Information Sciences, University of Delaware, Newark,2003
- [7] M. Gunes, U. Sorges, I. Bouazizi, "ARA- the ant colony based routing algorithm for MANET",ICPP Proc of the 2002 Workshop on Adhoc Networks.
- [8] Griselda Navarro Varela, Mark C.Sinclair, "Ant Colony optimization for virtual wave length pathrouting and wave length allocation", *Proc. Congress Evolutionary Computation*, pp1811-1815, 1999
- [9]. G. Di Caro and M. Dorigo, "AntNet distributed stigmergetic control for communication networks", *Journal on Artificial Intelligence Research*, 1998.

## Exploring Wireless Device Driver Vulnerabilities

Victor Agapov, MCSE  
Graduate Student

Syed (Shawon) M. Rahman, Ph. D.  
Assistant Professor

Department of Computer Science and Software Engineering  
University of Wisconsin – Platteville, 1 University Plaza, Platteville, WI 53818, USA  
Phone: (608) 342 1625, Fax: (608) 342-1965, Email: {agapov, rahmans}@uwplatt.edu

**Abstract** – IEEE 802.11 set of standards grew from a highly obscured and unsecured technology to a robust and reliable solution for many of us. Broad acceptance and relatively simple implementation also made 802.11 a perfect target for possible exploitation. Recently discovered wireless driver vulnerabilities exposed a new critical threat that legitimate users community is facing. What makes this vulnerability so dangerous to many of us is the fact that the actual exploitation cannot be prevented by commonly used tools such as firewalls, antivirus, and intrusion detection system's programs. Access to the victim's machine is gained through a flawed third-party wireless driver, giving full kernel level access rights to the attacker. While kernel-level exploitation domain covers everything running beneath the operating systems, this paper describes wireless device drivers vulnerability and proposes methods for vulnerability mitigations. This paper would assist IT professionals in evaluating, creating, and assessing security requirements associated with wireless networks connectivity.

*Index Terms* – 802.11, wireless driver, security, vulnerability

### 1.0 Introduction and a Brief History

Wireless technologies had their own share of success since the introduction of the first 802.11 legacy standard almost 10 years ago. Previously named as a wired networks killer technology, 802.11 had a very rough road in last a few years. Not many new technologies had so many admirers that turned back so fast because of the disastrous introduction and security issues of its earlier releases. Wireless networking in comparison to wired networks requires an additional set of requirements for companies and organizations on its implementation, maintenance and more importantly, on securing the boundaries of valuable resources and services.

Unsecured medium of communications and ubiquitousness of wireless-enabled clients provides great deal of opportunities for possible exploitations. The fact that the communication is going through the open airwaves accessible to anyone including unauthorized users poses the greatest risks to the networking processes. These risks are usually associated with loss of confidentiality and integrity and the threat of Denial of Service (DoS) attacks. All of the

exploits and vulnerabilities existing in regular land wired networks have presence in wireless settings. What makes the 802.11 technology special is the huge number of small portable devices holding sensitive information that can be compromised. Proper management and assessment of these resources requires an extra set of knowledge and effort from IT professionals in configuring and ongoing monitoring networks for anything unusual and suspicious as well as performing regular administrative duties such as backups, user management, and security assessments. The level of expertise of the tech support personal is very crucial as many of the commercially available products used in WLANs don't provide adequate protection, have weak default settings or poor software implementation.

### 2.0 Overview of Wireless Technology

Wireless technology is based on the usage of single carrier direct-sequence spread spectrum radio technology. When introduced in 1985, it was operating in an unlicensed spread spectrum, specifically designated by FCC for scientific, medical and other types of devices, such as some types of cordless phones and microwaves [1]. After the standardization of the wireless technologies under the umbrella of 802.11 standards, wireless technology was assigned to a limited spectrum in the 2450 MHz to 5800 MHz band range in United States. Use of other ranges of bands differs from country to country. In the pre and earlier implementations of the 802.11 standard, security was deliberately weak due to the export regulations of US government [2]. This was addressed in 2004 with the release of the 802.11i [3]. WEP and WAP because of their weak protection were superseded by WPA2 and use of AES block ciphers [4].

Below is a chronologically ordered list of modifications of 802.11 standards [5]:

- ◆ Pre 802.11 - 802.3 - released in 1980, and based on Xerox's PARC work in 1970s, it sets the standards for wired Ethernet networks. Defines the physical layer, and the media access control (MAC) sub layer of the data link layer;
- ◆ 802.11-1997 (802.11 legacy) - designed for 1 Mbps to 2 Mbps wireless transmissions;
- ◆ 802.11a - high-speed WLAN standard for the 5 GHz band and, supports 54 Mbps;

- ◆ 802.11b - operates in the 2.4 - 2.48 GHz band and supports 11 Mbps;
- ◆ 802.11g - operates in the 2.4 GHz waveband;
- ◆ 802.11n - significantly improves network throughput by adding multiple-input multiple-output (MIMO);
- ◆ 802.11y – adds operation in 3650 to 3700 MHz band by utilizing high power wireless technology equipment

Table 1 shows the main characteristics of 802.11 standards [5]:

Table 1: Characteristics of 802.11 standards

Protocol	Release Date	Op. Freq (GHz)	Throughput (Mbit/s)	Data Rate (Max) (Mbit/s)	Range (Indoor) in meter	Range (Outdoor) in meter
Legacy	1997	2.4	0.9	2	~20	~100
802.11a	1999	5	23	54	~35	~120
802.11b	1999	2.4	4.3	11	~38	~140
802.11g	2003	2.4	19	54	~38	~140
802.11n	2009 (est.)	2.4 - 5	74	24	~70	~250

## 2.1 Architecture

802.11 standards stipulate operation between wireless-enabled clients in two basic ways: an ad hoc or peer-to-peer communication, where clients use direct links to connect to each other, and infrastructure mode, where stationary Access Points (AP) hardware devices are used to provide gateway access services for wireless clients to access wired networks such as intranets or private LANs. The main structure components that can be found in any type of wireless networks are listed below [6]:

- ◆ Client - laptop or other mobile device with wireless enabled connectivity;
- ◆ Fixed wireless Access Point (AP) – hardware device that connects wireless communications, provides access to other networks (wired);
- ◆ Independent Basic Service Set (IBSS) – no APs are used, direct client to client networking communications
- ◆ Basic Service Set (BSS) - at least one AP is present and all of the communication is performed through that device.
- ◆ Extended Service Set (ESS) – multiple APs are connected to the same network segment through the router

Extended service set mode can be logically represented by collections of more than one BSS configured to use same

network, where clients can hop from one AP to another without losing connection to the network recourses. Configuration details of wireless clients and network depend on which mode is used.

## 2.2 Protocol

802.11 can be best described by using a seven layer Open Systems Interconnection (OSI) Basic Reference Model [7]. It operates on the lower two layers of this model and utilizes a Media Access Control (MAC) sub layer of the data-link layer and the physical layer. Logical Link Control (LLC) is another data link layer component of the OSI that is being used by 802.11 standards. Figures 1 and 2 represent layers responsible for wireless communications. The Media Access Control (MAC) layer known as CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance) is the first of the OSI model layers that is being utilized by 802.11, where peer-to-peer connectivity does not require a server.

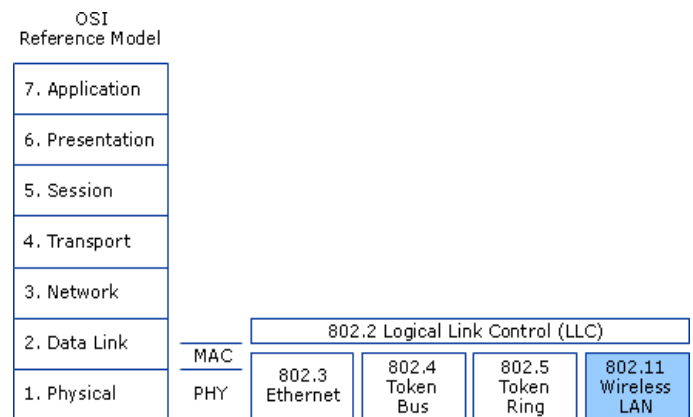


Figure 1 – OSI Basic Reference Model [6]

The 802.11 MAC frame consists of a MAC header, the frame body, and a Frame Check Sequence (FCS) as represented in an image below. The numbers in Figure 2 represent the number of bytes for each field.

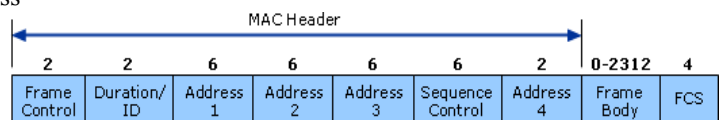


Figure 2 – MAC frame [6]

The Frame Control Field provides much of the information related to the wireless driver vulnerabilities. 802.11 MAC FCF subfield's main purpose is to tell the system how to process the frames in wireless communications. Figure 3 shows the responsibilities of each Frame Control Field's subfield:

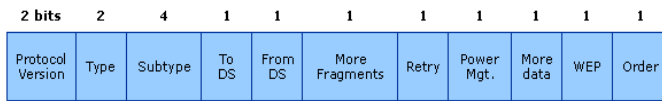


Figure 3 – Frame Control Field (FCF) [6]

Details and functionality of the FCF:

- ◆ Protocol Version provides protocol’s version type and used by client to determine if this version is supported;
- ◆ Type and Subtype field checks for the frame type (management, control, or data);
- ◆ To DS and From DS shows if the frame is coming or leaving from the distributed system;
- ◆ More Fragments indicates if more fragments of the frame will be received;
- ◆ Retry refers to the retransmitted frames;
- ◆ Power Management shows the status of the sending client (active vs. passive);
- ◆ More Data indicates that more data is coming;
- ◆ WEP shows if encryption and authentication are used in the frame;
- ◆ Order maintains the order of received frames.

Other parts of the MAC header contain information such as the address of source and destination, sequential number of each frame, actual data itself and the Frame Check Sequence algorithm for error prevention.

All of the existing flavors of 802.11 standards reside at the PHY sub layer. The physical layer of OSI model define a series of encoding and transmission schemes for wireless communications, such as the Frequency Hopping Spread Spectrum, Direct Sequence Spread Spectrum, and Orthogonal Frequency Division Multiplexing (OFDM) transmission schemes [8].

## 2.3 Wireless Communications

Wireless communication is established through the use of network frames defined by 802.11 standards. As described earlier, each frame’s fields contain information and describe the functions of that frame and thus can help classify frames into three categories: management, control, and data. Generally, in basic and/or extended service set mode, an access point device is broadcasting beacon management frames intended for any wireless clients. The clients are supposed to intercepts these frames and identify the correct AP and then join one of the available networks based on its setting. This process is a three stage communication where after identifying an AP, the client and the AP perform a mutual authentication by exchanging several management frames. At this point, the client is

already known to the AP, but does not yet have access to the network resources. Its status is “Authenticated and Unassociated.” The next stage of this process is “Authenticated and Associated,” where access is granted to the client. Most of the security risks arise before or during these processes and will be discussed in the next section.

## 3.0 Current Security Threats

In general, risks associated with wireless networks are the same as in the traditional wired networks environment, plus the new risks introduced by weaknesses in wireless protocols [12].

Jamming, deciphering and cryptanalysis, sniffing, injecting, masquerading, rogue access point insertion are the most common exploitation techniques found in wireless enabled environments today. Social engineering although not a unique wireless security problem, gets wider negative attention as well.

Great example of unacceptable negligence and inadequate qualification of IT people was biggest ever theft of personal data, T.J. MAX's failure to secure their wireless LAN [9] costing the company over a billion of US dollars. TJX was using WEP security algorithm on its network, nearly four years after it was proven to be insecure. Apparently the company was slow in implementing new standards. What's alarming that most of the other retail chains were you using WEP at that time and as a matter of fact some of them are still using it (Sears, Best Buy, Macys, etc.) [10]. It was recently reported that WEP 104-bit encryption can now be cracked in under a minute [11].

To mitigate all of these risks, all data and services involved in wireless networks should be evaluated and satisfy The Security Requirements Triad components of confidentiality, integrity and availability (CIA Triad) [13].

Continuous progress in developing more robust and secure operating systems in some way shifted the attention of hackers from finding the flaws in the operating systems to the less protected applications and in large numbers to the device drivers or wireless device drivers in particular.

## 4.0 Wireless Driver Vulnerabilities Overview

One of the most overlooked aspects of implementing secure wireless communication medium is the device driver’s vulnerabilities. Wireless driver exploitation falls into wider operation system’s kernel vulnerabilities domain and represents very fast growing and increasingly high risk trend in wireless security. The main reason for security professionals to be on alert is very high sophistication of the exploits.



Virtually all of the vulnerabilities are residing on a lower two levels of OSI model and are undetectable by the antivirus or intrusion detection systems. Historically, user written applications were separated from the operating systems Kernel, where device drivers reside. By exploiting the device, loaded into the memory, hackers receive unrestricted access to the whole system's resources. All of the existing OS as of today are at risk. Wide spread of the wireless technology with low default security settings and large number of old, legacy wireless clients creates a mess in situations where security needs to be tighten up to prevent unauthorized information access. Wireless drivers most often are exploited by using buffer overflow attack against particular fields of the MAC frames. One of the first known incidents that gained wide attention was Broadcom Wireless Device Driver vulnerability where SSID field was compromised.

### 4.1 Wireless Driver Exploits Details

Wireless security improved drastically in the last few years with many vendors and companies paying great deal of attention to producing solid secure products and implementing secure network environments. With introduction of more secure WPA/WPA2 encryption algorithms hackers are increasingly more often try to exploit client driver's vulnerabilities. Crafting malformed frames that trigger software vulnerabilities on a target machine was once very tedious work but with the introduction of the automated tools this will not be a problem in the near future. As described in "Wireless communication" section all clients go through a

three stage communication process before establishing connection to the network resources. More specifically, clients during first stage of that process perform mutual exchange of the requests in following order:

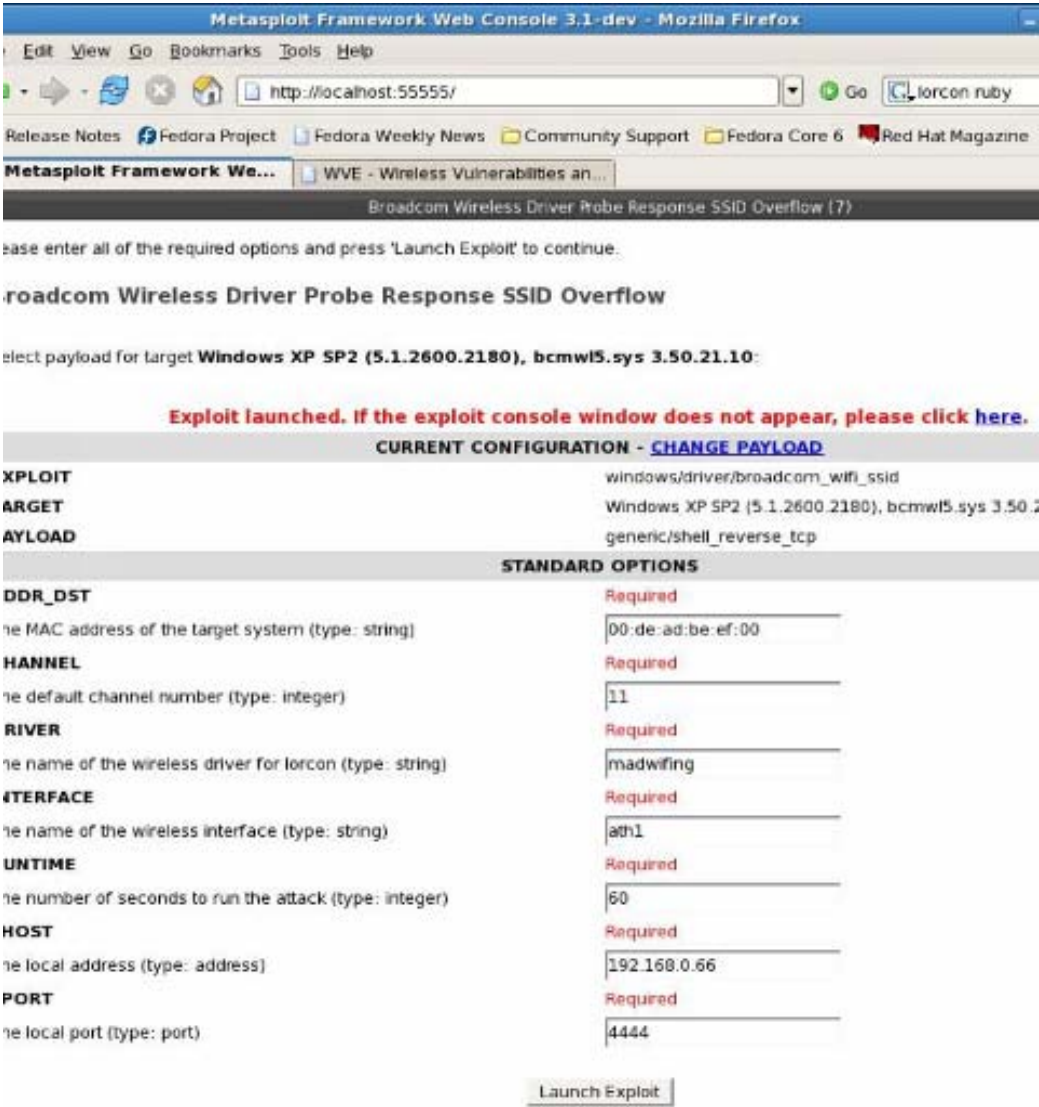


Figure 4 – Metasploit Framework web client interface

- ◆ Probe request;
- ◆ Probe response;
- ◆ Beacon;
- ◆ Authentication

```

< metasploit >

 \ (oo)_____\
 (__)_____) \
 ||--|| *

 =[msf v3.0-beta-dev
+ -- --=[178 exploits - 104 payloads
+ -- --=[17 encoders - 5 nops
 =[30 aux

msf > use windows/driver/broadcom_wifi_ssids
msf exploit(broadcom_wifi_ssids) > set PAYLOAD windows/adduser
PAYLOAD => windows/adduser
msf exploit(broadcom_wifi_ssids) > set INTERFACE wifi0
INTERFACE => wifi0
msf exploit(broadcom_wifi_ssids) > set DRIVER madwifi
DRIVER => madwifi
msf exploit(broadcom_wifi_ssids) > set PASS moo
PASS => moo
msf exploit(broadcom_wifi_ssids) > exploit
[*] Sending beacons and responses for 60 seconds...

```

Figure 5 – Exploit console, provided by Metasploit 3.1 [17]

Wireless discovery, whereas in passive or active mode, is done in a uniform way where clients are feeding probe requests to or listen to the beacon requests from AP or other wireless clients. In either way, network setting is embedded in the frames and broadcasted over the air. The nature of the wireless driver exploits is to produce wireless frames, very similar to the ones that are legit but containing slightly malformed subfield's data. Clients receiving malformed frames will behave differently due to the various network drivers' capabilities. Assuming the driver's weak code implementation is detected, and then the least risk scenario would be a temporary lockup of that driver network interface. The goal of the attacker would be not only to prevent clients from operations but to create a situation where detected driver vulnerability will allow intruders to gain an access to the victim's resources i.e. gain ring0 rights.

Ring0 or Kernel mode level access assumes that the code has full access to the operating systems resources and runs with supervisor mode [14] privileges. Weak driver's implementations would allow hackers to attain such rights through the process called fuzzing or automated driver's flaws testing. One of the available fuzzing tools is so called lorcon (loss of radio connectivity) [15] which inject 802.11 frames with malicious code that can be executed on client's machine.



Hello! This is an automatic message from the UWP campus network. It appears that your wireless driver is outdated. Your PC might be vulnerable to the remote exploits. You won't be able to access UWP resources until your PC is updated would you like to update your PC now?

Yes | No

Figure 6 – Example hot-spot default warning message

MAC layer frame's frame body subfield is a variable size field and is usually tested by supplying malformed data to it. The anticipated result of that activity is a code crash. While attack on this particular field is already known, other MAC frame fields can be compromised through more complicated fuzzing techniques. Results of successful fuzzing will expose client's vulnerability but the actual exploitation is not done yet. This is where the actual complication lies. The simple system or network interface denial of service also useful to some extent to intruders does not provide access to the victim's machine.

In order to fully exploit the wireless driver vulnerability capabilities, the buffer overflow technique is used to provide specifically constructed payload code that can be executed. This is when metasploit framework open source project comes into attention as a de facto standard framework with its capabilities to automate and produce highly sophisticated exploits tapping its Kernel mode flaws database. Figures 4 and 5 show how easy is the penetration of any system with outdated Broadcom wireless driver. Automated interface hides all the work producing valuable information in seconds.

Assuming, vulnerable target machine is identified by some tools such as fuzzing WiFIDEnum, network fingerprinting, or any other means of outdated wireless drivers detection, shell console window is launched, giving an attacker direct access to the victim's machine.

From here, attackers have the ability to issue commands and execute arbitrary code on victim's machine. As we can see, buffer overflow exploits, know for their high precision requirements, are even more complicated in terms of implementation in wireless driver's setting. Also automated tools as shown before exists, they are limited to the already precompiled sets of exposed vulnerabilities.

New exploits are just waiting to be found and with upcoming approval of 802.11n standard's complexity in driver's functionality, there is a huge area that is wide open for misuse. The skills level of the attacker has to be very high and this is why this type of exploits is still occurring very rarely. It also means that security professionals are suppose to be on constant alert, while continuing their education.

## 5.0 Mitigation

Today's arsenal of available prevention tools is very limited due to the nature of device driver exploits, it's highly advanced implementation status and similarities to kernel level root kits. They are almost impossible to detect using currently available security products and thus pose huge security risks to companies and regular end-users. While there is no single product available right now to detect, stop and prevent wireless drivers exploitation, complex multilayer approach should be taken by end users. Apparently, companies and other organizations would have different sets of objectives than the regular owners of wireless enabled devices. As the home end users are the ultimate consumers of the services provided by the third parties, main focus of the security burden should be placed on the wireless service providers. One biggest drawback of this and any other kinds of wireless medium common usage situations is the lack of regulations.

While this is the case, any size companies, providing wireless services should incorporate following techniques into their security requirements policies:

- ◆ Create an abstract set of security level risks associated with each wireless enabled end client based on its configuration, similar to the security clearance scheme of US government;
- ◆ Incorporate existing company-wise resources access group policies with created earlier security risks set
- ◆ Deploy proactive monitoring scheme to identify, notify, prevent, and upgrade wireless clients with inadequate device driver's software.
- ◆ Create an additional level of protection for hot-spots and any other mobile access point's installations by preventing access for any network clients with outdated device drivers.

In figure 6, UWP students trying to access intranet site with outdated device drivers are be warned of their high risk configurations and prevented from accessing any campus resources until the issue is resolved. In this scenario connection attempt also should be logged with any supporting information such as OS version, network card brand, driver's details, etc.

## 5.1 Available Tools

Exploitable conditions in a driver's code are usually detected using Black Box software testing methods. Most commonly used ones are fuzzing and driver fingerprinting where fuzzing mostly associated with discovering security flaws in the code. During the fuzzing process, as mentioned before, the wireless device is fed with random inputs. Inputs can be in a form of valid and/or malformed data. The goal is to make the driver fail, either by crashing the whole program or by misusing code assertions. Usually automated, fuzzing gives good return on the amount of time spent evaluating the code. As with any random data testing, it finds most obvious code bugs fairly fast. Another technique, wireless device driver fingerprinting, also not new, allows detection of the outdated drivers by sniffing the network traffic. Each wireless driver's type has its own network traffic pattern and can be used in detecting vulnerable clients.

An attacker's machine would passively listen in promiscuous mode to a wireless communication and accurately identifies wireless driver's type, version, and any other available information. Then this information would be matched against the precompiled database of vulnerable drivers. Such collections are freely available on the internet. As said before, there are numbers of fully automated software tools already available and no doubt more will be developed once more companies and organizations recognize the importance and seriousness of the wireless driver's vulnerabilities issue. Both fuzzing and fingerprinting techniques should be fully incorporated into proactive monitoring and prevention of possible exploitations by IT professionals.



## 6.0 Conclusion

There is a tremendous demand for more robust and secure wireless standards and availability of the relatively inexpensive wireless technologies. Constantly evolving technologies with their higher throughput rates and wider coverage areas makes network security professionals rethink how networks are designed, implemented, and secured. New 802.11n standards, along with increased performance gains, sets more complicated wireless drivers' requirements. There is a definite need for more comprehensive testing methods to discover flaws and vulnerabilities. Vendors should be responsible for providing timely updates and patched firmware to their hardware devices as well as users, who should be fully aware of the risks associated with using older products.

There is no magical ways to eliminate wireless driver vulnerabilities other than gradual replacement of legacy drivers on all wireless client machines, constant education of end users, monitoring and assessment of security policies.

## 7.0 References

- [1] "Trends in Unlicensed Spread Spectrum Devices." Presentation at FCC Commission Meeting 10 May 2001. Federal Communications Commission. April 28, 2008, [http://www.fcc.gov/Bureaus/Engineering\\_Technology/News\\_Releases/2001/net0104a.pdf](http://www.fcc.gov/Bureaus/Engineering_Technology/News_Releases/2001/net0104a.pdf).
- [2] U.S. Department of Commerce's Commercial Service. 802.11b Security overview. Intel.com. April 28, 2008, [ftp://download.intel.com/technology/itj/q22000/pdf/art\\_5.pdf](ftp://download.intel.com/technology/itj/q22000/pdf/art_5.pdf).
- [3] Perez, Eliot. "802.11i (How we got here and where are we headed)." The SANS<sup>TM</sup> Institute. April 28, 2008, <http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/10007/32124/01495946.pdf?arnumber=1495946>.
- [4] "Cisco WiFi Protected Access, WPA2 AND IEEE 802.11i." Cisco Systems, Inc. April 28, 2008 [http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps430/prod\\_gas0900aecd801e3e59.html](http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps430/prod_gas0900aecd801e3e59.html).
- [5] William Stallings, "IEEE 802.11: Wireless LANs from a to n," *IT Professional*, vol. 6, no. 5, pp. 32-37, Sept/Oct, 2004
- [6] "How 802.11 Wireless Works." TechNet. 23 Mar. 2003. Microsoft Corporation. April 29, 2008. <http://technet2.microsoft.com/windowsserver/en/library/370b019f-711f-4d5a-8b1e4289db0bcafd1033.mspx?mfr=true>.
- [7] "Open Systems Interconnection (OSI) Protocols." Cisco Systems, Inc. April 29, 2008 <http://www.cisco.com/en/US/docs/internetworking/technology/handbook/OSI-Protocols.html>.
- [8] "IEEE 802.11 Standard Overview." Pearson Education, Inform it. April 29, 2008 <http://www.informit.com/articles/article.aspx?p=19825>.
- [9] PEREIRA, JOSEPH. "How Credit-Card Data Went Out Wireless Door." Wall Street Journal 4 May 2007. April 28, 2008, [http://online.wsj.com/article\\_email/article\\_print/SB117824446226991797-1MyQjAxMDE3NzA4NDIwNDQ0Wj.html](http://online.wsj.com/article_email/article_print/SB117824446226991797-1MyQjAxMDE3NzA4NDIwNDQ0Wj.html).
- [10] "Retailers haven't learned from TJX - still running WEP." CNET Networks. April 28, 2008. <http://blogs.zdnet.com/Ou/?p=487&tag=rbxccnbzd1>.
- [11] "A history of WEP and RC4." Technical University Darmstadt, Germany. April 29, 2008, <http://www.cdc.informatik.tudarmstadt.de/aircrack-ptw/>.
- [12] FCA. Informational Memorandum. Managing Risks Associated with Wireless Networks and Customer Access. April 28, 2008. <http://www.fca.gov/apps/infomemo.nsf/59ab19ff3b724b73852569530076c529/0a7c7a98e6ff500985256c94006d2a99?OpenDocument>.
- [13] Stallings, William, and Lawry Brown. Computer Security: Principles and Practice. N.P.: Prentice Hall, 2007; 8-9.
- [14] "Glossary of Multicast acronyms and terms: Ring." Stratus Technologies. April 29, 2008. <ftp://ftp.stratus.com/vos/multics/tvv/mgr.html#ring>.
- [15] "Lorcon." 802.11 Mercenary. April 28, 2008. <<http://www.802.11mercenary.net/lorcon/>>.
- [16] "Fingerprinting Hackers: Technique Demonstrates Wireless Device Driver Vulnerabilities." Science Daily. 5 May 2008 <<http://www.sciencedaily.com/releases/2006/09/060912214943.htm>>.
- [17] "Wireless Driver Attacks." SANS Technology Institute. May 5, 2008 <[http://www.sans.edu/resources/securitylab/wireless\\_security\\_1.php](http://www.sans.edu/resources/securitylab/wireless_security_1.php)>.

# Energy-Efficient TDMA MAC Protocol for Wireless Sensor Networks Applications

G M Shafiullah<sup>1</sup>, Adam Thompson<sup>2</sup>, Peter J Wolfs<sup>2</sup>, Shawkat Ali<sup>3</sup>

<sup>1</sup> Centre for Railway Engineering, Faculty of Sciences, Engineering & Health,

<sup>2</sup> College of Engineering & Built Environment, Faculty of Sciences, Engineering & Health

<sup>3</sup> School of Computing Sciences, Faculty of Business & Informatics

Central Queensland University, Rockhampton, QLD-4702, Australia

Phone: +61749309313, Email: [g.shafiullah@cqu.edu.au](mailto:g.shafiullah@cqu.edu.au)

**Abstract** — The availability of low-powered and cheap microprocessors, radio frequency integrated circuits and the development of new wireless communication techniques, make the wireless sensor networks (WSN) one of today's most promising technologies. Minimizing energy consumption and maximizing the lifetime of the networks are key requirements in the design of sensor network applications. Optimally designed medium access control (MAC) and routing protocols minimize energy consumption and prolong the network life. In this study, we have investigated an energy-efficient adaptive TDMA (EA-TDMA) protocol for railway applications that used in communication between sensor nodes and the cluster-head (CH) placed in a railway wagon. This protocol is suitable for medium traffic applications and reduces energy consumption by shortening the idle period when devices have no data to transmit. We have developed an analytical model for EA-TDMA and compared its performance with conventional TDMA and bit-map-assisted (BMA) protocols.

**Key Words** – Wireless sensor networking; railway wagons; energy-efficiency; MAC protocol.

## I. INTRODUCTION

Recent emergence of micro-electro-mechanical systems (MEMS) technology, wireless communications and integrated circuit design has enabled the development of low-cost, low-power, multipurpose sensor networks. This low-power sensor network can provide new monitoring and control capability in architectural infrastructure, vehicle infrastructure, including monitoring of railway signalling systems, and rail tunnel and track monitoring. Each sensor has wireless communication capability and some level of intelligence to collect data and route it to the sink node. This enhances operational efficiency. Sensor network applications require long lifetime, data accuracy, and energy efficiency. In a sensor node, power is required for data sensing, communication and data processing [1-3].

Energy efficiency is a major issue in designing an efficient WSN; the main sources of energy loss are:

- *Idle listening* occurs when a node is waiting to receive data from its neighbour nodes. A radio keeps their receiver open

all the time as it does not know when a message will arrive from its neighbours. So it is seen that nodes are kept in idle for most of the time.

- *Collision* occurs when two nodes want to send data at the same time and interfere each other. Collision can be removed in schedule-based MAC protocols however; it is a concern issue in contention-based protocol.

- *Overhearing* takes place when a node receives packets that are destined to other nodes i.e. receiver node is not the exact destination. It is a dominant factor in high traffic load situations.

- *Over-emitting* happens when a node sends packet to its destination but its destination is not ready to receive that message.

- *Control packet overhead* is the energy required to send, receive and transmit control packet [2-3].

Optimal design of MAC and routing protocol ensures power-efficient WSN applications that minimise power consumption and hence, maximise battery lifetime of the networks. Owing to their poor energy conservation, traditional MAC and routing protocols are not suitable for WSN applications [4].

In this paper, we have investigated an energy-efficient TDMA MAC protocol using clustering techniques to monitor vertical acceleration and lateral instability of railway wagons. This protocol reduces energy dissipation in sensor nodes while transmitting information between two nodes. In this protocol each node wakes up in its allocated time slot but, turns off its radio immediately if it has no data to transmit, otherwise it transmits its data to the CH. A mathematical model has been developed to evaluate the performance of this cluster-based EA-TDMA protocol. This paper is organised as follows. Section II represents the review of existing literatures. The architecture of proposed EA-TDMA protocol is discussed in Section III. Analytical Model of EA-TDMA, TDMA and BMA protocol is discussed in Section IV. Performance evaluation of the protocol is discussed in Section V. Section VI concludes the paper with future direction.

## II. LITERATURE REVIEW

An efficient use of MAC and routing protocols plays a key role in achieving energy-efficient WSN applications. The major requirements of a MAC protocol are: energy-efficiency, scalability, latency, fairness and bandwidth

utilisation. MAC protocols generally focus on two different areas: contention based or random-access methods and schedule-based or fixed-assignment channel access methods.

Contention-based protocols are scalable and adaptable to node density or traffic load variations. However this scheme has a major limitation relating to an enormous amount of energy wasted due to collisions, overhearing and idle listening [4]. Overcoming these drawbacks are a primary concern of MAC protocol design and a few approaches that reduces these key problems have been introduced, such as S-MAC, T-MAC [4-6].

S-MAC [4] is a contention-based synchronous protocol, designed to reduce energy consumption from the sources of energy loss: *idle listening*, *collision*, *overhearing* and *control overhead*. Periodic sleep, virtual clustering and adaptive listening techniques are used in this protocol to achieve low power duty cycle that reduces energy consumption significantly. It uses virtual carrier sense techniques to reduce collision avoidance and in-channel signalling to implement overhearing avoidance. S-MAC fragments the long message into many small parts and transmits them in burst to reduce contention and communication overhead. An evaluation with the Mote developed by the UCB proved that S-MAC works 2-6 times more energy efficient than the IEEE 802.11 MAC. Simulation results show that power efficiency achieved with this protocol as it reduces collision and overhearing. As only active part of the frame is used for communication S-MAC reduces overall throughput. However, it has synchronization overhead of sending and receiving SYNC packets. It has high message latency and not suitable for a network with heavy traffic [4 - 5].

All messages are transmit in a burst of variable length in T-MAC protocol to reduce idle listening significantly. Nodes are periodically wakes up to transmit data and go to sleep again when there is no event for a specific time period. A node will be on active status until it is busy with listening and transmitting message and after waiting a period of  $T_A$  as idle it turns to sleep mode. It enhances the poor results of the S-MAC protocol under variable condition. Active period  $T_A$  is adjusted dynamically and data transmission occurs during the start of active period. Virtual clustering mechanism is used to synchronize node schedule. A back-off algorithm is used to reduce the collisions in high load and minimize latency during minimum load condition. The future-request-to-send (FRTS) mechanism is introduced to overcome early sleep problems. Determination of active period  $T_A$  is a complex work and sometime sending node may go to sleep even if the receiving node is alive. To overcome early sleeping problem FRTS and data-send (DS) packets are used which increases power dissipation [6].

Time-division multiple access (TDMA) is a scheduled-based MAC protocol in which the channel is divided into several time slots. Each node is assigned a time slot. The node wakes and transmits data only in their allocated slots and remains in sleep mode at other times [2-4, 13].

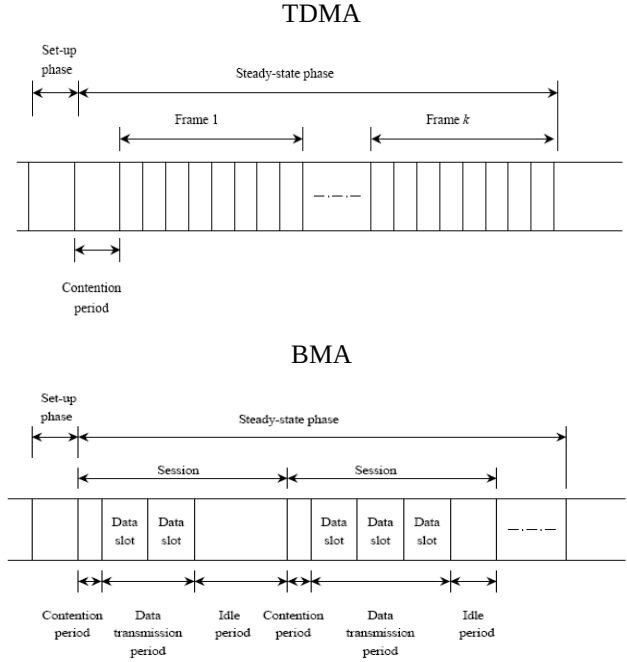


Fig. 1: Single round for TDMA and BMA protocol [8].

Conventional TDMA MAC protocols are used to prevent collisions by fixing individual time slots for each node. It also reduces energy consumption as all nodes are in sleep condition except for the transmitting nodes. Hence, TDMA protocol is a suitable choice for sensor networks. However, this protocol is suitable for a network with heavy traffic load. Usually in sensor networks there are fewer nodes in a cluster which have no data to transmit to the CH in their allocated time slots. Nodes with empty buffers still turn on their radio during their scheduled time and hence dissipate some of their remaining energy. The CH has to keep its radio on all the time in readiness to collect data from the nodes it serves. Hence, the CH does not use its energy efficiently as it wastes energy in its idle time. The energy-efficient TDMA (E-TDMA) reduces energy consumption due to idle listening. Sensor nodes keep their radios off when there is no data to transmit. However, CH has to keep on the radio during all the timeslots and hence waste energy [2-4, 8]. Fig. 1 illustrated a single round for TDMA protocol.

Low-energy adaptive clustering hierarchy (LEACH) [7] combines cluster-based routing and MAC protocol. This protocol architecture for wireless micro sensor networks achieves low energy dissipation and latency without sacrificing application-specific quality. The key design goals of LEACH are: adaptive cluster formation, energy-efficient TDMA MAC, application specific data processing. Nodes in a WSN organise themselves into local clusters with one node acting as the CH. Overall throughput of the network increases as latency is reduced and system life increases. Collision and interference are overcome by using TDMA and CDMA protocol respectively. Data aggregation mechanisms reduce the amount of data that must be transmitted to a BS, hence, reduces energy consumption. Strengths and weaknesses of highlighted MAC protocols are illustrated in Table-I.

TABLE - I: Comparison of MAC protocols

Article	Rationale	Strengths	Weaknesses
S-MAC [4, 5]	Techniques used in S-MACs are: nodes periodically listen and sleep, virtual clustering and adaptive listening.	Synchronization overhead decreases due to adaptive listening techniques. Energy consumption reduces by sleep scheduling	As sleep, listen periods are constant, may decrease the efficiency of the protocol. Increases latency because an event may occur during sleep time.
T-MAC [6]	Introduces an adaptive duty cycle and reduces energy loss by shortening the awake period during idle condition.	It enhances the poor results of the S-MAC protocol under variable traffic load by dynamic listen period. T-MAC can change the network condition.	High latency and overhead associated with synchronization and introduces early sleeping problem.
TDMA [2, 4]	Channel is divided into several time slots and one time slots is assigned to each node. Nodes awake and transmit data in their allocated slots.	Prevent collisions by fixing individual time slots for each node and reduce energy consumption as all nodes are in sleep condition except the transmitting nodes.	There may be fewer nodes in a cluster which have no data to transmit to the CH in their allocated time slots. CH has to keep its radio on all the time.
LEACH [7]	Distributed cluster formation enables randomized, adaptive, self configuring features.	Overall throughput of the network increases as latency is reduced and system life increases. Collision and interference are overcome by using TDMA and CDMA protocol respectively.	Due to its distributed cluster formation algorithm, it cannot ensure the coverage of entire network. TDMA schedule introduces time delay.
BMA [8]	An intra-cluster communication protocol reduces energy wastes due to idle listening and collisions.	BMA has low complexity, reduces energy wastes due to idle listening and low packet latency. BMA outperforms TDMA and E-TDMA in low and medium traffic loads.	BMA is suitable only for low traffic, i.e., relatively few sensor nodes per cluster. Data may arrive in empty node at any time during node to CH data transmission.

Bit-map assisted (BMA) [8-9] is a schedule-based MAC protocol that reduces energy wastage from idle listening and collisions. The cluster setup phase is identical to LEACH [16] protocol. In BMA sensor nodes forward data to the CH only if a significant event has been observed. BMA protocol reduces energy consumption occurring in conventional TDMA systems due to idle listening in the absence of data in any node in their allocated scheduled time slots. All nodes in the cluster keep their radios on in the contention period and transmit a 1-bit control message during its allocated slot if it has data to transmit, otherwise the slot remains empty. During data transmission period, each source node turns on its radio and sends its data to the CH over its allocated slot and turns off its radio all other times. Non-source nodes keep their radios off during the data transmission period. BMA has low complexity, reduces energy wastes due to idle listening and low packet latency. BMA outperforms TDMA and E-TDMA in low and medium traffic loads. It is an energy-efficient intra-cluster low-latency protocol with low complexity and reduced energy wastes in idle listening. BMA is suitable only for low traffic, i.e., relatively few sensor nodes per cluster [8-9]. Fig. 1 represents a single round for BMA protocol.

### III. ENERGY-EFFICIENT ADAPTIVE TDMA (EA-TDMA) PROTOCOL

Conventional TDMA-based MAC protocols are used to prevent collisions by fixing individual time slots for each node. This scheme also reduces energy consumption as all nodes are sleeping except the transmitting node. Nodes in the cluster send their sensed data to cluster head (CH) in their respective time slots. Nodes with empty buffers still turn on their radio during its scheduled time and the CH also keeps its radio on all the time to listen to the nodes in its cluster, this wastes energy. TDMA-performs better under high-traffic load conditions. A high traffic load means all nodes always have data to transmit. Conventional TDMA is the most suitable for an application that comprises with heavy traffic load and BMA is suitable for an application that comprises with low traffic load.

We have investigated train wagon body acceleration signals using sensor networking techniques to monitor vertical acceleration behaviour of railway wagons. In this application, accelerometer data was collected continuously from sensor nodes placed inside the wagons to investigate the train track irregularities and lateral instability. It has medium to high traffic loads. Therefore considering the application requirements, we have concentrated on developing a protocol suitable for medium to high data traffic and proposed energy-efficient EA-TDMA protocol. In the EA-TDMA protocol every node wakes up its allocated schedule but turns off its radio immediately if there are no data to transmit, otherwise it transmits its data to the CH. This protocol reduces energy consumption during data transmission by reducing idle listening period.

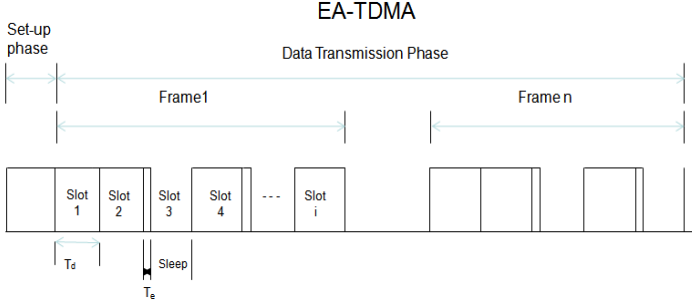


Fig. 2 Operation diagram of EA-TDMA protocol

We assume that the time slot duration in a TDMA schedule is  $T_d$ . The assigned node turns on its radio and transmits data to the CH during this period. If any node has no data to transmit it turns off its radio immediately. Therefore, we assume that the radio is turned off after time  $T_e$ , instead of remaining on for the entire time period of duration  $T_d$ , as illustrated in Fig 2.

To meet the requirements of WSNs, some assumptions of this approach are: nodes are static throughout the network; energy levels of the normal nodes are almost identical but die at almost the same time; CH has higher energy level than normal nodes; and all data slots in a TDMA schedule are the same size. We have divided the operation of the proposed protocol into rounds. Each round comprises of a set-up and a data transmission phase. Both cluster formation and CH selection occur in the setup phase, while data transmission from node to the CH and on to the BS occurs during data transmission phase, as described in the following details.

**Setup Phase:** Considering application area and simplicity in this study we assume the network consists of a multiple fixed cluster in which there is one CH is located in the centre of the cluster. Based on application and cluster size we have considered direct transmission for data communication between source nodes to CH instead of multihop data transmission. In setup phase, the CH builds a TDMA schedule and broadcasts the schedule to all nodes within the cluster. CH also informs all nodes about the start of current round, frame start/stop time, number of frame in a round.

**Data Transmission Phase:** The data transmission phase contains  $l$  frames. The size and duration of each of the frames are fixed. Nodes send their data to the CH once per frame during their allocated time slots. We assume that there is one CH node and  $N$  non-CH nodes in a cluster. During the data transmission period each node turns on its radio in its assigned time-slots and transmits data to the CH. If any node, after turning on its radio finds an empty buffer, i.e., there is no data to transmit, then the node turns off its radio immediately to save energy. The nodes turn into sleep mode instead of idle mode in the absence of data. Non-CH nodes transmit sensed data to the CH's buffer. After receiving all data from the nodes, the CH aggregates the data.

After a predefined time, the system begins the next round and the whole process is repeated. A timing diagram of the proposed EA-TDMA protocol is illustrated in Fig. 3.

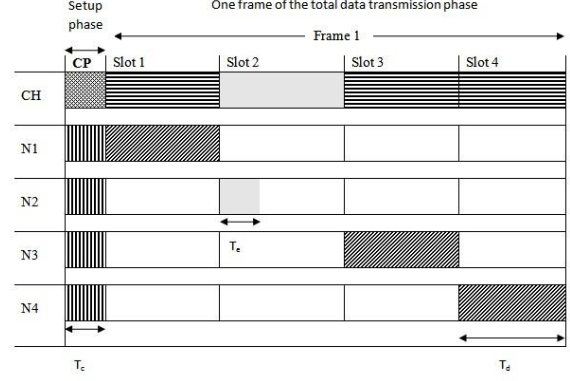


Fig. 3: Timing diagram of EA-TDMA protocol

#### IV. ANALYTICAL MODEL OF EA-TDMA PROTOCOL

We analyse our proposed protocol in a cluster-based network in which there are one CH and  $N$  non-CH nodes. Assuming that there are  $l$  frames in a round and  $n_k$  source nodes in the  $k_{th}$  frame. The nodes which have data to transmit are called source nodes. Bernoulli Trail, as used in BMA [9] protocol has been used to calculate the probability that a node has data to transmit. Hence,  $n_k$  is a Binomial random variable and source node in a frame would be  $E[n_k] = Np = n$ ,  $k = 1, 2, \dots, l$  where  $p$  = probability. The number of source nodes has varied from frame to frame. The power consumption in the transmit mode and the receive mode are respectively  $P_t$  and  $P_r$ . The typical power consumption in the proposed protocol occurred during turn on and empty checking period of the idle nodes:  $P_e$ . For simplicity, as stated in [4, 7], the energy required to turn on the radio by the source nodes for transmission or reception is negligible, hence ignored in the analysis here.  $T_d$  is the time required to transmit or receive a data packet, and  $T_c$  is the time required to transmit/receive a control packet. The time required for a node to switch on, check its buffer and turn off its radio is  $T_e$ . Power consumption and energy dissipation of an idle node is  $P_i$  and  $E_{in}$  respectively. The time required for the cluster-head to transmit a control packet for BMA is  $T_{ch}$ .

Initially, an analytical model was developed based on the energy model [9] for the EA-TDMA and then compared with the conventional TDMA and BMA protocol.

**A. EA-TDMA Protocol:** During setup phase, CH and all nodes keep their radios on and communication takes place between the CH and other nodes. In this period, the CH selects time slots for individual nodes for data communication and informs all nodes in the cluster. Therefore, energy consumption by CH to send a control packet is  $E_{ch} = P_t T_c$  and energy consumption by each node to receive control packet is  $E_n = P_r T_c$ . Total energy consumption during setup phase for all nodes is:  $E_C = P_t T_c + N P_r T_c$

Each node transmits a maximum of one packet per frame interval. During a frame, transmission energy dissipated in a data node is  $E_{dn} = P_t T_d$ . Energy dissipated in a node that has

no data to transmit to the CH is  $E_e = P_e T_e$ . The energy  $E_e$  is used to switch on, check transmit buffers and then turn off the radio module. The energy consumed by the CH is  $E_{ch-e} = n P_r T_d + (N - n) P_i T_d$ . Therefore, the system energy dissipation in the  $k_{th}$  frame is:

$$E_{Se} = n P_i T_d + (N - n) P_e T_e + (N - n) P_i T_d + n P_r T_d \quad (1)$$

Assume, there are  $l$  frames in a round. Therefore energy consumption of the system in a round, of the proposed protocol is:

$$E_{EA-TDMA} = P_t T_c + N P_r T_c + l[n P_i T_d + (N - n) P_e T_e + (N - n) P_i T_d + n P_r T_d] \quad (2)$$

**B. TDMA Protocol:** During the  $k_{th}$  frame, the energy dissipation in a source node is  $E_{dn} = P_i T_d$ . A non-source node, i.e., the empty-buffer node turns on its radio and keeps idle during its scheduled time slots consuming energy  $E_{in} = P_i T_d$ . Energy consumption in the CH is  $E_{ch} = n P_r T_d + (N - n) P_i T_d$

Energy consumption during setup phase of TDMA is the same as EA-TDMA. Therefore, the average energy dissipation in a round is

$$E_{TDMA} = P_t T_c + N P_r T_c + l[n P_i T_d + 2(N - n) P_i T_d + n P_r T_d] \quad (3)$$

**C. BMA Protocol:** In BMA there is a contention period in each session when all nodes keep their radios on. Each source node transmits a control packet during its scheduled slot and remains idle  $(N-1)$  slots. Total energy consumption by each source node during a single session is  $E_{dn} = P_t T_c + (N-1) P_i T_c + P_r T_{ch} + P_i T_d$ . Each non source node stays idle during the contention period and keeps its radio off during the data transmission periods and hence energy consumption is  $E_{in} = N P_i T_c + P_r T_{ch}$ . During the contention period the CH node receives  $n$  control packets and stays idle for  $(N-n)$  slots. Energy dissipation by the CH is  $E_{ch} = n (P_r T_c + P_i T_d) + (N-n) P_i T_c + P_t T_{ch}$

The average system energy dissipation in each round in BMA protocol is:

$$E_{BMA} = l[n(P_t T_c + (N-1) P_i T_c + P_r T_{ch} + P_i T_d) + (N-n)(N P_i T_c + P_r T_{ch}) + n(P_r T_c + P_i T_d) + (N-n) P_i T_c + P_t T_{ch}] \quad (4)$$

## V. RESULTS AND ANALYSIS

We compare EA-TDMA protocol with the BMA and conventional TDMA protocol in terms of energy dissipation. For analytical analyses we have used Rockwell's [10] WINS model as used in the BMA protocol. In this model the radio transceiver uses 462 mW for transmitting, 346 mW for receiving and 330 mW for idle listening. The data rate is 24 kbps. Assume a data packet size of 250 bytes and a control packet size of 18 bytes. In this experiment we assume  $P_i = P_e$ . We compare the above three protocol in terms of energy dissipation.

Fig. 4 is drawn by noting the above conditions that provide comparison of the EA-TDMA, TDMA and BMA protocols as a function of probability  $p$ . In this case  $N = 10$  and there are  $l$

$= 2$  frames in a round. It is observed that for very low traffic when  $p = 0.1$  to  $0.2$  the BMA protocol is superior to that of the EA-TDMA and TDMA protocols. However, energy consumption of BMA increases linearly when  $p \geq 0.3$  and EA-TDMA protocol outperforms both BMA and TDMA protocols. EA-TDMA protocol outperforms TDMA protocol except for heavy traffic. For heavy traffic energy dissipation of EA-TDMA and TDMA are same which is much less than the BMA protocol. Therefore, we may conclude that the EA-TDMA protocol is most suitable for medium traffic load.

Fig. 5 compares MAC schemes in terms of energy, where nodes  $N$  are variables and frame  $l = 10$  and probability  $p = 0.6$ . EA-TDMA protocol outperforms both BMA and TDMA protocols in terms of energy consumption. EA-TDMA protocol saves more energy if the number of nodes increases in the network. Hence, EA-TDMA is suitable for large networks.

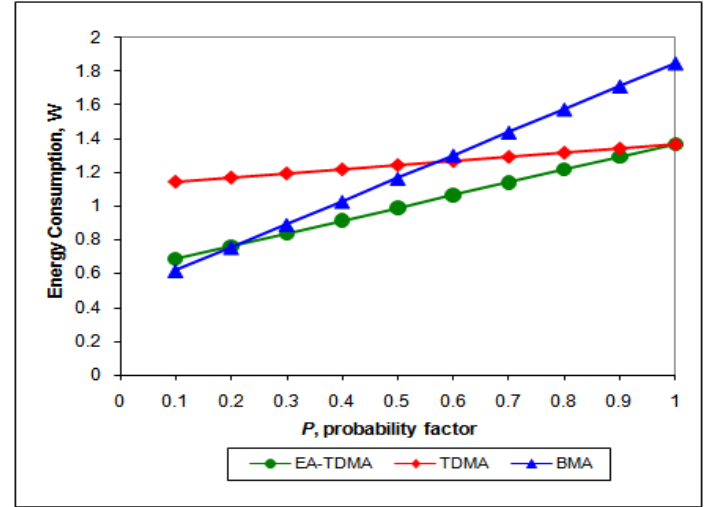


Fig. 4: Comparing the energy dissipation of EA-TDMA protocol with BMA and conventional TDMA protocol as a function of probability  $p$ .

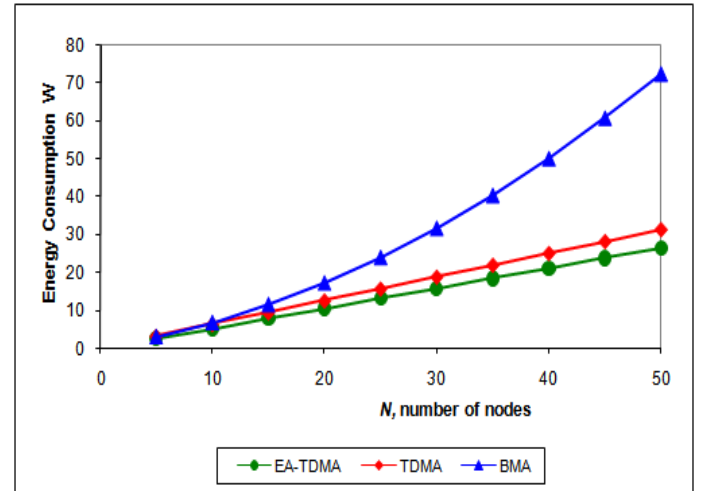


Fig. 5: Energy dissipation of EA-TDMA, BMA and TDMA protocols as a function of number of nodes in a cluster,  $N$ .

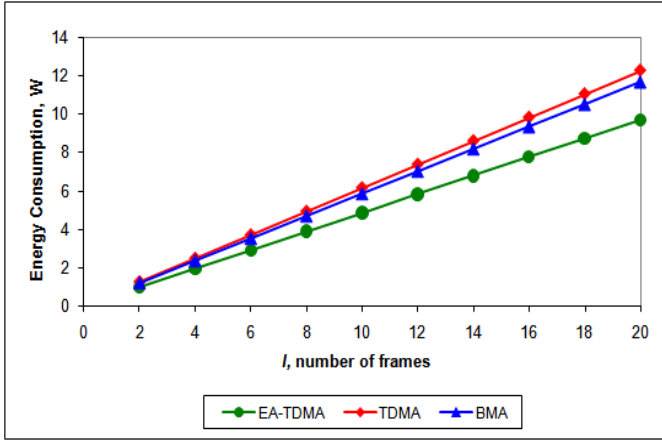


Fig. 6: Comparing the energy dissipation of EA-TDMA with BMA and TDMA protocol as a function of number of frame,  $l$

From Fig. 6, it is observed that EA-TDMA protocol saves more energy if the number of frames per round increases and outperforms BMA and TDMA protocol in terms of energy dissipation.

From the above analytical results we may conclude that EA-TDMA protocol is more energy-efficient than BMA and TDMA protocols in case of medium traffic applications. However, BMA protocol performs better than EA-TDMA and TDMA protocols in case of low traffic applications especially event driven applications. For high traffic load energy consumption of EA-TDMA and TDMA are the same and perform better than the BMA protocol. We have investigated a cluster-based TDMA protocol to develop a monitoring system for railway. This problem demanded medium to high traffic conditions, therefore EA-TDMA protocol is the most suitable for this application.

## VI. CONCLUSION

In this study, considering traffic load of the application, a cluster-based energy-efficient adaptive TDMA-based MAC protocol has been investigated for railway to monitor typical dynamic behaviour of railway wagons. A mathematical model has developed for the proposed EA-TDMA protocol and compares its performance with conventional TDMA and BMA protocols.

Analytical results prove that EA-TDMA protocol outperforms both TDMA and BMA protocols for medium to high traffic. From experimental results it has observed that BMA protocol is superior to EA-TDMA especially for low traffic or event driven applications. However, if the probability  $p \geq 0.3$ , EA-TDMA performs better than BMA protocol which is natural scenario for our railway applications.

This study is still on introductory stage and needs to be developed in different areas. Therefore, it deserves further investigation that focuses on some specific areas such as:

- investigate energy-efficient routing protocols and their integration with the proposed EA-TDMA algorithm.
- validate the analytical results stated in this paper using simulation software.
- reduce power consumption of the protocol by introducing duty cycle in CH.
- increase CHs lifetime by introducing rotational CH in the cluster.
- integrate the protocol with machine learning algorithm to develop monitoring signal.

## REFERENCES

- [1] G. M. Shafiullah, A. Gyasi-Agyei, and P. Wolfs, "Survey of wireless communications applications in the railway industry," in *Conf. on Wireless Broadband and Ultra Wideband Comm.*, Sydney, Aug'07.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey of Sensor Networks," *IEEE Communications Magazine*, vol. 40, August, 2002, pp.102-114
- [3] D. S. Sokwoo Rhee and S. Liu, "Techniques for minimizing power consumption in low data-rate wireless sensor networks," *Wireless Communications and Networking Conference*, vol. 3, March 2004, pp. 1727-1731.
- [4] Wei Ye, John Heidemann, and Deborah Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the IEEE Infocom*, pp. 1567-1576. New York, NY, USA, USC/Information Sciences Institute, IEEE. June, 2002
- [5] Wei Ye, John Heidemann, and Deborah Estrin, "Medium Access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM trans. on Networking*, vol. 12, pp. 493-506, Jun'04.
- [6] K. L. Tijs Van Dam, "An adaptive energy-efficient MAC protocol for Wireless Sensor Networks," in *SenSys'03*, Nov'03, pp. 171 - 180.
- [7] W.R. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE on Wireless Communications Trans.*, vol.1, No.4, Oct'02, pp. 660-670.
- [8] Jing Li, "A bit-map assisted energy-efficient MAC scheme for wireless sensor networks," Master's thesis, Electrical Engineering, Mississippi State University, May 2004.
- [9] Jing Li, and Georgios Y. Lazarou, "A bit-map-assisted energy-efficient MAC scheme for Wireless Sensor Networks," in *Proceedings of 3rd International Symposium on Information Processing in Sensor Networks (IPSN '04)*, pp. 55-60, Berkeley, Calif, USA, April 2004.
- [10] V. Raghunathan, C. Schurgers, S.Park, and M.B. Srivastava, "Energy-aware Wireless Microsensor Networks," *IEEE signal Processing Magazine*, Mar'02, pp.40-50.



## Extensive Video Quality Evaluation: A Scalable Video Testing Platform

Nick Vercammen

Nicolas Staelens

Alexis Rombaut

Brecht Vermeulen

Piet Demeester

Ghent University

IBBT, Department of Information Technology (INTEC)

Gaston Crommenlaan 8, Bus 201, 9050 Ghent, Belgium

{nick.vercammen, nicolas.staelens, alexis.rombaut, brecht.vermeulen, piet.demeester}@intec.ugent.be

### Abstract

*With the advent of new upcoming online video services such as IPTV, Video on Demand (VoD) and Peer-to-Peer (P2P) video streaming, content providers are gaining more and more interest in measuring and monitoring video quality as perceived by end-users; also known as Quality of Experience (QoE). Objective video quality metrics provide a means of measuring visual quality degradations but in order to be able to measure QoE, these objective metrics should incorporate all quality affecting parameters such as encoding bitrate, network impairments and error concealment techniques. As a consequence, in order to construct or validate a proper objective video quality metric, extensive video evaluation tests must be performed. In this paper we present a scalable video testing platform that simplifies the management and execution of such video quality evaluation tests. Results indicate that the use of our testing platform drastically reduces overall experiment duration.*

### 1. Introduction

With the advent of new upcoming online video services such as IPTV, Video on Demand (VoD) and Peer-to-Peer (P2P) video streaming, content providers are gaining more and more interest in measuring and monitoring video quality as perceived by end-users; also known as Quality of Experience (QoE). According to the definition in [2], QoE includes all end-to-end system effects which means that in the case of video streaming the influence of both encoding, streaming, decoding and viewing must be taken into account. During the encoding process, encoding artefacts such as blockiness and blurring can be introduced. While streaming the video, packet loss can occur which result in the loss of information. Finally, during decoding and viewing, different error concealment techniques can be used to mask the visual impairments. These are only a few

of the possible system effects that influence the way human observers perceive visual quality. Objective video quality metrics, such as Structural Similarity (SSIM) [7] and Video Quality Metric (VQM) [1], provide a means of measuring visual quality degradation based on a comparison between the original and the received video sequence. Results indicate that these metrics correlate well with scores of subjective quality assessment but they were developed mainly to evaluate video codec quality. As already stated, an end-to-end objective video quality metric should incorporate all possible system effects. As a consequence, in order to construct or validate a proper objective video quality metric extensive video evaluation tests must be performed using different values for each of the quality affecting parameters such as encoding bitrate, packet loss rate and error concealment technique.

In this paper we present a scalable video testing platform that simplifies the management and execution of such video quality evaluation tests. Our proposed platform enables the parallel execution of different test cases at the same time and thus reducing the overall time needed to perform all quality evaluation tests.

The remainder of this paper is structured as follows. In section 2, the individual building blocks of a general video quality experiment are described which, in turn, are used to derive the requirements for a video testing platform. Section 3 provides an overview of the physical testbed and the management software which enable scalable video quality experiments. The performance of our scalable platform is evaluated in section 4 and finally, conclusions and future work are described in section 5.

### 2. Video Quality Evaluation

When a video is streamed across a network, quality can degrade due to packet loss, jitter and delay. The influence



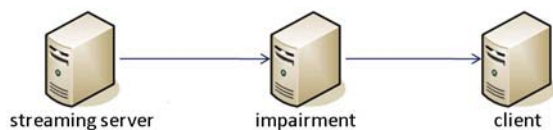
of these network impairments on the video quality can be researched in different ways: through theoretical models, by simulating the network impairments or by emulating them in a real network environment. The last approach is the most realistic one since it enables us to stream video using real-life network conditions and hardware. This is the way we perform our video quality research. In this section, we first analyse a general video quality experiment and break it up into different building blocks.

## 2.1. Video quality experiment: building blocks

A basic video quality experiment, as described in [6], consists of several different test cases which in turn consists of different test steps: *encoding* a reference sequence, *impairing* the encoded sequence and *analysing* the video quality of the impaired sequence.

In the encoding step a raw video sequence is compressed by means of an encoder. An encoder is an implementation of a specific codec like MPEG-2 or H.264. There exist a variety of different encoding standards and encoding tools, each with their own performance and characteristics. Because the quality of a video sequence as experienced by the end user for the same artifacts depends strongly on the type of content used [4, 5], a test should be performed across a number of sequences, all different in terms of spatial and temporal complexity.

Network impairments can be generated through simulation as is done in [6] or by emulating the real network with real streamers. Within our research group we put a strong emphasis on emulation, because we want to test the behaviour and influence of different hardware, real network setups and real streaming servers and packetisation.



**Figure 1. Basic streaming setup: streamer streams video to the client over a network emulated by the impairment node**

A basic video streaming setup for generating network impairments is shown in Figure 1. This basic setup consists of three machines: a streaming server, an impairment node and a client. The streamer streams the video to the client. Different streaming protocols (e.g. UDP, RTP and MPEG-TS) can be used to deliver the video to the client. The behaviour of the network under test is emulated by an impairment node. Input for the impairment node can be: packet

loss rate, types of packets to impair, variable packet delay and so on. The client receives the video stream and captures it. Each client has its own error concealment techniques, so it is important to test with different client software or to control the error concealment technique by means of a custom developed receiver.

In the last step, the actual video quality evaluation is performed. By using one or more quality metrics like Peak Signal-to-Noise Ratio (PSNR), SSIM [7] and VQM [1] the quality degradation of the video can be determined. This test step is the most time consuming step in the experiment due to the complex algorithms and byte by byte comparisons that are used.

Video quality evaluation usually involves extensive testing using different encoding techniques, different amounts of network impairments and different error concealment techniques. As a simple example, the basic video test experiment from [6] involves 4 different video sequences, encoded with 9 different quantization parameters. Each of these encoded sequences was then impaired with 7 different packet loss rates. This resulted in a total number of 252 different test cases. As can be seen, the input parameter space for video quality testing can be extensive and therefore calls for parallel execution of test cases. This parallel execution should be managed by a flexible and scalable video testing management platform.

## 2.2. Video test platform requirements

Based on the description of the general test experiment in the previous section, we can extract the basic requirements for the management software of the video test platform:

- **Distributed application management:** A variety of applications on distributed machines have to be configured, started and stopped. It should not matter to the platform which operating system is running on the machines (e.g. videoclient on Linux or Windows).
- **Distributed file handling:** Video data files and application configuration files have to be distributed to the different servers or other devices (e.g. set-top boxes, traffic generators) before the start of each test. After a test case execution the resulting data files need to be collected to a central repository. The platform should be able to handle this reliably and efficient.
- **Repeatable testing:** We want to be able to repeat a test with slightly different input parameters. For example after doing a test with a set of files with resolution X, the test should be easily repeatable with files with resolution Y. So it should be easy to recall an already executed experiment and tweak some parameters in it.

- **Scalable:** The experiment duration will further increase when we increase the duration of our input sequences or if we want to do more advanced quality analysis (see Section 4.1). As a typical experiment consists of several hundreds of tests, and we still want to perform this in a reasonable amount of time, we will have to come up with a scalable solution that will parallelize some processes and eliminate dead time.
- **Flexible:** A basic setup can be as simple as depicted in Figure 1, but it can be more complex with more devices. The platform should be flexible in terms of the number and types of hardware devices that can be managed. It should be possible to easily create different network situations.

### 3. Video testing management platform

In this section we first describe the architectural solutions that were developed to execute video experiments. Then, we elaborate further on the design of the test management software.

#### 3.1. Video testbed

To execute video experiments we have created a general video testbed as shown in Figure 2. Extra machines can be easily plugged in.

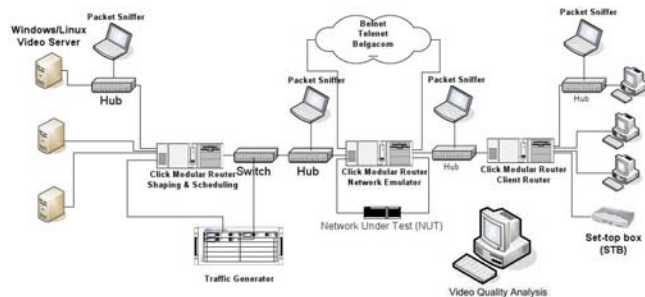


Figure 2. The video test bed

This testbed consists of a number of fixed machines which can be classified based on their specific function.

**streamers** The streamer streams the video across the network. Our testbed supports open source streamers (Darwin Streaming Server and VLC), commercial streamers (Windows Media Server) and our in-house developed streamer (xstreamer which will be described in future papers).

**clients** The client receives a video stream and possibly displays it on screen and/or captures it to disk. We can receive the video with different players: Windows Media Player, Quicktime, VLC, in-house developed receiver, ... Next to streaming the video to regular desktop computers, the testbed also supports streaming video to set-top boxes which in turn are connected to TV sets.

**Click Modular Routers** The Click Modular routers in the testbed are servers which emulate the network. Each of them runs the Click Modular Router software package [3]. Through this software (enhanced with our own models) we can drop packets, introduce jitter and packet delay. Packets can be dropped according to specific models (random, burst, Gilbert, ...). The router near the streamers can act as a traffic shaper, the one near the clients emulates the home gateway and the middle router emulates the core network. We can also choose to stream over a real-life network of an ISP or over a real wireless network.

**sniffers** At specific demarcation points in our testbed we have placed network sniffers. These capture the video streamed over the network. After testing, sniffer files can be used to extract network and failure statistics.

**traffic generator** Additional network traffic can be generated by a dedicated traffic generator like the Spirent SmartBits <sup>1</sup>.

**video quality analyser** One high performant machine will do all the encoding, decoding and quality analysis.

**controller** All the machines in the network are controlled by one machine running the testbed management software. This machine controls the other machines through a separate control network.

This testbed is used for general video quality experiments or for evaluating third party hardware/software. It is clear that this testbed lacks scalability due to its single instantiation, flexibility and complexity. Therefore an adapted solution will be explained in the next section.

#### 3.2. Large scale video testing using the iLab.t Virtual Wall

To perform realistic scalability tests of networks, servers and networked applications the iLab.t Virtual Wall testbed<sup>2</sup> (see Figure 3) was built.

In a first phase we have built a testbed with 100 server machines (see Table 1), with 4 or 6 gigabit Ethernet ports

<sup>1</sup><http://www.spirent.com/>

<sup>2</sup><http://ilabt.ibbt.be/>



**Figure 3. The iLab.t Virtual Wall consists of 100 identical server machines, 20 of them are connected to a monitor**

each connected to a huge nonblocking switch (Force 10 networks E1200) which can be configured to make the interconnections between the servers dynamically and point to point. Twenty of our server machines are also connected to a display, so we can visualize the results of video streaming tests.

CPU	2 x dual core 2.0Ghz
Memory	4GB RAM
Storage	4 SATA disks in RAID 0 configuration

**Table 1. Virtual Wall node: hardware specifications**

By using the Emulab software [8] of the University of Utah<sup>3</sup> it is possible to configure specific network setups and dynamically load operating system images on the allocated servers through a web interface.

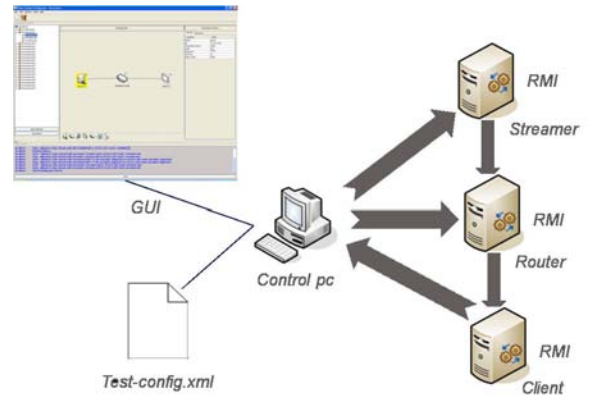
With this testbed we can organize our scalable video quality experiments. With the 100 machines we have available, we can create 33 basic streaming setups (see Figure 1) controlled by 1 machine. This way 33 streaming steps can run concurrently. The encoding, decoding and analysing processes can be distributed amongst the 99 machines. And with the 4 cores per machine, even 396 of these processes can run in parallel.

To achieve this we first created a video node disk image. This disk image contains an operating system, all the necessary software tools like streamers, clients, encoders, decoders and quality analysers for performing each individ-

<sup>3</sup><http://www.emulab.net>

ual step in the video quality test process. Within the emulab software we then created an experiment setup with a parameter containing the number of basic streaming setups we want to allocate. When we want to execute a series of video quality tests we just call this setup script with the number of desired setups and allocate them through the emulab software. The emulab software then allocates the server machines, loads the disk image on it and creates the necessary network connections between them. The next step is then the configuration and execution of the different tests. This is the task of our management software and will be discussed in the next section.

### 3.3. Management software



**Figure 4. Management software**

To configure, manage and execute video experiments we have developed a management application in Java (see Figure 4). The central application runs on the control machine and is in charge of the test configuration, control of the distributed processes, distributing and collecting the data files and scheduling the different test steps. On each machine in the test setup a Java service is running and the central application communicates with them through Java Remote Method Invocation (RMI).

Java was chosen for its platform independency, this way we can easily plug in additional devices like set-top boxes or even mobile phones. In the following paragraphs the functionality of the management software will be described.

#### 3.3.1 Configuration

Since repeatability was one of the requirements for our testing platform (Section 2.2), we should be able to reload an executed experiment or reconstruct it. To achieve this, we store our experiment configuration data into xml files. Actually, this is a serialization of our data model into xml. There are two ways to create an experiment configuration: through

a graphical user interface or through the scripting engine. With the interface as shown in Figure 4, one can create a configuration and start up the experiment. During execution the user receives visual feedback about the actions the software takes and the possible errors that occur.

One can imagine that creating several 100 different test cases in this interface is quite a hassle. To facilitate this we have created a script toolbox in JRuby<sup>4</sup>. Within the script we can build up our experiment data model over the ranges of a set of input parameters. We can then serialize the data model to xml, load it in the management software and run the experiment. We can also directly start the experiment from within the script.

### 3.3.2 Application Management

Configuring, starting and stopping the applications on the different devices in the test setup is controlled by the central management application. Configuration of an application is usually done by uploading a configuration file, that is stored locally on the control pc and possibly generated through scripting, to the device. Some applications need to be started in a predefined order, because for example a client cannot request a video stream from a server that is not running, this order is stored in the configuration. The Java service will start the application when it receives an execute request with a process object as argument. This process object contains the command to execute and the arguments to that command. The call to the service is made by means of RMI. The service keeps track of every process that is started. Some applications will terminate automatically, e.g. encoding or decoding, whereas others will run until they are stopped after a predefined time, e.g. servers. This is done by calling the stop method on the service, again with the process object as parameter.

The following code listings present the service and the process objects interface.

```
public interface RemoteRuntime extends Remote {

 Object execute(VTProcess process);
 void stopProcess(VTProcess process);
 Object registerLogServer(InetAddress ip, int port);
 Object unregisterLogServer(InetAddress ip, int port);
}

public interface VTProcess
 extends Serializable, Comparable<VTProcess> {

 public void run();
 public void stop();
 public String getName();
 public void setName(String name);
 public void setWaitFor(boolean b);
 public boolean isWaitFor();
 public String getDescription();
 public void setDescription(String description);
 public List<CopyFileInfo> getFilesToUploadInfo();
}
```

<sup>4</sup><http://jruby.codehaus.org/>

```
public Set<String> getFilesToDownload();
public boolean endsAutomatic();
public VTProcess copy();
public String getCommandToExecute();
}
```

### 3.3.3 File Distribution

The central management software distributes and collects files to and from the devices in the testbed. By default this is done through FTP, because this is the fastest transfer solution. For this, each device will need a ftp server. If it is not possible to install one on a device, we can choose to use sftp (secure ftp), which is a lot slower than ftp, or we can setup sockets in the central application and the services and transfer the files through them.

### 3.3.4 Test Scheduling

The central management application is responsible for scheduling the different test steps in an experiment: encoding, streaming and quality analysis. Steps belonging to the same category can be executed concurrently. But streaming will only start when all encoding processes are finished and quality analysis cannot start before all streaming processes are finished.

This scheduling is defined at configuration time. When configuring an experiment, each step is assigned to a machine or in case of the streaming step assigned to a specific setup. Steps that can run concurrently are then put into batches, these batches will be executed sequentially. We illustrate this with an example. Suppose we have one basic streaming setup (Figure 1) allocated, so we have three testbed servers available and one control machine. We need to encode 1 video sequence 2 times and stream the resulting files each 2 times. The streamed files will each be analysed. So we will have 2 encoding steps that can run in parallel, 4 streaming steps that will run sequentially and 4 analysis steps that will be divided into a batch of 3 and a batch of 1 (we only have three servers in use). Figure 5 shows this test configuration as it appears in the interface.

When a test step is scheduled and executed, the application performs 4 consecutive tasks:

1. Upload input files and configuration files to the devices
2. Start up the processes
3. Stop the processes
4. Download all captured and resulting files



**Figure 5. Experiment configuration: batches are executed sequentially, elements in batches are run in parallel**

## 4. Performance evaluation

### 4.1. Theoretical approach

The processing time of a single test case, that consists of one video sequence that must only be encoded and streamed once, can be calculated using equation (1), where  $t_{enc}$  represents the time needed to encode the video sequence,  $t_{str}$  represents the time needed to stream the encoded sequence and  $t_{an}$  represents the time needed for analysing the received video stream.

$$t_{single} = t_{enc} + t_{str} + t_{an} \quad (1)$$

In case the video sequence in a single test case must be encoded  $n$  different times and streamed using  $m$  different network settings, the total processing time can be calculated using equation (2).

$$t_{multicase} = n \times t_{enc} + n \times m \times t_{str} + n \times m \times t_{an} \quad (2)$$

Furthermore, if the test case can be run in parallel on  $x$  different testbed setups, each consisting of  $\alpha$  machines, the total processing time becomes:

$$t_{process} = \frac{n \times t_{enc}}{\alpha \times x} + \frac{n \times m \times t_{str}}{x} + \frac{n \times m \times t_{an}}{\alpha \times x} \quad (3)$$

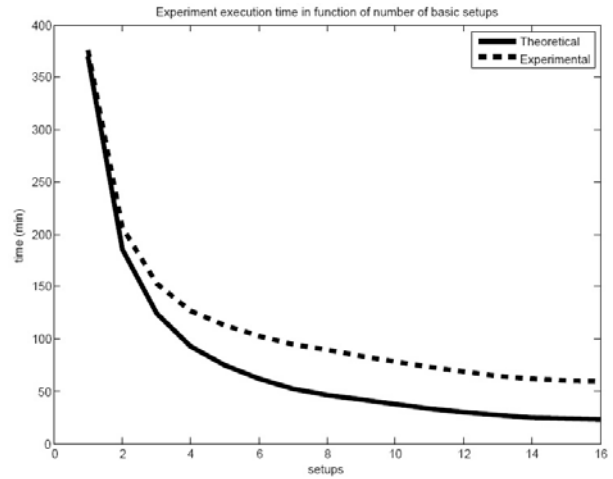
Besides processing there is also some time spent in downloading and uploading files:

$$t_{ftp} = t_{upload} + t_{download} \quad (4)$$

Where  $t_{upload}$  is the time to distribute files like raw video files, encoded video files and application configuration files to the servers. The  $t_{download}$  is the time to collect result files like encoded video files and streamed video files to the control machine. The  $t_{ftp}$  is fixed and depends on the number and size of the files to be distributed and collected. Increasing the number of setups will not diminish the time spent downloading and uploading files.

The total execution time for an experiment can thus be computed by combining equations (3) and (4):

$$t_{exp} = t_{ftp} + t_{process} \quad (5)$$



**Figure 6. Experiment execution time in function of the number of setups**

Figure 6 shows the experiment execution time in function of the number of setups used. The dashed line represents the real execution times (See 4.2). The full line represents the theoretically determined values by using equation (5) for the same experiment.

The graph shows that there is a diminishing performance gain with increasing number of setups. When starting from one setup, the performance can be doubled by doubling the number of setups. To have the same relative performance gain the number of setups need to be increased by 4,8,16 and so on. The difference between the theoretical and real durations will be explained in Section 4.2.

The optimal time  $t_{opt}$  is achieved when  $x = \frac{m \times n}{\alpha}$ . At this point the maximum parallelisation of individual test steps is achieved: all encoding steps are run concurrently, then all streaming steps and finally every quality evaluation step is parallelized. The time  $t_{opt}$  is then the total ftp time ( $t_{ftp}$ ) and the sum of the maximum durations of the individual test steps: encoding, streaming and quality evaluation. Increasing the number of setups any further beyond this point would not result in further time gains.



# setups	encoding	streaming	analysing	ftp	total	experiment time	overhead
1	1.21	40.47	318.90	12.58	373.15	376.08	2.93
2	0.61	20.27	174.76	8.39	204.02	206.49	2.47
4	0.33	10.43	105.28	7.75	123.78	126.53	2.75
8	0.33	5.45	72.48	7.90	86.17	89.28	3.11
16	0.33	2.79	45.26	7.66	56.05	59.37	3.33

**Table 2. Profiling data: durations (in minutes) of individual steps in an experiment**

## 4.2. Experiment monitoring

To back up our theoretical model with some real data a video quality evaluation experiment was executed on a different number of setups and the duration of each single test step was measured. One video sequence, with a duration of 20 seconds, was used in the experiment. The sequence had a resolution of 672x384 and its filesize was 184MB. This sequence was encoded using ffmpeg<sup>5</sup> with following parameters: `-vcodec libx264 -b #value`. With `#value` having following bitrate values : [100k, 300k, 500k, 700k, 900k, 1100k, 1300k, 1500k, 1700k, 1900k]. The resulting files were streamed over a network containing random packet loss in % [0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5]. This way, the experiment consisted of 100 different test cases. The experiment was executed on one, two, four, eight and 16 setups. Each setup is configured as a basic streaming setup as shown in Figure 1.

The measurements are shown in Table 2, each duration is presented in minutes. The execution of the test steps was profiled within the management application to determine the duration of each step. The "total" column is the sum of encoding, streaming, analysing and ftp durations. The "experiment time" is the duration of the execution of the experiment within our application. The difference between these two is the overhead that is created by our management software to schedule the different steps and to perform its management job. We notice that this overhead only slightly increases with the number of setups used.

When the number of setups reaches four (12 machines), all ten encoding steps are run concurrently. At this point we reach the lowest achievable encoding time, increasing the number of setups will not result in a lower encoding time as is shown in the table. Notice that the encoding and the streaming time basically halves when doubling the number of setups, but the analysis duration doesn't. This is because the duration of an analysis step strongly depends on the artefacts in the video sequence and when a batch of concurrent processes is started we have to wait until all processes are finished before a new batch can be started. So the execution time of a batch will be the duration of the longest analysis step within that batch. It could be that within a batch of 10

<sup>5</sup><http://ffmpeg.mplayerhq.hu/>

analysis steps, nine have finished and still need to wait for a few minutes for the last one to finish. It is clear that there is a need for a more dynamic scheduling to more optimally use the available machines in the experiment.

On the contrary to what we expected the ftp time is not constant. It is the highest for one setup, then drops to about 8 minutes for two setups and remains roughly the same when further increasing the number of setups. Transferring files across a network should take up the same amount of time whether they are sent sequential or concurrent. One of the possible causes is that the ftp connection setup consumes some time and when it is run in parallel this time is averaged out. We will need some deeper monitoring to really determine the cause of this.

In this experiment only one video sequence was used as input, if we would use more than one video sequence this will have a few consequences. First of all the number of test cases will lineary increase with the number of input sequences, as a consequence the total experiment duration will also increase lineary. Secondly the durations  $t_{enc}$ ,  $t_{str}$  and  $t_{an}$  will be different for each sequence. So the optimal time we can achieve with our current scheduling will be strongly dependent on the maximum values of these durations.

## 5. Conclusion and future work

Performing video quality evaluation experiments provides challenges to an automated testing platform like an extensive input parameter space and long processing times. We described the solutions we've developed to provide a flexible and scalable platform that allows us to easily repeat different video quality evaluation experiments. Management software was developed to automate test execution. This software handles configuration, distributed application management and file management in a generic and platform independent way. The iLab.t Virtual Wall allows us to drastically reduce the experiment execution time by creating a number of similar network setups and running the different steps in the quality evaluation process concurrently. This performance improvement was proven both theoretically and practically by means of experiment monitoring.

Until now, the Virtual Wall was used to create basic

streaming setups, like the one in Figure 1. In future we want to create and test more complex setups. We could for example create a network with several servers connected through one router to several clients. Some of these servers will act as streaming servers, others as file servers or web servers. Some of the clients can be configured to display the video, others will download files or send requests to the web servers. This way we will be able to monitor the behaviour of a video stream in a more realistic network environment.

With the static scheduling mechanism currently implemented, we do not reach an optimal usage of the resources within the testbed. This is shown by the experiment monitoring, e.g. the quality evaluation duration does not halve when doubling the number of setups in the testbed. In future we will look at implementing a more dynamic scheduling mechanism. This scheduling should happen at run time and keep track of test step dependencies and resource (setup or machine) usage. This way further performance gains can be made.

As the analysis step is the most time consuming, we could look at breaking this process down into different sub-processes so that we can parallelize them.

By now, we reach an acceleration factor of 6 (376 vs 59 minutes) by using 48 servers instead of 3.

## 6. Acknowledgements

The research activities that have been described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT) and the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT). Nicolas Staelens would like to thank the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT) for financial support through his Ph.D. grant. Part of this work has been funded by the IBBT GBO project Video Q-SAC.

## References

- [1] ITU-T Rec. J.144. Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference. Technical report, International Telecommunication Union, March 2001.
- [2] ITU-T Rec. P.10/G.100. Vocabulary for performance and quality of service. Technical report, International Telecommunication Union, 2007.
- [3] E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. F. Kaashoek. The click modular router. *ACM Transactions on Computer Systems*, 18(3):263–297, August 2000.
- [4] A. Reibman, S. Kanumuri, V. Vaishampayan, and P. Cosman. Visibility of individual packet losses in mpeg-2 video. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 1, pages 171–174, October 2004.
- [5] M. Ries, O. Nemethova, and M. Rupp. Video quality estimation for mobile h.264/avc video streaming. *Journal of Communications*, 3(1):41–50, January 2008.
- [6] O. Verscheure, P. Frossard, and M. Hamdi. Mpeg-2 video services over packet networks: Joint effect of encoding rate and data loss on user-oriented qos. *NOSSDAV*, 98:257–264, July 1998.
- [7] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13:600–612, April 2004.
- [8] B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar. An integrated experimental environment for distributed systems and networks. In *Proc. of the Fifth Symposium on Operating Systems Design and Implementation*, pages 255–270, Boston, MA, Dec. 2002. USENIX Association.

# **DMAI'08**

## **IEEE International Workshop On Data Mining and Artificial Intelligence**

24 December, 2008  
Khulna University of Engineering and Technology (KUET), Bangladesh

## **WORKSHOP PROCEEDINGS**



**Organized by**

Faculty of Electrical and Electronic Engineering (EEE)  
**Khulna University of Engineering and Technology (KUET)**

Khulna-9203, Bangladesh

<http://www.kuet.ac.bd/>



# **DMAI'08**

## **IEEE International Workshop On Data Mining and Artificial Intelligence**

### **Editors**

Prof. Dr. M.M.A Hashem  
Dr.M.M. Hafizur Rahman  
Pintu Chandra Shill  
Rushdi Shams  
Md. Jahidul Islam

### **Published by**

Faculty of Electrical and Electronic Engineering (EEE)  
Khulna University of Engineering and Technology (KUET)  
Khulna-9203, Bangladesh

### **Additional Copies of the Proceedings can be obtained from**

ICCIT 2008 Secretariate  
Department of Computer Science & Engineering (CSE)  
Khulna University of Engineering and Technology (KUET)  
Khulna-9203, Bangladesh

**IEEE Catalog Number: CFP0817D-PRT**

**ISBN: 978-1-4244-2135-0**

**Library of Congress: 2008900703**

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Operations Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved. Copyright ©2008 by the Institute of Electrical and Electronics Engineers.

# A New Approach of Modified Transaction Reduction Algorithm for Mining Frequent Itemset

<sup>1</sup>Ramaraj Eswara Thevar

<sup>2</sup>Rameshkumar Krishnamoorthy

<sup>1</sup>Director, Computer Centre, Alagappa University, Karaikudi Email: [dr\\_ramaraj@yahoo.co.in](mailto:dr_ramaraj@yahoo.co.in)

<sup>2</sup>Ph.D (CSE), Department of CSE, Alagappa University, Karaikudi, Email: [rameshkumar\\_phd@yahoo.co.in](mailto:rameshkumar_phd@yahoo.co.in)

## Abstract

*Association rule mining is to extract the interesting correlation and relation between the large volumes of transactions. This process is divided into two sub problem: first problem is to find the frequent itemsets from the transaction and second problem is to construct the rule from the mined frequent itemset. Frequent itemsets generation is the requirement and most time vast process for association rule mining. Nowadays, most efficient Apriori-like algorithms rely heavily on the minimum support constraints to prune the vast amount of non-candidate itemsets. These algorithms store many unwanted itemsets and transactions. In this paper propose a novel frequency itemsets generation algorithm called MTR-FMA (Modified Transaction Reduction based Frequent Itemset Mining algorithm) that maintains its performance even at relative low supports. The experimental reports also show that proposed MTR-FMA algorithm on an outset is faster than High Efficient AprioriTid and other some algorithms.*

## Key words

Association rule mining, Frequent itemsets, Modified Transaction Reduction, MTR-FMA

## 1. Introduction

Data Mining refers to extracting or mining knowledge from large amounts of data. In 1993, Agarwal, Imielinski, and Swami introduced a class of regularities, association rules and gave an algorithm for finding such rules [3]. The typical example for association rule mining is Market Basket analysis. The Bar code technology has made it possible for retail organizations to collect and store massive amounts of sales data. A record in it typically consists of itemsets bought in each transaction in each row. Finding association rules for such basket data will improve the

sales and cross marketing. Other applications include store layout, add on sales and customer segmentation.

In general, the work of association rule can be decomposed into two phases:

(1) Frequent itemsets generation - find out all itemsets that sufficiently exceed the minsup, and (2) Rules construction - from the frequent itemsets, generate all association rules having confidence higher than the minconf. Since the second phase is straightforward and less expensive, This paper concentrate only on the first phase for finding all frequent itemsets.

This paper has proposed a new algorithm called MTR-FMA (Modified Transaction Reduction based Frequent Itemset Mining algorithm) to efficiently mine frequent itemsets. This paper has also proved that the proposed algorithm MTR-FMA has an edge over existing algorithms like Apriori, AprioriTID, HEA (High Efficient AprioriTid) and FP-Growth in terms of transaction reduction and speed. The rest of the paper is organized as follows. A review of the previous works regarding AprioriTid and HEA are given in the next section. Then this paper will describe the proposed theorem with proof and the proposed algorithm called MTR-FMA for finding frequent item sets. Empirical evaluations and results of the proposed algorithm (MTR-FMA) over other algorithms using FIMI's dataset (retail.dat) and other synthetic data sets are described in the next section. Conclusions are stated in the last section.

## 2. Previous Works

### 2.1. Apriori Algorithm

The *Apriori* [3-6] algorithm is a level wise search algorithm for mining frequent itemsets for boolean association rules. The large itemsets are computed through iterations. In each iteration, the database is scanned once and all large itemsets of same size are computed. The large itemsets are computed in the

ascending order of their sizes. In the first iteration, the size 1- large itemsets are computed by scanning the database once. Subsequently, in the  $k$ th iteration ( $k > 1$ ), a set of candidate sets  $C_k$  is created by applying the candidate set generating function Apriori-gen on  $L_{k-1}$ , where  $L_{k-1}$  is the set of all large  $(k-1)$  - itemsets found in the iteration  $k-1$ . Apriori-gen generates only those  $k$ -itemsets whose every  $(k-1)$ -itemsets subset is in  $L_{k-1}$ . The support counts of the candidate itemsets in  $C_k$  are then computed by scanning the database once and the size- $k$  large itemsets are extracted from the candidates. The drawback of Apriori is that when the cardinality of the longest frequent itemsets is  $k$ , Apriori needs  $k$  passes of database scans. In addition, this algorithm is computation intensive in generating the candidate itemsets and computing the support values, especially for application with very low support threshold and / or vast amount of items. In this algorithm, if the number of first itemsets element is  $k$ , DB will be scanned  $k$  times at least. So it will have low efficiency. It is a base point to reduce the number of itemset for improving the Apriori algorithm.

## 2.2. AprioriTID Algorithm

The *AprioriTID* [2] is another variant of Apriori which reduces the time needed for the frequency counting procedure by replacing every transaction in the database by the set of candidate sets that occurs in that transaction. This is done repeatedly at every iteration  $k$ . The adapted transaction database is denoted by  $\bar{C}_k$ . Although the AprioriTID algorithm is much faster in later iterations, it performs much slower than Apriori in early iterations. This is mainly due to the additional overhead that is created when  $\bar{C}_k$  does not fit into main memory and has to be written to disk [7-10]. If a transaction does not contain any candidate  $k$ -sets, then  $\bar{C}_k$  will not have an entry for this transaction. Hence, the number of entries in  $\bar{C}_k$  may be smaller than the number of transactions in the database, especially at later iterations of the algorithm. The AprioriTid algorithm also uses the Apriori-gen function to determine the candidate itemsets before the pass begins.

The transaction in  $\bar{C}_k$  will be of the form  $\langle TID, \{X_k\} \rangle$ , where each  $X_k$  is a potentially large  $k$ -itemset present in the transaction with identifier TID. The member of  $\bar{C}_k$  corresponding to the transaction  $t$  is defined as

$$\bar{C}_k = \langle t.TID, \{c \in C_k \mid c \text{ contained } t\} \rangle$$

The drawbacks of AprioriTID are that the modified data structure can be much larger than the initial database and only faster in the later stage of the passes.

## 2.3. High Efficient AprioriTid Algorithm

*HEA* [1] is the improved algorithm of AprioriTID.

Here, the member of  $\bar{C}_k$  corresponding to the transaction  $t$  is defined as

$$\bar{C}_k = \langle t.TID, \{c \in C_k \mid c \text{ contained } t, c.support \geq minsup\} \rangle$$

This considerably decreases the size of stored data in  $\bar{C}_k$  [1]. Moreover, searching data scale is reduced when computing support of itemsets in  $C_k$ . At the same time, it reduces much time of I/O and running. The process is described as follows: (a). Confirming itemset  $c$  in  $C_k$ , then the transaction set  $T_c$  presented with TID including the item of  $c$ , is computed. (b). The number of entries computed in  $T_c$ , defined as the  $|T_c|$ , which is support of itemset  $c$ . (c). If  $|T_c| \geq minsup$ ,  $c$  is included into  $L_k$  and  $\bar{C}_k$ , otherwise deleting  $c$ . Through the above process, with the computed support in  $C_k$ , void itemset can be directly deleted from  $\bar{C}_k$  or added to  $\bar{C}_k$  and  $L_k$ . After analyzing this algorithm, this paper found that HEA can be further improved by reducing unwanted transactions and itemsets which is explained in next section through the proposed new algorithm MTR-FMA.

## 3. Proposed MTR-FMA Algorithm

MTR-FMA-Modified Transaction Reduction based Frequent Itemset Mining algorithm in Figure 1 is an improvement algorithm of AprioriTid and HEA whose idea is based on the theorem described in next sub section. Here relationship of transaction  $t$  with the entry in  $\bar{C}_k$  is defined as

$$\bar{C}_k = \langle t.TID, \{c \in C_k \mid c \text{ contained } t, c.support \geq minsup, T_{items} > k, t.no-of \text{ combinations} > 1\} \rangle$$

Each member of the set  $\bar{C}_k$  will be of the form  $\langle TID, \{X_k\} \rangle$  where each  $X_k$  is a potentially large  $k$ -itemset present in the transaction with identifier TID.

### 3.1. Theorem and Proof

**Theorem:** If  $c \in C_{k-1}$  and  $c.support < minsup, T_{items} \leq k, m = 1$ , then  $c$  is useless in  $\bar{C}_{k-1}$  where  $T_{items}$  is total itemcount in each transaction and  $m$  is the no. of combinations.

**Proof: 1. [For  $T_{\text{items}} \leq k$ ]**

From the Apriori-gen algorithm, it is known that  $C_k$  is formed from  $L_{k-1}$ . Consider a transaction  $T_i = \{I_1, I_2, I_3, \dots, I_m\}$ . The candidate generation would lineup from  $C_1 = \langle \{I_1\}, \{I_2\}, \{I_3\}, \dots, \{I_m\} \rangle$ ;  $C_2 = \langle \{I_1, I_2\}, \{I_1, I_3\}, \dots, \{I_1, I_m\}, \{I_2, I_3\}, \dots, \{I_{m-1}, I_m\} \rangle$ ;

$C_3 = \langle \{I_1, I_2, I_3\}, \{I_1, I_2, I_4\}, \dots, \{I_2, I_3, I_4\}, \dots, \{I_{m-2}, I_{m-1}, I_m\} \rangle$ ;

and it will go up to  $C_m = \langle \{I_1, I_2, I_3, \dots, I_m\} \rangle$ . So  $\{I_1, I_2, I_3, \dots, I_m\}$  will not appear in  $C_{m+1}$ . That is to say that this algorithm need not compute the  $C_{m+1}$  in  $T_i$ .

Hence,  $\forall T_i \in \bar{C}_{m+1}$ ,  $T_i$  need not be taken into account. So,  $T_i$  can be rejected in  $\bar{C}_{m+1}$ . Hence the proof.

**2. [For  $m = 1$ , where  $m$  is no. of combinations]**

Now consider the same transaction  $T_i = \{I_1, I_2, I_3, \dots, I_m\}$ . In general the candidate generation would have  $\langle \{I_1\}, \{I_2\}, \dots, \{I_1, I_2\}, \dots, \{I_1, I_m\}, \dots, \{I_2, I_m\}, \dots, \{I_1, I_2, \dots, I_m\} \rangle$

If all the itemsets in  $C_k$  except one is rejected during  $L_{m-1}$  due to above hypothesis (say  $A_b$  remains) then  $\bar{C}_{k+1}$  will contain only  $\langle \{A_b\} \rangle$ . So  $\langle \{A_b\} \rangle$  alone will not able to form a candidate in  $C_{k+1}$ . Hence in Such  $\forall T_i \in \bar{C}_{k+1}$ ,  $T_i$  can be rejected in  $\bar{C}_{k+1}$ . Hence the proof.

**Figure 1. MTR-FMA Algorithm**

*INPUT* :  $D, \text{minsup}$   
*OUTPUT*:  $L(D, \text{minsup})$

- 1)  $L_1 = \{\text{large 1-itemsets}\}$ ;
- 2)  $\bar{C}_1 = \text{Database } D$ ; (With all items not in  $L_1$  and  $\forall T_i \ T_{\text{items}} = 1$  removed )
- 3) For ( $k=2$ ;  $\bar{C}_{k-1} \neq \Phi$ ;  $k++$ ) do begin
- 4)  $C_k = \text{Apriori-gen}(L_{k-1})$  // New candidates
- 5)  $\bar{C}_k = \Phi$ ;
- 6) For all  $c \in C_k$  do begin
- 7)  $\bar{C} = \Phi$ ;
- 8)  $T_c = \{t.TID \mid t \in \bar{C}_{k-1}, (c - c[k]) \in t.\text{set-of-itemsets} \wedge (c - c[k]) \in t.\text{set-of-itemsets}\}$
- 9) If  $|T_c| \geq \text{minsup}$  then begin
- 10)  $L_k = \bigcup_k \{c\}$ ;
- 11) For all  $p \in T_c$  do

- 12) If ( $T_{\text{items}} > k$ ) then begin
- 13)  $\bar{C} = \bigcup_k < p, c >$ ;
- 14) end
- 15) end
- 16) If  $|\bar{C}| \neq 1$  then begin
- 17)  $\bar{C}_k = \bar{C}$ ;
- 18) end
- 19) end
- 20) end
- 21) end
- 22)  $\text{Ans} = \bigcup_k L_k$ ;

**3.2. MTR-FMA Algorithm Details**

Let consider the following example in Table 1 to show how MTR-FMA works. The algorithm starts by scanning the database to generate 1- frequent itemset  $L_1$ .  $\bar{C}_1$  is generated with itemsets that are present in  $L_1$ . For  $k = 1$ ,  $\bar{C}_1$  corresponds to the database  $D$  although conceptually each item  $I$  is replaced by the itemset  $\{i\}$ . For  $k > 1$ ,  $C_k$  is generated by the apriori-gen algorithm. Confirming itemset  $c$  in  $C_k$ , then the transaction set  $T_c$  presented with TID, including itemset  $c$  is computed. The number of entries is computed in  $T_c$  defined as  $|T_c|$  which is support of itemset  $c$ . If  $|T_c| \geq \text{minsup}$ ,  $c$  is included into  $L_k$  and  $\bar{C}_k$ , otherwise deleting  $c$ .

If the total itemcount of a transaction  $T_{\text{items}} \leq k$ , then that transaction  $t$  is deleted from  $\bar{C}_{k-1}$ . Also, if  $|\bar{C}_k| = 1$ , then it is also not included in  $\bar{C}_k$  and that transaction is deleted from  $\bar{C}_{k-1}$ . Thus the entries in  $\bar{C}_k$  may be smaller than the transaction in database especially for large values of  $k$ . The corresponding  $L_k$  and  $\bar{C}_k$  are shown in Tables 2, 3, 4, 5, 6 and 7.

**Table 1. Sample Database**

<i>Tid</i>	<i>Items</i>
T1	a, e, f, h
T2	b, d, h
T3	d, e, g
T4	b, c
T5	e, f, g
T6	b, c, d
T7	b, f, g, i
T8	e

T9	h
T10	c, e, g
T11	c, e, g
T12	e, f, h
T13	b, d, f, g
T14	a, c, e, g
T15	b, c, i

**Table 2. L1**

Item set	Frequency
b	6
c	6
d	4
e	8
f	5
g	7
h	4

**Table 3.  $\bar{C}_1$**

Tid	Items
T1	{e},{f},{h}
T2	{b},{d},{h}
T3	{d},{e},{g}
T4	{b},{c}
T5	{e},{f},{g}
T6	{b},{c},{d}
T7	{b},{f},{g}
T10	{c},{e},{g}
T11	{c},{e},{g}
T12	{e},{f},{h}
T13	{b},{d},{f},{g}
T14	{c},{e},{g}
T15	{b},{c}

**Table 4. L2**

Item set	Frequency
{b, c}	3
{b, d}	3
{c, e}	3
{c, g}	3
{e, f}	3
{e, g}	5
{f, g}	3

**Table 5.  $\bar{C}_2$**

TID	Items
T5	{e, g},{e, f},{f, g}
T6	{b, c},{b, d}
T10	{c, e},{c, g},{e, g}
T11	{c, e},{c, g},{e, g}
T13	{b, d},{f, g}
T14	{c, e},{c, g},{e, g}

**Table 6. L3**

Item set	Frequency
{c, e, g}	3

**Table 7.  $\bar{C}_3$**

TID	Items
-	-

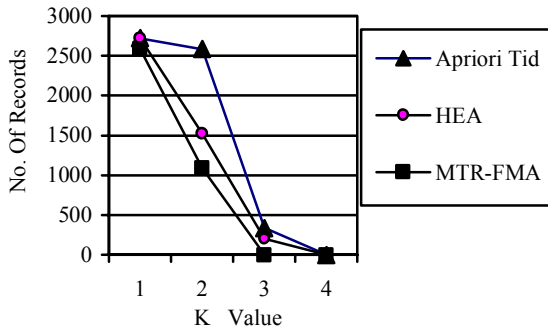
#### 4. Result and Discussion

The proposed algorithm have tested various datasets like FIMI data set [16] and synthetic dataset and compared them with Apriori, and FP growth whose implementations were used from [17] and [18]. Since no public implementation of AprioriTID and HEA is available this paper used the own implementations to compare with MTR-FMA algorithm. Using a sample Database of 1020 transactions with 9 items this paper tested MTR-FMA with AprioriTID and HEA. With the minimum support as 20%, the results are as shown in the Table 8 [11 – 15] and figure 2.

**Table 8. Contrast of Record numbers**

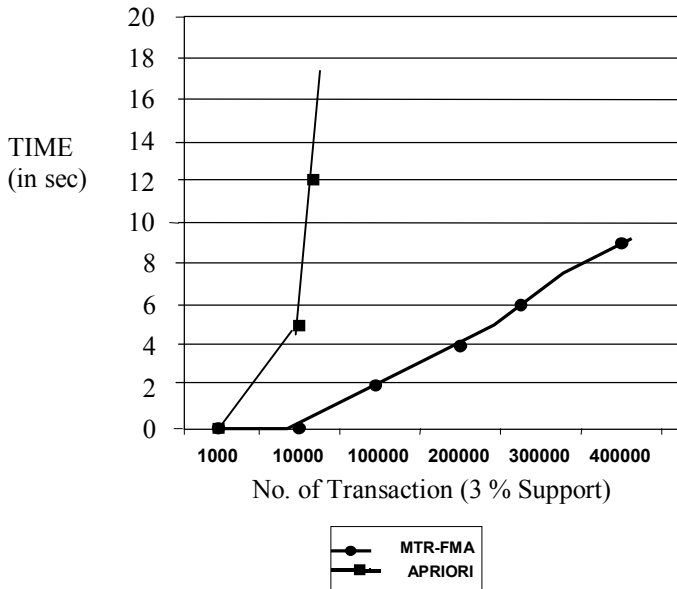
Records	AprioriTID	HEA	MTR-FMA
$\bar{C}_1$	2720	2720	2584
$\bar{C}_2$	2584	1524	1088
$\bar{C}_3$	340	204	0
$\bar{C}_4$	0	0	0

**Figure 2. Comparison of Apriori , HEA and MTR-FMA**



From the Table 8 it is observed that number of records is averagely reduced to 35 % when compared to AprioriTID and 18 % when compared to HEA. Hence the execution will also be considerably reduced. The graph in the Figure 2 is plotted with the help of the Table 8.

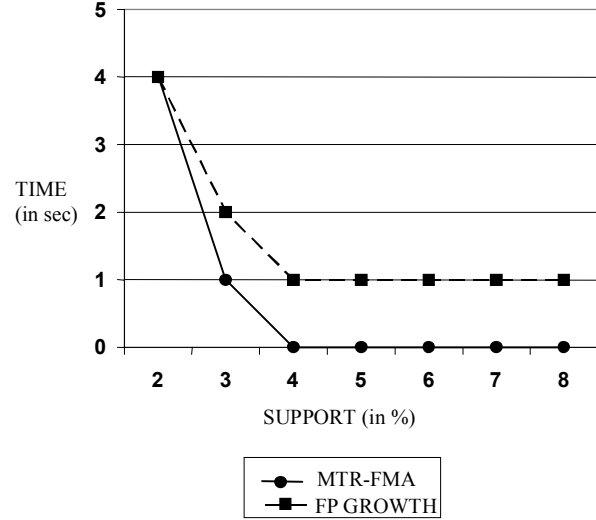
**Figure 3.Execution times of Apriori and MTR-FMA**



This paper also compared proposed MTR-FMA algorithm with Apriori using the synthetic dataset whose transaction varies from 1000 to 500000 and with a constant minimum support of 3% in Figure 3.

This paper used the dataset retail.dat from FIMI site [16] with 17822 transactions and 9999 items. The experiment is performed on HP dx 2280 with processor clock cycle 2.0 GHz, 512 MB of main memory. The Figure 4 shows the result up to the minimum support of 8%.

**Figure 4.Execution Times of FP- Growth and MTR-FMA on retail.dat**



## 5. Conclusion

This paper proposed a novel data structure, an MTR-FMA algorithm to mine frequent itemsets. The advantages of MTR-FMA over existing algorithms include (1) Interactive mining with different supports (2) Faster execution time (3) Infrequently used item are not stored and hence improves the size of the query data. This paper has implemented MTR-FMA algorithm and compared proposed approach with other known algorithms. From the experimental results, MTR-FMA is shown to be more efficient and scalable to large amount of transactions and out perform other algorithms.

## 6. References

- [1] Zhi-Chao Li, Pi-Lian Hi and Ming Lei, "A High Efficient AprioriTID Algorithm for mining Association rule", Proceedings of 4<sup>th</sup> International Conference on machine learning and cybernetics, 18-21 AUG 2005.

- [2] He Li-jian, Chen Li-chao, Liu shuang-ying, "Improvement of AprioriTid Algorithm for Mining Association Rules", Journal of Yantai University, Vol.16, No.4, 2003.
- [3] R. Agrawal, J. Shafer, "Parallel mining of association rules", IEEE Transactions on Knowledge and Data Engineering, 8(6), December 1996.
- [4] Agrawal, T. Imielinski, A. Swami, "Mining association rules between sets of items in large databases", Proceedings of the ACM SIGMOD Conference on Management of Data, 1993, pp. 207-216.
- [5] R.C. Agarwal, C.C. Aggarwal, and V.V.V. Prasad. Depth first generation of long patterns. In Ramakrishnan et al. [32], pp. 108-118.
- [6] R. Agrawal, T. Imielinski, and A.N. Swami. Mining association rules between sets of items in large databases. In P. Buneman and S. Jajodia, editors, Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, volume 22(2) of SIGMOD Record, ACM Press, 1993, pp. 207-216.
- [7] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A.I. Verkamo. Fast discovery of association rules. In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, editors, Advances in Knowledge Discovery and Data Mining, MIT Press, 1996, pp. 307-328.
- [8] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In J.B. Bocca, M. Jarke, and C. Zaniolo, editors, Proceedings 20th International Conference on Very Large Data Bases, Morgan Kaufmann, 1994, pp. 487-499.
- [9] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. IBM Research Report RJ9839, IBM Almaden Research Center, San Jose, California, June 1994.
- [10] C. Borgelt and R. Kruse. Induction of association rules: Apriori implementation. In W. Härdle and B. Ronz, editors, Proceedings of the 15th Conference on Computational Statistics, 2002, pp. 395-400.
- [11] R. Agrawal and R. Srikant. Quest Synthetic Data Generator. IBM Almaden Research Center, San Jose, California, <http://www.almaden.ibm.com/cs/quest/syndata.html>.
- [12] Basket and Sequence Analysis Benchmarks, XELPOES Java Version 1.1.7, Prudsys AG, Germany.
- [13] Survey on Frequent Pattern Mining, Bart Goethals, HIIT Basic Research Unit, University of Helsinki, Finland.
- [14] Ja-Hwung Su, Wen-Yang Lin: CBW: An efficient algorithm for Frequent Itemset Mining, Proceedings of 37<sup>th</sup> Hawaii International Conference on System Science – 2004.
- [15] Jiawei Han, Jian Pei and Yiwen Yin, Mining Frequent Pattern without Candidate Generation, School of Computing Science, Simon Fraser University.
- [16] FIMI dataset – <http://fimi.cs.helsinki.fi/>
- [17] Apriori public implementation <http://fuzzy.cs.uni-magdeburg.de/~borgelt/software>.
- [18] FPGrowth public implementation – [www.csc.liv.ac.uk/frans/KDD/Software/FPGrowth/fpGrowth.html](http://www.csc.liv.ac.uk/frans/KDD/Software/FPGrowth/fpGrowth.html)
- [19] Data Mining – Concepts and Techniques, Jiawei Han, Micheline Kamber.

## A Study of White Matter and Skull Inhomogeneous Anisotropic Tissue Conductivities on EEG Forward Head Modeling

Md. Rezaul Bashar<sup>1,3</sup>, Yan Li<sup>1</sup> and Peng Wen<sup>2</sup>

<sup>1</sup> Department of Mathematics and Computing

<sup>2</sup> Faculty of Surveying and Engineering

Centre for Systems Biology

University of Southern Queensland, Toowoomba, QLD 4350, Australia

<sup>3</sup>Department of Information and Communication Engineering

Islamic University, Kushtia, Bangladesh

Email: {bashar, liyan, pengwen}@usq.edu.au

### Abstract

*The aim of this study is to investigate the effects of white matter (WM) and skull inhomogeneous anisotropic tissue conductivities on human head modeling. The inhomogeneity of WM and skull is included using fractional anisotropy (FA) method and the anisotropy is included according to Volume constraint in the head model construction. A five-layered spherical head model implemented using finite element method (FEM) is used as a volume conductor with a known current source to measure the electroencephalogram (EEG) on the head surface. Statistical measurement techniques are applied to analyze the EEGs obtained from inhomogeneous anisotropic head models and a homogeneous isotropic model. This study finds that the effects of WM and skull inhomogeneous anisotropy on EEG are significant.*

**Keywords:** Fractional anisotropy, finite element method, inhomogeneous anisotropic conductivity, EEG, and forward problem

### 1. Introduction

The electroencephalogram (EEG) is the measurement of the electric potential differences on scalp caused by neural activity inside the brain. The estimation of these electrical potential differences is known as forward problem [1]. The neural activity is modeled by a distribution of current sources and its implementation technique is known as inverse problem or source localization [1][2][3]. The inverse problem, solved by iterative solving forward problem, is a

valuable tool in pre-surgical evaluation of patients suffering from epilepsy and other neurological disorders [3]. Therefore, the importance of the solution of forward problem is significant. The accuracy of this solution depends on assigning the appropriate conductivity to each tissue which is still a challenging research because of its complicated inhomogeneous and anisotropic properties [1]. However, many head model implementations neglect the inhomogeneous anisotropic conductivities inherent to brain tissues, such as the white matter (WM) and skull. WM has the mean resistivity (reciprocal of conductivity) of 700  $\Omega\text{cm}$ , with 350  $\Omega\text{cm}$  and 1050  $\Omega\text{cm}$  values for lower and upper bounds, respectively, and having the variation of  $\pm 50\%$  [4]. Skull resistivity varies between 1360  $\Omega\text{cm}$  and 21400  $\Omega\text{cm}$ , with a mean of 7560  $\Omega\text{cm}$  and a standard deviation of 4230  $\Omega\text{cm}$  [1]. In WM, the conductivity ratio between longitudinal and transversal directions is 10:1 while it is 1:10 between radial and tangential directions for the skull layer. However, there are no reliable algorithms to measure this ratio non-invasively. Baser *et al* [5] proposed the relation between the electrical conductivity tensors and effective water diffusion tensors measured by diffusion tensor magnetic resonance imaging (DT-MRI). The authors proposed that the tissue conductivity tensors share the same eigen vectors with the effective diffusion tensor. In DT-MRI, the conductivity tensors are not derived directly but inferred from diffusion tensors which describe the movement of water molecules and electrically charged particles (ions) [6]. Later on, Tuch *et al* [7] applied a differential effective medium approach and derived a linear relationship among the eigen values of DT-MRI and the conductivity tensors. However, the eigen values are not equal in the same tissue layer and vary according to



direction. These make the tissue layer as inhomogeneous and anisotropy. The anisotropy depends on the volume of tissues and directions of tissue fibers, which are affected by some conductivity tensor reconstruction obstacles. To overcome these obstacles, C. H. Wolters [2] proposed Volume constraint which restricts the geometric mean with the conductivity tensor.

C. H. Wolters [2] investigated the effects of anisotropic conductivities on EEG for WM and skull tissue layers. The author applied transversal: longitudinal conductivity ratio with the values of 1:1 (isotropic), 1:2, 1:5, and 1:10 for WM tissues. He also applied the same values for radial: tangential conductivity ratio for skull tissues. By applying these conductivity ratios, he found some effects of anisotropy on EEG for Volume constrained WM and skull tissues. Gullmar *et al* [6] also used the same anisotropy ratios for Volume constrained WM on a rabbit head model and found the same results like Wolters. However, they neglected the inhomogeneity of WM anisotropic properties. Li *et al* [8] proposed inhomogeneous anisotropy properties using fractional anisotropy (FA) by step and linear functions for Volume constrained WM. Hallez *et al* [3] implemented five-layered spherical head model to investigate the Volume constrained WM and skull anisotropic tissue conductivities for source localization. They found that neglecting the skull anisotropy resulted on an average of 13.73 mm with a maximum of 24.51 mm dipole localization errors. Most recently, Hallez *et al* [9] implemented a realistic head model incorporating Volume constrained anisotropic WM conductivity and found that WM anisotropy plays a vital role in EEG source localization. Anwender *et al* [10] investigated the influence of WM and skull anisotropic conductivities and found some significant effects on source localization. In our previous paper [11], we investigated the influence of WM inhomogeneous anisotropic conductivity on EEG using conductivity ratio approximation (CRA) and statistical conductivity approximation (SCA) techniques. We found that there are strong effects of WM inhomogeneous anisotropic conductivity on the scalp EEG. In the paper, we didn't consider skull tissue inhomogeneity and anisotropy. From an extensive literature review, we realize, most of the researchers investigate effects of tissue anisotropy on EEG, however, they didn't mention about the effects of tissue inhomogeneity.

The goal of this study is to evaluate the effects of WM and skull inhomogeneous anisotropic tissue conductivities on the scalp EEG by means of finite element method (FEM). We classify the inhomogeneity

by the discrepancy of the ratio longitudinal over transverse eigen values ( $R_{lt}$ ). We construct a spherical head model by means of varying FA and Volume constraint to simulate the head volume conductor with inhomogeneous anisotropic conductivity settings. We model a current dipole inside gray matter (GM) to simulate brain electrical activity. And we assess the effects of WM and skull inhomogeneous anisotropic tissue conductivities on the scalp EEG by comparing the potential differences generated by the inhomogeneous anisotropic and homogeneous isotropic head models.

The organization of this paper is as follows. Introduction section describes background and the importance of this study with literature review in Section I. Section II illustrates the methods of fractional anisotropy, head model construction, homogeneous and inhomogeneous anisotropic conductivity allocation, forward solution using FEM and similarity measurement techniques. Simulation setup is described in Section III. Experimental results are presented in Section IV. Finally, Section V concludes our research findings.

## 2. Methods

### 2.1 Fractional anisotropy (FA)

FA is a technique to measure the extent of the anisotropy property for each voxel (element). Let us suppose that  $\lambda_1, \lambda_2$ , and  $\lambda_3$  ( $\lambda_1 \geq \lambda_2 \geq \lambda_3$ ) are the three eigen values of a diffusion tensor matrix and  $\lambda$  is the average eigen value. Then FA is defined as [12]:

$$FA = \frac{\sqrt{3}}{\sqrt{2}} \frac{\sqrt{(\lambda_1 - \lambda)^2 + (\lambda_2 - \lambda)^2 + (\lambda_3 - \lambda)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad \dots \quad (1)$$

The FA is in the range of 0 to 1 [12]. A fully anisotropic tissue has a factor FA=1, and an isotropic tissue has a factor FA=0.

### 2.2 Head model

Anatomically, human head consists scalp, skull, cerebrospinal fluid (CSF), gray matter (GM) and WM. Based on this anatomical concept and different literature [2][3], we consider a five-layered spherical head model. We segment the five-layered sphere and perform mesh generation to produce tetrahedral elements for FEM model. The conductivity tensor in WM or skull of a multi-sphered head model is a diagonal matrix [3]. We assume that the conductivity

tensors share the same eigen vectors with the effective diffusion tensors measured by (DT-MRI) [2][5][6][8]. Then, we consider the conductivity tensor for a WM finite element as [2] [6]:

$$\sigma_{WM} = S \begin{pmatrix} \sigma_{long} & 0 & 0 \\ 0 & \sigma_{trans} & 0 \\ 0 & 0 & \sigma_{trans} \end{pmatrix} S^{-1} \quad \dots\dots\dots(2)$$

where  $S$  is orthogonal matrix of unit length eigenvectors of the measured diffusion tensor at the barycentre of the WM finite element,  $\sigma_{long}$  is the parallel (longitudinal) eigen value and  $\sigma_{trans}$  is the perpendicular (transverse) eigen value where  $\sigma_{long} \geq \sigma_{trans}$ . We also consider the conductivity tensor for a skull finite element as [2] [13]:

$$\sigma_{Skull} = S \begin{pmatrix} \sigma_{tan} & 0 & 0 \\ 0 & \sigma_{tan} & 0 \\ 0 & 0 & \sigma_{rad} \end{pmatrix} S^{-1} \quad \dots\dots\dots(3)$$

where  $\sigma_{tan}$  is the tangential (parallel) eigen value and  $\sigma_{rad}$  is the radial (perpendicular) eigen value where  $\sigma_{rad} \geq \sigma_{tan}$ . Assigning the conductivity, we apply FEM for a known current source to solve forward problem.

### 2.3 Homogeneous anisotropic conductivity allocation

Modeling of the conductivity tensor eigen values is a difficult task involving layer's boundary segmentation, inhomogeneous resistivity, tissue type, and fibre orientation. Therefore, we use isotropic conductivity to simulate the anisotropic conductivity using  $\sigma_{trans}$ :  $\sigma_{long}$  ratios according to literature [2][13]. Then, we calculate the longitudinal and transversal eigen values based on Volume constraint.

#### 2.3.1 Volume constraint

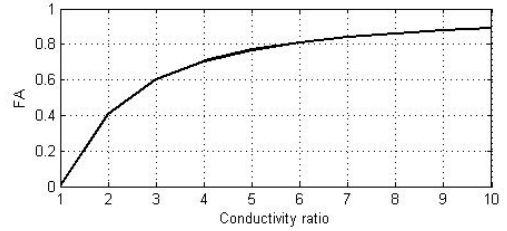
The Volume constraint remains constant of the geometric mean and the volume of the conductivity tensor. The volume constraint is defined as [2][6],

$$\frac{4}{3}\pi\sigma_{long}(\sigma_{trans})^2 = \frac{4}{3}\pi\sigma_{iso}^3 \quad \dots\dots\dots(4)$$

where  $\sigma_{iso}$  is isotropic conductivity.

### 2.4 Inhomogeneous anisotropic conductivity allocation

For the homogeneous anisotropic model, all elements share the same conductivity ratio ( $R_{lt}$ ) between longitudinal and transversal conductivities. However,  $R_{lt}$  varies for inhomogeneous anisotropy.  $R_{lt}$  reflects the extent of anisotropy property as FA does, so we set  $R_{lt}$  as a variable of FA. We establish the relation between FA and  $R_{lt}$  by implementing eq(1). The results are shown in Figure 1. Though the values of FA lie between 0 and 1 [12], we found the values of FA ranging from 0 to 0.9. Considering Figure 1, we define  $R_{lt}$  using multi-steps FA function as presented in eq(5).



**Figure 1:** Conductivity ratio Vs fractional anisotropy (FA).

$$R_{lt} = \begin{cases} 10, & FA > 0.8 \\ 6, & FA > 0.7 \\ 4, & FA > 0.4 \\ 2, & otherwise \end{cases} \quad \dots\dots\dots(5)$$

Using  $R_{lt}$ , we generate longitudinal and transversal inhomogeneous conductivities for Volume constrained WM and skull tissue layers.

### 2.5 Forward solutions using finite element method

The electric potential field generated by a current source in a head is described by the Poission's equation as [2][14]:

$$\nabla \cdot \sigma (\nabla \phi) = -I_{sv} \text{ in } \Omega \quad \dots\dots\dots(6)$$

where  $\sigma$  is conductivity tensor,  $\phi$  is electric potential field, and  $I_{sv}$  is internal current source per unit volume due to current dipoles placed within the brain.

The FEM provides a convenient way to assign the conductivity tensor for a small region. Using Galerkin approximation [15] in the framework of FEM and boundary conditions, the solution of eq(6) can be written in a matrix form as [14]:

$$[K]\{\phi\} = \{F\} \quad \dots\dots\dots(7)$$

where K is the stiffness matrix [15],  $\phi$  is the vector of nodal electric potentials, and F is the force vector including the boundary conditions and the current sources. The matrix K contains the head geometry and conductivity information for the forward problem and has the properties being sparse, symmetric, and positive definite. As K is symmetric and positive definite, we use Choleski preconditioner [16] to accelerate the rate of convergence of iterative solvers. We apply the preconditioned conjugate method (pcg) to iteratively solve the linear equation stated in eq(7).

## 2.6 Similarity measurements

We calculate the electric potential differences produced by a single dipole measured using 64 electrodes positioned at different places on a head surface. The forward computed data obtained from the homogeneous isotropic and inhomogeneous anisotropic model are analyzed by calculating relative difference measure (RDM) for the topology error (minimum error: RDM=0) and magnitude difference (MAG) values (minimum error: MAG=1) of the scalp electric potentials. The RDM and MAG values are calculated as follows [2][6][11][13]

$$RDM = \sqrt{\sum_{i=1}^n \left( \frac{ref_i}{\sqrt{\sum_{i=1}^n ref_i}} - \frac{meas_i}{\sqrt{\sum_{i=1}^n meas_i}} \right)^2} \quad \dots\dots\dots (8)$$

$$MAG = \sqrt{\frac{\sum_{i=1}^n meas_i^2}{\sum_{i=1}^n ref_i^2}} \quad \dots\dots\dots (9)$$

where the values obtained from the homogeneous isotropic model are interpreted as reference (ref) and the values obtained from the inhomogeneous anisotropic model are defined as measurement (meas) [11]. The index i represents the number of electrodes.

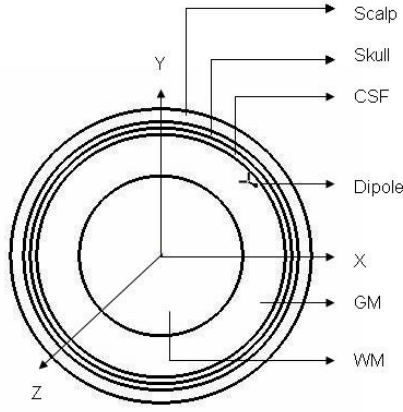
## 3. Simulation setup

First, we implement a five-layered spherical head model according to the radii of Table 1 (row 2) using Matlab [16]. We segment the head model into surfaces,

perform tessellation for mesh generation and then, apply a constrained Delaunay tessellation [16] using Tetgen® package provided by BrainStorm [17]. The details of head model construction are described in our previous paper [11]. The mesh generation produces 112 K tetrahedral elements from 19 K nodes as shown in Table 1 (last row). We use these tetrahedral elements for FEM modeling. For homogeneous isotropic model, we assign the mean conductivity to each tissue layer (row 3 in Table 1). However, we assign the conductivities produced by implementing multi-steps FA function (described in eq(5)) to individual elements of WM and skull having other tissue layers isotropic for the inhomogeneous anisotropic head model. We also implement multi-steps FA technique stated in eq(5) using Matlab coded software. After assigning conductivities, we perform EEG forward computation using the adopted FEM tool provided by BrainStorm [17] for a fixed current dipole inside the GM. We assume that the dipole is located in axial, coronal and sagittal planes with the azimuth angle of  $\pi/4$  and elevation angle of  $\pi/5$ , and having the unit magnitude. A dipole can be decomposed into three orthogonal dipoles along the main axes [9]. For each dipole location, we consider three orthogonal orientations: along left-right (X orientation), along back-front (Y orientation) and along bottom-top (Z orientation). Figure 2 shows the logical head model with different tissue compartments and dipole orientations. We measure the EEGs by putting 64 electrodes on head surface. Finally, we apply RDM and MAG techniques to analyze the results. We perform this computation using an Intel® dual core 2.0 Ghz processor with 2GB RAM. It takes approximately an hour to execute each computation.

**Table1:** Head model parameters

	Scalp	Skull	CSF	GM	WM
Radii (cm)	9.2	8.4	8.0	7.6	5.0
Means (S/m)	0.33	0.0042	1.0	0.14	0.33
Elements	19397	24563	21379	20674	26841



**Figure 2:** Head model construction showing different head compartments and dipole (source) location.

#### 4. Experimental results

To study the influences of inhomogeneous anisotropic WM and skull tissue conductivities, we carry out four independent experiments.

1. We measure EEG from the reference model.
2. We measure EEG from the measurement model where inhomogeneous anisotropic conductivity is assigned to WM while other tissue layers are homogeneous and isotropic.
3. We measure EEG by assigning inhomogeneous anisotropic conductivity to skull while other layers are homogeneous and isotropic.
4. We measure EEG by assigning the WM and skull inhomogeneous anisotropic conductivities considering other layers homogeneous and isotropic.

Table 2 shows the RDM and MAG errors caused by the inhomogeneous anisotropic WM conductivity generated using the Volume constraint. We found that the RDM (1.59% ~ 18.87%) and MAG (0.95 ~ 1.12) values are far from their ideal values, 0 and 1, respectively. These results indicate that WM inhomogeneous anisotropy affects the scalp EEG strongly. The longitudinal inhomogeneous conductivity produces fewer errors than those of transversal conductivity. We found that the obtained results for WM inhomogeneous anisotropic models are consistent with those published from different literature for WM anisotropic model neglecting inhomogeneity [2][10][13]. We conclude that WM inhomogeneous anisotropic conductivity has significant effects on EEG and WM transversal inhomogeneous conductivity has

more effects on EEG than longitudinal inhomogeneous conductivity.

**Table 2:** RDM and MAG values generated by inhomogeneous anisotropic WM conductivity

Conductivity	Dipole orientation	RDM	MAG
Longitudinal	X	4.04%	1.02
	Y	5.91%	1.12
	Z	4.21%	1.03
Transversal	X	18.87%	1.07
	Y	1.59%	0.97
	Z	7.3%	0.95

Table 3 shows the topological and magnitude difference errors due to anisotropic inhomogeneous skull conductivity. Note that the RDM values range from 4.37% to 17.19% and MAG values are between 0.84 and 1.11. It demonstrates that the effects of inhomogeneous anisotropic skull tissue conductivity on EEG are significant. Radial inhomogeneous conductivity results more errors than tangential inhomogeneous conductivity. These results are also consistent with different literature where inhomogeneity was absent [2] [13].

**Table 3:** RDM and MAG values generated by inhomogeneous anisotropic skull conductivity

Conductivity	Dipole orientation	RDM	MAG
Radial	X	17.19%	0.89
	Y	7.93%	0.84
	Z	7.17%	0.96
Tangential	X	8.18%	1.09
	Y	4.37%	1.11
	Z	4.64%	1.03

Table 4 shows the RDM and MAG errors generated by combining the inhomogeneous anisotropic WM and skull tissue conductivities. The longitudinal conductivity produces 1.23% to 5.9% RDM and 0.95 to 1.01 MAG errors while the transverse conductivity produces 7.03% to 20.39% RDM and 1.04 to 1.09 MAG errors. It clearly demonstrates that the combination of inhomogeneous anisotropic WM and skull conductivities have significant effects on EEG. By contrasting Table 2 and Table 4, we found the similar results, namely, EEGs are more affected by transversal conductivity than longitudinal conductivity.

**Table 4:** RDM and MAG values generated by inhomogeneous anisotropic WM and skull conductivities.

Conductivity	Dipole orientation	RDM	MAG
Longitudinal	X	5.9%	0.95
	Y	1.23%	1.01
	Z	4.48%	0.99
Transversal	X	20.39%	1.09
	Y	8.96%	1.08
	Z	7.03%	1.04

## 5. Conclusion

In the present study, we have investigated the influences of WM and skull inhomogeneous anisotropic tissue conductivities on EEG forward computing using a spherical head model by means of Volume constraint. We have implemented the multi-steps FA to generate Volume constrained anisotropic conductivity and various anisotropy ratios to generate inhomogeneity. From our experiment results, we find that WM inhomogeneous anisotropic tissue conductivity results maximum 18.87% RDM and 1.12 MAG errors and skull inhomogeneous anisotropic tissue conductivity results maximum 17.17% RDM and 1.11 MAG errors. Combining WM and skull inhomogeneous anisotropic conductivities result maximum 20.39% RDM and 1.09 MAG errors. This study also finds that the WM inhomogeneous transversal and skull inhomogeneous radial conductivities have more effect on EEG.

## 6. Acknowledgement

This project is supported by ARC Discovery Program: DP0665216.

## 7. References

- [1] P. Wen, *Human Head Modelling and Computation for the EEG Forward Problem*, PhD dissertation, The Flinders University of Western Australia, Australia, 2000.
- [2] Wolters, C. H., *Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain*, PhD dissertation, University of Leipzig, France, 2003.
- [3] Hallez, Hallez, Bart Vanrumste, Peter Van Hese, Yuves D'Asseler, Ignace Lemahieu and Rik Van de Walle, "A finite difference method with reciprocity used to incorporate anisotropy in electroencephalogram dipole source localization", *Journal of Physics in Medicine and Biology*, 2005, vol 50, pp. 3787-3806.
- [4] Hauseisen, J., Ramon, C., Eiselt, M., Brauer, H. and Nowak, H., "Influence of Tissue Resistivities on Neuromagnetic Fields and Electric Potentials studied with a Finite Element Model of the Head", *IEEE Transactions on Biomedical Engineering*, 1997, vol 44, no. 8, pp.727-735.
- [5] Bassar, P., Mattiello, J. and LeBihan, D., "MR diffusion tensor spectroscopy and imaging", *Biophysical Journal*, 1994, vol 66, pp. 259-267.
- [6] Gullmar, D., Haueisen, J., Wiselt, M., Giebler, F., Flemming, L., Anwander, A. Thomas, R. K., Wolters, C. H., Dumpelmann, M., David, S. T. and Jorgen, R. R., "Influence of Anisotropic Conductivity on EEG source Reconstruction: Investigations in a Rabbit Model", *IEEE Transactions on Biomedical Engineering*, 2006, vol 53, no. 9, pp.1841-1850.
- [7] David S. Tuch, Van J. Wedeen, Anders M. Dale, John S. George, and John W. Belliveau, "Conductivity tensor mapping of the human brain using diffusion tensor MRI", *Proceedings of the National Academy of Sciences, USA*, 2001, vol 98, no. 20, pp. 11697-11701.
- [8] Li, L., Wang, K., Zhu, S., Mueller, K., Lim, K., Liu, Z. and He, B., "A Study of White Matter Anisotropic conductivity on EEG Forward Solution", *IEEE Proceedings of Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging (NFSI & ICFBI)*, 2007, pp.130-132.
- [9] Hallez, H., Vanrumste, B., Hese, P. V., Delputte, S. And Lemahieu, I., "Dipole estimation errors due to differences in modeling anisotropic conductivities in realistic head models for EEG source analysis", *Journal of Physics in Medicine and Biology*, 2008, vol 53, pp. 1877-1894.
- [10] Anwandwer, C.H. Wolters, M. Dumpelmann and T. Knosche, "Influence of realistic skull and white matter on the inverse problem in EEG/MEG- source localization", *Proceedings of International Conference on Biomagnetism*, Germany, 2002.
- [11] R.Bashar, Y. Li and P.Wen, "Influence of white matter inhomogeneous anisotropy on EEG forward computing", *Australasian Physical & Engineering Sciences in Medicine*, 2008, vol 31, no. 2, pp. 122-130.

- [12] Hai Li, Tianming Liu, Geoffrey Young, Lei Guo, and Stephen TC Wong, "Brain Tissue Segmentation based on DWI/DTI Data", *3<sup>rd</sup> IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2006.
- [13] Wolters, C. H., Anwander, A., Tricoche, X., Weinstein, D., Koch, M. A. and MacLeod, R. S., "Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modelling", *Journal of NeuroImage, Elsevier*, 2006, vol 30, pp. 813-826.
- [14] Klepfer, R. N., Johnson, C. R. and Robert, S. M., "The effects of Inhomogeneities and Anisotropies on Electrocardiographic Fields: A 3-D Finite -element Study", *IEEE Transactions on Biomedical Engineering*, 1997, vol 44, no. 8, pp. 706-719.
- [15] Young W. Kwon and Hyochoon C. Bang, *The finite element method using MATLAB*, Second edition, CRC Press, 2000.
- [16] Mathwork software, The Matlab.
- [17] Baillet, S., Mosher, J. C. and Leahy R. M., "Electromagnetic Brain Imaging using Brainstorm", *IEEE International Symposium on Biomedical Engineering: Macro to Nano*, 2004, pp. 652-655.

# Selecting Signature Optical Emission Spectroscopy Variables Using Sparse Principal Component Analysis

Beibei Ma, Seán McLoone, John Ringwood

Department of Electronic Engineering

National University of Ireland, Maynooth, Ireland

beibei.ma, sean.mcloone, john.ringwood@eeng.nuim.ie

Niall Macgearailt

Department of Electronic Engineering

Dublin City University, Ireland

niall.macgearailt2@mail.dcu.ie

## Abstract

*Principal component analysis (PCA) is a widely used technique in optical emission spectroscopy (OES) sensor data analysis for the low dimension representation of high dimensional datasets. While PCA produces a linear combination of all the variables in each loading, sparse principal component analysis (SPCA) focuses on using a subset of variables in each loading. Therefore, SPCA can be used as a key variable selection technique. This paper shows that, using SPCA to analyze 2046 variable OES data sets, the number of selected variables can be traded off against variance explained to identifying a subset of key wavelengths, with an acceptable level of variance explained. SPCA-related issues such as selection of the tuning parameter and the grouping effect are discussed.*

## 1. Introduction

Principal component analysis (PCA) is widely known as a dimension reduction technique. Using PCA, a high dimensional data set can be decomposed into the sum of a small number of principal components (PCs). Given an  $n \times m$  data matrix  $\mathbf{X}$ ,  $n$  being the number of observations and  $m$  being the number of variables, there is [3]

$$\mathbf{X} = \mathbf{TP}^T \quad (\mathbf{T} \in \mathbb{R}^{n \times p}, \mathbf{P} \in \mathbb{R}^{m \times p}), \quad (1)$$

where  $\mathbf{P}^T \mathbf{P} = \mathbf{I}_p$ .  $\mathbf{T}$  and  $\mathbf{P}$  are referred to as score and loading matrices, respectively. Since the loading of each PC is generally a linear combination of *all* the original variables, PCA cannot be used directly for variable selection.

Research on obtaining sparse components has been conducted for over two decades. The earliest method, proposed in 1958, is referred to as varimax [7]. Using varimax rotation, a number of the coefficients of the loading vectors can be adjusted to have greater values than the remaining coefficients. Such adjustment can help in the selection of key

variables, but it is hard to quantify the distinction between small and large coefficients.

Jeffers [5] proposed a straight-forward method for achieving PCA sparsity. For each loading, any coefficients that are less than 70% of the greatest one are set to zero, regardless of their sign. This method can lead to a selection deficiency in two cases, one where the variables have small coefficients and the other where the variables have high mutual correlations [1].

In [10], the 'simple principal components' is proposed. This focuses on restricting the coefficients of the loadings to have integer values, such as -1, 0 and 1, to help simplify variable selection.

The first true algorithmic method for achieving sparse loadings was proposed in 2003 by Jolliffe *et al.* [6] and is known as SCoTLASS (Simplified Component Technique for Least Absolute Shrinkage and Selection). This employs a penalty term referred to as the Least Absolute Shrinkage and Selection Operator (LASSO) [9] to force loadings to be sparse. Nevertheless, it is not practical due to the relatively high computational cost [12].

A recently proposed algorithm, known as semidefinite programming, is described in [2]. Using this method, the normal loadings are constrained by a cardinality condition, that is, a limit on the number of the nonzero elements in each loading. By relaxing this constraint, the problem is converted into a convex optimization problem and hence, can use semidefinite programming as a solution. The generated PCs are shown to be able to explain larger variance than competing algorithms, but the computational cost is high.

In 2004 Zou *et al.* [12] proposed an alternative approach to solving the SPCA problem, which they referred to as elastic net for SPCA (EN-SPCA). EN-SPCA can be implemented in two forms. One is similar to an approach used to solving the LASSO problem, details of which will be provided in the next section, and the other is the so called soft thresholding algorithm, designed for handling large data sets (thousands of variables). Both EN-SPCA

implementations are computational alternatives to semidefinite programming, but the later implementation has the key advantage that it can scale to much larger problems than the semidefinite programming algorithm.

This paper explores the application of SPCA to key variable selection in optical emission spectroscopy (OES). OES is an optical emission detection technique, used for detecting the optical intensity of the chemicals in a plasma as a function of wavelength and time. Because the wavelengths are the ‘fingerprints’ of the corresponding chemical species, OES data can be used to trace the chemical reactions in a plasma chamber. Consequently, OES is increasingly being used by semiconductor manufacturing to assist with the monitoring and control of plasma etch processes. However, in practice direct use of OES spectra is limited due to the difficulties with handling and interpreting the large number of variables that are associated with such spectra.

The remainder of the paper is organized as follows. Section 2 gives a theoretical description of SPCA. Section 3 describes the methods used to estimate SPCA model accuracy. Section 4 shows experimental results on the application of SPCA to Optical Emission Spectroscopy (OES) data from a semiconductor manufacturing plasma etch process. Finally, the conclusions are presented in Section 5.

## 2 Theoretical Framework and Numerical Solution

EN-SPCA employs a penalised regression estimator to solve the SPCA problem. This exploits the fact that PCA can be formulated as a least squares regression problem, thereby facilitating the inclusion of the LASSO penalty, which is known from regression theory to yield sparse solutions. The theory of LASSO and EN-SPCA are quite involved, hence only a brief overview will be provided here. For a more complete development of the algorithms see [9],[12] and [11].

### 2.1 PCA as a LS problem

Given a data matrix  $\mathbf{X}$  (as defined in Eq. (1)) and defining its transpose as the matrix  $\mathbf{Z} = \mathbf{X}^T$ , the LS estimation of the PCA loading matrix  $\mathbf{P} \in \mathbb{R}^{m \times p}$  can be expressed as

$$(\hat{\mathbf{A}}, \hat{\mathbf{P}}) = \arg \min_{\mathbf{A}, \mathbf{P}} \left\{ \sum_{i=1}^n \|\mathbf{z}_i - \mathbf{A}\mathbf{P}^T \mathbf{z}_i\|_2^2 \right\}, \quad (2)$$

$$\text{s.t. } \mathbf{A}^T \mathbf{A} = \mathbf{I},$$

Where  $\mathbf{z}_i$  are the columns of  $\mathbf{Z}$ . The constraint on matrix  $\mathbf{A}$  ensures that the orthogonality of  $\mathbf{P}$  is guaranteed. While theoretically the LS estimate is the best unbiased estimate of  $\mathbf{P}$  [9], in practice better mean square error performance can

be obtained by biasing the regression coefficients towards zero. This is typically achieved by adding either an  $L_1$  or an  $L_2$  norm penalty to the LS cost function. The  $L_2$  implementation, referred to as Ridge regression or regularisation, is frequently employed to address data ill-conditioning and singularity issues and benefits from a straightforward algebraic solution. In contrast, the  $L_1$  implementation has traditionally been avoided due to the associated computational issues, but in recent years it has been receiving increasing attention as a variable selection method, due to its tendency to yield sparse solutions. In this context it is referred to as LASSO [9].

### 2.2 Ridge Estimation

The Ridge estimate of  $\mathbf{A}$  and  $\mathbf{P}$  (denoted by  $\hat{\mathbf{A}}^R$  and  $\hat{\mathbf{P}}^R$ ) is expressed as

$$(\hat{\mathbf{A}}^R, \hat{\mathbf{P}}^R) = \arg \min_{\mathbf{A}, \mathbf{P}} \left\{ \sum_{i=1}^n \|\mathbf{z}_i - \mathbf{A}\mathbf{P}^T \mathbf{z}_i\|_2^2 + \gamma_2 \sum_{j=1}^p \|\mathbf{p}_j\|_2^2 \right\}, \quad \text{s.t. } \mathbf{A}^T \mathbf{A} = \mathbf{I}, \quad (3)$$

where  $\|\cdot\|_2$  represents the  $L_2$  norm,  $\gamma_2$  is the Ridge tuning parameter ( $\gamma_2 \geq 0$ ) and  $\mathbf{p}_j$  is the  $j$ th column vector of  $\mathbf{P}^R$ . The main benefit of employing the Ridge estimator is that it can handle data matrices in which the number of observations is less than the number of variables. Unlike regression this comes at no cost since the Ridge estimate of  $\mathbf{P}$  is simply a scaled version of the LS estimate and scaling does not affect the PCA decomposition.

### 2.3 LASSO Estimation

Mathematically, the LASSO estimate of  $\mathbf{A}$  and  $\mathbf{P}$ , denoted by  $\hat{\mathbf{A}}^L$  and  $\hat{\mathbf{P}}^L$  can be expressed as

$$(\hat{\mathbf{A}}^L, \hat{\mathbf{P}}^L) = \arg \min_{\mathbf{A}, \mathbf{P}} \left\{ \sum_{i=1}^n \|\mathbf{z}_i - \mathbf{A}\mathbf{P}^T \mathbf{z}_i\|_2^2 + \gamma_1 \sum_{j=1}^p \|\mathbf{p}_j\|_1 \right\}, \quad \text{s.t. } \mathbf{A}^T \mathbf{A} = \mathbf{I}, \quad (4)$$

or equivalently as

$$(\hat{\mathbf{A}}^L, \hat{\mathbf{P}}^L) = \arg \min_{\mathbf{A}, \mathbf{P}} \left\{ \sum_{i=1}^n \|\mathbf{z}_i - \mathbf{A}\mathbf{P}^T \mathbf{z}_i\|_2^2 \right\}, \quad (5)$$

$$\text{s.t. } \sum_{j=1}^p \|\mathbf{p}_j\|_1 \leq c_1 \text{ and } \mathbf{A}^T \mathbf{A} = \mathbf{I},$$

where  $\|\mathbf{p}_j\|_1 = \sum_{i=1}^m |b_{ij}|$ ,  $\gamma_1$  is the LASSO tuning parameter and  $c_1$  is a corresponding upper bound. For each  $\gamma_1$ ,



there exists a  $c_1$  that gives an equivalent constraint on the regression coefficients. The second formulation (Eq. (5)) has the attraction that it is in the form of a quadratic programming problem with linear inequality constraints making its solution mathematically tractable.

Although the Ridge and LASSO penalty both cause the regression coefficients to shrink towards zero a powerful feature of the LASSO is that it is more likely to drive coefficients to exactly zero, hence generating sparse solutions. This arises because of the distinctive shape of the LASSO penalty as illustrated in Fig.1 for the 2-D case. This shows the elliptical contours of a LS quadratic cost function with the feasible region corresponding to the  $L_1$  (LASSO) constraint indicated in grey. The  $L_1$  norm produces a diamond shaped region with corners aligned with the axes. In contrast, as illustrated in Fig.1(b), the  $L_2$  norm generates a circular boundary. Increasing the constraint penalty causes the feasible regions to shrink towards the origin. As this happens there is a tendency for the optimum elliptical contour to intersect the diamond region at the corners and since these are aligned with the axes this leads to sparse solutions. This is not the case with the circular boundary of the Ridge penalty.

## 2.4 Naive Elastic Net and Elastic Net

The drawback of LASSO is that if the number of observations ( $n$ ) is less than the number of the variables ( $m$ ), LASSO will at most choose  $n$  variables. In contrast, Ridge can extend the selection to  $m$  variables. However, Ridge cannot provide sparse solutions. The naive elastic net method was developed to address these problems. Mathematically, the naive elastic net estimate is defined as:

$$(\hat{\mathbf{A}}^N, \hat{\mathbf{P}}^N) = \arg \min_{\mathbf{A}, \mathbf{P}} \left\{ \sum_{i=1}^n |\mathbf{z}_i - \mathbf{A} \mathbf{P}^T \mathbf{z}_i|^2 + \gamma_2 \sum_{j=1}^k \|\mathbf{p}_j\|_2^2 + \sum_{j=1}^k \gamma_{1j} \|\mathbf{p}_j\|_1 \right\}, \text{ s.t. } \mathbf{A}^T \mathbf{A} = \mathbf{I}, \quad (6)$$

where  $\gamma_{1j}, j=1 \dots k$  are scalars used to individually penalize the loadings.

In [11], Zou and Hastie argue that naive elastic net estimation introduces a double bias, but that the reduction in estimation variance is no greater than with LASSO or Ridge estimation alone. To correct for this double shrinkage, the authors propose re-scaling the naive elastic net estimate by a factor  $(1 + \gamma_2)$ , giving the final elastic net SPCA estimate as:

$$(\hat{\mathbf{A}}^{\text{EN}}, \hat{\mathbf{P}}^{\text{EN}}) = (1 + \gamma_2)(\hat{\mathbf{A}}^N, \hat{\mathbf{P}}^N). \quad (7)$$

Because the constraint function  $\gamma_2 \|\mathbf{p}_j\|_2^2 + \gamma_{1j} \|\mathbf{p}_j\|_1$  is strictly convex, elastic net estimation retains a special property of least squares regression known as the grouping effect

[12]. This means that highly correlated variables will be assigned similar regression coefficients. While this property can be beneficial in many applications, as will be demonstrated later, it is not ideal for variable selection.

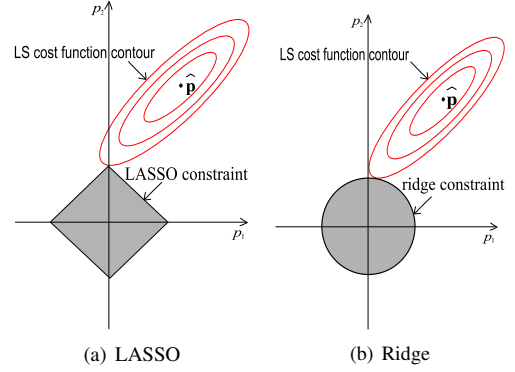


Figure 1. Solution for different estimators

## 2.5 Numerical Solution

Simple matrix transformations can be used to estimate the least squares or Ridge regression estimates of  $\mathbf{P}$ , but efficient computation of the LASSO estimate is much more challenging, and was only effectively addressed with the development of the least angle regression (LARS) algorithm in 2002 [8] (S refers to its close relation to LASSO and stagewise regression).

In LARS, all the coefficients are first set to zero and then revised successively until the least squares solution is reached. LARS-EN is a modified implementation of LARS designed to fit the elastic net framework [8]. In particular, it includes a special algorithm, called soft-thresholding, that is scalable to high dimensional data sets [12].

## 3 Estimating Model Accuracy

### 3.1 Variance Explained

SPCA employs a similar approach to PCA, known as adjusted variance [12], to measure the estimation accuracy. Like PCA, the sparse scores ( $\hat{\mathbf{T}}^S$ ) are used to calculate the variance as:

$$\hat{\mathbf{T}}^S = \mathbf{X} \hat{\mathbf{P}}^{\text{EN}}, \quad (8)$$

where  $\hat{\mathbf{P}}^{\text{EN}}$  denotes  $\hat{\mathbf{P}}^{\text{EN}}$  normalized to have unit length columns. However, unlike PCA,  $\hat{\mathbf{P}}^{\text{EN}}$  is not orthogonal, hence the variances explained by individual PCs are not independent of each other and hence not additive. To correct for this the sparse scores matrix must first be orthogonalised leading to the following variance estimation algorithm [4, 8].

- Orthogonalize  $\hat{\mathbf{t}}_j^S$  (the  $j$ th column vector of  $\hat{\mathbf{T}}^S$ ) by applying the recursion

$$\hat{\mathbf{t}}_j^{S*} = \hat{\mathbf{t}}_j^S - \hat{\mathbf{T}}_{(j-1)}^S [(\hat{\mathbf{T}}_{(j-1)}^S)^T (\hat{\mathbf{T}}_{(j-1)}^S)]^{-1} (\hat{\mathbf{T}}_{(j-1)}^S)^T \hat{\mathbf{t}}_j^S,$$

for  $j = 1, \dots, p$  ( $p$  is the number of sparse principal components), where  $\hat{\mathbf{T}}_{(j)}^S = [\hat{\mathbf{t}}_{(1)}^S, \dots, \hat{\mathbf{t}}_{(j)}^S]$ .

- Collect the orthogonalized vectors into a matrix  $\hat{\mathbf{T}}^{S*}$ , *i.e.*

$$\hat{\mathbf{T}}^{S*} = [\hat{\mathbf{t}}_1^{S*}, \dots, \hat{\mathbf{t}}_j^{S*}, \dots, \hat{\mathbf{t}}_p^{S*}]. \quad (9)$$

- Compute the variance explained ( $V_e$ ) by the  $p$  sparse components as

$$V_e = \text{trace}\{(\hat{\mathbf{T}}^{S*})^T \hat{\mathbf{T}}^{S*}\}. \quad (10)$$

The  $j$ th diagonal entry of  $(\hat{\mathbf{T}}^{S*})^T \hat{\mathbf{T}}^{S*}$  corresponds to the variance explained by the  $j$ th sparse PC.

### 3.2 SPMSE

When using sparse PCs to reconstruct a data set, many columns of the reconstructed data are in fact zero, because of the zero elements in the sparse loadings. Therefore, a fairer assessment of the accuracy of reconstruction is to only compare the reconstruction against the original data over the regions where the reconstruction exists. Here, a sparse mean square error measure is proposed, denoted SPMSE, where S stands for the sparse, P for percentage and MSE for mean square error. This is given by

$$SPMSE = \frac{\|\hat{\mathbf{X}}_s - \mathbf{X}_s\|_f^2}{\|\mathbf{X}_s\|_f^2} \times 100\%, \quad (11)$$

where  $\|\cdot\|_f$  is the Frobenius norm,  $\hat{\mathbf{X}}_s$  consists of the nonzero columns of the reconstructed data matrix  $\hat{\mathbf{X}}$  and  $\mathbf{X}_s$  is the corresponding subset of the original data matrix  $\mathbf{X}$ . Based on SPMSE,  $SV_e$ , variance explained by the sparse components, is proposed as

$$SV_e = 100\% - SPMSE. \quad (12)$$

In contrast to  $V_e$  (Eq. 10),  $SV_e$  only calculates the variance for the non-zero reconstructed-channels, so  $SV_e$  can more effectively reflect the reconstruction accuracy. Note, that since the sparse components are not orthogonal, the reconstruction of  $\mathbf{X}$  is defined as

$$\hat{\mathbf{X}} = \hat{\mathbf{T}}^S (\hat{\mathbf{P}}^{\text{EN}})^T [\hat{\mathbf{P}}^{\text{EN}} (\hat{\mathbf{P}}^{\text{EN}})^T]^{-1}. \quad (13)$$

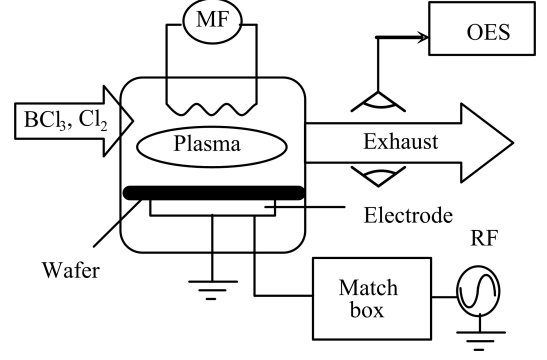


Figure 2. Diagram of OES data collection

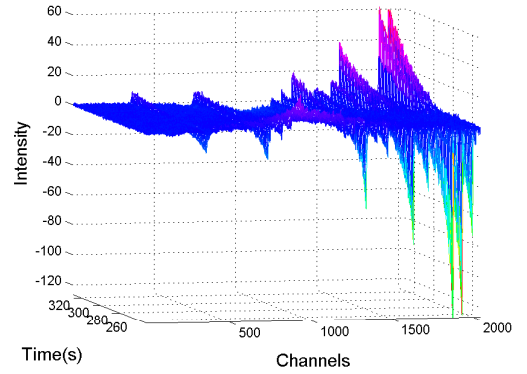


Figure 3. Sample OES data set

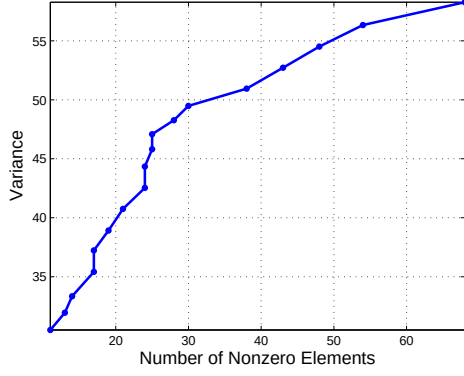
## 4 Results

### 4.1 Data Description

The OES data under consideration is from a plasma etch chamber used in the manufacture of semiconductor chips (Fig. 2). It is collected for the exhaust plasma leaving the chamber and consists of 2046 channels (each channel corresponds to one wavelength) recorded at a sampling interval of 0.76s. Fig. 3 shows a typical data set collected for a 77s etch step on a single wafer. The data, which has been mean-centred, is clearly highly redundant. A PCA analysis reveals that a single PC can capture 95.8% of the variance in the data.

### 4.2 Selecting the EN-SPCA Tuning Parameters

As noted previously when the number of variables is large (in this case 2046) the soft-thresholding EN-SPCA algorithm developed in [12] can be used to efficiently compute components. In this formulation the  $L_2$  penalty tuning



**Figure 4.** The relationship between  $N_{NE}$  and  $V_e$

parameter ( $\gamma_2$ ) is set to infinity leaving only the  $L_1$  penalty parameters ( $\gamma_{1j}$  - one for each computed PC,  $j=1 \dots p$ ) to be determined experimentally. These parameters essentially determine the sparseness of the corresponding PCs.

For example, Table 4.2, shows how the number of nonzero elements ( $N_{NE}$ ) varies as a function of  $\gamma_{11}$ , for the first sparse component of the OES data. The table also shows how the variance explained by the component (expressed as a percentage of the total data variance). As might be expected  $V_e$  decreases as more and more components are forced to zero. In contrast, the  $SV_e$  value remains large for all values of  $\gamma_{11}$  demonstrating that the sparse PC achieves good accuracy for those channels where a reconstruction exists.

Fig. 4 demonstrates the corresponding relationship between  $N_{NE}$  and  $V_e$  and provides a useful guide for making a judgment call on the trade-off between sparsity and variance explained. As can be seen, in this instance there is a 'knee' in the graph at  $N_{NE} = 25$ , beyond which there is a marked decrease in the rate of variance increase with included variables. This corresponds to choosing  $\gamma_{11}$  as 29000.

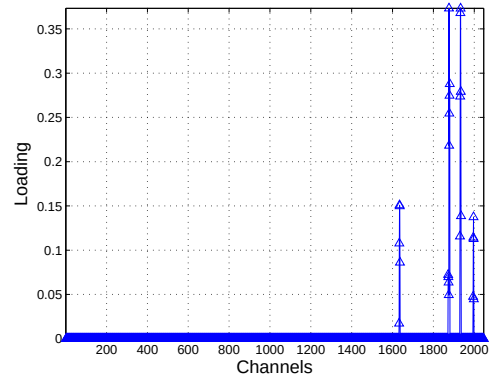
The variance explained by the sparse PC for  $\gamma_{11} = 29000$  is 47% which compares to 95.8% for the unrestricted principal component. The corresponding  $SV_e$  value, which measures the reconstruction accuracy on the non-zero reconstructed-channels only, is more than 91%, by the 25 channels used in its computation.

### 4.3 The Grouping Effect

Figure 5 shows the distribution of the first sparse loading computed with  $\gamma_{11} = 290000$ . For comparison purposes the distribution of the first PCA component loading is included in Figure 6. The PCA loading elements are all

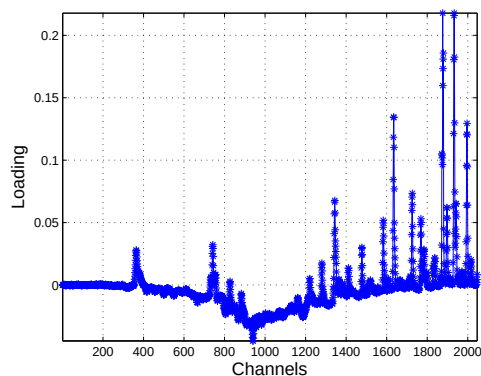
$\gamma_{11}$	$N_{NE}$	$V_e$ (%)	$SV_e$ (%)
0	2046	95.8057	95.8057
5000	1564	95.6068	96.129
10000	1373	95.0602	95.9379
50000	611	84.1086	89.6853
100000	189	65.0638	85.9841
150000	68	58.2831	90.3285
170000	54	56.3396	90.88
190000	48	54.5187	89.956
210000	43	52.7259	89.0393
230000	38	50.9482	88.6207
250000	30	49.4871	91.1976
270000	28	48.2702	90.5483
<b>290000</b>	<b>25</b>	<b>47.0887</b>	<b>91.127</b>
310000	25	45.8107	88.6538
330000	24	44.3460	86.9806
350000	24	42.5365	83.4315
370000	21	40.7546	84.3206
390000	19	38.9135	84.0687
410000	17	37.2303	84.522
430000	17	35.4131	80.3966
450000	14	33.3417	83.5658
470000	13	31.9432	83.0192
490000	11	30.4860	86.5150

**Table 1.**  $N_{NE}$ ,  $V_e$  and  $SV_e$ , corresponding to different  $\gamma_{11}$  values for the first sparse PC

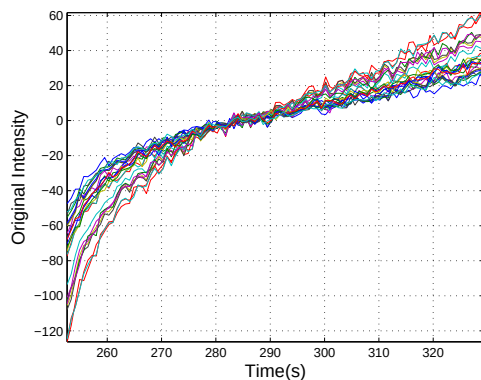


**Figure 5.** The loading of the first sparse principal component obtained for the sample OES spectra in Fig. 3 using SPCA with  $\gamma_{11} = 29000$

non-zero with several clusters of large values centered on the active OES channels. These clusters arise because of the spectral bleed between adjacent channels. In contrast, the SPCA loading has only 25 non-zero entries and these are in four distinct clusters of points. Analysis of these four sets of points show that they are all highly correlated as can be seen in Fig. 7 (correlation coefficient  $> 0.99$ ). This is a direct consequence of the grouping effect that is a feature of EN-SPCA, that is, EN-SPCA has a tendency to give equal weighting to strongly correlated variables and, as such, selects all the correlated variables as a group, rather than selecting a single representative example. This is useful when trying to identify groups of related variables, but is not ideal for a variable selection algorithm.



**Figure 6. The loading of the first principal component obtained for the sample OES data in Fig. 3 using PCA**



**Figure 7. Intensity changes of the nonzero loadings (over time) for the first sparse PC of the OES data**

## 5 Conclusions

This paper introduces SPCA as a variable selection tool for the identification of key variables in large data sets. SPMSE has been proposed as a measure that better reflects the estimation accuracy of SPCA, given the sparse structure of the model. Using analysis of OES data from a plasma etch chamber the main features of SPCA have been illustrated, particularly in relation to how it provides a trade-off between variance explained and a sparse representation. The existence of a grouping effect in the selection of variables has also been highlighted as a weakness of the method.

## Acknowledgments.

The authors gratefully acknowledge the financial support of Enterprise Ireland (grant IP/2006/0325).

## References

- [1] J. Cadima and I. T. Jolliffe. Loadings and Correlations in the Interpretation of Principal Components. *Journal of Applied Statistics*, 22(2):203–214, 1995.
- [2] A. D’Aspremont, L. Ghaoui, M. I. Jordan, and G. Lanckriet. A Direct Formulation for Sparse PCA Using Semidefinite Programming. *SIAM Review*, 49(3):434–448, 2007.
- [3] P. Geladi and B. R. Kowalski. Partial Least-squares Regression: a tutorial. *Analytica Chimica Acta*, 185:1–17, 1986.
- [4] D. Gervini and V. Roushon. Criteria for Evaluating Dimension-reducing Components for Multivariate Data. *American Statistician*, 58(1):72–76, 2004.
- [5] J. N. R. Jeffers. Two Case Studies in the Application of Principal Component Analysis. *Applied Statistics*, 16(3):225–236, 1967.
- [6] I. T. Jolliffe, N. T. Trendafilov, and M. Uddin. A Modified Principal Component Technique Based on the LASSO. *Journal of Computational and Graphical Statistics*, 12(3):531–547, 2003.
- [7] F. H. Kaiser. The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psychometrika*, 23:187–200, 1958.
- [8] K. Sjöstrand, M. B. Stegmann, and R. Larsen. Sparse Principal Component Analysis in Medical Shape Modeling. In *Proceedings of SPIE*, volume 6144, page 61444X, March 2006.
- [9] R. Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of Royal Statistical Society B*, 58(1):267–288, 1996.
- [10] S. K. Vines. Simple Principal Components. *Applied Statistics*, 49(4):441–451, 2000.
- [11] H. Zou and T. Hastie. Regularization and Variable Selection via the Elastic Net. *Journal of Royal Statistical Society B*, 67(2):301–320, 2005.
- [12] H. Zou, T. Hastie, and R. Tibshirani. Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, 15(2):265–286, 2006.

## A Data Mining Approach for Finding Optimal Discount of Retail Assortments

Maryam Nafari

Department of Industrial Engineering  
Amirkabir University of Technology  
424, Hafez Avenue, Tehran, Iran  
Maryam\_nafari@aut.ac.ir

Jamal Shahrabi

Department of Industrial Engineering  
Amirkabir University of Technology  
424, Hafez Avenue, Tehran, Iran  
Jamalshahrabi@aut.ac.ir

### Abstract

*Due to the recent competition in the retailing industry, retailers are striving to improve their operations in order to run their stores more efficiently. One of the most important factors that encourages customers to buy products is discount. The effects of discount on sales have rarely been dealt with academically. Moreover, in few previous researches in this case, the temporal characteristics of product discount have not been noticed. The problem addressed in this paper is the consideration of products' discounts in discovering association rules in different time intervals that a specific discount appears on a specific product and finding optimal discount for each product with the aim of maximizing total profits. Additionally, experiments on real world data demonstrate the effectiveness of the proposed approach.*

### 1. Introduction

Recently, the capabilities of generating and collecting data have been increasing rapidly. Widespread use of bar codes for most commercial products, the computerization of many business, and the advance in data collection tools have provided us with huge amount of retail data. This explosive growth in data and databases has generated an urgent need for data mining techniques and tools [7, 11] that can extract implicit, previously unknown and potentially useful information from data in data storages. One of the most popular data mining approaches is "association rules", which is commonly applied to analyze market baskets to help managers to determine which items are frequently purchased together by customers [5, 15]. It is, originally, presented by Agrawal et al. [1]. It studies the frequency of items occurring together in transactional database. Generally,

there are two phases for mining association rules. In the first phase, we find large itemsets, whose supports are larger than or equal to the user-defined minimum support. In the second phase, we use large itemsets generated in the first phase to generate effective association rules. An association rule is effective, if its confidence is larger than or equal to the user-defined minimum confidence. The problem of discovering association rules progressively followed by refinements, generalizations, improvements and extensions, including generalized association rules and multi-level patterns [10, 14, 19, 24]; finding association rules with categorical and numeric attributes [4, 12, 13, 21, 22, 23]; qualified association rules [16]; temporal association rules [3, 9, 20]; spatial association rules [18, 25].

Standard association rules algorithms [2, 17] applied in market basket analysis mainly determines what products customers purchase together in a static fashion. In fact, dynamic changes in discounts rates on products and its effect on sales have been disregarded. In previous work [8], patterns with discount information are discovered. They, however, assume that each product may have different discounts over different time periods, but the periods are fixed for all products. That is, if our pre-specified period is 2 weeks, the discount rate on items A and B will not be changed in corresponding period. But, in practical situations, the period in which item A has a fixed discount rate could differ from item B. For example, in first week, item A could be sold with 10% discount and in the next week, with 20% discount. On the other hand, item B could be sold with 30% discount in the entire 2 weeks. Moreover, in previous studies [6, 8], optimal discount for assorted products has not been extracted.

In this paper, the best discount for each product is selected based on the total profitability that it generates, including profits from cross-selling.

To deal with this problem a more sophisticated algorithm is developed. Thus, our approach needs a differently formatted data warehouse that must have discount information and temporal content. Such information is available, since transactions not only record purchased items, but also their discounts and the time of purchase. This approach could help reveal reasons behind purchasing certain items together frequently. It may be due to the discount rate on each item or other customer behavior.

The proposed procedure of finding optimal discount for each product begins with an algorithm that discovers temporal association rules with discount information. Next, the procedure proceeds to product discount selection in which the profits of frequent itemsets are considered.

The rest of the paper is organized as follows. In section 2, we define the problem. In section 3, a new algorithm for mining association rules with consideration of product discount is proposed. In section 4, the product discounts selection optimization model is developed. In section 5, the results of the empirical study are presented. Finally, in section 6 we draw conclusions.

## 2. The problem and definitions

Let  $I = \{i_1, i_2, \dots, i_n\}$  denote the set of items. Consider a transaction database  $D$ , which consists of a set of transactions  $\langle Tid_i, s_i, t_{Tid_i} \rangle$ , where  $Tid_i$  is a transaction identifier associated with each transaction,  $s_i$  is a set of items appearing in a transaction ( $s_i \in I$ ), and  $t_{Tid_i}$  is a time epoch expressing the time that  $Tid_i$  occurs (see Table 1).

In our approach, we partition the discounts rates into intervals and denote each discount rate included in  $i$ th interval by  $dr_i$ . For example, assume that we have five discount intervals as follows: [0%, 10%], (10%, 20%], (20%, 30%], (30%, 40%], (40%, 50%] and (50%, 100%]. Therefore, we have a set of discounts rates denoted by  $Dr = \{dr_1, dr_2, \dots, dr_n\}$ . Also, if  $i < j$  then  $dr_i < dr_j$ .

**Table 1: Transaction table D**

Tid	Item	Time
1	ABCD	2
2	ABE	7
3	AC	15
4	BED	20
5	ABCD	21
6	C	23
7	AC	26

The discount list of products is recorded in table  $P$ , which consists of  $\langle i_i, dr_p, T_{i,dr_p} \rangle$ , and expresses the period  $T_i$  in which item  $i_i$  has  $dr_p$  discount (see Table 2).

**Definition2.1:** Let  $X$  be an itemset which is a subset of  $I$ . A set of transactions in  $D$  containing  $X$  indicated by  $V(X, D) = \{s | s \in D \wedge X \subseteq s\}$ . If  $|X| = k$ , we call it  $k$ -itemset.

**Definition2.2:** Let  $PD = \{ps_1, ps_2, \dots, ps_d\}$  be transaction database  $D$  incorporated with discount information of discount table  $P$ . In fact, each  $ps_i$  indicates that each item in  $s_i$  has appeared with how much discount when it was purchased.

Let  $d$ -itemset  $X_d$  be an itemset carrying discount information and denoted by a set of pairs  $(i_x, dr_p)$ ; we can define two separate sets:  $X_d^1 = \{i_i | (i_x, dr_p) \in X_d\}$  and  $X_d^2 = \{dr_p | (i_x, dr_p) \in X_d\}$ . If  $|X_d| = k$ , we call it  $k$ - $d$ -itemset. The set of transactions in  $PD$  containing  $d$ -itemset  $X_d$  indicated by  $(X_d, PD) = \{ps | ps \in PD \wedge X_d \subseteq ps\}$ .

**Table 2: Discount list P**

Item	Discount rate	Time
A	$dr_2$	[0,8]
A	$dr_1$	[9,20]
A	$dr_2$	[21,30]
B	$dr_3$	[0,10]
B	$dr_1$	[11,30]
C	$dr_2$	[0,15]
C	$dr_4$	[16,22]
C	$dr_5$	[23,30]
D	$dr_4$	[0,8]
D	$dr_2$	[9,22]
D	$dr_1$	[23,30]
E	$dr_1$	[0,9]
E	$dr_3$	[10,20]
E	$dr_2$	[21,30]

**Definition2.3:** As we mentioned in definition 2.1, the time interval in which item  $i_i$  has  $dr_p$  discount, denoted by  $T_{i,dr_p}$ . It is simply an interval  $[t_i, t_j]$ , where  $t_i < t_j$ . Therefore, we can express the lifespan of  $k$ - $d$ -itemset  $X_d$ , with  $k > 1$ , by  $T_{X_d}$  and define it as follows; suppose  $T_{X_d}$  is an interval indicated by  $[t_{x_{d_i}}, t_{x_{d_j}}]$ , where  $t_{x_{d_i}} = \max\{t_i | [t_i, t_j] \text{ is the time period of an item } i_i \text{ in } X_d\}$  and  $t_{x_{d_j}} = \min\{t_j | [t_i, t_j] \text{ is the time period of an item } i_i \text{ in } X_d\}$ . It shows the period that  $X_d$  may occur. If  $t_{x_{d_j}} < t_{x_{d_i}}$ , means items in  $X_d$  do not have time intersection and this itemset is not

feasible. We also should define  $PD_{X_d} = \{ps | ps \in PD \wedge \text{the time of } ps \text{ is in time interval } T_{X_d}\}$  as a set of transaction in  $PD$  where  $X_d$  may occur.

For example, by considering Table 2, we have  $T_{A,dr_2} = [0,8] \cup [21,30]$ ,  $T_{B,dr_3} = [0,10]$ , and  $T_{C,dr_2} = [0,15]$ . Suppose  $X_d = \{(A, dr_2)(B, dr_3)(C, dr_2)\}$ , so we have  $T_{X_d} = [0,8]$  and  $Tid_{X_d} = \{1, 2\}$ .

**Definition2.4:** The global support of itemset  $X$  in database  $D$ , denoted by  $s(X, D)$ , is the fraction of transactions in  $D$  that contains  $X$ . Hence, we have  $s(X, D) = |V(X, D)| / |D|$ . Given a global support threshold  $\sigma_g \in [0, 1]$ ,  $X$  is a frequent itemset if  $s(X, D) \geq \sigma_g$ .

**Definition2.5:** The local support of extended itemset  $X_d$  in database  $D$ , denoted by  $ls(X_d, PD)$ , is the fraction of extended transactions in  $PD$  that contains  $X_d$ . Thus, we have  $ls(X_d, PD) = |V(X_d, PD)| / |PD|$ . Given a local support threshold  $\sigma_l \in [0, 1]$ ,  $X_d$  is frequent d-itemset if  $ls(X_d, PD) \geq \sigma_l$ .

**Definition2.6:** The real support of extended itemset  $X_d$ , denoted by  $rs(X_d, PD)$ , is the fraction of extended transactions in  $PD_{X_d}$  that contains  $X_d$ . Hence, we have  $rs(X_d, PD) = |V(X_d, PD_{X_d})| / |PD_{X_d}|$ . In order to determine the effect of discounts on sales we use  $EF = rs(X_d, PD) / s(X, D)$ . If this fraction results in a value which is greater than the threshold  $gf$ , it means this discount combination encourages sales. On the other hand, if  $EF$  is smaller than the threshold  $sf$ , it implies that this discount combination decrease sales. Note that  $sf$  and  $gf$  specified by users or experts.

**Definition2.7:** In this paper we need to remove  $d$ -itemsets  $X_d$  with very short  $T_{X_d}$ . Since a set of items may appear with a specific discount combination in a very short period of time, it is not reasonable to consider such itemset. For instance, assume that  $X_d = \{(A, dr_2)(C, dr_4)\}$ , so we have  $T_{X_d} = [21, 22]$ . Therefore with the help of definition 2.6, its real support is 100%. Without consideration of very short period that items A and C appeared with discounts rates  $dr_2$  and  $dr_4$ , respectively, we may mislead about its positive effect on sales. Therefore, we need a temporal support, denoted by  $ts(X_d)$ , and represented by  $ts(X_d) = |T_{X_d}|$ . Given a temporal support threshold  $\sigma_t$ ,  $T_{X_d}$  is big enough to judge about the real support of  $X_d$  if  $ts(X_d) \geq \sigma_t$ .

**Definition2.8:** If  $X_d$  is a large- $d$ -itemset and  $X'_d \subset X_d$ ,  $X'_d \neq \emptyset$ , The confidence of a rule  $X'_d \Rightarrow X_d - X'_d$ , where  $X_d$  and  $X'_d$  are set of pairs  $(i_x, dr_p)$ , is computed as follows:

$$C(X'_d \rightarrow X_d - X'_d) = \frac{ls(X_d - X'_d, T_{X_d}, PD)}{ls(X'_d, T_{X_d}, PD)},$$

Where  $ls(X_d - X'_d, T_{X_d}, PD)$  and  $ls(X'_d, T_{X_d}, PD)$  computed in time intervals  $T_{X_d}$  which is an intersection of periods of every pairs in  $X_d$  that they could possibly occur. We call a rule an interesting one, if  $C(X_d \rightarrow X'_d) \geq \sigma_c$ .

### 3. The Apriori-Td algorithm

The algorithm proposed in this paper differs from Apriori [2] and Apriori-like [8] algorithms for mining frequent itemsets. However, It takes advantage of approach represented in Chen et al.[8] and makes changes in algorithm in such a way that temporal concept is taken into account. We introduce two filtering operators,  $ts$  and  $ls$ . The former filters d-itemsets occurred in a very short period of time, since there is no sufficient data to judge on its support and consequently its effects. The later removes d-itemsets which is not frequent enough. Otherwise, we will face problem that a mining process produces a very large number of rules, thus proving to be of little value to the user.

Function Apriori-Td

- 1)  $GL_1 = \text{get-global-large-1-itemsets}(D);$
- 2) For( $k=1$ ;  $LL_{k-1} \neq \text{null}; k++$ ) {
- 3)   If ( $k==1$ ) {
- 4)      $LC_1 = \text{get-local-candidate-1-d-itemset}(GL_1, P);$
- 5)      $LL_1 = \text{get-large-local-k-d-itemset}(LC_1);$  }
- 6)   Else {
- 7)      $LC_k = \text{get-local-candidate-k-d-itemset}(LL_{k-1});$
- 8)      $LL_k = \text{get-large-local-k-d-itemset}(LC_k);$  }
- 9) }  $\cup LL_k;$

**Figure 1: Apriori-Td algorithm**

This approach needs more input data than Apriori-like algorithm. They would be discount list  $P$  and thresholds  $\sigma_l$  and  $\sigma_t$ .

We present the framework of the algorithm in Figure1. In the first step, we find large 1-itemsets, that exceed the global support threshold  $\sigma_g$ , labeled  $GL_1$ . In contrast with Apriori-like algorithm, we only find large 1-itemset in very first step and there is no need to extract  $k$ -itemsets,  $GL_k$ , in future phases. Since the main reason for extracting  $k$ -itemsets in Apriori-like algorithm is finding frequent  $k$ -d-itemsets, using Apriori-Td algorithm proposed in this paper we can acquire frequent  $k$ -d-itemsets directly from  $(k-1)$ -d-itemsets and we do not have to scan database once again.

Then in line 4 we join discount information to frequent items by the use of discount list  $P$ . The



description of the function illustrated in Figure 2. The function of making local candidate 1-itemsets,  $LC_1$ , differs from achieving  $LC_k$ , which is described later. In contrary, finding  $LL_1$  and  $LL_k$  are done by the same function. Hence, interesting discount patterns,  $LL_k$ , are discovered by applying *Function get-large-local-k-d-itemset*. As illustrated in Figure 3, in order to find a period in which k-d-itemset could possibly occur, we compute  $T_{x_d}$  (see definition 2.3). All itemsets which do not satisfy temporal condition  $|T_{x_d}| \geq \sigma_t$  and local threshold  $ls(X_d, PD) \geq \sigma_l$  are eliminated. Then, the number of transactions in corresponding period is determined and the computation of real support  $X_d$  and  $EF$  is done to find out the effect of such pattern on sales (see Figure 3).

Function	get-local-candidate-set-1-d-itemset( $GL_1, P$ )
1)	For each item $i_i$ in $X \in GL_1$
2)	From table $P$ Determine every different discount levels for $i_i$ ;
3)	Determine the time interval in which item $i_i$ appears with each discount level;

**Figure 2: Function of get local 1-d-itemsets candidates**

In line 7, we generate candidate of k-d-itemset from large  $(k-1)$ -d-itemset. The method is similar to Apriori algorithm, but it has some extra conditions. Inasmuch as our itemsets carrying extra information about their discount combination, when two  $(k-1)$ -d-itemset is joined, we have to consider whether this combination has actually happened. If  $T_{x_d''} \cap T_{x_d'} = \emptyset$ , it means  $x_d''$  and  $x_d'$  cannot occur simultaneously. This condition automatically satisfies the fact that one item could not occur with different discounts rates in a specific period. Moreover, the set of first components of pairs  $(i_x, dr_p)$  in  $x_d''$  and  $x_d'$  should not be equal (see Figure 4).

Function	get-large-local-k-d-itemset( $LC_k$ )
1)	For each pairs in $LC_k$ Find $T_{x_d}$
2)	If $ T_{x_d}  \geq \sigma_t$
3)	If $(ls(X_d, PD) \geq \sigma_l)$
4)	Insert $X_d$ to $LL_k$ ;
5)	Compute the number of transaction in $T_{x_d}$ ;
5)	Compute the real support $T_{x_d}, rs$ ;
6)	Compute $s(X_d^1, D)$ ;
7)	Compute the $EF = rs(X_d, PD) / s(X_d^1, D)$ ;
8)	If $(EF \geq gf)$ this discount combination increases sales;
9)	If $(EF \leq sf)$ this discount combination decreases sales;

**Figure 3: Function of get large k-d-itemset**

After finding large k-d-itemsets, the set of association rules can be derived from  $\cup LL_k$ . Results obtained from this approach are more complete than previous work because time information has been included. In this algorithm, generating association rules has a kind of time constrain and the rules is derived by the use of definition 2.8.

Function	get-local-candidate-set-k-d-itemset( $LL_{k-1}$ )
Each two $(k-1)$ -d-itemset $x_d', x_d'' \in LL_{k-1}$ join to construct a k-d-itemset $x_d$ , providing that (1) $x_d'^1 \neq x_d''^1$ , (2) $x_d''$ and $x_d'$ share $(k-2)$ pairs items $(i_x, dr_p)$ , (3) $T_{x_d''} \cap T_{x_d'} \neq \emptyset$ , (4) all $(k-1)$ -d-subsets of $x_d$ should be in $LL_{k-1}$ .	

**Figure 4: Function of get local k-d-itemsets candidates**

Example 3.1: In order to demonstrate the proposed Apriori-Td algorithm, a simple example is presented. Assume the transaction database  $D$  and discount list  $P$  are as shown in Tables 1 and 2, respectively. Also suppose that the predefined parameters are as follows:  $\sigma_g=0.4$ ,  $\sigma_l=0.2$ ,  $\sigma_t=3$ ,  $\sigma_c=0.5$ ,  $sf=1.3$ , and  $gf=0.7$ .

The large 1-itemsets satisfying  $\sigma_g=0.4$  is derived in the first step. As a result, we have  $A, B, C$ , and  $D$  as frequent global items with global supports 0.71, 0.57, 0.71, and 0.42, respectively. After each frequent item incorporated to their discount level, the computation of  $T_{x_d}$  and  $ls$  are done to find large 1-d-itemsets for the next step (see Table 3).

**Table 3: Generate  $LC_1$  from  $GL_1$**

$LC_1$	Time	$ T_{x_d} $	$ls$
$(A, dr_2)$	$[0,8] \cup [21,30]$	19	0.57
$(A, dr_1)$	$[9,20]$	12	0.14
$(B, dr_3)$	$[0,10]$	11	0.28
$(B, dr_1)$	$[11,30]$	20	0.28
$(C, dr_2)$	$[0,15]$	16	0.28
$(C, dr_4)$	$[16,22]$	7	0.14
$(C, dr_5)$	$[23,30]$	8	0.28
$(D, dr_4)$	$[0,8]$	9	0.14
$(D, dr_2)$	$[9,22]$	14	0.28
$(D, dr_1)$	$[23,30]$	8	0.0

We must calculate the real support of items with that specific discount level in the corresponding time interval so that we are able to determine their effects. As shown in Table 4, when item  $B$  appears with  $dr_3$  discount may further encourage customers to purchase it rather than other discount levels. On the other hand,



its  $dr_1$  discount may lead to the sales reduction. Item  $A$  with  $dr_2$  discount does not have significant influence on sales and this combination is neutral.

**Table 4: generate  $LL_1$  from  $LC_1$**

$LL_1$	Time	$rs$	$s$	$EF$	Effect
$(A, dr_2)$	$[0,8] \cup [21,30]$	0.8	0.71	1.13	
$(B, dr_3)$	$[0,10]$	1	0.57	1.75	+
$(B, dr_1)$	$[11,30]$	0.4	0.57	0.7	-
$(C, dr_2)$	$[0,15]$	0.67	0.71	0.94	
$(C, dr_5)$	$[23,30]$	1	0.71	1.4	+
$(D, dr_2)$	$[9,22]$	0.67	0.42	1.59	+

**Table 5: Generation of large 2-d-itemsets**

$LC_2$	Time	$ T_{x_d} $	$ls$	$s$
$(A, dr_2; B, dr_3)$	$[0,8]$	9	0.28	0.42
$(A, dr_2; B, dr_1)$	$[21,30]$	10	0.14	0.42
$(A, dr_2; C, dr_2)$	$[0,8]$	9	0.14	0.57
$(A, dr_2; C, dr_5)$	$[23,30]$	8	0.14	0.57
$(A, dr_2; D, dr_2)$	$[21,22]$	2		
$(B, dr_3; C, dr_2)$	$[0,10]$	11	0.14	0.28
$(B, dr_3; D, dr_2)$	$[9,10]$	2		
$(B, dr_1; C, dr_2)$	$[11,15]$	5	0.0	0.28
$(B, dr_1; C, dr_5)$	$[23,30]$	8	0.0	0.28
$(B, dr_1; D, dr_2)$	$[11,22]$	12	0.28	0.42
$(C, dr_2; D, dr_2)$	$[9,15]$	7	0.0	0.28

$LL_2$	Time	$rs$	$s$	$EF$	Effect
$(A, dr_2; B, dr_3)$	$[0,8]$	1	0.42	2.38	+
$(B, dr_1; D, dr_2)$	$[11,22]$	0.67	0.42	1.59	+

From  $LL_1$ , we get 2-d-itemsets candidates. As earlier mentioned, owing to the fact that there exist situations where specific discount combinations of itemsets do not simultaneously fall within the same database time interval (lack of intersection), hence, they do not satisfy the temporal conditions to join together. For example, there might not be a situation where  $B$  is sold with discount level  $dr_3$  and  $C$  with

$dr_5$ . So  $(B, dr_3; C, dr_5)$  is considered obsolete. It is illustrated in Table 5,  $(A, dr_2; D, dr_2)$  and  $(B, dr_3; D, dr_2)$  may occur in very short period of time, so for these two 2-d-itemsets, the program does not continue computing their local support.

Finally, large 2-d-itemsets, which their local supports exceed local threshold, are extracted from  $LC_2$  and the result reveals that when item  $A$  is sold with  $dr_2$  discount and item  $B$  with  $dr_3$ , customers purchase these two more frequently. Also,  $(B, dr_1; D, dr_2)$  has a positive influence on sales.

It is the time for finding interesting rules, which their confidence is greater than  $\sigma_c$ . For example, consider the case when  $(A, dr_2) \Rightarrow (B, dr_3)$ , the confidence is  $\left(\frac{2}{2}\right) / \left(\frac{2}{2}\right) = 1$ . Note that local support for  $(A, dr_2; B, dr_3)$  and  $(A, dr_2)$  are computed in time intervals  $[0, 8]$ , which this combination could possibly occur.

After the execution of this algorithm, we have frequent  $d$ -itemsets and also their effects on sales. In the next step, the zero-one programming is presented to find the optimal product discount with the aim of maximizing the profits of frequent  $d$ -itemsets.

#### 4. The product discount selection procedure

In this study, by the use of Apriori-Td algorithm, we can specify which products with how much discount are frequently bought by customers at the same market baskets. In this section, a microeconomic zero-one integer programming model is introduced for product discount selection. That is, the proposed model provides us the products that should be on sale and of course, the discount rates of these selected products. The results of the optimization problem may express that some products should not be sold with discount.

In the first step of the procedure, the gross margins of frequent itemsets with consideration of the product discount are estimated. The profit estimation follows the idea of Brijs et al.[6]. In their approach, besides considering individual profit generated by a specific product, the cross-selling effects with other products in the assortment are taken into account. In fact, Brijs et al. addressed the product assortment problem. However, in this paper, we assume that the products are assorted. So, we deal with the problem of finding best combination of discounts for assorted products. The constructed model is able to select an optimal discount for each product which yields the maximum overall profit. By using frequent  $d$ -itemsets, the model takes into account cross-selling effects.

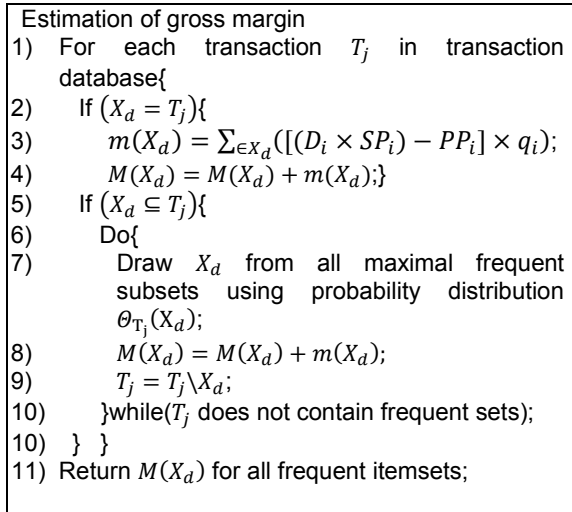
Before specifying the microeconomic optimization model, we describe the method of estimating profit of frequent itemsets [6].

#### 4.1. Estimating the gross margin

The parameters of the method are as follows:

$T_j$	Items and their discount level included in the $j$ th transaction
$F_l$	The collection of all frequent $d$ -itemsets of $T_j$
$SP_i$	The selling price of product $i$
$PP_i$	The purchase price of product $i$
$D_i$	The discount rate of product $i$
$q_i$	Quantity of product $i$ that was purchased in $T_j$
$X_{dmax}$	The maximal frequent $d$ -itemset in the $j$ th transaction
$Y_{dmax}$	The second maximal frequent $d$ -itemset in the $j$ th transaction
$\theta_{T_j}(X_d)$	The probability of selecting $X_d$ in $T_j$ to allocate gross margin, $\theta_{T_j}(X_d) = \frac{localSupport(X_{dmax})}{\sum_{Y_{dmax} \in T_j} localSupport(Y_{dmax})}$
$T_j/X_d$	Items and their discount level included in $j$ th transaction after eliminating frequent itemset $X_d$
$m(X_d)$	The profit of products in frequent $d$ -itemsets $X_d$ , $m(X_d) = \sum_{i \in X_d} [(D_i \times SP_i) - PP_i] \times q_i$
$M(X_d)$	The summation of $m(X_d)$

The method of estimating the profit of frequent itemset is illustrated in Figure 5.



**Figure 5: Function of estimation of gross margin**

#### 4.2. The discount selection model

In this section, we propose an optimization problem for discovering the best discount combination for products.

The goal of the optimization problem maximizes the overall profitability of the frequent  $d$ -itemsets. The key point in this paper is that not only are the frequency and quantity of a purchased product with a specific discount level taken into account, but also the effect of such a discount level on selling other products is considered.

The parameters of the model are given in below:

$M_i(X_d)$	The estimated gross margin of $i$ th frequent $d$ -itemsets
$LL_i$	The set of items incorporated with their discount level in the $i$ th frequent $d$ -itemset

Decision variables:

$$p_i = \begin{cases} 1, & \text{if any item in frequent d-itemset } LL_i \text{ is selected;} \\ 0, & \text{otherwise.} \end{cases}$$

$$dis_{jl} = \begin{cases} 1, & \text{if item } j \text{ with discount level } l \text{ is selected;} \\ 0, & \text{otherwise.} \end{cases}$$

The mathematical model of product discount selection is as follows:

Maximize  $\sum_i M_i(X_d) p_i$

Subject to:

- (1)  $dis_{jl} \geq p_i \quad \forall i, \forall j_l \in LL_i$
- (2)  $\sum_i dis_{jl} \leq 1 \quad \forall j$
- (3)  $p_i \in \{0,1\}, dis_{jl} \in \{0,1\}$

Constraint (1) ensures that once a frequent  $d$ -itemset is selected, the product with their discount level in this frequent itemset have to be selected. Constraint (2) forces the model to select just one discount level for each assorted product. Constraint (3) limits the decision variables to be binary.

If no specific discount is selected for an item, it means that it does not exist in any frequent itemsets. One probable reason for this fact is that it is sold rarely with any discount rate. So, it does not appear in frequent itemsets. The second likely reason is that its sell is not low in whole, but it is not sensitive about its discount and it is sold uniformly in almost low rate. So, any discount rate could not have great influence on its

sale. So, it would be better not to put this item on sale and sell it with its original price.

## 5. The experiment

In this section, we use a real dataset to assess the performance of proposed model for finding optimal product discount. The Apriori-Td algorithm was implemented in VB 2000. The real data set is derived from the retail data of chain supermarket in Iran over a period of 4 months in Microsoft SQL Server 2000. Each transaction records purchase data, purchase time, purchased items, prices, discounts, and quantities. These transactions consist of 1241 different products and totally there are 7305 sales transactions.

In the first phase of the procedure, in order to discover frequent  $d$ -itemsets, we partition the discounts rates into 6 intervals, as shown in Table 6. These frequent itemsets work as input for optimization model for finding best product discount in the second phase. The minimum global support, local support, and confidence are set to 0.3, 0.2, 0.1. The minimum temporal support is set to 3 days. These thresholds are set very low in order to discover enough association rules for further analysis. By the use of Apriori-Td algorithm, 1871  $d$ -frequent itemsets were obtained.

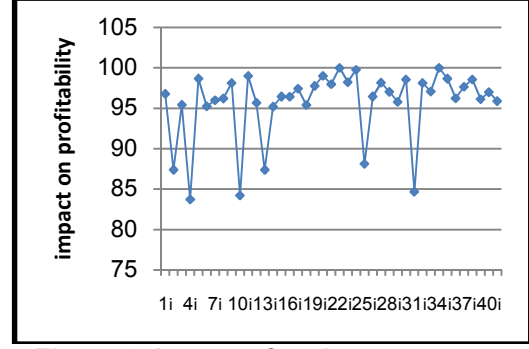
This approach uses only  $d$ -frequent itemsets. The discovery of association rules, however, will be helpful for interpreting the output of the optimization model.

**Table 6: Discount level**

Level	dr <sub>1</sub>	dr <sub>2</sub>	dr <sub>3</sub>	dr <sub>4</sub>	dr <sub>5</sub>	dr <sub>6</sub>
Dis. rate(%)	0	1-19	20-39	40-59	60-79	>80

The optimization model can exploit cross-sales potential between items with different discount rates. The effect of a product with a selected discount rate on total profitability can be assessed by means of sensitivity analysis. That is, a specific discount rate, which is assigned to a product by the use of proposed optimization model, is replaced by the best discount rate outside the result of the optimization model and its impact on profitability can be observed through its contribution to increasing the objective function. The best discount rate for a product outside the optimization model is the discount rate that has the greatest impact on cross-selling.

Figure 6 illustrates the profit impact of the replacement of each selected discount rate. It indicates that the discount rates for some of the products such as  $i_2$ ,  $i_4$ ,  $i_{10}$ ,  $i_{13}$ ,  $i_{25}$ , and  $i_{31}$  have great influence on profitability and should not be replaced by other discount rates.



**Figure 6: Impact of replacement on profitability**

In this approach, although some of the selected discount rates have low total profit margin individually, they have considerable cross-selling opportunities with some of the other selected discount combinations. For instance, the empirical results indicate that 30% discount rate on *chips* does not provide high profit with regard to the total margin and its  $EF = 1.53$  (if  $fg = 1.4$ , its  $EF$  does not show strong positive effect on its sale). It is, however, selected as the optimal discount combination for *chips*. The reason is that 30% discount for *chips* encourages costumers to purchase other items (*yogurt*, 10%; *ketchup*, 0%) with lower discounts rates more frequently and in high quantity. So, it makes the total profit increase significantly. It, also, can easily be obtained from evaluation of the association rules.

$(Chips, 30\%) \Rightarrow (Yogurt, 10\%; ketchup, 0\%)$

Minimum local support=0.39

Minimum confidence=0.95

$(Yogurt, 10\%; ketchup, 0\%) \Rightarrow (Chips, 30\%)$

Minimum local support=0.39

Minimum confidence=0.83

The rules illustrate that when a customer buys *chips* with 30% discount he will often also buy *Yogurt* with 10% discount and *ketchup* with no discount. Other direction of the rule demonstrate that when a customer purchases *Yogurt* with 10% discount and *ketchup* with no discount in most of the time (0.85) he also purchases *chips* with 30%.

## 6. Conclusions and future work

In the first phase of this paper, we presented an algorithm for discovery association rules in market baskets with consideration of discount information and their temporal characteristics. In previous studies, this problem is not viewed as a temporal association rules mining. In real world, the discount rates on a product

could, however, change with different periods and these periods differ from product to product. In the second phase of the paper, we integrated the notion of frequent itemsets with some important microeconomic parameters and the best discount rate selected by means of zero-one integer programming. Finally, we used a real dataset from a retail data of chain supermarket to empirically validate our approach. The results indicate the effectiveness and usefulness of the approach.

In the future, we plan to improve the pruning conditions of the Apriori-Td algorithm. Also, we are going to extend our approach to shelf allocation problem.

## 7. References

- [1] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases", in *Proc. ACM SIGMOD International Conference on Management of Data*, Washington DC, USA, 1993, pp. 207-216.
- [2] R. Agrawal and R. Srikant, "Fast algorithm for mining association rules", in *Proc. 20<sup>th</sup> International Conference on very large data bases*, Santiago, China, 1994, 487-499.
- [3] J.M. Ale and G.H. Rossi, "An approach to discovering temporal association rules", in *Proc. ACM Symposium on Applied Computing*, Como, Italy, 2000, pp. 294-300.
- [4] R.J. Bayardo and R. Agrawal, "Mining the most interesting rules", in *Proc. 15<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1999, pp.145-154.
- [5] M. Berry, G. Linoff, "Data mining techniques: for marketing, sales, and customer support", *New York: Wiley*, 1997, pp. 464.
- [6] T. Brijs, B. Goethals, G. Swinnen, K. Vanhoof, and G. Wets, "A data mining framework for optimal product selection in retail supermarket data: the generalized PROFSET model", in *KDD-2000*, Boston, MA, USA, 2000, pp. 300-304.
- [7] M.-S. Chen, J. Han, and P.S. Yu, "Data mining: an overview from database perspective", *IEEE Transactions on Knowledge and Data Engineering*, 1996, pp. 866-883.
- [8] Y.-L. Chen, T.C.-K. Huang, and S.-K. Chang, "A novel approach for discovering retail knowledge with price information from transaction databases", *Journal of Expert Systems with applications*, 2008, pp.2350-2359.
- [9] X. Chen, I. Petrounias, "Mining temporal features in association rules", in *Proc. 3<sup>th</sup> European conference on Principles and Practice on Knowledge Discovery in Databases*, 1999, pp. 295-300.
- [10] S. Fortin and L. Liu, "An object-oriented approach to multi-level association rule mining", in *Proc. 15<sup>th</sup> International Conference on Information and Knowledge Management*, Rockville, USA, 1996, pp. 65-72.
- [11] W.J. Frawley, G. Piatetsky-Shapiro, and C.J. Matheus, "Knowledge discovery in databases: an overview", AAAI/MIT Press, 1991.
- [12] T. Fukuda, Y. Morimoto, S. Morishita, and T. Tokuyama, "Mining optimized association rules for numeric attributes", in *Proc. 15<sup>th</sup> ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, Canada, 1996, pp. 182-191.
- [13] T. Fukuda, Y. Morimoto, S. Morishita, and T. Tokuyama, "Data mining with optimized two dimensional association rules", in *Proc. ACM Transaction on Database Systems*, 2001, pp. 179-213.
- [14] J. Han and Y. Fu, "Discovery of multiple-level association rules from large databases", in *Proc. 21<sup>st</sup> International Conference on very large data bases (VLDB'95)*, Zurich, Switzerland, 1995, pp. 420-431.
- [15] J.W. Han, M. Kamber, "Data mining: concepts and techniques", San Francisco: Morgan Kaufmann, 2006, pp.770.
- [16] N. Jukic and S. Nestorov, "Comprehensive data warehouse exploration with qualified association-rule mining", *Journal of Decision Support Systems*, 2006, pp. 859-878.
- [17] M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, and A.I. Verkamo, "Finding interesting rules from large sets of discovered association rules", in *Proc. Third International Conference on Information and Knowledge Management*, 1994, pp. 401-407.
- [18] K. Koperski, J. Han, "Discovery of spatial association rules in geographic information databases", in *Proc. 4<sup>th</sup> International Symposium on Advance in spatial Databases*, 1995, pp. 47-66.
- [19] Y.-C. Lee, T.-P. Hong, and T.-C. Wang, "Multi-level fuzzy mining with multiple minimum supports", *Journal of Expert Systems with Applications*, 2008, pp. 459-468.
- [20] Y. Li, P. Ning, X.S. Wang, and S. Jajodia, "Discovering calendar-based temporal association rules", *Data & Knowledge Engineering*, 2003, pp. 193-214.
- [21] J.S. Park, M.-S. Chen, and P.S. Yu, "An effective hash-based algorithm for mining association rules", in *Proc. ACM SIGMOD International Conference on Management of Data*, San Jose, CA, 1995.
- [22] R. Rastogi, and K. Shim, "Mining optimized association rules with categorical and numeric attributes", in *Proc. IEEE Transactions on Knowledge and Data Engineering*, 1998, pp. 503-512.
- [23] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables", in *Proc. ACM SIGMOD International Conference on Management of Data*, Montreal, Canada, 1996, pp. 1-12, ACM Press.
- [24] S. Tsur, J. Ullman, S. Abiteboul, C. Clifton, R. Motwani, S. Nestorov, and A. Rosenthal, "Query flocks: a generalization of association-rule mining", in *Proc. ACM SIGMOD Conference*, Seattle, Washington, USA, 1998, pp. 1-12.
- [25] L. Wang, K. Xie, T. Chen, and X. Ma, "Efficient discovery of multilevel spatial association rules using partitions", *Journal of Information and Software Technology*, 2005, pp. 829-840.

## Identifying Stock Similarity Based on Episode Distances

Abhi Dattasharma  
Intermedia Softech Pvt. Ltd.  
Bangalore, India  
abhi@intermediasoftech.com

Praveen Kumar Tripathi  
Applied Research Group,  
Satyam Computer Services Ltd.  
Bangalore, India  
praveen\_tripathi@satyam.com

Sridhar Gangadharpalli  
Applied Research Group,  
Satyam Computer Services Ltd.  
Bangalore, India  
sridhar\_gangadharpalli@satyam.com

### Abstract

*Predicting stock market movements is always difficult. Investors try to guess a stock's behavior, but it often backfires. Thumb rules and intuition seems to be the major tools. One approach suggested that instead of trying to predict one particular stock's movement with respect to the whole market, it may be easier to predict a stock  $A$ 's movement based on another stock  $B$ 's movement, because  $A$  may get affected by  $B$  after  $B$ 's movement. This may provide the investor invaluable time advantage. It would be very useful if a general framework can be introduced that can predict such dependence between stocks based on any user defined criterion. This article attempts to lay down one such framework, where the stock time series is encoded as a binary string. This binary representation depends on the user defined criterion. The string distances between two such encoded time series has been used as a measure of dependence. Further, this technique has been used in the 'Pairs Trading strategy'; in fact, it is more powerful as varied user defined criterion can be handled while detecting similarity. The presented technique has been demonstrated with one typical user defined criterion.*

### 1. Introduction

Algorithmic trading [1] tries to identify possible "windows of opportunities" automatically, using algorithms, for performing stock trading, from the historical data. Stock data is a time series data [7], describes how the stock values behave over time. The most frequently used values for visu-

alizing and predicting stock movement over a time window (Week/Day/Time) are, price at trading opening time, maximum value on that window, lowest value on that window, price at closing time and total volume of transaction within that window.

One interesting related problem is to find out how one stock affects another. It is quite important from the viewpoint of investors; as it gives the investor an idea about how the markets are moving and also provides some unforeseen insights. As a simple example, if stock  $B$  has been found to closely follow stock  $A$  within three days period, steady increase in  $A$ , may provide the investor some idea about  $B$ 's growth. This  $A - B$  relation might provide some useful information along with  $B$ 's individual pattern.

In the well known *Pairs Trading strategy* [4], [21], two similar stocks are chosen so that, if the behavior of one stock outperforms the average performance it is sold short, whereas, if the behavior of one stock falls below average, it is bought in the hope that it will reach average behavior sometime. The key concept behind pairs trading strategy is that a well chosen pair tend to cancel each other's deviations and thus, a loss in one will be gained in the other. Evidently, it would be very useful if a general framework can be introduced that can predict such dependence between stocks, that could be used for pairs trading strategy. For pairs trading, the most commonly used index is a correlation coefficient [19], and investors often also check other metrics like relative strength index [22, 3] and Bollinger band [5, 3].

There have been some attempts to evaluate the stocks' similarity *e.g.*, indicators, price ratio and price comparison [3]. However, these have been largely for the purpose of finding strong stocks rather than dependences between

them. Researchers have used clustering techniques [12], association rule mining [20, 18, 16], and geometric properties [9, 6] to find stock similarity. The metric distances and methods based on topological properties and random matrix theory [8, 13] have also been used for the purpose. However, these methods are complicated and could not provide simple yet practical indicators useful for investors which are applicable both for short term and long term dependencies. Also, all these metrics invariably try to find similarity between stocks in terms on one fixed distance measure. In [10], this problem has been addressed and three simple metrics are introduced which effectively predicts two stocks' dependence on each other. However, those metrics were applicable only for a certain class of dependence (for example, if the opening price pattern of one stock follows the opening price pattern of another). In practice, an investor may need completely different criteria for measuring distances between two stocks depending on the requirement, time and market conditions. Naturally, it would be more useful if some generic framework could be devised that would enable an investor check for dependence between stocks with respect to a criterion defined by the investor.

In this article, a simple framework based on the concepts of events and episodes, first introduced in [17] is developed. It is shown that, given a user defined criterion, it is possible to encode any stock's time series data using a finite alphabet, and then the distance between these two encoded series efficiently measures dependence. Some practical results for one particular user defined criterion is also provided. It has been shown that the proposed technique can be used effectively for pairs trading; in fact, this technique is more powerful as it can model different user requirements while finding similarity. The data set has been taken from Standard and Poor 500 [2]. This data shows daily price and volume fluctuations (opening price, closing price, maximum price, minimum price and volume) over an year for some well-known stocks. Most of the stocks have a series of 252 numbers; thus we have a time series of length 252. This paper is organized as follows. In Section 2, we define events and episodes. Section 3 describes the framework. Section 4 gives results, and Section 5 concludes the paper.

## 2 Events and Episodes

As defined in [17], an event is an occurrence of a particular type which has a time stamp attached to it. Given a set  $\bar{E}$  of event types, an event  $\alpha$  is a pair  $(A, t) \in \bar{E} \times \mathcal{N}$ ,  $\mathcal{N}$  set of natural numbers, where  $A \in \bar{E}$  is an event type and  $t$  is an integer, the occurrence time of the event. Mathematically, a time series can be seen as an ordered sequence of events  $\{\alpha_i\} = \{(e_i, t_i)\}$  where  $e_i$  is an occurrence of a particular type and  $t_i$  is the time stamp. For example, for a stock time

series, the set of event types can be (daily opening price, daily closing price, daily trading volume, daily maximum price and daily minimum price) and the time series can be described by any of these four event types by giving the value associated with the day. An episode is a sequence of events. An episode is a triple  $(V, <, g)$  where  $V$  is a set of nodes,  $<$  is a partial order on  $V$ , and  $g : V \rightarrow \bar{E}$  is a mapping associating each node with an event type. In the above article, authors gave algorithms for identifying frequently occurring episodes in a sequence of events.

Therefore, each stock data time series can be viewed as a series of events, where each event is a value (the opening price, or closing price, or volume traded and so on) associated with a time. Now, a user can define an episode as a predefined sequence of events. For example, the user can define "three successive days of trading volume increase" as an episode  $(T_i < T_{i+1} < T_{i+2}, T_i = \text{Trading Volume value on } i^{\text{th}} \text{ day})$ .

Note that the definition implicitly assumes that an episode is defined completely only by the event type and their partial ordering. However, it may not be enough for defining somewhat complicated episodes where conditionals or temporal behavior comes into play, *e.g.*, the episode "closing price lower than opening price, occurring at least twice a week". The problem here lies in the fact that the episode does not only need a sequence of events, but also a quantifier on the time.

To evaluate the similarities between stock prices based on their time series data, the episodes in that data should be found in a time sequential order. In this article an episode is defined as a sequence of events together with a restriction on the values as well as the time stamps of the events. Thus, an episode is

$$E = \{(e_i, t_i) \in \bar{E} \times \mathcal{N} : f(e_i), g(t_i)\},$$

where  $f(e_i)$  and  $g(t_i)$  are functions of  $e_i$  and  $t_i$  respectively, and events are always found from left to right. Suppose  $O_i$  is the opening price on  $i^{\text{th}}$  day for the stock data. An example of  $f$  and  $g$  can be seen from the episode "Opening price of stock is more than 10 USD for two days in succession", which is equivalent to  $O_i > 10, O_j > 10, j = i + 1$ . Here the event type is opening price,  $f$  defines the relationship between opening price and the value 10, and  $g$  defines the relationship between  $i$  and  $j$ . Note that, as we are defining an episode in terms of a set of events detected from left to right, each event's position in time gets fixed and the ordering is automatic. We will call a sequence of events which satisfy the complementary conditions as a complementary episode.

It is possible that a wrong choice of  $f()$  and  $g()$  may generate very poor results. However, as we are relying on

experienced investor inputs, we expect  $f()$  and  $g()$  to be good. In fact, generating well behaved  $f()$  and  $g()$  may be a good future research direction.

Given a definition of an episode, we can find all occurrences of the episode from a time series. Below, we give an outline of the algorithm that finds all episodes, and its complements from a time series.

```

ALGORITHM FIND_EPISODES_FROM_STOCK_TIME_SERIES
Read time series $S(t)$, $t = 1, \dots, N$
Read episode def. $E = \{(e_i, t_i) : f(e_i), g(t_i)\}$
episode_index = 1
complementary_episode_index = 1
for $i = 1, \dots, N$ do
 Check if $(S(t_i), t_i)$ satisfies the conditions
 or complementary conditions
 If neither, push $(S(t_i), t_i)$ to candidate_list
 If satisfies conditions
 push candidate_list and $(S(t_i), t_i)$ to
 episode[episode_index]
 set episode_index += 1
 endif
 If satisfies complementary conditions
 push candidate_list and $(S(t_i), t_i)$ to
 complementary_episode_index[episode_index]
 set complementary_episode_index += 1
 endif
enddo

```

With this definition, the episode “closing price is lower than opening price, occurring at least twice in the same week” will be:

$$\{(C_i, t_i), (O_i, t_i), (C_j, t_j), (O_j, t_j) \text{ such that ,} \\ C_i < O_i, C_j < O_j, \\ |i - j| \leq 7, i, j \text{ not part of a member already}\}. \quad (1)$$

Where  $C_i$  and  $O_i$  are the closing and opening prices, respectively on  $i^{th}$  day. Similarly, an episode “the stock opening price increases by  $X$  and does not decrease by  $Y$  or more anywhere in the same period” will be described as

$$\{(O_k, t_k), \quad k = i, i + 1, \dots, j \text{ such that ,} \\ O_j - O_i \geq X, O_m - O_l < X, \forall m \neq j, \forall l \neq i, m > l \\ O_m - O_l < Y, \forall \{m, l\} \in (i, i + 1, \dots, j), m < l, \\ \text{no } k \in \{i, i + 1, \dots, j - 1\} \\ \text{a part of a member already}\}. \quad (2)$$

Note the definition of  $k$ ;  $j$  can be a part of two members and exactly two members which share the boundary.

### 3 Framework

Given such an episode definition  $E$ , it is possible to encode a stock time series  $S(t)$ ,  $t = 1, \dots, N$  as a binary

string. Suppose the user defined episode is  $E$ , viz., the set of events and their time stamps, subject to  $f(.)$  and  $g(.)$  as defined by the user. Now, consider every episode  $E$  present in the time series, which is a set  $\{(e_i, t_i)\}$ . To encode the stock time series in binary, for every event member of an episode, the corresponding value of the time series gets encoded to ‘1’, where as complementary event member gets ‘0’. Mathematically, suppose  $S(t)$ ,  $t = 1, \dots, N$  is the stock time series, and  $E'$  is the set of all episodes in  $S(t)$ , each episode  $E$  defined as  $\{(e_i, t_i) : f(e_i), g(t_i)\}$  found from  $S(t)$ .

$$\begin{aligned} &\text{Define } h : \{1, \dots, N\} \rightarrow \{0, 1\} \text{ such that,} \\ &\forall \text{ event } e = (e_i, t_i) \in E, \forall E \in E' \\ &h(t_i) = 1, h(t) = 0 \text{ otherwise.} \end{aligned} \quad (3)$$

Note that this encoding removes all dependence on amplitude and therefore, no normalization is needed.

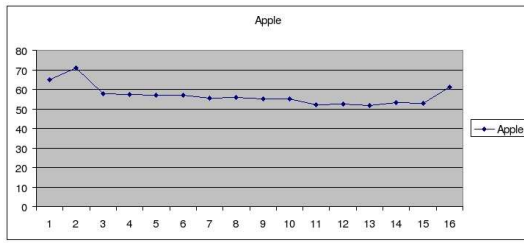
Further after binary encoding, string size is shortened by replacing any contiguous block of 1s in a single episode with a single 1, and any contiguous block of 0 in a single complementary episode with a single 0. This is meaningful in the following way. Consider the episode “the stock opening price increases by  $X$  and does not decrease by  $Y$  or more anywhere in the same period”. Thus, intuitively, the entire period of movement over which the price increased is one single episode instance, and the whole set should be seen as ‘1’. The function  $h$  as defined above replaces each time stamp of that period by a 1, and so replacing the entire block by a single 1 gives us the intuitive picture. Also note that, in case of two successive instances of episodes, we get two successive 1s.

```

ALGORITHM FIND_ENCODING_OF_STOCK_TIME_SERIES
Read episode description
 $E = \{(e_i, t_i) : f(e_i), g(t_i)\}$, $B = \phi$
for each $i = 1, \dots, N$,
($N = \text{length of time series}$)
do
 if t_i belongs to a member of episode
 description
 $B = B. '1'$
 else $B = B. '0'$
 enddo
for each episode do
 replace contiguous block of ‘1’ by a single
 ‘1’
enddo
for each complementary episode do
 replace contiguous block of ‘0’ by a single
 ‘0’
enddo

```

Figure 1 shows a snapshot of Apple’s stock opening value data over 16 days. Let us try to encode this with re-



**Figure 1. Stock time series snapshot for Apple**

spect to the episode “the stock opening price increases by 1 USD and does not decrease by 1 USD or more anywhere in the same period”. Note that the first two points show a significant rise, so  $h$  will return 11. Then, points 2 and 3 show a marked decrease, so it will return 00. Points 3 to 7 show another marked dip, returning 00000. Points 7 to 11 shows yet another large dip, giving 00000. Points 11 to 14 show a rise, giving 1111. Points 14 to 16 show a large rise, giving 111. Overall,  $h$  returns the following binary string: 1100000000000111111. Finally, we collapse each member, which gives 100011. Intuitively, this means that Apple’s stock opening value showed a marked rise, followed by three successive significant drops, followed by two marked rises.

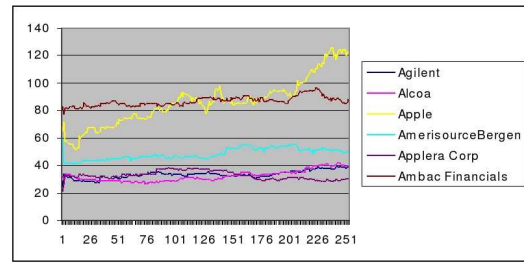
Now, after binary encoding of the stock time series, the similarity between two stocks  $S_1$  and  $S_2$  with respect to the episode definition can simply be seen as  $d(S_1, S_2)$ , where  $d(.,.)$  is a string distance measure, like Hamming distance. We use Levenshtein distance [15] and Jaro string distance [14] metric for measuring the distance between two strings.

## 4 Results

Quite clearly, the results are the most important part for evaluating a framework like the one defined above. We try to do that in this section. We have computed results for all pairs of companies across Standard and Poor 500 data set using the full range of 252 days’ data. However, lack of space prohibits a complete discussion across all pairs, so we give a few typical examples.

We define an episode similar to that used for the above example. We look at the opening prices of each stock, and define an episode as “the stock opening price increases by at least 1 USD and does not decrease by 1 USD or more anywhere in the same period”. This is a typical investor cue: is the stock price increasing, and is it increasing sufficiently? If it is, then it may be a good idea to follow it up and invest.

The process of finding the results is as follows. We took



**Figure 2. Stock time series for 6 stocks**

Standard and Poor 500 data set and selected all the companies which had 252 data points. Then we run our algorithm for every pair of companies with the given episode. The data spans over one year. As not all companies in the original data set had full 252 days’ data, the effective number of companies used in the test is 420. Also, observe that the same technique can be successfully used for finding similarities for a group of companies instead of just a pair.

On the basis of the above episode the stock series are first encoded in binary. The similarity between these encoded series is then estimated using Levenshtein and Jaro distances. In this article the threshold for similarity has been taken as 20 and 0.20 for Levenshtein and Jaro distances, respectively. Since the time series that has been used has 252 values for each stock data, when encoded, a binary string of the length of the same order is expected. Therefore, a distance of 20 or less will imply that the two strings are less than 10% different. Jaro metric is somewhat stronger as it considers transpositions, so the condition has been relaxed and a threshold of 0.2 is taken.

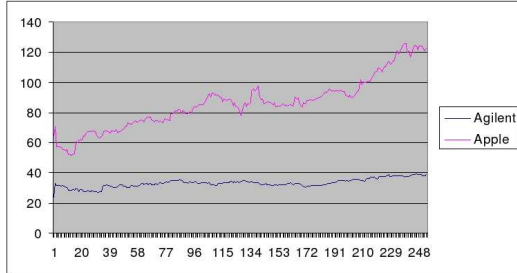
All the experiments have been conducted on Standard and Poor 500 data set. However, because of lack of space, we give only a selected set of results here. First, we take six companies which have the first six ticker symbols in S&P 500 data. Second, we take Starbucks, and show the results for some companies whose stock time series follow Starbucks closely. Third, we take Microsoft, and show the results for a set of companies whose stock time series follow Microsoft closely. The aim in this article is to show that hidden similarities often exist in completely unrelated areas, which an investor may not be aware about. In order to highlight the significance of the concept formulated in this article, the results of the proposed approach together with the correlation coefficients are provided.

Figure 2 shows the stock time series for 6 companies viz., Series 1: Agilent Technologies (IT), Series 2: Alcoa Inc. (Materials), Series 3: Apple (IT), Series 4: AmerisourceBergen Corp (Healthcare), Series 5: Applera Corp-Applied Biosystems Group (Healthcare), and Series 6: Ambac Financial Group (Financial).



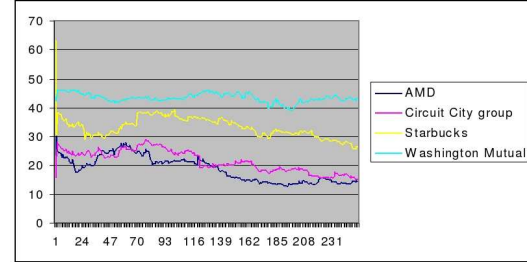
**Table 1. Values of Levenshtein, Jaro distance and correlation coefficient in an  $X/Y/Z$  form**

	Agilent	Alcoa	Apple	Amerisource	Applera	Ambac
Agilent	0/0/1	12/0.10/0.67	81/0.28/0.86	18/0.14/0.46	13/0.16/ - 0.15	33/0.25/0.59
Alcoa		0/0/1	78/0.30/0.79	14/0.11/0.63	17/0.18/ - 0.60	32/0.22/0.62
Apple			0/0/1	72/0.27/0.65	90/0.36/ - 0.29	51/0.22/0.65
Amerisource				0/0/1	20/0.22/ - 0.41	26/0.18/0.65
Applera					0/0/1	42/0.29/ - 0.24
Ambac						0/0/1

**Figure 3. Stock time series for Agilent and Apple stocks**

From Figure 2, we see that series 1, 2 and 5 are very close, series 4 is slightly off, and series 3 and 6 are quite far away. Interestingly, Agilent Technologies closely follows Alcoa, which are in completely different areas and therefore an investor might not have any cue of this fact unless specifically pointed out. Similarly, Apple seem to be very far from Agilent even though both are in IT sector. The string distances for these set of series are given in Table 1.

As we can see, the pairs (1, 2), (1, 5), (1, 4), (2, 4), (2, 5) are close. (4, 5) is just slightly off (expected, given that 4 is deviating from 5. This is an exact match with the graphically observed result. Note that there are some significant differences if we consider the correlation coefficient, e.g., according to the proposed method, investor should not look at Agilent-Apple pair, while the correlation coefficient suggests that they are close. From Figure 3 it is evident that Agilent is more or less stable in its movement, i.e., it moves up and down without too much fluctuations, as compared to Apple. Recall that the episode definition says that the stocks should increase by 1 USD *without any substantial fall in the same period*. Thus, the investor is looking for only *stable, not fluctuating* stocks that rise or fall steadily. In that respect, even though Apple and Agilent tends to show a high degree of correlation, they are *not good* candidates for a pair trading, and that is exactly what is reflected in the measure proposed. Thus, the string metric helped a cautious in-

**Figure 4. Stock time series for Starbucks stock vs few others**

vestor to detect candidates for a pairs trading with respected to the investor's requirement and not necessarily by a standard, fixed criterion. The other differences are explained similarly.

Figure 4 shows the time series for Starbucks and some similar companies. As is evident, AMD, Circuit City Group, Washington Mutual and Starbucks are very alike in graphs and so are their values in Table 2. However, the correlation coefficients show some disagreement. Consider the plots of AMD and Starbucks as in Figure 5. Recall that the investor is looking for stable movements without too much fluctuations. Note that in the initial period, the two stocks follow each other closely, and around 80th time step, AMD falls but Starbucks gains, making these a perfect pairs trading candidate. Later, the stocks move steadily without much fluctuations and they follow each other closely. Thus, the investor's criterion is well captured even though the correlation metric is ambivalent.

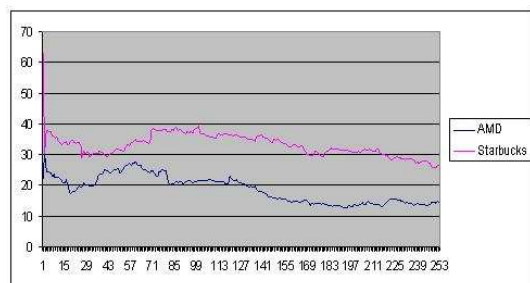
Figure 6 depicts the closeness of Bank of NY to Century Telephones, and that of Microsoft, Symantec and Time Warner to each other. The computed values also reflect these similarities faithfully in Table 3. Note that Time Warner and Microsoft work in different areas, and thus such a strong similarity, though exists and can be very helpful, is not easy to find for an investor. Interestingly, in this case also, the correlation coefficients differ from the string distance measure at times. Consider Figure 7. Here, for both

**Table 2. Values of Levenshtein, Jaro distance and correlation coefficient in an  $X/Y/Z$  form for Figure 4**

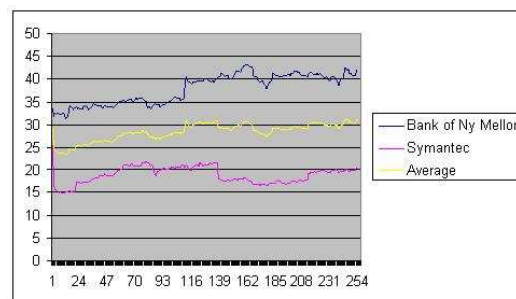
	AMD	Circuit City Group	Starbucks	Washington Mutual
AMD	0/0/1	12/0.098/0.82	10/0.12/0.50	11/0.14/0.31
Circuit City Group		0/0/1	12/0.085/0.59	10/0.11/0.29
Starbucks			0/0/1	14/0.135/0.29
Washington Mutual				0/0/1

**Table 3. Values of Levenshtein, Jaro distance and correlation coefficient in an  $X/Y/Z$  form, for Figure 6**

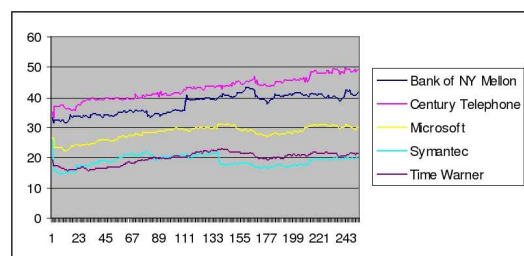
	Bank of NY Mellon	Century Telephone	Microsoft	Symantec	Time Warner
Bank of NY Mellon	0/0/1	08/0.16/0.89	07/0.16/0.76	06/0.12/0.04	10/0.21/0.84
Century Telephone		0/0/1	11/0.20/0.79	09/0.20/0.16	14/0.27/0.79
Microsoft			0/0/1	03/0.08/0.51	04/0.12/0.92
Symantec				0/0/1	06/0.15/0.33



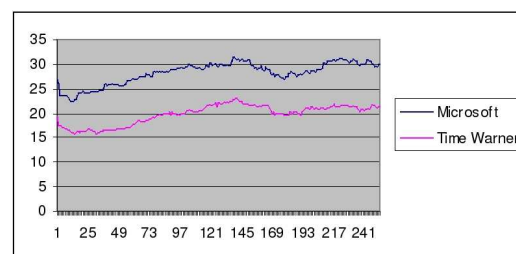
**Figure 5. Stock time series for AMD and Starbucks stocks**



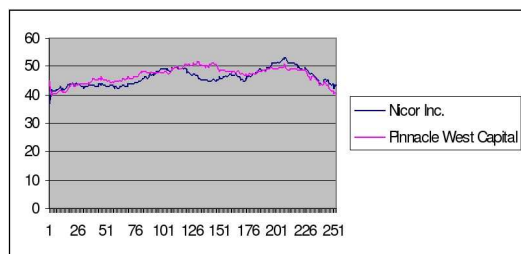
**Figure 7. Stock time series for Bank of NY and Symantec**



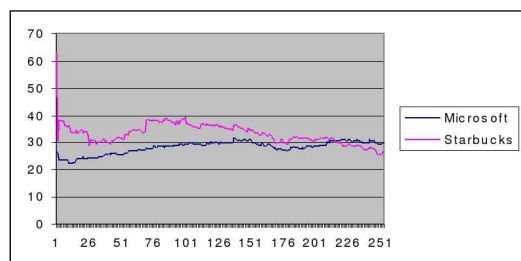
**Figure 6. Stock time series for Microsoft stock vs few other stocks**



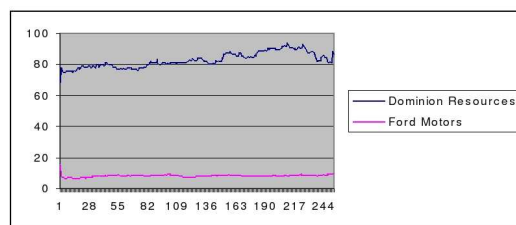
**Figure 8. Stock time series for Microsoft and Time Warner stocks, Distance/Correlation Computed 4/0.12/0.93**



**Figure 9. Stock time series for Nicor Inc. and Pinnacle West Capital stocks, Distance/Correlation Computed 12/0.11/0.729**



**Figure 10. Stock time series for Microsoft and Starbucks stocks, Distance/Correlation Computed 26/0.30/ - 0.075**



**Figure 11. Stock time series for Dominion Resources and Ford Motors stocks, Distance/Correlation Computed 54/0.38/0.18**

the stocks the steady fluctuations parts match; and when they diverge they do so by about the same amount from the average, making them an ideal pair for a pairs trade, which is exactly what the string metric suggested, while the correlation coefficient stayed conservative. For a better visualization, two one-one relationships plots are provided. Microsoft vs Time Warner group's similarity is shown in Figure 8, where as that of Nicor Inc. vs Pinnacle West's stocks in Figure 9. Microsoft vs Time Warner, seems to follow each other's pattern quite faithfully, and with a high correlation coefficient of 0.93 they are strong contenders for a pairs trading unit. Nicor vs Pinnacle West has a correlation coefficient of 0.729, and is therefore a sufficiently good pair for pairs trading. Observe the movement of Nicor and Pinnacle West; as Nicor falls Pinnacle grows and they tend to balance each other, making a pairs trading justified from an investor's perspective. In both these cases the similarities are pronounced even though the companies are from different areas.

To illustrate that the algorithm identifies difference between stocks too, the graphs in Figure 10 and 11 show two pairs of very different time series, namely, for Microsoft vs Starbucks, and Dominion Resources vs Ford Motors, respectively. The calculations also show a large separation. Microsoft vs Starbucks show a negative (though close to zero) correlation coefficient, while Dominion and Ford shows a low correlation, making them unsatisfactory pair for a pairs trading.

Finally, we take one example from an expert opinion [11]. The stocks under consideration are Wells Fargo bank and Washington Mutual. In [11], it was shown that these two were good candidates for a pairs trading though it may not be obvious. Our method gave the distance between these two as 16/0.28, whereas the correlation coefficient is 0.47. Thus, the correlation coefficient is not very positive, whereas in our case the indication given by Levenshtein measure is strong, helping the investor to get a better idea about the possible pair.

## 5 Conclusions

The similarity computation between the stocks could be very useful as the dependence between two stocks can reveal some unforeseen insights. This may be more useful if such similarities are computed on the basis of some user defined criterion. It has been shown in this article that such criterion may be defined mathematically. On the basis of such a definition the stock time series is encoded into binary followed by the binary distance computation between them. This computation gives the measure of similarity between the stocks. One such criterion has been used in this article for the illustration using the Standard and Poor 500 data. The results have been found quite significant. This strategy has been explored for the pair trading with some very significant results. Here, it is to be noted that we find the relationship between two stocks without determining if there is a cause and effect relationship. We plan to look at that in future. It is possible to include some other events not dependent on stock time series explicitly, but have significant impact on its growth, e.g., company policy. Such a case is under consideration.

## 6 Acknowledgments

The authors wish to thank anonymous reviewers for their useful comments.

## References

- [1] [http://en.wikipedia.org/wiki/algorithmic\\_trading](http://en.wikipedia.org/wiki/algorithmic_trading).
- [2] <http://kumo.swcp.com/stocks/>.
- [3] <http://www.incrediblecharts.com>.
- [4] D. R. Aronson. *Evidence Based Technical Analysis: Applying the Scientific Method and Statistical Inference to Trading Signal*. Wiley Trading, 2006.
- [5] J. A. Bollinger. *Bollinger on Bollinger Bands*. McGraw Hill, 2001.
- [6] B. Bollobas, D. G. Gautam Das, and H. Mannila. Time-Series Similarity Problems and Well-Separated Geometric Sets. In *13<sup>th</sup> ACM Symp. Computational Geometry*, pages 454–456, 1997.
- [7] P. J. Brockwell and R. A. Davis. *Intro. to Time Series and Forecasting*, volume 2nd edition. Springer-Verlag, NY, 2002.
- [8] C. Coronnello, M. Tumminello, F. Lillo, S. Micciche, and R. N. Mantenga. Sector identification in a set of stock return time series traded at the london stock exchange. *Acta Physica Polonica B*, 36(9):2653–2679, 2005.
- [9] G. Das, D. Gunopulos, and H. Mannila. Finding Similar Time Series. *Principles of Data Mining and Knowledge Discovery*, pages 88–100, 1997.
- [10] A. Dattasharma and P. K. Tripathi. Practical Inter-stock Dependency Indicators using Time Series and Derivatives. In *The 6th ACS/IEEE Int'l Conf. on Comp. Sys. and Appln. (AICCSA-08)*, 2008.
- [11] D. S. Ehrman. Trading pairs. *Active Trader Magazine*, June 2006.
- [12] M. Gavrilov, D. Anguelov, P. Indyk, and R. Motwani. Mining the Stock Market: Which Measure is Best? In *6th ACM Int'l Conf. on Knowledge Discovery and Data Mining*, pages 487–496, 2000.
- [13] P. Gopikrishnan, B. Rosenow, V. Pleroul, and E. Stanley. Quantifying and interpreting collective behavior in financial markets. *Physical Review E*, 64(3), 2001.
- [14] M. A. Jaro. Advances in Record Linking Methodology as Applied to the 1985 Census of Tampa Florida. *Journal of the American Statistical Society*, 84(406):414–420, 1989.
- [15] V. I. Levenshtein. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics Doklady*, 10:707–710, 1966.
- [16] H. Lu, J. Han, and L. Feng. Stock Movement Prediction and N-dimensional Inter Transaction Association Rules. In *ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*, pages 12.1–12.7, 1998.
- [17] H. Mannila and a. A. I. V. Hannu Toivonen. Discovery of frequent episodes in event sequences. *Data Mining and Knowledge Discovery*, 3(1):259–289, 1997.
- [18] R. J. Miller and Y. Y. Y. Association Rules over Interval Data. In *ACM SIGMOD Conf. on Mgmt. of Data*, pages 452–461, Tucson, Arizona, USA, May 1997.
- [19] J. Rodgers and W. Nicewander. Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42:59–66, 1988.
- [20] R. Srikant and R. A. R. Mining Quantitative Association Rules in Large Relational Tables. In *ACM SIGMOD Conf. on Mgmt. of Data*, pages 1–12, Montreal, Canada, June 1996.
- [21] G. Vidyamurthy. *Pair Trading: Quantitative Methods and Analysis*. Wiley Finance, 2004.
- [22] J. W. Wilder. *New Concepts in Technical Trading Systems*. Trend Research, 1978.

# Behavior-based Robotics And The Reactive Paradigm

## A Survey

L. De Silva<sup>1</sup> and H. Ekanayake<sup>2</sup>

University of Colombo School of Computing

35, Reid Avenue, Colombo 07

Sri Lanka

lalindra84@yahoo.com<sup>1</sup>, hbe@ucsc.cmb.ac.lk<sup>2</sup>

### Abstract

*Behavior-based robots have come under the spotlight in recent years by making their presence known in diverse fields of human interest. Applications in behavior-based robotics have continued to grow in areas such as demining, search and rescue, office automation, health care, etc and continue to replace human beings in risky and menial tasks. With this inspiration, this paper investigates a variety of aspects in behavior-based robotics and the reactive paradigm in the context of their origins, concepts, applications and current research and is intended to provide a comprehensive overview about this area in robotics. The paper will review several central issues including a brief history of robotics, transition of robotic systems from hierarchical paradigm to reactive paradigm, biological/ethological inspiration for behavior-based robots, fundamentals of the reactive paradigm, architectures for controlling robotic behavior and the hybrid deliberative/reactive paradigm.*

### 1. Introduction

The human desire to mechanically model living creatures dates back to pre-Aristotle era (before 350BC), as far as the times where a Greek mathematician, Archytas, is thought to have built a mechanical bird, the first mechanical artifact of the sort [1]. History reveals a large number of examples where humans have attempted to build animal-like artifacts and instances like Leonardo's robot (a knight-like mechanical artifact designed by Leonardo Da Vinci around 1495AD) are mere examples of a larger lot [1]. Thousands of years of such experiments have led us to the present era in which we deal with cognitively advanced robots.

The field of behavior inspired robotics began shaping in the 1950s with the inspiring work done by Grey Walter with

his invention of the electronic tortoise, the first robot to display reactive behavior [7]. Even without any knowledge of its environment or any representation of its world state, the electronic tortoise was able to produce reactive behavior by responding to various forms (the intensity) of light in different ways. Such inventions gave way to a more structured approach to design and implement robots, namely, the hierarchical paradigm. Section 2 covers in detail the characteristics of the hierarchical paradigm and why, after few decades of prevalence, it gave way to the reactive paradigm.

The extent of inspiration from ethological studies in shaping the reactive paradigm has been immense. Thus, section 3 delves into identifying the similarities between animal behavior and robots that are built on reactive behavior. Section 3 also considers how animal behavior relates to perception and action, how multiple behaviors are controlled and coordinated in animals and how schema theory can be adopted to represent these behaviors. Section 4 catalogs the reactive paradigm in detail and discusses how it emerged from ethological studies and became the leading underlying architecture for robotic systems for many years ahead. Several other related topics including how this paradigm deviated from the hierarchical paradigm, its prominent methods for controlling behaviors, the mechanisms to integrate competing behaviors and the shortcomings of this paradigm will be discussed in detail.

Section 5 describes the hybrid deliberative/reactive paradigm. This paradigm addressed the shortcomings of the purely reactive systems by incorporating the planning element into the methodology and has been identified currently as the most favorable architecture to design and implement the robotic systems. A large number of control architectures that are built upon the hybrid paradigm have appeared overtime and section 5 will take into account some of these architectures in order to exemplify the use of the hybrid paradigm.

## 2. Hierarchy Matters

### 2.1. The Unseparable Trilogy

In designing the architectures for robotic control, the earliest robotic systems used a high level structured approach known as the hierarchical paradigm. In systems that followed this paradigm, the robots followed a strict sequence of distinct states known as the SENSE, PLAN and ACT cycle [19]. In the SENSE stage, the robot would acquire information about its environment using its sensors and use this information to create a representative data structure, known as the world model. According to Murphy (2000), the typical world model would comprise of priori information of the environment, sensing information from the robot's sensors and any other cognitive knowledge that would assist the robot in carrying out the specified task. The PLAN stage receives this world model and devises an action plan for achieving the task. In cases when the final goal comprised of complex situations and world states, the planner would break the goal into sub goals and account for each in turn to achieve the final goal. The ACT stage represents the execution of actuator commands that are generated according to the directives received from the planning stage. When the robot's actuators have carried out the task specified, the cycle begins again and continues until the goal is reached.

Among the representative architectures and methodologies that were built on the hierarchical paradigm, the strips method was a forerunner and was used in Shakey, a robot by the Stanford Research Institute [5]. Strips method measured the difference between the goal state and the current state and chose the best action that would minimize the difference. Constructing the world model was imperative and the action to be chosen at a certain point was selected from a descriptive table called the difference table in which differing actions corresponding to varying differences were listed beforehand. During the planning stage of each cycle a difference-evaluator would measure the difference between the current state and the goal state, enabling the planner to select the corresponding commands from the difference table and pass them on to the actuators.

### 2.2. Two Challenging Problems

Among the biggest issues that made the hierarchical paradigm less inspiring with time were the close world assumption and the frame problem [19], [9]. The close world assumption assumes that the robot has obtained all the information it needs about the environment within the world model. Such an assumption presents a significant problem when the environment changes with time since the planner cannot keep track of all the changes of the environment in a continuous manner.

Hierarchical paradigm also suffered what is known as the frame problem [6]. Frame problem is the inability to represent all the world information that was needed by the robot in a computationally viable method. The required intricacy and the detail of the information of the robot's environment were considerable even for the robot that was meant to achieve the simplest task. Consequently, addressing uncertainty in the event of a bigger problem was too tedious and was not worth the effort.

### 2.3. Relaxing The Hierarchies

The nested hierarchical controller (NHC) proposed by Meystel (1986) suggested a method to resolve the aforementioned issues by enhancing the planner by decomposing it into three distinct components, namely, the mission planner, navigator and the pilot [18]. NHC's most obvious deviation from a pure hierarchical architecture was that the sensors would continuously update the world model while the actuators were carrying out the commands. This gave a more realistic response in terms of the receptiveness to the changes in the environment.

A similar architecture was proposed by Albus through his real-time control system and was an architecture that was heavily influenced by the NHC architecture [8]. Shananhan presented a logic-based high level robotic control approach which also used a hierarchical approach to control the robot's behavior [20].

Even with these proposed strategies to enhance the hierarchical paradigm, the fact that this paradigm fails to address uncertainty has made people look for different paradigms for robotic control. Thus, with the advances in the field of artificial intelligence, cognitive psychology and ethology, people turned to studying animal behavior and understanding how behavior coordination is handled biologically.

## 3. Animals To The Rescue

In the early days of robotics, the notion of using animal behavior as models in developing robotic systems was treated with much detest due to several reasons. The strongest argument against using animal behaviors was the inadequate knowledge of the intricate biological intelligence that was available at the time. Secondly, the available technology at the time was hardly sufficient to model even the simplest animal behaviors that had been identified.

However, the advances in cognitive psychology and ethology combined with the growing limitations faced in the hierarchical paradigm gave the roboticists an incentive to return to these biological studies and to search for new paradigms for robotic control. Still, the findings in these areas were not comprehensive enough to produce a model

for robotic behavior that would directly map to a biological behavior. As a result, roboticists tried to come around this issue by going in the reverse direction, i.e. to build models for robotic behavior and to find confirmation of those behavioral models in biological and ethological findings. Furthermore, animals interact with a much more relaxed, open environment and thus the study of their behaviors provided a foundation to tackle the close world assumption which proved to be the biggest obstacle in the hierarchical paradigm [19].

### 3.1. The Basic Building Block

The foundation of behavior-based architecture for robotics lies in the definition of a behavior. In its simplest form, a behavior is a mapping of sensory inputs to a pattern of actions that in turn carries out a specific task [19]. Animal behaviors have been studied in three broad categories, namely, reflexive behaviors, reactive behaviors and conscious behaviors [19]. Reflexive behaviors are direct stimulus-response mappings in which the response to a particular sensory input is directly wired with a response and which is carried out without any cognitive involvement of the subject. An example for this type of behavior can be the reaction of a person's knee when it is tapped. The second category concerns with the behaviors that are learned over time and are ultimately embodied in the system such as learning to ride a bicycle or learning to walk in early childhood. Every other behavior that requires conscious thought throughout the time of the action is categorized as conscious behavior.

In behavior-based robotics, reflexive behavior prevails over any other behavior type. Therefore, this paper will confine only to the reactive systems and will describe in detail how reactive behavior is introduced in robotics. However, there is a slight difference in the connotations between robotic literature and ethology. The word 'reactive' actually means reflexive behaviors in robotics while in other disciplines the word 'reactive' refers to the learned behaviors.

### 3.2. Behavior Diversity

In studying the reflexive behavior of animals, ethologists have been able to identify three forms of reflexive behaviors [19]. These forms can be explained as follows.

**Reflexes:** Where the magnitude of response is directly proportional to the intensity of the stimulus and lasts as long as the stimulus lasts. This form of behavior is visible in domestic millipedes and certain snails where they would contract their muscles in response to a disturbance [9].

**Taxes:** Where the response is an orientation towards or away from the stimulus. The best examples for this kind of

behavior is the Chemotaxis [2] behavior of trail following ants and the behavior of sea turtles who would head directly to the sea once they are born. It has also been discovered that certain animals react to the magnetic field of the earth and display behavior accordingly [4].

**Fixed Action Patterns:** Where the response lasts for a longer time than the stimulus and where the behavior can be influenced from multiple stimuli rather than from a simple stimulus [3]. Examples are drawn for this sort of behavior in animals fleeing away from predators, the singing of crickets and locust fly patterns [9].

It should not be misunderstood from the above description that a pair of stimulus and response is all that needs in reflexive behavior. A mechanism to release the behavior in the presence of a stimulus should also exist in order to execute the behavior.

Innate Releasing Mechanism (IRM) was a phrase coined by Lorentz and Tinbergen (1973) to denote the mechanism by which a specific behavior is initiated [9]. The underlying explanation is that a behavior will not be initiated even in the presence of sensory inputs unless some mechanism was present to initiate that behavior. One of the simplest examples from the animal world is that most predators would not hunt prey unless they are hungry, even if the prey wanders close by.

A releaser, which can be presumed to be a boolean value, has to be set in order to initiate the behavior. In most occasions, compound releasers, probably both internal and external, have to be present in order to trigger a behavior. This essentially means that IRMs act as control switches for releasing behaviors and provide a low level coordination mechanism for triggering and inhibiting basic behaviors.

### 3.3. Controlling Behaviors

Another advantage of studying animal behavior is the ability to understand how the concurrent, competing behaviors are handled biologically. Ethologists have identified several biological mechanisms that animals portray when they are faced with competing behaviors. These mechanisms that make choices between competing behaviors can assist roboticists in modeling robots with concurrent behaviors and several of such control mechanisms are listed below.

**Equilibrium:** Occurs when a dominant behavior cannot be identified and no behavior takes precedence over the other behaviors. Murphy (2000) illustrates the example of a squirrel on a park bench that sights a bread crumb near a person sitting on the edge of the same bench [19]. Most times it can be seen that the squirrel seems to be lost in a dilemma whether to go for the bread crumb (execute the

feed behavior) or to flee from the situation (execute the flee behavior).

**Dominance:** Occurs when one behavior clearly dominates the others. The obvious example is that animals cannot eat and sleep at the same time.

**Cancellation:** Occurs when the competing behaviors cancel each other out so that a completely new behavior is displayed. Murphy (2000) draws an example from male stickleback fish. Sticklebacks get caught up in between defending their nests and attacking the enemies when their territories are threatened and end up building a new nest altogether in a different place [19].

### 3.4. A Novel Representation

Perception is a crucial factor in animal behavior as it plays a vital role in coordination and selection of different behaviors. Arkin (1998) identifies that perception can be two-fold at the highest level: Proprioception, which refers to perceptions arising from stimuli that are internal to the animal's biological system (For e.g. hunger) and exteroception that relates to perceptions arising from external stimuli such as the cue of a predator nearby [9].

An interesting theory presented by Gibson (1979) that has assisted the way roboticists think in terms of behaviors is the theory of affordances [15]. Affordances differ from stimuli in that affordances take into account the environment in which the stimuli was produced. A better explanation is that an affordance is a potentiality for inherent action [19], [9]. This analysis of perception and affordances leads to the intriguing issue of representing this information in robots. One successful way around this problem was presented by Arbib (1992) through his famous schema theory. Schema theory facilitates the representation of perceptual and motor-action information in a similar way to that of object oriented concepts. A schema is an entity comprising of perceptual data and computational activity that result in motor actions. Accordingly, a behavior is defined as comprising of perceptual schemas and motor schemas, one from each at a bare minimum [21]. Schema theory also backs the concept that complex behaviors can be comprised of much simpler behaviors, thus facilitating the behavior subdivision in robotic systems.

### 3.5. From Animals To Robots

Several concrete implementations stemming from these studies have been surfacing over the years and it appears necessary that a few of those examples be drawn out here. The chemotaxis display of ants has been modeled using robots that can sense a chemical trail left by other robots and these experiments have been a success in the study of multi-robot environments and how behavior coordination is

handled in such environments [9]. The locomotion patterns of the insects such as cockroaches have provided the substance to another research area and this area has concentrated on aspects such as path following in robots and other navigational behaviors. The artificial insect project gives examples for such experiments in great detail [13].

## 4. Planning: To Do Or Not To Do

The highlight of the reactive paradigm was its basic building block - a behavior [9]. As mentioned earlier a behavior is a direct mapping of sensory inputs to the corresponding motor actions. The biggest difference the reactive paradigm had introduced from the earlier hierarchical paradigm is the removal of the planning stage from the Sense-Plan-Act triad. This was due to the popular belief among the roboticists that maintaining world state, history information and planning of actions, no matter how efficiently designed, only adds to the deterioration of the performance of the system. Thus, the reactive paradigm relied on behaviors and the fact that emergent behavior of the system can be produced by different mechanisms that are available to coordinate these behaviors. The concept of a behavior facilitated solid architectures that allowed extension of older behaviors, inhibition of behaviors, parallel execution of behaviors and many other coordination mechanisms that produced the emergent behavior of the complete system [19].

Two other important considerations in behavior-based robotics are the local sensing techniques and the ecological niche. In the case of sensing, reactive paradigm advocates the use of a specific sensor input for each behavior. This effectively means that each behavior takes sensory inputs from a sensor that is local to it and passes this information to the perceptual schema. However, latter systems were built with architectures that can share sensory information where several behaviors would use sensory information from a single sensor.

The other important consideration in reactive robotics is the concept of an ecological niche. An ecological niche, from a purely biological point of view, means the environment in which the animal lives in and relative to which its behavior can be defined. In terms of robotics, this means that the roboticists should design the robot targeting a specific environment since the robot's interaction in a different environment would be meaningless and would not result in the expected outcomes.

### 4.1. Status Representation

In developing reactive architectures, several representational mechanisms have been introduced [19], [9]. Some of these representation mechanisms can be listed as follows.



**Stimulus-Response diagrams:** Presents a very primitive way to represent the collection of behaviors that are coordinated. This mechanism simply portrays each behavior as a mapping of sensory stimuli and the action response.

**Functional Notation:** Presents the behaviors in a procedural manner that is similar to a pseudocode of a general program. Murphy (2000) explains a similar method named scripts that functions as a general script of a play [19].

**Finite State Acceptors:** Presents a general approach that is used in many other areas in computer science that is used to express the situational information of a system. A finite state automata is defined as a tuple  $M = \{Q, \Sigma, \delta, q_0, F\}$  with the symbols representing the set of states, the input alphabet, the transition function, the start state and the set of acceptable final states respectively. This definition is best used in its graphical form and presents a clear method of understanding the interaction of different behaviors.

## 4.2. Building Upon The Reactive Paradigm

Behaviors by themselves are of little importance to a reactive system unless they are coordinated to produce the emergent behavior of the whole system. Several architectures have emerged over the years and have presented varying approaches to coordinate behaviors in robotic systems. This paper focuses on the two most tried architectures of all, namely, subsumption and potential-fields methodology.

In 1986, Brooks presented an inspiring paper titled “Intelligence Without Representation”, in which he questioned the viability of representing the world state and planning the sequence of actions in robotic systems [11]. He vehemently questioned the practicability of using specific data structures and attempting to model the robot’s environment. Instead, he stressed that ‘the world is its own representation’, i.e. that any other representation of the world that was to be used in the robotic systems would be futile since no model can perfectly account for the uncertainties of the real world.

Consequently, Brooks (1987) presented an architecture for controlling and coordinating the robotic behaviors, for which he gave the name, subsumption. Subsumption builds on layers of competencies (or simply, a collection of behaviors that achieved a specific task) that are vertically arranged in the order of their sophistication. Thus, the basic survival behaviors such as collision avoidance are aligned in the lowest layer while more cognitive behaviors such as path mapping are aligned in the higher levels [19]. The higher layers have the ability to inhibit or suppress the behaviors of the lower layers, hence the word, subsumption.

The layers of competences are supposedly executed independently with each behavior having its own sensorial

system and each behavior being unaware of what’s happening in the other layers. Each layer executes a dedicated behavior and this avoids the need for them to know the complete scenario they are trying to solve.

However, this separation of layers calls for a mechanism to control and coordinate the behaviors when several of them are trying to execute simultaneously. Subsumption resolves this issue by always giving the higher layer the ability to take control and override the behaviors of the lower layers. Allowing such superimposing behavior requires that the higher level behaviors embody the competencies of the lower layers.

Subsumption does not come without its shortcomings. First, the inclusion of lower-layer-competencies in higher levels means a waste of resources. Second, there often arises the requirement to pass information between layers and pure subsumption architecture acts in contrast to this by precluding the layers from passing information between themselves [24]. Finally, the subsumption architecture fails to take into account any advantage a planning module can introduce to the system.

Several solutions to address these rigidities of the subsumption architecture were proposed by many and they introduced several alterations to the original architecture. One such proposition (Flanagan et al, 1995) brought in three variations to the original architecture [24]. This method substituted the competences of the subsumption architecture by individual processes with each process carrying out a defined task. The methodology also added a global data structure where each process can post and read interesting data to and from this data structure. Finally, they included a priority scheme that allowed the behaviors to compete for control and made this priority scheme time varying so that no single behavior can hold up the robots control for a long time.

Yongjie, Qidan and Chengtao (2006) presented another modification to the subsumption architecture by assuming a hybrid control approach that was based on subsumption [12] (hybrid control architectures will be discussed in detail in section 5). In addition to the use of processes as in the previous approach, this methodology also introduced a behavior-managing layer. The task of the behavior-managing layer was to adjust the behavior parameters (for e.g. priorities, behavior arithmetic etc) according to a number of factors that effected the robotic system at a certain point in time. Finally, they added a scheduler to the system to select which behavior process will take control of the system and provide the necessary actuator command.

No matter how hard it is tried to reduce the rigidity of the purely reactive subsumption architecture, there always is a tradeoff in performance when planning and its constituent data structures are introduced to the system. However, this tradeoff will be justified in complex systems where purely

reactive systems are insufficient to provide the complete behavior.

Another approach taken by the proponents of the reactive paradigm is a methodology known as the Potential Fields Methodology (PFM). PFMs build on the groundwork that each behavior is represented as a vector and thus the methodology is inherently regarded to be confined to the navigational robots. These behaviors, represented as vectors, are combined in vector summation to produce the emergent behavior.

For presentational purposes, these behaviors are assumed to exert on the robot in the form of force fields by the objects in the environment and are illustrated as magnetic or gravitational fields. For e.g. an obstacle in the environment may be illustrated as emitting a repulsive field from its vicinity as much as a bulb is radiating light whereas a goal in the environment can be illustrated by an attractive field on the surrounding area. In representing these force fields, the general representation of the vectors, where the magnitude of the vector is given by the length of the arrow and the orientation given by the angle, can be used (For e.g. in the case of a navigational robot, the length of the vector represents the velocity and the angle represents the direction of the robot). The robot is assumed to be a particle entered into the force field area and the behavior of the robot is the path the robot takes as a result of the multiple potential fields [19]. Magnitude profiles are used to control the strength of the potential field to avoid the unpleasant outcomes caused by having a constant strength throughout the field (For e.g. to avoid over shooting a goal, the field will be calculated such that the strength of the field is a reducing function of the distance between the robot and the goal. As the robot moves closer to the goal, the magnitude of the force reduces so as to allow collision-free arrival at the goal).

One common misconception is that the PFM requires the calculation of the entire field in one occurrence. However, this is not the case. The robot only needs to calculate the force fields that exert on it in each Sense-Act cycle [19]. Thus the computational load that would appear to exist in designing the complete force field at once is greatly reduced.

The greatest drawback of the PFM is the possibility of local minima existing in the force field environment. A local minimum occurs where the strength and orientation of the vectors cancel each other out so that the robot caught up in such a situation will be stuck there forever. Masoud (1996) proposed a way around this issue by introducing Harmonic Potential Fields [17] while Iianmiao and Bo (1992) introduced time-varying potential fields [9]. Arkin (1987) suggested an approach where random noise can be injected to the fields to shake off the robot out of a local minimum trap [9].

Koren and Borenstein (1991) draw a detailed mathemat-

ical analysis of several other limitations and drawbacks of the PFM and proposed an altogether different methodology for obstacle avoidance in navigational robots in real time environments [10].

Reactive paradigm emerged as the ultimate solution for the rigidities put forth by the hierarchical paradigm. Though this newly found paradigm proved more than satisfactory in robots executing simple tasks, the obvious need for planning in more complex tasks emerged with time. Hence, roboticists looked into new ways of combining the planning process with the reactive behavior of the robots and thus a new breed of architectures under the name Hybrid Deliberative/Reactive paradigm was born.

## **5. One Step Ahead: The Hybrid Deliberative/Reactive Paradigm**

### **5.1. The Need For Planning**

The reactive paradigm started showing its limitations when moving into domains with complex tasks that required cognitive capabilities rather than just reactive behavior. These requirements were presented in the form of efficient path planning, map making and performance evaluation of the robot itself, etc, and the reactive paradigm proved incapable of satisfying these requirements.

The proponents of the hybrid architecture advocate the use of the advantageous aspects of each of the previous two architectures and combining them to produce the desired architecture for the complex scenarios. Effectively, this meant that the planning aspect of the hierarchical paradigm was to be integrated with the rapid execution capabilities of the reactive paradigm. The technological foundations for such an architecture was emerging as well and these technological aspects facilitated such integration in the two paradigms (For e.g. the emergence of multithreaded architecture allowed the planning process to be executed simultaneously with the reactive behaviors but as a separate module from that of the reactive modules).

### **5.2. Integrating Plans**

The hybrid paradigm altered the Sense-Plan-Act trilogy by allowing the Plan stage to be executed separately from the Sense-Act sequence. Arkin (1998) mentions three broad methods for integrating planning with reactive behavior [9]. Hierarchical integration makes the planning layer and the reactive layer to be aligned vertically in which the planning layer provides the information on which the reactive layer acts. Second method involves planning layer working prior to or concurrently with the reactive behaviors, updating the robots behavioral parameters or changing the world state.

Finally, coupled planning-reacting method makes the planning and reactive activities to occur concurrently, allowing each one to interact with the other in real-time.

Another interesting deviation of hybrid architectures from the reactive paradigm is that the sensors that were unique to each behavior in the reactive paradigm are now shared with the planning modules. Planning needs to create world models and to represent the state of the environment and thus the use of sensorial information is imperative.

### 5.3. The Hybrid Features

Though the architectures that fall under the hybrid paradigm vary considerably in their designs, several common features can be identified in most of them [19], [9].

**Sequencer:** The task of the sequencer is to generate the set of behaviors that should be executed and their sequence in order to carry out a specific task. A dependency network or a finite state machine is generally used in coordinating these behaviors.

**Resource Manager:** Introduction of this module was primarily due to the fact that resources such as sensors are now been shared between the planning and reactive modules. Resource manager handles the allocation of sensors and other resources between different modules and selects the best resource that should prove practical for each module.

**Cartographer:** The cartographer module's main application is to create, store and maintain maps of the robotic environment by interacting with the external world. This module facilitates the knowledge representing scheme of the robot by maintaining and updating the data structures pertaining to the world model.

**Mission Planner:** This module's task is two fold. Initially, it presents an interface to the outer world and secondly it transforms the external commands into robotic terms, subdivides the tasks and forwards the subtasks into actuator modules to carryout.

**Performance Monitor:** Robot's performance in executing a given task should be evaluated in order to verify if the robot is carrying out its tasks as it should be. The performance monitor measures the robots performance in comparison to previously set targets by the mission planner and prompts any unfavorable activity.

### 5.4. Hybrid Architectural Styles

Murphy (2000) draws out three architectural styles that are visible in the hybrid paradigm.

Managerial Style architectures pose the general managerial style that is evident in an organizational hierarchy by subdividing the tasks in each module and allowing each

task to be carried out by a subordinate module. Secondly, State Hierarchies style uses the robot's state to determine the appropriate action. The reactive behaviors are thought to maintain no state while the deliberative modules keep track of the world model and the past actions of the robot. Finally, Module-Oriented styles follow architectures that closely resemble the ones in hierarchical paradigm in which most modules have access to the world models individually.

### 5.5. Representative Architectures

In discussing the varied architectures that were based on the hybrid paradigm, Autonomous Robot Architecture (AuRA) proposed by Arkin (1998) draws special interest. In this architecture the previously mentioned five models of a hybrid paradigm were clearly visible and the architecture had clear division between the deliberative and the reactive layers.

Stenzel (2000) presents a utility-based approach to handle the problem of combining commands of different behavior mechanisms [23]. Additionally, his approach uses a priority management system to asses the importance of goals and this system uses subsumption to inhibit the low priority goals giving way to the high priority ones. Egerstedt and others (1999) used hybrid automata to depict each behavior of the robot and allowed the behaviors to fuse rather than to allow one behavior to overpower the other [16]. Li and Feng (1994) proposed a fuzzy logic based approach to coordinate the behaviors of navigational robots. In this approach, they used fuzzy reasoning to weigh the reactive behaviors rather than to inhibit the lower level behaviors as in the subsumption process [14]. A similar behavior-based neuro-fuzzy controller was introduced by Rusu and others (2003) where the coordination system was organized in a top-to-bottom hierarchy and in which command fusion using neuro-fuzzy subsystems were used to handle multiple low level behaviors [25].

A comprehensive study of behavior coordination mechanisms in behavior-based robotics has been presented by Piranian (1999) and this paper covers in detail all the alternative methods in addition to the primary methods discussed above [22].

## 6. Conclusion

The reactive paradigm answered many problems that were persistent in the early robots due to the rigidity that was present in the hierarchical paradigm. Robots that were based on reactive behavior proved much more efficient and showed rapid execution in performing its specified tasks.

However, the inability to cater into much more complex task domains has made the pure reactive behavior-based robots to decline in popularity with the emergence of the

hybrid paradigm. The obvious need for a planning module to provide the necessary information required to handle the complex behaviors and complex environments has made the latter paradigm much popular towards the recent years.

Even the inclusion of the planning element does not make the robots as intelligent as we would expect them to be. It may be true that a robot is able to find the most efficient path to a goal, to measure its performance, to produce efficient representations of the environment, etc but in order to reap the real robotic performance, a learning process has to be integrated to the system. The field of machine learning has been producing a lot of results that cater to this need but the extent to which the robotic world has made use of these mechanisms is still questionable.

Research areas such as cognitive modeling has been devising models to produce intelligent behavior in machines at a conceptual level but the implementation of such models in robots still seems to be at a very rudimentary level. However, it can be expected that with the advancements of such research which try to model human intelligence in machines could take the behavior-based robotics to the next level, a level in which the robots can cope up with the uncertainties of its environment at a highly acceptable level.

Therefore, the future that we are looking at in the context of behavior-based robotics, to a large extent, would focus on hybrid architectures with robotic learning playing an important role. The extent to which robotics can benefit from areas such as machine learning, cognitive modeling, etc leaves us with optimism, where we can hope for the newest robots to be cognitively much more advanced than their precursors.

## References

- [1] A Brief History of Robotics. <http://robotics.megagiant.com/history.html>.
- [2] Chemotaxis. <http://en.wikipedia.org/wiki/Chemotaxis>.
- [3] Fixed action pattern. [http://en.wikipedia.org/wiki/Fixed\\_action\\_pattern](http://en.wikipedia.org/wiki/Fixed_action_pattern).
- [4] Movement: Taxis and Kinesis. <http://www.sparknotes.com/biology/animalbehavior/orientationandnavigation/section1.html>.
- [5] Shakey. <http://www.ai.sri.com/shakey/>.
- [6] The Frame Problem. <http://plato.stanford.edu/entries/frame-problem/>.
- [7] William Grey Walter. [http://en.wikipedia.org/wiki/William\\_Grey\\_Walter](http://en.wikipedia.org/wiki/William_Grey_Walter).
- [8] J. S. Albus. The NIST Real-time Control System (RCS): an approach to intelligent systems research. *Journal of Experimental and Theoretical Artificial Intelligence*, 9:157–174, Apr. 1997.
- [9] R. C. Arkin. *Behavior-based Robotics*. The MIT Press Cambridge, Massachusetts, 1998.
- [10] J. Borenstein and Y. Koren. Potential field methods and their inherent limitations for mobile robot navigation. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pages 1398–1404, 1991.
- [11] R. A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.
- [12] C. Chengtao, Y. Yongjie, and Z. Qidan. Hybrid Control Architecture of Mobile Robot Based on Subsumption Architecture. In *Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, pages 2168–2172, Luoyang, Henan, June 2006.
- [13] H. J. Chiel, L. S. Sterling, and R. D. Beer. A biological perspective on autonomous agent design. In *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*. The MIT Press Cambridge, Massachusetts, 1990.
- [14] X. Feng and W. Li. Behavior fusion for robot navigation in uncertain environments using fuzzy logic. In *Proceedings of the 1994 IEEE International Conference on Systems, Man, and Cybernetics*, volume 2, pages 1790–1796, 1994.
- [15] J. G. Greeno. Gibson's Affordances. *Psychological Review-New York-*, 101:336, 1994.
- [16] J. Lygeros, S. Sastry, M. Egerstedt, and K. Johansson. Behavior Based Robotics Using Regularized Hybrid Automata. In *Proceedings of the 38th IEEE Conference on Decision and Control*, volume 4, pages 3400–3405, 1999.
- [17] A. A. Masoud. Using hybrid vector-harmonic potential fields for multi-robot, multi-target navigation in a stationary environment. In *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, pages 3564–3571, 1996.
- [18] A. Meystel. Planning in a hierarchical nested controller for autonomous robots. *25th IEEE Conference on Decision and Control*, 25:1237–1249, Dec. 1986.
- [19] R. R. Murphy. *Introduction to AI Robotics*. The MIT Press Cambridge, Massachusetts, 2000.
- [20] S. Murray. Reinventing Shakey. *Workshop on Logic-Based Artificial Intelligence*, 1999.
- [21] G. Pezzulo. Schemas and Schema-based Architectures, 2007. [http://www.eucognition.org/wiki/index.php?title=Schemas\\_and\\_Schema-based\\_Architectures](http://www.eucognition.org/wiki/index.php?title=Schemas_and_Schema-based_Architectures).
- [22] P. Pirjanian. Behavior coordination mechanisms – state-of-the-art. CMPSCI Tech. Rep. IRIS-99-375, Institute of Robotics and Intelligent Systems, School of Engineering, University of Southern California, 1999.
- [23] R. Stenzel. A behavior-based control architecture. In *Proceedings of the 2000 IEEE International Conference on Systems, Man, and Cybernetics*, volume 5, pages 3235–3240, 2000.
- [24] B. Strunz, C. Flanagan, and D. Toal. Subsumption Control of a Mobile Robot. In *Proceedings Polymodel 16, Applications of Artificial Intelligence, Sunderland*, pages 150–158, 1995.
- [25] T. E. Whalen, A. Cornell, H. J. W. Spoelder, P. Rusu, and E. M. Petriu. Behavior-based neuro-fuzzy controller for mobile robot navigation. In *IEEE Transactions on Instrumentation and Measurement*, volume 52, pages 1335–1340, 2003.

## Reversible Data Hiding using Increased Peak Histogram

T.C Thanuja  
Asst. Professor  
Dept. of TCE  
R V College of Engg.  
Bangalore, India-560 059  
tc.thanuja@gmail.com

R Nagaraj  
Director, R.V. Centre for  
Cognitive Technologies.  
Bangalore  
India-560 059  
rnrv@rediffmail.com

M.Uttara kumari  
Professor, Dept. of ECE,  
R V College of Engg  
Bangalore  
India-560 059  
[uttarakumari@hotmail.com](mailto:uttarakumari@hotmail.com)

### Abstract

*In this paper, we present a reversible watermarking algorithm which provides higher embedding capacity and higher PSNR value compared to the existing algorithms. In the proposed algorithm, we utilize a peak point of image histogram and increase the height of the peak by transferring the pixels from the neighbouring gray values to achieve the higher embedding capacity.*

### 1. Introduction

Development of computer technology and widespread use of internet have driven this world into fast-changing digital place. With digitization of multimedia contents, everybody can access multimedia contents more easily than in analog age. Even if digitization of the multimedia contents provides more opportunities to media contents, it also provide easy access paths to copy and distribution of digital contents, because of characteristics of the digital data, represented by 0 and 1. As the copy and distribution of digital contents are widely conducted illegally in internet environment, the copyright holders began to pay attention to copyright protection technologies. Off the technologies that can protect copyright of digital contents, watermarking technology has received keen interests from research communities.

Watermarking technology hides copyright information into original contents to protect copyright, which eventually leads to modification of original contents[1]. However, in order to achieve the goal of imperceptibility, the watermarking technology modifies original contents so that the modification is not perceptible to naked eyes using Human Visual

System (HVS) modelling. As a result, the original image cannot be recovered when the image is watermarked. In the applications where a slight modification can lead to significant difference in final decision making process, such as medical images or military images systems, there has been a requirement of recovery to original contents[2]. For the reason, there have been research efforts of reversible watermark that can recover to original images from watermarked images.

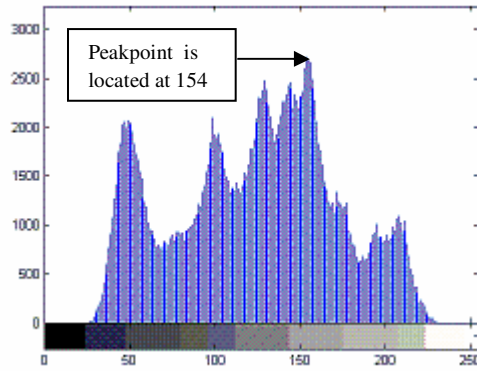
This paper is organised as follows. The overview of the existing scheme and the proposed scheme is presented in section 2. The procedure for the proposed technique is explained in section 3. The section 4 deals with simulation results and analysis with conclusion in section 5.

### 2. Increased Peak Histogram

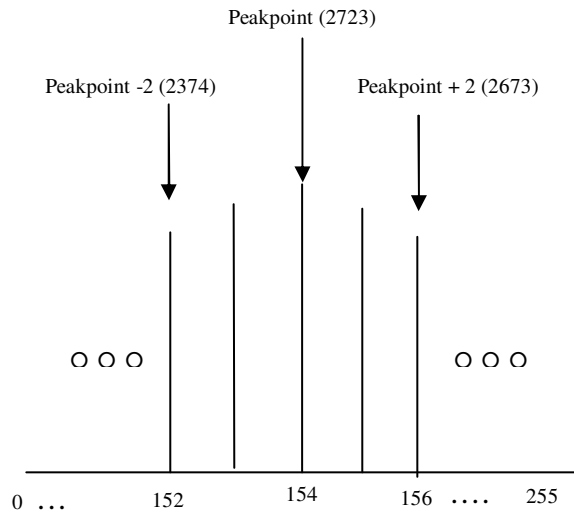
The Ni et al. [3] utilizes zero or minimum point histogram. If the peak is lower than the zero or minimum point in the histogram, it increases pixel values by one from higher than the peak to lower than the zero or minimum point in the histogram. While embedding, the whole image is scanned. Once a peak-pixel value is encountered, if the bit to be embedded is '1' the pixel is added by '1', else it is kept intact. Alternatively, if the peak is higher than the zero or minimum point in the histogram, the algorithm decreases pixel values by one from lower than the peak to higher than the zero or minimum point in the histogram, and to be embed bit is '1', the encountered peak-pixel value is subtracted by 1. The decoding process is quite simple and opposite of the embedding process.

In this algorithm the maximum amount of data that can be embedded is limited by the peak value of the histogram of the cover image. In the proposed algorithm efforts has been made to increase the peak

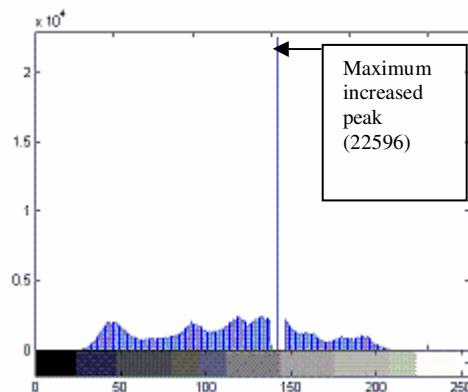
to the maximum without changing the perception.



**Figure 1. Host Image (Lena) Histogram.**



**Figure 2. Histogram of Lena (zoomed around peakpoint)**



**Figure 3. Increased Peak Histogram**

The proposed algorithm makes use of the assumption that most of the gray values around the peak has relatively large number of pixels[5][6]. So the embedding capacity is increased by increasing the number of pixels at the peak.

From the histogram of the Lena image of 512 x 512 x 8 size, as shown in the fig 1. it is observed that the maximum peak point or the gray value is at 154( maximum number of pixels) and is represented as 'Peakpoint'. Neighbouring gray values towards the left of Peakpoint such as ...152, 153 and the gray values towards the right of Peakpoint such as 155, 156 ..... are represented as 'Peakpoint-2', 'Peakpoint-1' and 'Peakpoint+1', 'Peakpoint+2' respectively. The number of pixels for various peak points of Lena image is shown in Figure 2.

By transferring the pixels from 'Peakpoint-2', 'Peakpoint-1', 'Peakpoint+1', 'Peakpoint+2' and so on, we could observe the maximum peak point is increased to 22596 bits. This process is continued until there is no perceptual disturbance in the watermarked image. Using this method the number of pixels in the maximum peak can be increased by approximately ten times. The Figure 3. is scaled down, to show the increased peak.

The increase in capacity results in a slight decrease in the PSNR with out changing visual appearance of the image i.e. the image is highly imperceptible.

### 3. Data embedding and extraction

This algorithm makes use of the peak of the histogram of the host image, since the maximum capacity is limited by the peak value of the histogram of the host image, this peak value is modified as discussed in the section 2. to suit for maximum embedding capacity. The proposed algorithm is summarised as follows:

#### 3.1 Data Embedding

**Input:** Host image, Watermark data.

**Output:** Watermarked image.

1. The histogram of the host image is scanned to find the peak point (i.e. the gray level having maximum number of pixels).
2. Then maximum increased peak is achieved as explained in section 2.

3. To embed data the host image is scanned to find 'Peakpoint'. The data to be embedded is then scanned.
  - (i) If a binary value '1' is to be embedded then the pixel gray value is changed from 'peakpoint' to 'peakpoint+1'..
  - (ii) If a binary value '0' is to be embedded then there is no change in the pixel gray value.
4. Step 3 is repeated till the entire data is embedded into the host image to produce the watermarked image.

### 3.2 Extraction

**Input:** Watermarked image

**Output:** Recovered original host image and watermark

1. Watermarked image is scanned to retrieve the information such as size of watermark and the 'Peakpoint'.
2. Once again the watermarked image is scanned pixel by pixel. Each time when we encounter a pixel with gray value 'peakpoint+1' we extract a binary value '1' or when we encounter a pixel with gray value 'peakpoint' we extract binary value '0'. In this way the hidden data is extracted. For example if the watermarked image is Lena, each time when we come across gray level 155 a '1' is extracted and when we come across gray level 154 a '0' is extracted.
3. In order to recover the original host image the reverse of the embedding process is carried out.

The performance of the algorithm is determined by calculating the Peak Signal to Noise Ratio (PSNR)

$$\text{PSNR} = 10 \log_{10}((255 \times 255) / \text{MSE}) \quad (1)$$

Where MSE is the Mean Square Error between the original image and the watermarked image.

### 4. Simulation results

The simulation is carried out using MATLAB. Standard test images of 512 X 512 x 8 grayscale were used in studying the effects of perceptibility and quantity of information hiding or payload, for the reversible data hiding technique.

Watermarked and extracted Lena image is shown in the Figure 4., Figure 5 and 6 show the 193 x 117 size watermark and extracted watermarked image. The above figures show no visual disturbance with this algorithm. The PSNR and payload is calculated for the test images shown in Figure 7.

The comparison between the existing reversible watermarking scheme and proposed scheme in terms of embedding capacity and the PSNR is presented in Table I. The PSNR of minimum 57.13 dB to 61.79 dB is achieved for various test images. There is a 10dB improvement from the existing algorithm.

The embedding capacity for Airplane test image is 58,679 bits using our method, 16,171 bits are embedded using Ni. et al. algorithm whereas 8,155 bits can be embedded without modifying the peak of the original histogram. Data embedding capacity is increased by 3.5 times with Ni.'s algorithm and 7 times with unmodified original histogram.

From the Table I. it is observed that payload for other test images is around 22000 bits with improvement of 300 percent. from the results we can conclude that embedding capacity depends on the nature of the histogram.

Table II shows the variation of PSNR with respect to the brightness of the image. For darker watermark image the PSNR is 64.81 dB and it is decreased to 60.05 dB for brighter image. If the image is a brighter one then PSNR can be improved by negating the watermark before embedding.



**Figure 4. Lena Image (a) Watermarked image (b) Extracted image**



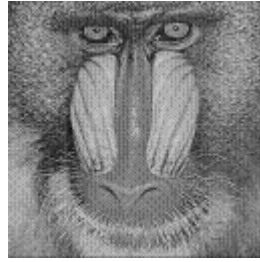
**Figure 5. Watermark**



**Figure 6. Extracted Watermark**



(a) Sailboat



(b) Baboon



(c) Pepper



(d) Airplane

**Figure 7. Watermarked Images**

**Table I**  
**Comparison between existing and proposed reversible watermarking method**

Test images	Existing Method		Proposed method		
	PSNR of marked image (dB)	Embedding Capacity (bits)	PSNR of marked image (dB)	Embedding Capacity (bits)	% increase
Airplane	48.3	16,171	57.54	58,679	262
Pepper	48.2	5,449	61.85	22,820	318
Lena	48.2	5,460	61.79	22,596	313
Sailboat	48.2	7,301	60.80	29,382	302
Baboon	48.2	5,421	60.67	27,665	410

**Table II**  
**Brightness of the watermark Vs PSNR**

Number of 1's(%)	Number of 0's(%)	PSNR (dB)
25	75	64.81
40	60	62.78
50	50	61.80
60	40	61.01
75	25	60.05

## 5. Conclusions

Our proposed technique is able to embed about 22kb - 59 kb into a 512 x 512 x 8 gray scale image with the average PSNR of 57 dB. Our algorithm does not depend on the minimum or zero peak point of the histogram and also it is independent of shifting of range cells of the histogram, whereas in the existing algorithm the maximum embedding capacity is limited by the frequency of peak-pixel value in the histogram and also the range of bins towards the minimum peak point gets shifted was contributed for lower PSNR value.

In the existing reversible data embedding methods it has been observed that there is a trade off between higher data embedding capacity and degree of perceptibility. The proposed algorithm achieves both the above mentioned criteria, i.e., high data embedding capacity as well as imperceptibility.

## 6. References

- [1] Ingemar J Cox, Mathew L. Miller, Jeffery A. Bloom, "Digital Watermarking", Morgan Kaufmann publisher, 2002.
- [2] J. Feng, I.Lin, C.Tsai, Y Chu, "Reversible Watermarking: Current Status and Key Issues", in international Journal of Network Security, vol.2, No.3, pp 161-171, May 2006.
- [3] Z. Ni, Y. Q. Shi, N.Ansari, and S. Wei, "Reversible data Hiding," in ISCAS Proceedings of the 2003 International Symposium on Circuits and Systems, vol. 2, pp. II-912-II-915, Thailand, May 2003.
- [4] Mohammad Awrangjed, "An Overview of Reversible Data Hiding", ICCIT 2003, 19-21 Dec, pp 75-79
- [5] J. Fridrich, J. Goljan, and R. Du, "Invertible authentication," in SPIE Proceedings of Security and Watermarking of Multimedia Content, pp. 197-208, San Jose, Jan 2002
- [6] J.Fridrich and M. Goljan, "Lossless data embedding for all image formats," in SPIE Proceedings of Photonics West, Electronic Imaging, Security and Watermarking of Multimedia Contents, vol. 4675, pp. 572-583, San Jose, Jan 2002.



## Web Document Clustering Approach using WordNet Lexical Categories and Fuzzy Clustering

Tarek F. Gharib<sup>1</sup>, Mohammed M. Fouad<sup>2</sup>, Mostafa M. Aref<sup>1</sup>

<sup>1</sup> Faculty of Computer and Information Science, Ain Shams University, Egypt

<sup>2</sup> Computer Science Department, Akhbar El-Yom Academy, Egypt

[tgharib@asunet.shams.edu.eg](mailto:tgharib@asunet.shams.edu.eg), [mmdoha200@yahoo.com](mailto:mmdoha200@yahoo.com), [aref\\_99@yahoo.com](mailto:aref_99@yahoo.com)

### Abstract

*Web mining is defined as applying data mining techniques to the content, structure, and usage of Web resources. The three areas of Web mining are commonly distinguished: content mining, structure mining, and usage mining. In all these areas, a wide range of general data mining techniques, in particular association rule discovery, clustering, classification, and sequence mining, are employed and developed further to reflect the specific structures of Web resources and the specific questions posed in Web mining. In this paper, we introduced a web document clustering approach that uses WordNet lexical categories and fuzzy c-means algorithm to improve the performance of clustering problem for web document. Experiments show that Fuzzy c-means algorithm achieves great performance optimization with comparison with the recent algorithms for document clustering.*

### 1. Introduction

The Web is a huge collection of documents that comprises some features like (1) semi-structured (HTML, XML) information, (2) hyper-link information, (3) access and usage information and (4) dynamic page generation. The Web has made cheaper the accessibility of a wider audience to various sources of information. The advances in all kinds of digital communication have provided greater access to networks. It has also created free access to a large publishing medium. These factors have allowed people to use the Web and modern digital libraries as a highly interactive medium [3].

Web mining is the application of data mining techniques to the content, structure, and usage of Web resources. It is thus “the nontrivial process of identifying valid, previously unknown, and potentially

useful patterns” [8]. Web mining application works on huge amount of Web data to extract some patterns that can describe this data. Like other data mining applications, Web mining can profit from given structure on data (as in database tables), but it can also be applied to semi-structured or unstructured data like free form text. This means that Web mining is an invaluable help in the transformation from human-understandable content to machine-understandable semantics.

Web content mining can be used to detect co-occurrences of terms in texts. For example, co-occurrences of terms in newswire articles may show that “gold” is frequently mentioned together with “copper” when articles concern Canada, but together with “silver” when articles concern the US. Trends over time may also be discovered, indicating a surge or decline in interest in certain topics such as the programming languages “Java”. Another application area is event detection: the identification of stories in continuous news streams that correspond to new or previously unidentified events [1].

Web content mining analyzes the content of Web resources. Today, it is mostly a form of text mining that can be defined as the process of extracting interesting information and knowledge from unstructured text. Text mining can be also defined as the application of algorithms and methods from the machine learning and statistics fields to texts with the goal of finding useful patterns [4]. In this paper, we deal with web content mining as an application of text mining process for input web documents.

One main reason for applying data mining methods to text document collections is to structure them. A structure can significantly simplify the access to a document collection for a user. Well known access structures are library catalogues or book indexes. However, the problem of manual designed indexes is the time required to maintain them. Therefore, they are

very often not up-to-date and thus not usable for recent publications or frequently changing information sources like the World Wide Web. The existing methods for structuring collections either try to assign keywords to documents based on a given keyword set (classification or categorization methods) or automatically structure document to find groups of similar documents (clustering methods).

Text document clustering can be defined as the process of grouping of text documents into semantically related groups [2]. Clustering methods aims to form the input documents into a set of clusters. Each cluster consists of number of documents that are similar to each other and dissimilar to documents of other clusters.

Many well-known methods of text mining have two problems: first, they don't consider semantically related words/terms (e.g., synonyms or hyper/hyponyms) in the document. For instance, they treat {Vehicle, Car, and Automobile} as different terms even though all these words have very similar meaning. This problem may lead to a very low relevance score for relevant documents because the documents do not always contain the same forms of words/terms.

Second, on vector representations of documents based on the bag-of-words model, text mining methods tend to use all the words/terms in the documents after removing the stop-words. This leads to thousands of dimensions in the vector representation of documents; this is called the "Curse of Dimensionality". However, it is well known that only a very small number of words/terms in documents have distinguishable power on clustering documents [9] and become the key elements of text summaries. Those words/terms are normally the concepts in the domain related to the documents.

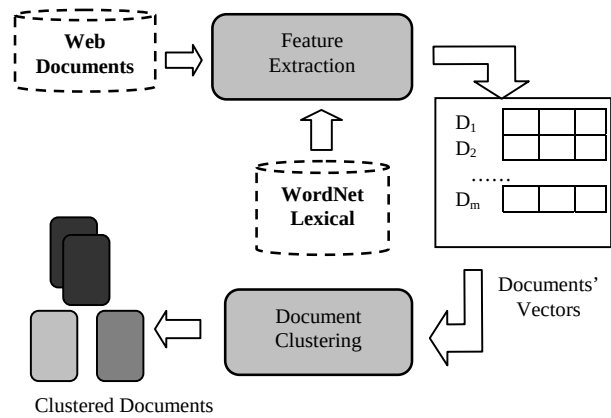
Recent studies for fuzzy clustering algorithms [25, 26, 27, 28] have proposed a new approach for using fuzzy clustering algorithms in document clustering process. But these studies neglect the lexical information that can be extracted from text as we used WordNet lexical categories in our proposed approach.

In this paper, we introduced an efficient approach web document clustering approach that uses WordNet lexical categories and fuzzy c-means algorithm to improve the performance of clustering problem for web document. The proposed approach uses WordNet words lexical categories information [12] to reduce the size of vector space and present semantic relationships between words. The generated document vectors will be input for the clustering process. In the clustering process we use fuzzy c-means algorithm to increase the clustering accuracy as shown in the experiments.

The rest of this paper is organized as following; section 2 show the proposed web document clustering approach. In section 3 a set of experiments is presented to compare the performance of the proposed approach with current document clustering algorithms. Related work is discussed and presented in section 4. Finally, conclusion is given in section 5.

## 2. Web Document Clustering Approach

In this section we describe in details the components of the proposed fuzzy web document clustering approach. There are two main processes: first is Feature Extraction that generates output document vectors from input text documents using WordNet lexical information as illustrated in figure 1.



**Figure 1: Fuzzy Web Document Clustering Approach**

The second process is Document Clustering that applies fuzzy c-means algorithm on document vectors to obtain output clusters.

### 2.1. Feature Extraction

The first step in the proposed approach is feature extraction or documents preprocessing which aims to represent the input web documents collection (corpus) into vector space model. In this model a set of words (terms) is extracted from the corpus called "bag-of-words" and represent each document as a vector by the words (terms) it contains and their weights regardless of their order. Documents preprocessing step contains five sub-steps: Text Extraction, PoS Tagging, Stopword Removal, Stemming, and WordNet Lexical Category Mapping.

**2.1.1. Text Extraction.** The first step of the feature extraction process is to extract textual data from the

web pages. Then convert each page into individual text document to apply text preprocessing techniques on it.

This step is applied on input Web documents dataset by scanning the web pages and categorizing the HTML tags in each page. Then exclude the tags that contain no textual information like formatting tags and imaging tags (i.e. <HTML>, <BODY>, <IMG>, etc.). Also exclude all the scripts and codes that are found in the page like JavaScript and VBScript. Then extract the textual data from other tags (like paragraphs, hyperlinks, and metadata tags) and store it into individual text documents as input for next steps.

**2.1.2. PoS Tagging.** The second step is to PoS tag the corpus. The PoS tagger relies on the text structure and morphological differences to determine the appropriate part-of-speech. This requires the words to be in their original order. This process is to be done before any other modifications on the corpora. For this reason, PoS tagging is the first step to be carried out on the text documents as proposed in [2].

**2.1.3. Stopwords Removal.** Stopwords, i.e. words thought not to convey any meaning, are removed from the text. In this work, the proposed approach uses a static list of stopwords with PoS information about all tokens. This process removes all words that are not nouns, verbs or adjectives. For example, stopwords removal process will remove all the words like: he, all, his, from, is, an, of, and so on.

**2.1.4. Stemming.** The stem is the common root-form of the words with the same meaning appear in various morphological forms (e.g. player, played, plays from stem play). In the proposed approach, we use the morphology function provided with WordNet for stemming process. Stemming will find the stems of the output terms to enhance term frequency counting process because terms like “learners” and “learning” come down from the same stem “learn”. This process will output all the stems of extracted terms.

The frequency of each stemmed word across the corpus can be counted and every word occurring less often than the pre-specified threshold (called Minimum Support) is pruned, i.e. removed from the extracted words list, to reduce the document vector dimension. In our implementation we use minimum support value set to 10%, which means that the words found in less than 10% of the input documents is pruned.

**2.1.5. WordNet Lexical Category Mapping.** As proposed in [5], we use WordNet lexical categories to map all the stemmed words in all documents into their lexical categories. We use WordNet 2.1 that has 41

lexical categories for nouns and verbs. For example, the word “dog” and “cat” both belong to the same category “noun.animal”. Some words also has multiple categories like word “Washington” has 3 categories (noun.location, noun.group, noun.person) because it can be the name of the American president, the city place, or a group in the concept of capital.

Some word disambiguation techniques are used to remove the resulting noise added by multiple categories mapping which are concept map and disambiguation by context as discussed with more details in [5].

## 2.2. Document Clustering

After generating the documents' vectors for all the input documents using feature extraction process, the proposed approach continues with the clustering process as shown in figure 1.

The problem of document clustering is defined as follows. Given a set of  $n$  documents called  $DS$ ,  $DS$  is clustered into a user-defined number of  $k$  document clusters  $D_1, D_2, \dots, D_k$ , (i.e.  $\{D_1, D_2, \dots, D_k\} = DS$ ) so that the documents in a document cluster are similar to one another while documents from different clusters are dissimilar.

There are two main approaches to document clustering, hierarchical clustering (agglomerative and divisive) and partitioning clustering algorithms [7]. In this process we apply three different clustering algorithms which are  $k$ -means (partitioning clustering), bisecting  $k$ -means (hierarchical clustering) and fuzzy  $c$ -means (fuzzy clustering).

**2.2.1. K-means and Bisecting k-means.** We have implemented the  $k$ -means and bisecting  $k$ -means algorithms as introduced in [7]. We will state some details on bisecting  $k$ -means algorithm that begins with all data as one cluster then perform the following steps:

1. Choose the largest cluster to split.
2. Use  $k$ -means to split this cluster into two sub-clusters. (Bisecting step)
3. Repeat step 2 for some iterations (in our case 10 times) and choose the split with the highest clustering overall similarity.
4. Go to step 1 again until the desired  $k$  clusters are obtained.

**2.2.2. Fuzzy c-means Algorithm.** Fuzzy  $c$ -means is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade while other classical clustering algorithms assign each data point to exactly one

cluster. This technique was originally introduced by Bezdek [16] as an improvement on earlier clustering methods. It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters. Most fuzzy clustering algorithms are objective function based: they determine an optimal (fuzzy) partition of a given data set  $c$  clusters by minimizing an objective function with some constraints [25].

In the proposed approach, we use the implementation of fuzzy c-means algorithm in MATLAB (fcm function). This function takes the generated document vectors as a matrix and the desired number of clusters then outputs the clusters centers and the optimal objective function values.

### 3. Experimental Results

Some experiments are performed on some real text documents to compare the performance of three text clustering algorithms which are  $k$ -means, bisecting  $k$ -means and fuzzy c-means. There are two main parameters to evaluate the performance of the proposed approach which are clustering quality and running time.

Fuzzy c-means algorithm is implemented in MATLAB and  $k$ -means, bisecting  $k$ -means algorithms are implemented in Java. Feature extraction process is also implemented in Java using Java NetBeans 5.5.1 and Java API for WordNet Searching (JAWS Library) to access WordNet 2.1.

#### 3.1. Silhouette Coefficient (SC) for Clustering Quality Evaluation

For clustering, two measures of cluster "goodness" or quality are used. One type of measure allows us to compare different sets of clusters without reference to external knowledge and is called an internal quality measure. The other type of measures lets us evaluate how well the clustering is working by comparing the groups produced by the clustering techniques to known classes which called an external quality measure [7].

In our application of document clustering, we don't have the knowledge of document classes in order to use external quality measures. We will investigate one of the main internal quality measures which is silhouette coefficient (SC Measure).

To measure the similarity between two documents  $d_1$  and  $d_2$  we use the cosine of the angle between the two document vectors. This measure tries to approach the semantic closeness of documents through the size of the angle between vectors associated to them:

$$\text{dist}(d_1, d_2) = \frac{d_1 \bullet d_2}{|d_1| \cdot |d_2|}$$

Where  $(\bullet)$  denotes vector dot product and  $(| |)$  is the dimension of the vector. A cosine measure of 0 means the two documents are unrelated whereas value closed to 1 means that the documents are closely related [5].

Let  $D_M = \{D_1, \dots, D_k\}$  describe a clustering result, i.e. it is an exhaustive partitioning of the set of documents  $DS$ . The distance of a document  $d \in DS$  to a cluster  $D_i \in D_M$  is given as:

$$\text{dist}(d, D_i) = \frac{\sum_{p \in D_i} \text{dist}(d, p)}{|D_i|}$$

Let further consider  $a(d, D_M) = \text{dist}(d, D_i)$  being the distance of document  $d$  to its cluster  $D_i$  where  $(d \in D_i)$ .  $b(d, D_M) = \min_{d \notin D_i} \text{dist}(d, D_i) \forall D_i \in D_M$  is the distance of document  $d$  to the nearest neighbour cluster. The silhouette  $S(d, D_M)$  of a document  $d$  is then defined as:

$$S(d, D_M) = \frac{b(d, D_M) - a(d, D_M)}{\max(b(d, D_M), a(d, D_M))}$$

The silhouette coefficient as:

$$SC(D_M) = \frac{\sum_{p \in DS} S(p, D_M)}{|DS|}$$

The silhouette coefficient is a measure for the clustering quality that is rather independent from the number of clusters. Experiences, such as documented in [11], show that values between 0.7 and 1.0 indicate clustering results with excellent separation between clusters, viz. data points are very close to the center of their cluster and remote from the next nearest cluster. For the range from 0.5 to 0.7 one finds that data points are clearly assigned to cluster centers. Values from 0.25 to 0.5 indicate that cluster centers can be found, though there is considerable "noise". Below a value of 0.25 it becomes practically impossible to find significant cluster centers and to definitely assign the majority of data points.

#### 3.2. Web Documents Datasets

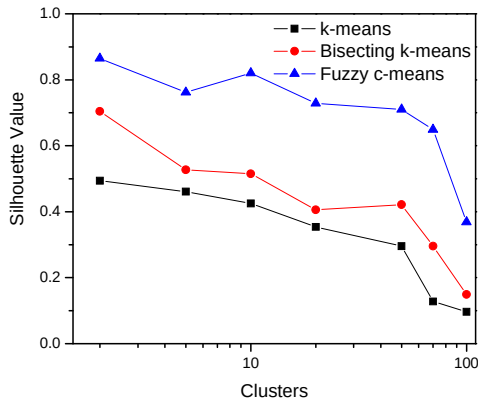
We evaluate the proposed web document clustering approach on three different web document datasets: SCOTS, EMail1200, and WebKB web documents corpuses. **SCOTS** corpus (Scottish Corpus Of Text and Speech) contains about 1,177 written and spoken texts, with about 4 million words of running text. 80%

of this total is made up of written texts and 20% is made up of spoken texts and can be obtained from [15]. **EEmail1200** corpus contains about 1,245 test email documents for spam email detection and can be downloaded from [14]. **WebKB** dataset contains about 8,282 web pages collected from computer science departments of various universities in January 1997. WebKB dataset is collected by the World Wide Knowledge Base (WebKb) project of the CMU text learning group and can be downloaded from [18].

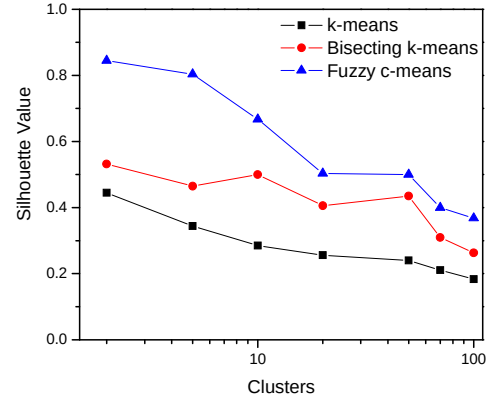
### 3.3. Results

First, we pass the three document datasets into the feature extraction process to generate the corresponding document vectors. The output vectors are used as an input for each clustering algorithm. We used the following list (2, 5, 10, 20, 50, 70, and 100) as the desired number of clusters for each algorithm. Then output clusters for each dataset are measured using silhouette coefficient (SC Measure) and report the total running time of the whole process.

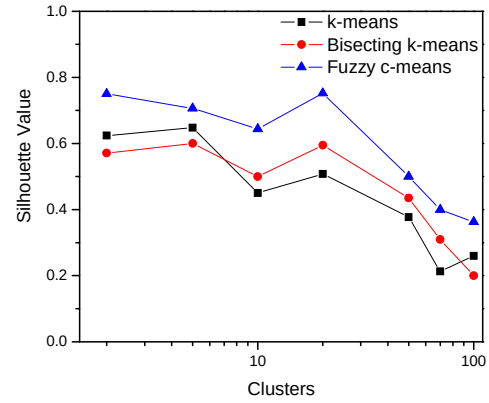
**3.3.1. Clustering Quality.** Figures 2, 3, and 4 show the silhouette values for the three datasets respectively. In all experiments fuzzy c-means outperforms bisecting *k*-means and *k*-means algorithms in the overall clustering quality using silhouette measure. This shows that the fuzzy clustering is more suitable for the unstructured nature of the text document clustering process itself.



**Figure 2: Silhouette values comparing all clustering algorithms using WebKB dataset**



**Figure 3: Silhouette values comparing all clustering algorithms using EMail1200 dataset**

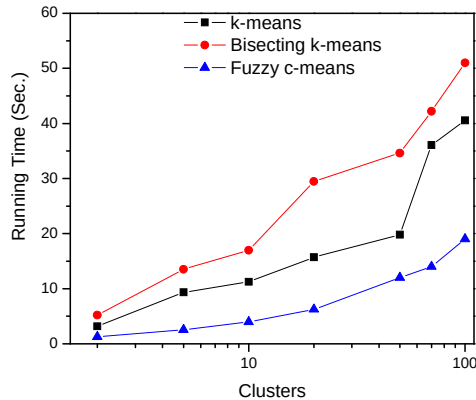


**Figure 4: Silhouette values comparing all clustering algorithms using SCOTS dataset**

**3.3.2. Running Time.** WebKB dataset, as mentioned in section 3.2, contains about 8,282 web documents. This is considered a real challenge task that faces any clustering approach because of “**Scalability**”. Some clustering techniques that are helpful for small data sets can be overwhelmed by large data sets to the point that they are no longer helpful.

For that reason we test the scalability of our proposed approach with the different algorithms using WebKB dataset. This experiment shows that the fuzzy c-means performs a great running time optimization with comparison to other two clustering algorithms. Also, according to the huge size of WebKB dataset, the proposed approach shows good scalability against document size.

Figure 5 depicts the running time of the different clustering algorithms using WebKB dataset with respect to different values of desired clusters.



**Figure 5: Scalability of all clustering algorithms on WebKB dataset**

## 4. Related Work

In the recent years, text document clustering has been introduced as an efficient method for navigating and browsing large document collections and organizing the results returned by search engines in response to user queries [13]. Many clustering techniques are proposed like k-secting k-means [17] and bisecting k-means [7], FTC and HFTC [19] and many others. From the performed experiments in [7] bisecting k-means overcomes all these algorithms in the performance although FTC and HTFC allows to reduce the dimensionality if the data when working with large datasets.

WordNet is used by Green [20, 21] to construct lexical chains from the occurrences of terms in a document: WordNet senses that are related receive high higher weights than senses that appear in isolation from others in the same document. The senses with the best weights are selected and the corresponding weighted term frequencies constitute a base vector representation of a document.

Other works [22, 23] have explored the possibility to use WordNet for retrieving documents by carefully choosing a search keyword. Dave [24] uses WordNet synsets as features for document representation and subsequent clustering. But the word sense disambiguation has not been performed showing that WordNet synsets decreases clustering performance in all the experiments. Hotho et al. [10] uses WordNet in an unsupervised scenario taking into account the WordNet ontology and lexicon and some strategy for word sense disambiguation achieving improvements of the clustering results.

Another technique for feature selection by using WordNet to discover synonymous terms based on

cross-referencing is introduced in [6]. First, terms with overlapping word senses co-occurring in a category are selected. A signature for a sense is a synset containing synonyms. Then, the list of noun synsets is checked for all senses for signatures similarity. The semantic context of a category is aggregated by overlapping synsets of different terms senses. The original terms from the category that belongs to the similar synsets will be finally added as features for category representation.

In [2] the authors explore the benefits of partial disambiguation of words by their PoS and the inclusion of WordNet concepts. They show how taking into account synonyms and hypernyms, disambiguated only by PoS tags, is not successful in improving clustering effectiveness because the noise produced by all the incorrect senses extracted from WordNet. Adding all synonyms and all hypernyms into the document vectors seems to increase the noise.

Reforgiato [5] presented a new unsupervised method for document clustering by using WordNet lexical and conceptual relations. In this work, Reforgiato uses WordNet lexical categories and WordNet ontology in order to create a well structured document vector space whose low dimensionality allows common clustering algorithms to perform well. For the clustering step he has chosen the bisecting k-means and the Multipole tree algorithms for their accuracy and speed.

Friedman et al. [26] introduced FDCM algorithm for clustering documents that are represented by vectors of variable size. The algorithm utilizes fuzzy logic to construct the cluster center and introduces a fuzzy based similarity measure which provided reasonably good results in the area of web document monitoring.

Rodrigues and Sacks [28] have modified the Fuzzy c-means algorithm for clustering text documents based on the cosine similarity coefficient rather than on the Euclidean distance. The modified algorithm works with normalized  $k$ -dimensional data vectors that lie in hyper-sphere of unit radius and hence has been named Hyper-spherical Fuzzy c-means (H-FCM). Also they proposed a hierarchical fuzzy clustering algorithm ( $H^2$ -FCM) to discover relationships between information resources based on their textual content, as well as to represent knowledge through the association of topics covered by those resources [27].

The main problem about these studies is that they neglect any lexical information about the textual data in the documents. This information helps in improving clustering quality as shown in section 3.3.

## 5. Conclusion

In this paper we proposed a fuzzy web document clustering approach based on the WordNet lexical categories and fuzzy c-means clustering algorithm. The proposed approach generates documents vectors using the lexical category mapping of WordNet after preprocessing the input web documents. We apply three different clustering algorithms, *k*-means, bisecting *k*-means and fuzzy *c*-means, to the generated documents vectors for clustering process. Output clusters of each algorithm are evaluated using silhouette coefficient as an internal cluster quality measure.

We have obtained some points that appeared in this work which are:

- Using WordNet lexical categories in feature extraction process reduces the dimension of the generated documents vectors and achieves higher clustering quality.
- Fuzzy *c*-means clustering algorithm achieves higher clustering quality than classical clustering algorithms like *k*-means (partitioning clustering), and bisecting *k*-means (hierarchical clustering).

From the shown experimental results, we found that the proposed approach shows good scalability against the huge number of documents as in WebKB dataset along with different values of desired clusters.

## 6. References

- [1] B. Berendt, A. Hotho, and G. Stumme, "Towards semantic web mining". In I. Horrocks and J.A. Hendler, editors, International Semantic Web Conference, volume 2342 of Lecture Notes in Computer Science, Springer 2002, pages 264–278.
- [2] J. Sedding, and D. Kazakov, "WordNet-based Text Document Clustering", COLING 3rd Workshop on Robust Methods in Analysis of Natural Language Data, 2004, pp 104–113.
- [3] M. Lan, C.L. Tan, H.B. Low, and S.Y. Sung, "A Comprehensive Comparative Study on Term Weighting Schemes", 14th International World Wide Web (WWW2005) Conference, Japan 2005.
- [4] G. Stumme, A. Hotho, and B. Berendt, "Semantic Web Mining: State of the art and future directions", Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 4, No. 2, 2006, pp. 124–143.
- [5] D. Reforgiato, "A new unsupervised method for document clustering by using WordNet lexical and conceptual relations." *Journal of Information Retrieval*, Vol (10), 2007, 563–579.
- [6] S. Chua, and N. Kulathuramaiyer, "Semantic feature selection using wordnet". In Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence, 2004, pp. 166–172.
- [7] M. Steinbach, G. Karypis, and V. Kumar "A Comparison of Document Clustering Techniques", Department of Computer Science and Engineering, University of Minnesota, Technical Report #00-034, 2000.
- [8] U.M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. "From data mining to knowledge discovery". *Advances in Knowledge Discovery and Data Mining*, AAAI / MIT Press, Cambridge, MA, 1996, pp 1-34.
- [9] B.B. Wang, R.I. McKay, H.A. Abbass, and M. Barlow, "Learning text classifier using the domain concept hierarchy." In Proceedings of International Conference on Communications, Circuits and Systems, China, 2002.
- [10] A. Hotho, S. Staab, and G. Stumme. "Wordnet improves text document clustering". ACM SIGIR Workshop on Semantic Web, 2003.
- [11] L. Kaufman, and P.J. Rousseeuw, "Finding Groups in Data: an Introduction to Cluster Analysis", John Wiley & Sons, 1999.
- [12] WordNet project available at <http://wordnet.princeton.edu/>
- [13] O. Zamir, O. Etzioni, O. Madani, and R.M. Karp, "Fast and intuitive clustering of web documents". KDD 97, 1997, pp. 287–290.
- [14] EMail1200 dataset, available at: <http://boole.cs.iastate.edu/book/acad/bag/data/lingspam>
- [15] SCOTS Project, Scottish Corpus Of Text and Speech, <http://www.scottishcorpus.ac.uk/>
- [16] J.C. Bezdec, "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York, 1981.
- [17] B. Larsen, and C. Aone, "Fast and effective text mining using linear-time document clustering". In Proc. of the 5th ACM SIGKDD international conference on knowledge discovery and data mining, 1999, pp. 16–22.
- [18] WebKB (The 4 Universities Dataset), available at: <http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-20/www/data/>
- [19] F. Beil, M. Ester, and X. Xu, "Frequent term-based text clustering". KDD 02, 2002, pp. 436–442.
- [20] S.J. Green, "Building hypertext links in newspaper articles using semantic similarity". NLDB 97, 1997, pp. 178–190.
- [21] S.J. Green, "Building hypertext links by computing

semantic similarity”. TKDE, 11(5), 1999, pp. 50–57.

[22] D.I. Moldovan, and R. Mihalcea, “Using wordnet and lexical operators to improve internet searches”. IEEE Internet Computing, 4(1), 2000, pp.34–43.

[23] E.M. Voorhees, “Query expansion using lexical-semantic relations”. In Proc. of ACM-SIGIR 1994, pp.61–69.

[24] K. Dave, “Mining the peanut gallery: Opinion extraction and semantic classification of product reviews”. WWW 03 ACM, 2003, pp. 519–528.

[25] C. Borgelt, and A. Nummerger, “Fast Fuzzy Clustering of Web Page Collections”. PKDD Workshop on Statistical Approaches for Web Mining, 2004.

[26] M. Friedman, A. Kandel, M. Schneider, M. Last, B. Shapka, Y. Elovici and O. Zaafrany, “A Fuzzy-Based Algorithm for Web Document Clustering”. Fuzzy Information, Processing NAFIPS '04, IEEE, 2004.

[27] M. E. S. Mendes Rodrigues, L. Sacks, “A Scalable Hierarchical Fuzzy Clustering Algorithm for Text Mining”. In Proceedings of the 5<sup>th</sup> International Conference on Recent Advances in Soft Computing, 2004.

[28] M. E. S. Mendes Rodrigues, L. Sacks, “Evaluating fuzzy clustering for relevance-based information access”. In Proceedings of the 12th IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2003.



## A Unifying Viewpoint of some Clustering Techniques Using Bregman Divergences and Extensions to Mixed Data Sets

Cécile Levasseur, Brandon Burdge, Ken Kreutz-Delgado  
University of California, San Diego  
Jacobs School of Engineering  
La Jolla, California, USA  
{clevasseur, bburdge, kreutz}@ucsd.edu

Uwe F. Mayer  
University of Utah  
Department of Mathematics  
Salt Lake City, Utah, USA  
mayer@math.utah.edu

### Abstract

*We present a general viewpoint using Bregman divergences and exponential family properties that contains as special cases the three following algorithms: 1) exponential family Principal Component Analysis (exponential PCA), 2) Semi-Parametric exponential family Principal Component Analysis (SP-PCA) and 3) Bregman soft clustering. This framework is equivalent to a mixed data-type hierarchical Bayes graphical model assumption with latent variables constrained to a low-dimensional parameter subspace. We show that within this framework exponential PCA and SP-PCA are similar to the Bregman soft clustering technique with the addition of a linear constraint in the parameter space. We implement the resulting modifications to SP-PCA and Bregman soft clustering for mixed (continuous and/or discrete) data sets, and add a nonparametric estimation of the point-mass probabilities to exponential PCA. Finally, we compare the relative performances of the three algorithms in a clustering setting for mixed data sets.*

### 1. Introduction

We present a general point of view that relates the exponential family Principal Component Analysis (exponential PCA) technique of [7] to the Semi-Parametric Principal Component Analysis (SP-PCA) technique of [20] and to the Bregman soft clustering method presented in [4]. The proposed standpoint is then illustrated with a clustering problem in mixed data sets.

The three techniques considered here all utilize Bregman divergences and can all be explained within a single hierarchical Bayes graphical model framework shown in Figure 1. They are not separate unrelated algorithms but different manifestations of parameter choices taken within a common

framework. The proposed model is mathematically equivalent to equation (6) and we will demonstrate that various parametric choices symbolically shown in Figure 2 determine the three algorithms. Because of this insight, we will readily extend the algorithms to deal with the important mixed data type case.

There are two ways in which Bregman divergences are important. First, they generalize the squared Euclidean distance to a class of distances that all share similar properties. Second, there exists a bijection between Bregman divergences and exponential family distributions [3, 4]. Recently, researchers have shown that many important algorithms can be generalized from Euclidean metrics to distances defined by a Bregman divergence [7, 20, 4, 6], i.e., from Gaussian distributed data components to data components distributed according to an exponential family, such as binary- or integer-valued.

Data mining techniques seek potentially useful information contained in complex data sets. Complexity of these data generally comes from a high number of components and also from the fact that the components usually are of mixed data types (categorical, count, continuous, etc.). In order to address the latter, techniques should allow for the components to have different parametric forms by using the large range of exponential family distributions and their associated Bregman divergences. While a modification for mixed data sets was presented in [12] for exponential PCA, we implement here a modification for SP-PCA and Bregman soft clustering. The issue of a high number of components of the original data is addressed by exponential PCA and SP-PCA through dimensionality reduction by projecting the data to a low-dimensional parameter subspace.

Finally, we consider synthetic data examples of mixed types. We demonstrate that exponential PCA, with the addition of a nonparametric estimation tool, rivals SP-PCA and Bregman soft clustering in terms of clustering performance.

## 2. Theoretical background

To motivate theoretical developments, the hierarchical Bayes graphical model for hidden or latent variables shown in Figure 1 is considered.

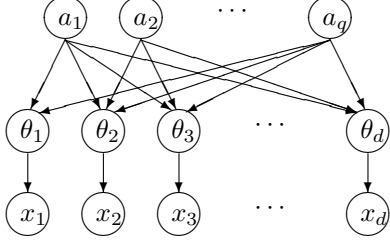


Figure 1. Graphical model for our framework.

The row vector  $\mathbf{x} = [x_1, \dots, x_d] \in \mathbb{R}^d$  consists of observed features in a  $d$ -dimensional space. It is assumed that training points can be drawn from populations having class-conditional probability density functions

$$p(\mathbf{x}|\boldsymbol{\theta}) = p_1(x_1|\theta_1) \cdot \dots \cdot p_d(x_d|\theta_d), \quad (1)$$

where, when conditioned on the random parameter vector  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_d] \in \mathbb{R}^d$ , the components of  $\mathbf{x}$  are independent. It is further assumed that  $\boldsymbol{\theta}$  can be written as

$$\boldsymbol{\theta} = \mathbf{a}\mathbf{V} + \mathbf{b}$$

with the hidden or latent variable  $\mathbf{a} = [a_1, \dots, a_q] \in \mathbb{R}^q$  random and unknown with  $q < d$  (and ideally  $q \ll d$ ),  $\mathbf{V} \in \mathbb{R}^{q \times d}$  and  $\mathbf{b} \in \mathbb{R}^d$  deterministic and unknown. The latent variable  $\mathbf{a}$  in some way explains part (or all) of the random behavior of the observed variables. The subscript  $i$  on  $p_i(\cdot|\cdot)$  serves to indicate that the marginal densities can all be different, allowing for the possibility of  $\mathbf{x}$  containing categorical, discrete, and continuous valued components. Also, the marginal densities are each assumed to be one-parameter exponential family densities, and  $\theta_i$  is taken to be the natural parameter (or some simple bijective function of it) of the exponential family density  $p_i$ . Hence, each component density  $p_i(x_i|\theta_i)$  in (1) for  $x_i \in \mathcal{X}_i, i = 1, \dots, d$ , is of the form

$$p(x_i|\theta_i) = \exp(\theta_i x_i - G(\theta_i)), \quad (2)$$

where  $G(\cdot)$  is the cumulant generating function defined as

$$G(\theta_i) = \log \int_{\mathcal{X}_i} \exp(\theta_i x_i) \nu(dx_i), \quad (3)$$

with  $\nu(\cdot)$  a  $\sigma$ -finite measure that generates the exponential family. It can be shown, using Fubini's theorem [11], that  $G(\boldsymbol{\theta}) = \sum_{i=1}^d G(\theta_i)$ .

The maximum likelihood identification of the blind random effect model

$$p(\mathbf{x}) = \int p(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \int \prod_{i=1}^d p_i(x_i|\theta_i)\pi(\boldsymbol{\theta})d\boldsymbol{\theta}, \quad (4)$$

where  $\pi(\boldsymbol{\theta})$  is the probability density function of  $\boldsymbol{\theta}$ , is quite a difficult problem. It corresponds to identifying  $\pi(\boldsymbol{\theta})$ , which, under the condition  $\boldsymbol{\theta} = \mathbf{a}\mathbf{V} + \mathbf{b}$ , corresponds to identifying the matrix  $\mathbf{V}$ , the vector  $\mathbf{b}$ , and a density function on the random effect  $\mathbf{a}$  via a maximization of the likelihood function  $p(\mathbf{X})$  with respect to  $\mathbf{V}$ ,  $\mathbf{b}$ , and the random effect density function, where

$$\begin{aligned} p(\mathbf{X}) &= \prod_{k=1}^n p(\mathbf{x}[k]) = \prod_{k=1}^n \int p(\mathbf{x}[k]|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} \\ &= \prod_{k=1}^n \int \prod_{i=1}^d p_i(x_i[k]|\theta_i)\pi(\boldsymbol{\theta})d\boldsymbol{\theta}, \end{aligned} \quad (5)$$

and  $\mathbf{X}$  is the  $(n \times d)$  observation matrix

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}[1] \\ \mathbf{x}[2] \\ \vdots \\ \mathbf{x}[n] \end{pmatrix} = \begin{pmatrix} x_1[1] & \dots & x_d[1] \\ x_1[2] & \dots & x_d[2] \\ \vdots & \ddots & \vdots \\ x_1[n] & \dots & x_d[n] \end{pmatrix}.$$

This difficulty can be avoided by NonParametric Maximum Likelihood (NPML) estimation of the random effect distribution, concurrently with the structural model parameters. The NPML estimate is known to be a discrete distribution on a finite number of support points or “atoms” [9, 10, 13]. Finding the NPML estimate is widely regarded as computationally intensive, the particular difficulty being the location of the atoms [1].

As shown in [12], with  $\boldsymbol{\theta} = \mathbf{a}\mathbf{V} + \mathbf{b}$ , with  $\mathbf{V}$ ,  $\mathbf{b}$  fixed and  $\mathbf{a}$  random, the single-sample likelihood (4) is equal to

$$p(\mathbf{x}) = \sum_{l=1}^m p(\mathbf{x}|\boldsymbol{\theta}[l])\pi_l = \sum_{l=1}^m p(\mathbf{x}|\mathbf{a}[l]\mathbf{V} + \mathbf{b})\pi_l$$

and the data likelihood (5) is equal to

$$p(\mathbf{X}) = \prod_{k=1}^n \sum_{l=1}^m p(\mathbf{x}[k]|\boldsymbol{\theta}[l])\pi_l = \prod_{k=1}^n \sum_{l=1}^m p(\mathbf{x}[k]|\mathbf{a}[l]\mathbf{V} + \mathbf{b})\pi_l, \quad (6)$$

with point-mass probability estimates  $\pi_l$ , unknown point-mass support points  $\mathbf{a}[l]$ , and the linear predictor  $\boldsymbol{\theta}[l] = \mathbf{a}[l]\mathbf{V} + \mathbf{b}$  in the  $l$ th mixture component,  $l = 1, \dots, m$ .

The data likelihood is thus approximately the likelihood of a finite mixture of exponential family densities with unknown mixture proportions or point-mass probability estimates  $\pi_l$  and unknown point-mass support points  $\mathbf{a}[l]$ ,

with the linear predictor  $\theta[l] = \mathbf{a}[l]\mathbf{V} + \mathbf{b}$  in the  $l$ th mixture component [2]. The combined problem of Maximum Likelihood Estimation (MLE) of the parameters  $\mathbf{V}$ ,  $\mathbf{b}$ , the point-mass support points (atoms)  $\mathbf{a}[l]$  and the point-mass probability estimates  $\pi_l, l = 1, \dots, m$ , is known as the Semiparametric Maximum Likelihood mixture density Estimation (SMLE) problem [14, 16]. It can be attacked by using the Expectation-Maximization (EM) algorithm [8, 10, 13, 19, 15, 2, 5, 16, 18], as done in particular in the Semi-Parametric Principal Component Analysis (SP-PCA) technique proposed in [20] and discussed below. Note that, historically, [10] appears to be generally acknowledged as the first paper that proposed the EM algorithm for NPML estimation in the mixture density context; then, [13] improved upon the theoretical foundations of the NPML estimation approach and later [15] further explored some of the fundamental issues raised in [13].

However, the SMLE problem can also be attacked by simply considering the special case of uniform point-mass probabilities, i.e.,  $\pi_l = 1/m$  for  $l = 1, \dots, m$ , for which the number of support points equals the number of data samples, i.e.,  $m = n$ . It was demonstrated in [12] that this special uniform case corresponds to exponential PCA.

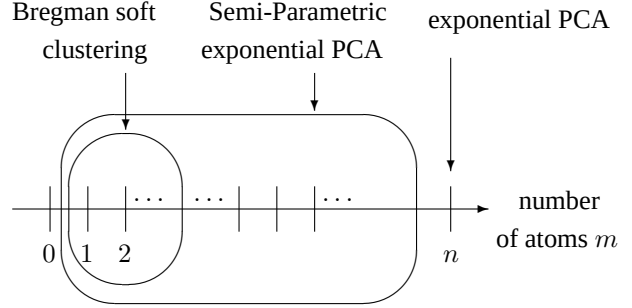
### 3. A unifying framework

Within the proposed hierarchical Bayes graphical model framework, exponential PCA, SP-PCA and Bregman soft clustering are not separate uncorrelated algorithms but different manifestations of parameter choices.

Figure 2 considers the number of atoms as a common characteristic for comparison purposes. It symbolizes how various parametric choices determine the three algorithms. Whereas the exponential PCA approach requires a number of atoms equal to the number of data points and hence can be seen as an extreme case of the NPML technique, SP-PCA deals with a smaller number of atoms. Finally, the Bregman soft clustering approach considers an even smaller number of atoms, viewed as cluster centers, since its primary goal is clustering. Furthermore, both exponential PCA and SP-PCA impose a low-dimensional (unknown) latent variable subspace in their structure. However, Bregman soft clustering does not impose this lower dimensional constraint and hence can be seen as a degenerate case.

It becomes clear while looking at Table 1, Table 2 and Table 3 shown below that both SP-PCA and Bregman soft clustering utilize the EM algorithm for estimation purposes whereas exponential PCA does not. This is because exponential PCA assumes uniform point-mass probabilities and does not need to estimate them. Both exponential PCA and SP-PCA impose the low-dimensional parameter subspace

constraint, hence the need for the Newton-Raphson iterative algorithm to find it as discussed below.



**Figure 2. General point of view on SP-PCA, exponential PCA and Bregman soft clustering based on the number of NPML atoms.**

## 4. Extensions to mixed data sets

Now that the relationship between exponential PCA, SP-PCA and Bregman soft clustering has been demonstrated within a unifying mixed data type framework, we can extend SP-PCA and Bregman soft clustering to deal with the important mixed data case. Recalling the modifications on exponential PCA presented in [12], we also introduce a penalty function that improves exponential PCA convergence efficiency. Towards these ends, following the notation in [5], we use the *mixing distribution*  $\mathcal{Q} = \{\theta[l], \pi_l\}_{l=1}^m$  which contains the parameters  $\theta[l]$  and their associated point-mass probabilities  $\pi_l$ . The mixing distribution needs to be estimated in all three algorithms.

### 4.1. Semi-parametric exponential PCA

The Semi-Parametric exponential family Principal Component Analysis (SP-PCA) approach presented in [20] attacks the Semiparametric Maximum Likelihood mixture density Estimation (SMLE) problem by using the Expectation-Maximization (EM) algorithm [8]. As done previously in [12] for exponential PCA, we present an SP-PCA modified approach for mixed data types. For simplicity of presentation, we consider that the  $f$  first attributes are distributed according to the exponential family distribution  $p^{(1)}$  and the  $(d - f)$  last attributes are distributed according to the exponential family distribution  $p^{(2)}$ . The following notation is used:

$$\begin{aligned} \mathbf{x}[k] &= [x_1[k], \dots, x_f[k], x_{f+1}[k], \dots, x_d[k]] \\ &= [\mathbf{x}^{(1)}[k] | \mathbf{x}^{(2)}[k]], \end{aligned}$$

for  $k = 1, \dots, n$ , and similarly for the observation matrix,

$$\mathbf{X} = (\mathbf{X}^{(1)} | \mathbf{X}^{(2)}).$$

The EM approach introduces a *missing* (unobserved) variable  $\mathbf{z}_k = [z_{k1}, \dots, z_{km}]$ , for  $k = 1, \dots, n$ . This variable is an  $m$ -dimensional binary vector whose  $l$ th component equals 1 if the observed variable  $\mathbf{x}[k]$  was drawn from the  $l$ th mixture component and 0 otherwise; its value is estimated during the E-step. Using this information, a *complete* log-likelihood function is defined as follows:

$$L^{(c)}(\mathcal{Q}, \{\mathbf{z}_k\}_{k=1}^n) = \log \prod_{k=1}^n \prod_{l=1}^m p(\mathbf{x}[k]|\boldsymbol{\theta}[l])^{z_{kl}} \pi_l^{z_{kl}},$$

with  $\boldsymbol{\theta}[l] = \mathbf{a}[l]\mathbf{V} + \mathbf{b}$ . Its maximization during the M-step yields parameters  $\mathbf{A}$ ,  $\mathbf{V}$  and  $\mathbf{b}$  estimates. Then,

$$p(\mathbf{x}[k]|\boldsymbol{\theta}[l]) = p^{(1)}(\mathbf{x}^{(1)}[k]|\boldsymbol{\theta}^{(1)}[l]) \cdot p^{(2)}(\mathbf{x}^{(2)}[k]|\boldsymbol{\theta}^{(2)}[l]) \quad (7)$$

using the assumption from (1), where, for  $l = 1, \dots, m$ ,

$$\boldsymbol{\theta}[l] = [\theta_1[l], \dots, \theta_f[l], \theta_{f+1}[l], \dots, \theta_d[l]] = [\boldsymbol{\theta}^{(1)}[l] | \boldsymbol{\theta}^{(2)}[l]],$$

$$\boldsymbol{\Theta} = \begin{pmatrix} \boldsymbol{\theta}[1] \\ \boldsymbol{\theta}[2] \\ \vdots \\ \boldsymbol{\theta}[n] \end{pmatrix} = (\boldsymbol{\Theta}^{(1)} | \boldsymbol{\Theta}^{(2)}).$$

The following decompositions arise:

$$\mathbf{V} = (\mathbf{V}^{(1)} | \mathbf{V}^{(2)}),$$

$$\mathbf{B} = (\mathbf{B}^{(1)} | \mathbf{B}^{(2)}),$$

where  $\mathbf{B}^{(1)} = [\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(1)}]^T$  and  $\mathbf{b}^{(1)} = [b_1, \dots, b_f]$ ,  $\mathbf{B}^{(2)} = [\mathbf{b}^{(2)}, \dots, \mathbf{b}^{(2)}]^T$  and  $\mathbf{b}^{(2)} = [b_{f+1}, \dots, b_d]$ . Hence,

$$\boldsymbol{\Theta} = (\mathbf{A}\mathbf{V}^{(1)} + \mathbf{B}^{(1)} | \mathbf{A}\mathbf{V}^{(2)} + \mathbf{B}^{(2)}).$$

Note that there is no such split for  $\mathbf{A}$ . The complete log-likelihood function becomes:

$$\begin{aligned} L^{(c)}(\mathcal{Q}, \{\mathbf{z}_k\}_{k=1}^n) &= \sum_{k=1}^n \sum_{l=1}^m z_{kl} \log \pi_l \\ &+ \sum_{k=1}^n \sum_{l=1}^m z_{kl} \log p^{(1)}(\mathbf{x}^{(1)}[k]|\boldsymbol{\theta}^{(1)}[l]) \cdot p^{(2)}(\mathbf{x}^{(2)}[k]|\boldsymbol{\theta}^{(2)}[l]). \end{aligned} \quad (8)$$

The E-step yields for  $k = 1, \dots, n$  and  $l = 1, \dots, m$ :

$$\begin{aligned} \hat{z}_{kl} &= E\{z_{kl}|\mathbf{x}[k], \pi_1, \dots, \pi_m\} \\ &= \frac{p^{(1)}(\mathbf{x}^{(1)}[k]|\boldsymbol{\theta}^{(1)}[l]) \cdot p^{(2)}(\mathbf{x}^{(2)}[k]|\boldsymbol{\theta}^{(2)}[l]) \pi_l}{\sum_{r=1}^m p^{(1)}(\mathbf{x}^{(1)}[k]|\boldsymbol{\theta}^{(1)}[r]) \cdot p^{(2)}(\mathbf{x}^{(2)}[k]|\boldsymbol{\theta}^{(2)}[r]) \pi_r}. \end{aligned}$$

For all  $l$  and all  $k$ , each data point  $\mathbf{x}[k]$  has an estimated probability  $\hat{z}_{kl}$  of belonging to the  $l$ th mixture component.

The M-step first yields the estimates for the point-mass probabilities:

$$\pi_l^{(new)} = \frac{\sum_{k=1}^n \hat{z}_{kl}}{n}.$$

The second part of the M-step, i.e., the estimation of the parameters  $\mathbf{V}$ ,  $\mathbf{b}$ , and the point-mass support points  $\mathbf{A} = [\mathbf{a}[1]^T, \dots, \mathbf{a}[m]^T]^T \in \mathbb{R}^{m,q}$ , is affected by the mixed data type assumption. It consists of maximizing the complete log-likelihood function (8) with respect to these parameters:

$$\arg \max_{\mathbf{A}, \mathbf{V}, \mathbf{b}} E \left\{ L^{(c)}(\{\boldsymbol{\theta}[l], \pi_l^{(new)}\}_{l=1}^m, \{\hat{\mathbf{z}}_k\}_{k=1}^n) \right\}.$$

Following the notation in [20], we set, for  $l = 1, \dots, m$ :

$$\tilde{\mathbf{x}}[l] = \frac{\sum_{k=1}^n \hat{z}_{kl} \mathbf{x}[k]}{\sum_{k=1}^n \hat{z}_{kl}},$$

the  $l$ th mixture component center. It can be shown, using exponential family properties, that the loss function is:

$$\begin{aligned} L(\mathbf{A}, \mathbf{V}, \mathbf{b}) &= \sum_{l=1}^m \left\{ G^{(1)}(\mathbf{a}[l]\mathbf{V}^{(1)} + \mathbf{b}^{(1)}) - (\mathbf{a}[l]\mathbf{V}^{(1)} + \mathbf{b}^{(1)})\tilde{\mathbf{x}}[l]^T \right\} \\ &+ \sum_{l=1}^m \left\{ G^{(2)}(\mathbf{a}[l]\mathbf{V}^{(2)} + \mathbf{b}^{(2)}) - (\mathbf{a}[l]\mathbf{V}^{(2)} + \mathbf{b}^{(2)})\tilde{\mathbf{x}}[l]^T \right\}, \end{aligned} \quad (9)$$

where  $G^{(1)}(\cdot)$ ,  $G^{(2)}(\cdot)$  respectively, is the cumulant generating function associated with the exponential family distribution  $p^{(1)}(\cdot)$ ,  $p^{(2)}(\cdot)$  respectively, as seen in equation (3).

Following the derivations in [12], the Newton-Raphson method is used for the iterative minimization of the loss function (9) and the resulting update equations are as follows. First at iteration  $t$ , for  $l = 1, \dots, m$ ,

$$\begin{aligned} \mathbf{a}^{(t+1)}[l]^T &= \mathbf{a}^{(t)}[l]^T - \alpha_{\mathbf{a}}^{(t+1)} \\ &\cdot \left\{ \mathbf{V}^{(1)(t)} G^{(1)'}(\mathbf{a}^{(t)}[k]\mathbf{V}^{(1)(t)} + \mathbf{b}^{(1)(t)})\mathbf{V}^{(1)(t),T} \right. \\ &+ \left. \mathbf{V}^{(2)(t)} G^{(2)'}(\mathbf{a}^{(t)}[k]\mathbf{V}^{(2)(t)} + \mathbf{b}^{(2)(t)})\mathbf{V}^{(2)(t),T} \right\}^{-1} \\ &\cdot \left\{ \mathbf{V}^{(1)(t)} (G^{(1)'}(\mathbf{a}^{(t)}[k]\mathbf{V}^{(1)(t)} + \mathbf{b}^{(1)(t)}) - \tilde{\mathbf{x}}[k]^T) \right. \\ &+ \left. \mathbf{V}^{(2)(t)} (G^{(2)'}(\mathbf{a}^{(t)}[k]\mathbf{V}^{(2)(t)} + \mathbf{b}^{(2)(t)}) - \tilde{\mathbf{x}}[k]^T) \right\}. \end{aligned} \quad (10)$$

For the second step, the two sets of row vectors  $\{\mathbf{v}_j^{(1)}\}_{j=1}^q$  and  $\{\mathbf{v}_j^{(2)}\}_{j=1}^q$  are updated separately. The update equations can then be used for  $\{\mathbf{v}_j^{(1)}\}_{j=1}^q$  and  $\{\mathbf{v}_j^{(2)}\}_{j=1}^q$  by changing  $\mathbf{v}_j$  to  $\mathbf{v}_j^{(1)}$ , respectively to  $\mathbf{v}_j^{(2)}$ ,  $\mathbf{b}$  to  $\mathbf{b}^{(1)}$ , respectively to  $\mathbf{b}^{(2)}$ ,  $G(\cdot)$ ,  $G'(\cdot)$ , and

$G''(\cdot)$  to  $G^{(1)}(\cdot)$ ,  $G^{(1)'}(\cdot)$ , and  $G^{(1)''}(\cdot)$ , respectively to  $G^{(2)}(\cdot)$ ,  $G^{(2)'}(\cdot)$ , and  $G^{(2)''}(\cdot)$ . For  $j = 1, \dots, q$ :

$$\begin{aligned} \mathbf{v}_j^{(t+1),T} &= \mathbf{v}_j^{(t),T} \\ &- \alpha_{\mathbf{v}}^{(t+1)} \left( \sum_{l=1}^m \underline{a}_j^{(t+1)} [l]^2 G''(\underline{\mathbf{a}}^{(t+1)}[l] \mathbf{V}^{(t)} + \mathbf{b}^{(t)}) \right)^{-1} \\ &\cdot \left( \sum_{k=1}^n \underline{a}_j^{(t+1)} [l] \{ G'(\underline{\mathbf{a}}^{(t+1)}[l] \mathbf{V}^{(t)} + \mathbf{b}^{(t)}) - \tilde{\mathbf{x}}[k]^T \} \right). \end{aligned} \quad (11)$$

And finally for the last step, the update equations can then be used for  $\mathbf{b}^{(1)}$  and  $\mathbf{b}^{(2)}$  by changing  $\mathbf{b}$  to  $\mathbf{b}^{(1)}$ , respectively to  $\mathbf{b}^{(2)}$ ,  $\mathbf{V}$  to  $\mathbf{V}^{(1)}$ , respectively to  $\mathbf{V}^{(2)}$ ,  $G(\cdot)$ ,  $G'(\cdot)$ , and  $G''(\cdot)$  to  $G^{(1)}(\cdot)$ ,  $G^{(1)'}(\cdot)$ , and  $G^{(1)''}(\cdot)$ , respectively to  $G^{(2)}(\cdot)$ ,  $G^{(2)'}(\cdot)$ , and  $G^{(2)''}(\cdot)$ .

$$\begin{aligned} \mathbf{b}^{(t+1),T} &= \mathbf{b}^{(t),T} \\ &= \mathbf{b}^{(t),T} - \alpha_{\mathbf{b}}^{(t+1)} \left( \sum_{l=1}^m G''(\underline{\mathbf{a}}^{(t+1)}[l] \mathbf{V}^{(t+1)} + \mathbf{b}) \right)^{-1} \\ &\cdot \left( \sum_{l=1}^m \{ G'(\underline{\mathbf{a}}^{(t+1)}[l] \mathbf{V}^{(t+1)} + \mathbf{b}) - \tilde{\mathbf{x}}[k]^T \} \right). \end{aligned} \quad (12)$$

Table 1 summarizes the SP-PCA algorithm.

## 4.2. Exponential family PCA

The work proposed in [12] is based on the generalization of Principal Component Analysis to the exponential family technique presented in [7], often referred to as exponential Principal Component Analysis.

As stated earlier, instead of the fastidious estimation of the point-mass probabilities, exponential PCA considers the special uniform case where  $\pi_l = 1/m$  for  $l = 1, \dots, m$ , for which the number of support points equals the number of data samples ( $m = n$ ). Hence, the point-mass probabilities do not need to be estimated and the EM algorithm is unnecessary. Then, to each vector  $\mathbf{x}$  corresponds a vector  $\underline{\mathbf{a}}$ , i.e., a vector  $\underline{\boldsymbol{\theta}}$ , and they can share the same index  $k = 1, \dots, n$ .

The update equations were presented in [12]. It is easily noticed that they are the same as equations (10), (11) and (12), the only difference being that the exponential PCA update equations use  $\mathbf{x}$  instead of  $\tilde{\mathbf{x}}$ , i.e., data points instead of mixture component centers. Each data point  $\mathbf{x}$  is its own mixture component center.

As noted in [7], it is possible for the atoms obtained with exponential PCA to diverge since the optimum may be at infinity. To avoid such behavior, we introduce a penalty function that defines and places a set of constraints into the

---

### Algorithm: Semi-Parametric exp. family PCA [20]

---

**Input:** a set of observations  $\{\mathbf{x}[k]\}_{k=1}^n \subseteq \mathbb{R}^d$ , two exponential family distribution  $p^{(1)}, p^{(2)}$  defined by their cumulant generating functions  $G^{(1)}, G^{(2)}$ , a number of atoms  $m, q \ll d$  the dimension of the latent variable lower dimensional subspace.

**Output:** the NPML estimator that maximizes the complete log-likelihood function  $L^{(c)}(\mathcal{Q}, \{\mathbf{z}_k\}_{k=1}^n)$ :  $\hat{\mathcal{Q}} = \{\hat{\boldsymbol{\theta}}[l], \hat{\pi}_l\}_{l=1}^m$  with  $\hat{\boldsymbol{\theta}}[l] = \hat{\mathbf{a}}[l] \hat{\mathbf{V}} + \hat{\mathbf{b}}$  for all  $l$ ,  $\{\hat{\mathbf{a}}[l]\}_{l=1}^m \in \mathbb{R}^q$ ,  $\hat{\mathbf{V}} \in \mathbb{R}^{q \times d}$  and  $\hat{\mathbf{b}} \in \mathbb{R}^d$ .

---

#### Method:

Initialize  $\mathbf{V}$ ,  $\mathbf{b}$  and  $\{\mathbf{a}[l], \pi_l\}_{l=1}^m$  with  $\pi_l \geq 0$  for all  $l$  and  $\sum_{l=1}^m \pi_l = 1$ ;  $\boldsymbol{\theta}[l] = \mathbf{a}[l] \mathbf{V} + \mathbf{b} \in \boldsymbol{\Theta}$  for all  $l$ ;  $p(\mathbf{x}[k]|\boldsymbol{\theta}[l])$  as defined in (7) for all  $k$  and  $l$ ;

#### repeat

{The Expectation Step}

for  $k = 1$  to  $n$  do

for  $l = 1$  to  $m$  do

$\hat{z}_{kl} \leftarrow p(\mathbf{x}[k]|\boldsymbol{\theta}[l])\pi_l / \sum_{r=1}^m p(\mathbf{x}[k]|\boldsymbol{\theta}[r])\pi_r$

end for

end for

{The Maximization Step}

for  $l = 1$  to  $m$  do

$\pi_l \leftarrow (1/n) \sum_{k=1}^n \hat{z}_{kl}$

end for

{The Newton-Raphson iterative algorithm}

for  $l = 1$  to  $m$  do

$\mathbf{a}[l] \leftarrow$  update equation (10)

end for

for  $j = 1$  to  $q$  do

$\mathbf{v}_j \leftarrow$  update equation (11)

end for

$\mathbf{b} \leftarrow$  update equation (12)

until convergence;

return  $\hat{\mathcal{Q}} = \{\hat{\boldsymbol{\theta}}[l] = \hat{\mathbf{a}}[l] \hat{\mathbf{V}} + \hat{\mathbf{b}}, \hat{\pi}_l\}_{l=1}^m$ .

---

**Table 1. Semi-Parametric PCA algorithm.**

loss function via a penalty parameter in a way that penalizes any divergence to infinity. The penalty function

$$\begin{aligned} \psi(\boldsymbol{\theta}) &= \sum_{i=1}^d \left\{ \exp(-\beta_{\min}(\theta_i - \theta_{\min,i})) \right. \\ &\quad \left. + \exp(\beta_{\max}(\theta_i - \theta_{\max,i})) \right\}, \end{aligned}$$

is designed so that  $\psi(\boldsymbol{\theta})$  is close to zero for  $\theta_{\min} \leq \theta \leq \theta_{\max}$  and reaches infinity otherwise. The penalty function modifications on the update equations (10), (11) and (12), i.e., an additional additive term within each of the large parentheses, are omitted here for sake of brevity.

Table 2 summarizes exponential PCA algorithm.

---

**Algorithm:** Exponential PCA [7]

---

**Input:** a set of observations  $\{\mathbf{x}[k]\}_{k=1}^n \subseteq \mathbb{R}^d$ , two exponential family distribution  $p^{(1)}, p^{(2)}$  defined by their cumulant generating functions  $G^{(1)}, G^{(2)}$ , a number of atoms  $n, q \ll d$  the dimension of the latent variable lower dimensional subspace.

**Output:** the NPML estimator that maximizes the log-likelihood function  $L(\mathcal{Q})$ :  $\hat{\mathcal{Q}} = \{\hat{\boldsymbol{\theta}}[k]\}_{k=1}^n$  with  $\hat{\boldsymbol{\theta}}[k] = \hat{\mathbf{a}}[k]\hat{\mathbf{V}} + \hat{\mathbf{b}}$  for all  $k$ ,  $\{\hat{\mathbf{a}}[k]\}_{k=1}^n \in \mathbb{R}^q$ ,  $\hat{\mathbf{V}} \in \mathbb{R}^{q \times d}$  and  $\hat{\mathbf{b}} \in \mathbb{R}^d$ .

**Method:**

Initialize  $\mathbf{V}, \mathbf{b}$  and  $\{\mathbf{a}[k]\}_{k=1}^n$ ;  $\boldsymbol{\theta}[k] = \mathbf{a}[k]\mathbf{V} + \mathbf{b} \in \Theta$  for all  $k$ ;  $p(\mathbf{x}[k]|\boldsymbol{\theta}[k])$  as defined in (7) for all  $k$ ;

**repeat**

{The Newton-Raphson iterative algorithm}

**for**  $k = 1$  to  $n$  **do**

$\mathbf{a}[l] \leftarrow$  penalty-modified update equation (10)  
with  $\mathbf{x}$  instead of  $\tilde{\mathbf{x}}$

**end for**

**for**  $j = 1$  to  $q$  **do**

$\mathbf{v}_j \leftarrow$  penalty-modified update equation (11)  
with  $\mathbf{x}$  instead of  $\tilde{\mathbf{x}}$

**end for**

$\mathbf{b} \leftarrow$  penalty-modified update equation (12)  
with  $\mathbf{x}$  instead of  $\tilde{\mathbf{x}}$

**until convergence;**

return  $\hat{\mathcal{Q}} = \{\hat{\boldsymbol{\theta}}[k] = \hat{\mathbf{a}}[k]\hat{\mathbf{V}} + \hat{\mathbf{b}}\}_{k=1}^n$ .

---

**Table 2. Exponential PCA algorithm.**

### 4.3. Bregman soft clustering

The Bregman soft clustering approach presented in [4] utilizes an alternative interpretation of the EM algorithm for learning models involving mixtures of exponential family distributions. It is a simple soft clustering algorithm for all Bregman divergences, i.e., for all exponential family distributions. We choose here to present this technique without referring to the Bregman divergence as done in [4] but by using its corresponding exponential family probability distribution for the sake of comparison with SP-PCA and exponential PCA.

Given a data set of observations  $\{\mathbf{x}[k]\}_{k=1}^n$ , Bregman soft clustering aims at modeling the statistical structure of the data as a mixture of  $m$  densities of the same exponential family. The clusters correspond to the components of the mixture model and the soft membership of a data point in each cluster is proportional to the probability of the data point being generated by the corresponding density function. The Bregman soft clustering problem is based on a

---

**Algorithm:** Bregman Soft Clustering [4]

---

**Input:** a set of observations  $\{\mathbf{x}[k]\}_{k=1}^n \subseteq \mathbb{R}^d$ , two exponential family distribution  $p^{(1)}, p^{(2)}$  defined by their cumulant generating functions  $G^{(1)}, G^{(2)}$ , a number of atoms  $m$ .

**Output:** the NPML estimator that maximizes the complete log-likelihood function  $L^{(c)}(\mathcal{Q}, \{\mathbf{z}_k\}_{k=1}^n)$ :  $\hat{\mathcal{Q}} = \{\hat{\boldsymbol{\theta}}[l], \hat{\pi}_l\}_{l=1}^m$ .

**Method:**

Initialize  $\{\boldsymbol{\theta}[l], \pi_l\}_{l=1}^m$  with  $\pi_l \geq 0$  for all  $l$  and  $\sum_{l=1}^m \pi_l = 1$ ;  $p(\mathbf{x}[k]|\boldsymbol{\theta}[l])$  as defined in (7) for all  $k$  and  $l$ ;  $\boldsymbol{\theta}[l] \in \Theta$  for all  $l$ ;

**repeat**

{The Expectation Step}

**for**  $k = 1$  to  $n$  **do**

**for**  $l = 1$  to  $m$  **do**

$\hat{z}_{kl} \leftarrow p(\mathbf{x}[k]|\boldsymbol{\theta}[l])\pi_l / \sum_{r=1}^m p(\mathbf{x}[k]|\boldsymbol{\theta}[r])\pi_r$

**end for**

**end for**

{The Maximization Step}

**for**  $l = 1$  to  $m$  **do**

$\pi_l \leftarrow (1/n) \sum_{k=1}^n \hat{z}_{kl}$

$\boldsymbol{\theta}[l] \leftarrow$  solve for  $\boldsymbol{\theta}[l]$ :

$G'(\boldsymbol{\theta}[l]) = \sum_{k=1}^n \hat{z}_{kl}\mathbf{x}[k] / \sum_{k=1}^n \hat{z}_{kl}$

**end for**

**until convergence;**

return  $\hat{\mathcal{Q}} = \{\hat{\boldsymbol{\theta}}[l], \hat{\pi}_l\}_{l=1}^m$ .

---

**Table 3. Bregman soft clustering algorithm.**

maximum likelihood estimation of the cluster parameters  $\{\boldsymbol{\theta}[l], \pi_l\}_{l=1}^m$  satisfying the following mixture structure:

$$p(\mathbf{x}) = \sum_{l=1}^m p(\mathbf{x}|\boldsymbol{\theta}[l])\pi_l,$$

where  $p(\mathbf{x}|\cdot)$  is an exponential family distribution. The data likelihood function takes the following form:

$$p(\mathbf{X}) = \prod_{k=1}^n \sum_{l=1}^m p(\mathbf{x}[k]|\boldsymbol{\theta}[l])\pi_l. \quad (13)$$

The data likelihood function in (13) is similar to the data likelihood function in (6) without the linear constraint  $\boldsymbol{\theta}[l] = \mathbf{a}[l]\mathbf{V} + \mathbf{b}$  for  $l = 1, \dots, m$ . Hence, the Bregman soft clustering problem is similar to the SP-PCA problem without the lower dimensional subspace constraint and a simple EM algorithm is used to estimate the cluster parameters. We consider again the mixed data type case. The E-step and the first part of the M-step yield the same results as for SP-PCA. In the second part of the M-step, the component parameters

$\theta[l], l = 1, \dots, m$ , are estimated in the following way:

$$\begin{aligned}\theta[l]^{(new)} &= \arg \max_{\theta[l]} \sum_{k=1}^n \sum_{r=1}^m \hat{z}_{kr} \log p(\mathbf{x}[k]|\theta[r]) \\ &= \arg \max_{\theta[l]} \left\{ \sum_{k=1}^n \sum_{r=1}^m \hat{z}_{kr} \log p^{(1)}(\mathbf{x}^{(1)}[k]|\theta^{(1)}[r]) \right. \\ &\quad \left. + \sum_{k=1}^n \sum_{r=1}^m \hat{z}_{kr} \log p^{(2)}(\mathbf{x}^{(2)}[k]|\theta^{(2)}[r]) \right\},\end{aligned}$$

with  $\log p(\mathbf{x}[k]|\theta[r]) = \theta[r]\mathbf{x}[k]^T - G(\theta[r])$ . Using the convexity properties of  $G(\cdot)$ , it is easily shown that:

$$G'(\theta[l]^{(new)}) = \left( \sum_{k=1}^n \hat{z}_{kl} \mathbf{x}[k] \right) / \left( \sum_{k=1}^n \hat{z}_{kl} \right)$$

can be solved for  $\theta[l]^{(new),(1)}$  and  $\theta[l]^{(new),(2)}$  by changing  $\mathbf{x}$  to  $\mathbf{x}^{(1)}$ , respectively to  $\mathbf{x}^{(2)}$ ,  $G'(\cdot)$  to  $G^{(1)'}(\cdot)$ , respectively to  $G^{(2)'}(\cdot)$ .

Table 3 summarizes the Bregman soft clustering algorithm.

## 5. Experimental results on synthetic data

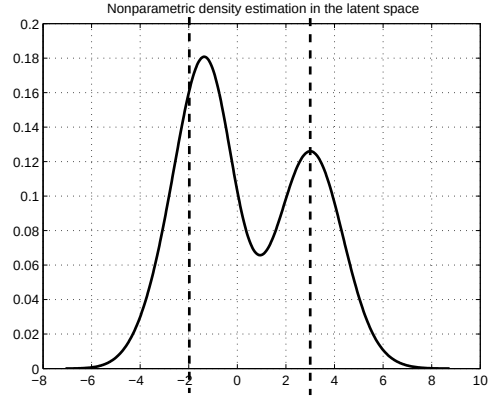
We compare the relative performances of exponential PCA, SP-PCA and Bregman soft clustering in a mixed data set clustering problem with two data types and demonstrate how exponential PCA with the addition of a nonparametric estimation of the point-mass probabilities exceeds SP-PCA in performance.

We first consider a synthetic  $d = 3$ -dimensional data set with a lower dimensional subspace of dimension  $q = 1$ . The first data feature is Poisson distributed, the second and third features are Gaussian distributed. The data has  $n = 500$  points and is composed of two mixture components with parameters  $\theta[1]$  and  $\theta[2]$  constrained to the lower dimensional subspace.

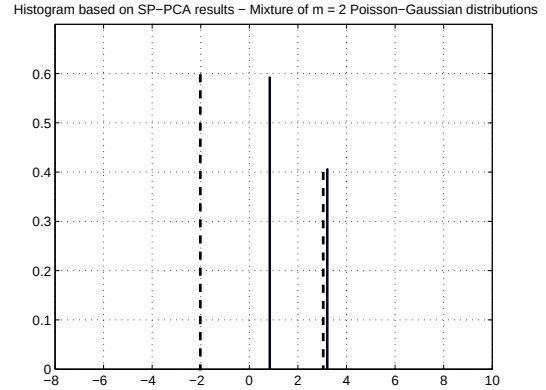
We first use exponential PCA. However, exponential PCA does not estimate point-mass probabilities. We use a nonparametric density estimation technique based on a kernel smoothing method to estimate the point-mass probabilities using the support points values  $\mathbf{a}[k]$ ,  $k = 1, \dots, n$ , obtained by exponential PCA. Figure 3 shows that the nonparametric density estimation exhibits a definite two-component shape. The dotted lines represent the correct values  $\mathbf{a}[1]$  and  $\mathbf{a}[2]$ . We can then estimate the values of  $\mathbf{a}[1]$  and  $\mathbf{a}[2]$  as well as their mixing distributions  $\pi_1$  and  $\pi_2$  using a simple kmeans algorithm, with the  $\pi_1 + \pi_2 = 1$  assumption.

Figure 4 presents the histogram of the estimated point-mass probabilities obtained with SP-PCA,  $m = 2$ .

Table 4 shows detailed results for this synthetic data setting (“modified” means the extension to mixed data sets of the algorithm): the mixing distributions or point-mass probabilities  $\pi_1$  and  $\pi_2$ , the latent variable or point of support values  $\mathbf{a}[1]$  and  $\mathbf{a}[2]$ , the parameter values  $\theta[1]$  and  $\theta[2]$  as well as the sine of the angle between the estimated lower dimensional subspace and the correct subspace. Bregman soft clustering does not have the lower dimensional subspace constraint, and hence does not exhibit a sine or the latent variables values in Table 4. The estimation quality of the  $\theta[1]$ ,  $\theta[2]$  and  $\pi_1$ ,  $\pi_2$  values defines the clustering performance. For this simple Poisson-Gaussian mixed data setting, both exponential PCA and Bregman soft clustering seem to perform better than SP-PCA: the SP-PCA obtained parameter values for  $\theta[2]$  are far from the original values, contrary to exponential PCA and Bregman soft clustering.



**Figure 3. Nonparametric estimation of the point-mass probabilities obtained with exponential PCA (dotted: correct cluster centers).**



**Figure 4. Histogram of the estimated point-mass probabilities obtained with SP-PCA (dotted: correct cluster values).**

Results for a second experiment are shown in Table 5 for a Binomial-Gaussian mixed data set created in a similar fashion as the Poisson-Gaussian mixed data set (the parameter  $N$  is set to 10 for the Binomial component). Again, exponential PCA exceeds SP-PCA in clustering performance.

	$\pi_1; \pi_2$	$\mathbf{a}[1]; \mathbf{a}[2]$	$\boldsymbol{\theta}[1]; \boldsymbol{\theta}[2]$	sin
correct model values	0.4 0.6	3 -2	[1.9404, 1.6148, 1.6210] [-1.2936, -1.0765, -1.0806]	
modified exponential PCA	0.4107 0.5893	3.0009 -1.3725	[1.6235, 1.8648, 1.7007] [-0.7425, -0.8529, -0.7778]	0.1368
modified SP-PCA	0.3724 0.6276	3.2170 0.8355	[2.1732, 1.5715, 1.7768] [0.5644, 0.4081, 0.4614]	0.058663
modified Bregman soft clustering	0.4069 0.5931		[1.9317, 1.7162, 1.5585] [-1.1061, -1.0802, -1.0304]	

**Table 4. Clustering results for a Poisson-Gaussian mixed data set.**

	$\pi_1; \pi_2$	$\mathbf{a}[1]; \mathbf{a}[2]$	$\boldsymbol{\theta}[1]; \boldsymbol{\theta}[2]$	sin
correct model values	0.4 0.6	1 -2	[0.8914, 0.1688, 0.4206] [-1.7828, -0.3375, -0.8412]	
modified exponential PCA	0.4475 0.5525	0.8559 -1.9972	[0.7796, 0.1166, 0.3334] [-1.8193, -0.2721, -0.7779]	0.049038
modified SP-PCA	0.3978 0.6022	-0.9548 -3.1821	[-0.9046, -0.0989, -0.2890] [-3.0148, -0.3296, -0.9633]	0.1455
modified Bregman soft clustering	0.3973 0.6027		[0.82252, 0.144, 0.41004] [-1.8072, -0.3089, -0.9816]	

**Table 5. Clustering results for a Binomial-Gaussian mixed data set.**

## 6. Conclusion

We presented a mixed data-type hierarchical Bayes graphical model framework that adds clarity and perspective to our understanding of exponential PCA, SP-PCA and Bregman soft clustering. We demonstrated that these techniques are not separate unrelated algorithms but different manifestations of parameter choices taken within a common framework. Because of this insight, we were able to extend the algorithms to readily derive novel extensions that deal with the important mixed data type case. Our framework has the critical advantage of allowing one to go from high-dimensional mixed-type data components to low-dimensional common-type latent variables that are then used to perform clustering in a much simpler manner using well-known continuous-parameter clustering techniques.

## References

- [1] M. Aitkin, A general maximum likelihood analysis of overdispersion in generalized linear models, *Statistics and Comput.*, vol. 6, pp. 251-262, 1996.
- [2] M. Aitkin, A maximum likelihood analysis of variable components in generalized linear models, *Biometrics*, vol. 55, pp. 117-128, 1999.
- [3] K. S. Azoury and M. K. Warmuth, Relative loss bounds for on-line density estimation with the exponential family of distributions, *Mach. Learning*, vol. 43, pp. 211-246, 2001.
- [4] A. Banerjee, S. Merugu, I. Dhillon and J. Ghosh, Clustering with Bregman divergences, *J. Mach. Learning Research*, vol. 6, pp. 1705-1749, 2005.
- [5] D. Boehning, *Computer-Assisted Analysis of Mixtures and Applications: Meta-Analysis, Disease Mapping, and Others*, Chapman and Hall/CRC, 2000.
- [6] L. Cayton, Fast nearest neighbor retrieval for Bregman divergences, *25th Int. Conf. Mach. Learning*, 2008.
- [7] M. Collins, S. Dasgupta and R. Shapire, A generalization of principal component analysis to the exponential family, *NIPS*, 2001.
- [8] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum likelihood estimation from incomplete data via the EM algorithm, *J. Royal Statistical Soc., B*, vol. 39, pp. 1-38, 1977.
- [9] J. Kiefer and J. Wolfowitz, Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters, *The Annals of Mathematical Statistics*, vol. 27, pp. 887-906, 1956.
- [10] N. Laird, Nonparametric Maximum Likelihood Estimation of a Mixing Distribution, *J. American Statistical Soc.*, vol. 73, 1978.
- [11] E. L. Lehmann and G. Castella, *Theory of Point Estimation*, Springer, 2nd edition, 1998.
- [12] C. Levasseur, U. F. Mayer, B. Burdige and K. Kreutz-Delgado, Generalized statistical methods for unsupervised minority class detection in mixed data sets, *IAPR Workshop on Cognitive Inf.*, 2008.
- [13] B. G. Lindsay, The geometry of mixture likelihoods: a general theory, *Annals of Stat.*, vol. 11, no. 1, pp. 86-94, 1983.
- [14] B. G. Lindsay and M. L. Lesperance, A review of semiparametric mixture models, *J. Statistical Planning and Inference*, vol. 47, pp. 29-99, 1995.
- [15] A. Mallet, A maximum likelihood estimation method for random coefficient regression models, *Biometrika*, vol. 73, no. 3, pp. 645-656, 1986.
- [16] G. McLachlan and D. Peel, *Finite Mixture Models*, Wiley-Interscience, 2000.
- [17] F. Nielsen, J.-. Boissonnat and R. Nock, On Visualizing Bregman Voronoi Diagrams, *Proc. of the 23rd ACM Conf. on Comput. Geometry*, 2007.
- [18] R. S. Pilla and B. Lindsay, Alternative EM methods for nonparametric finite mixture models, *Biometrika*, vol. 88, no. 2, pp. 535-550, 2001.
- [19] R. A. Redner and H. F. Walker, Mixture densities, maximum likelihood and the EM algorithm, *SIAM Review*, vol. 26, no. 2, pp. 195-239, 1984.
- [20] Sajama and A. Orlitsky, Semi-parametric exponential family PCA, *NIPS*, 2004.



## Making Good Choices of Non-Redundant N-gram Words

Maria Fernanda Moura  
Embrapa Informática Agropecuária  
PO Box 6041, Campinas, SP, Brazil  
fernanda@cnptia.embrapa.br

Bruno Magalhães Nogueira  
Merley da Silva Conrado  
Fabiano Fernandes dos Santos  
Solange Oliveira Rezende  
University of São Paulo  
Computer and Mathematics Science Institute  
PO Box 668, São Carlos, SP, Brasil  
{brunomn,merleyc,fabianof,solange}@icmc.usp.br

### Abstract

*A new complete proposal to solve the problem of automatically selecting good and non redundant n-gram words as attributes for textual data is proposed. Generally, the use of n-gram words is required to improve the subjective interpretability of a text mining task, with  $n \geq 2$ . In these cases, the n-gram words are statistically generated and selected, which always implies in redundancy. The proposed method eliminates only the redundancies. This can be observed by the results of classifiers over the original and the non redundant data sets, because, there is not a decrease in the categorization effectiveness. Additionally, the method is useful for any kind of machine learning process applied to a text mining task.*

### 1. Introduction

The amount of information in the digital universe corresponded to 161 exabytes according to 2006 estimates, which was about 3 million times the information in all books already written. In the period from 2006 to 2010, an increase in the digital universe has been estimated from 161 exabytes to 988 exabytes [6]. Facing this scenario, the human capability of analyzing and understanding those data has been exceeded. Therefore, much of the information and knowledge, potentially useful, can be wasted, hidden in data bases spread throughout the world. As a consequence, the need for developing new tools and knowledge extraction techniques from storage data is becoming more and more indispensable [5].

The Text Mining (TM) process, using computational techniques of knowledge extraction, acts in this scenario as a transformer agent, extracting from this enormous quan-

tity of textual data some useful knowledge that can be used as a competitive advantage or as a support to decision making. This process can be seen as a particular case of Data Mining and it is composed by five steps: problem identification, pre-processing, pattern extraction, pos-processing and knowledge use. These steps can be instanced according to the process goals [18].

Frequently, the pre-processing is considered a step of minor importance, or less interesting than the others, due to the lack of technical glamor and the excess of manual tasks. Basically, this step aims at transforming the text collection into a useful form for the learning algorithms, involving tasks as treatment, cleaning and reduction of the data. Additionally, in this step, the most representative features in the collection must be considered, looking for the elimination of the irrelevant features. Nevertheless, the difficulties and challenges imposed by this nontrivial step result in the rich or poor quality of the analyzed data and, consequently, in the efficiency of the used learning algorithms and the confidence of the obtained results. Moreover, it is important to consider the explicit knowledge contained in the attributes in a text mining task, since the final results can be subjectively analyzed [20].

In this way, the probability of considering an attribute as a domain term in a text collection can be taken as a quality and discriminative property. Although many authors consider the dimensionality increase more harmful than beneficial, the attributes representation using single and compound words seems to be more easily interpreted and, consequently, it is more suitable to the users of the results of a text mining task ([14], [15]). When working with single and compounded words, found after a statistical analysis, there are some observed redundancies and a lot of compounded words that can be discarded. Some initiatives to discard redundancies can be found in the work of Zhang and Zhu

[21] and of Dejun and Maosong [3]. According to Dejun and Maosong there is a kind of redundancy arising from feature overlapping in Chinese character n-gram and they have proposed a model to avoid the redundancies that they believed to be language independent, in which the overlapping is estimated and then absorbed by the attribute model; besides, the model is not confined to supervised learning. In the proposal of Zhang and Zhu, a window length between words that form a n-gram is estimated in a model to set a discriminative value to the n-gram words, based on the document frequency of the found n-gram words and in the class label value; consequently, it is better applied to supervised learning, but it is an interesting solution to remove words among the goal word grams.

In this work, a new complete proposal to solve the problem of automatically selecting good and non redundant n-gram words as attributes for textual data is detailed. The model is useful to supervised and unsupervised learning and it depends only on statistical tests over the attribute candidates. In the next section the methodology to select good and non redundant n-gram words is described. In the experiments and results, section 3, we can observe that the reduction of feature space did not sacrifice the categorization results. Finally, some the conclusions and future work are discussed in section 4.

## 2. Proposed Method

The general idea is to generate attributes and apply filters that have more reliable behavior for each kind of n-gram sequentially, using statistical methods only, so the whole process can be completely automated. First of all, there are some terms that has to be defined according to their use in this paper:

- **n-gram word**: the used n-gram word model removes the stopwords from the texts, identify simple words as tokens, apply a stemming process over the words and then combine all of them in a sequential way [13]. The stemming process was based on Porter's algorithm [17] and adapted to the Portuguese and Spanish languages, as well as English. For example, a n-gram word can be "artific", "artific-intelligenc", "intelligenc-conferenc" and "artific-intelligenc-conferenc".
- **stopwords**: always the habitual list of stopwords, as articles and other grammatical classes that are very common in the texts and can be discarded from the attribute set. The set of words discarded from a filter selection is also considered stopwords, and in this paper this set of discarded words is called **collection stopwords**, referring to the specific text collection used in the example processes.

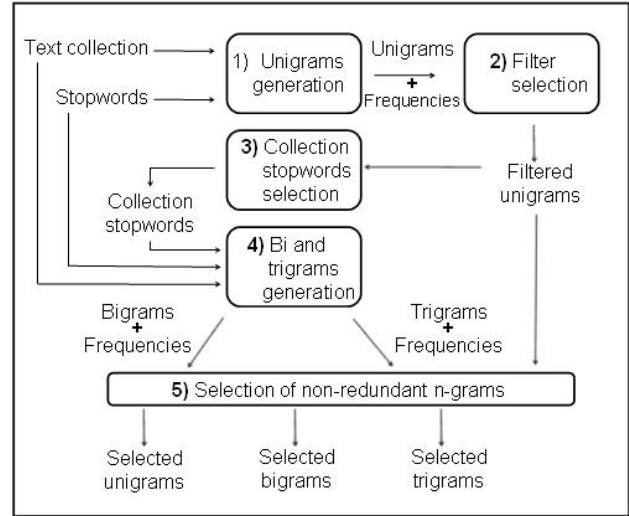


Figure 1. Steps of non-redundant n-gram word selection

- **attribute or term**: the two words refer to a selected n-gram word, that is, a n-gram word that is accepted as a valid attribute or a candidate attribute, because of its statistical discriminative potential.

The method is divided in five steps as illustrated in Fig.1. The text collection is supposed to be from a specific domain, because this method sees the text collection as a bag of n-gram words, so it is not able to treat polysemous words (or n-gram words) in different contexts; consequently this assertion decreases the probability of polisemy occurrence in the collection. For the first step, **unigram generation**, the text collection and the habitual stopwords list are considered to generate only the unigram word set of candidate attributes and their frequencies are counted; that is all possible unigram word forms is in this set, for example, "artific", "intelligenc" and "conferenc".

In the second step, a **filter selection** is applied to the generated unigram word set. The filters are based on the observed frequencies and their variance. The most common used filter attribute is the Luhn's method [11], that finds two cutoff points by ascending and plotting the attributes. So, two cutoff points are chosen close to the tendency curve inflexion points, considering the attributes which have very low or very high frequencies as irrelevant. Despite that, the elimination of low frequency terms is not a common-sense. For example, in the information retrieval area they have been favored, because of the wide use of the tf-idf indexer (tf-idf: term frequency - inverse document frequency) [19]. Some filters based on variance representativeness are also used: the Term Contribution (TC) [10] that can be defined as how much one specific term contributes to the simi-

larity among documents in a document collection; the Term Variance (TV) measure [9] uses the variance of all terms in the collection, attributing the highest scores to those terms that do not have low document frequency and have a non-uniform distribution through the collection; and, the Term Variance Quality (TVQ) [4] that is very similar to Term Variance and used to calculate the quality of a term through the total variance. Therefore, these filters provide an attribute ranking which implies in a cutoff choice. All of these filters are applicable for unsupervised attribute sets, because the most of interesting documents are unstructured and not classified. Besides, filters based on frequencies are more reliable to one gram sets, because they have a tendency on favoring terms that are not so frequent neither so rare. Experimentally it is easy to note that composed terms tend to have very low frequencies compared to simple terms (or unigram words). Moreover, the most interesting composed terms generally are the most frequent ones, although this observation does not guarantee that they can be considered domain terms or collocation candidates[12]. In this step, the filter that presents the best behavior is selected and the unigram word set selected by the filter is considered as a **valid attribute set**.

After obtaining the **filtered unigram word** set, in the third step, the **collection stopwords** are selected. They correspond to those single terms discarded from the text collection, that is, those words that do not help in the construction of the filtered unigram word set. The collection stopwords have to be discarded of the high order n-gram words because the interesting high order n-gram words are only those whose are formed by unigram words that are statistically discriminative. In this way, many of the possible n-gram words for the text collection have already been eliminated. In this work only **bigram and trigram words are being generated** in the fourth step so far.

Finally, there are three sets of attributes, where only the one gram set has already been evaluated by some statistical tests, and there are still a lot of redundant n-gram words. For example, in the text “*Textual data mining is data mining over textual data. Textual data are textual collections in several formats. Data mining has several definitions, but data are always data that are aimed at a special meaning and use.*” there are several possibilities of constructing n-gram words. Not all combinations of unigram words are interesting to discriminate the text or make any sense. In table 1 there are only the trigram, bigram and unigram words, presented in the text, that were subjectively considered to be useful and to have sense. In the **second column** there are the observed occurrence frequencies of each n-gram word in the text. As each low order n-gram word is also presented in a high order n-gram word, in the **third column** the number of occurrences was subtracted from the highest order n-gram word, in this first case from the trigram word; and,

Subjectively chosen n-grams	observed frequency	- 3gram occurrence	- 2gram occurrence
textual data mining	1		
data mining	3	3-1=2	
textual data	3	3-1=2	
textual collections	1	1-0=1	
several formats	1	1-0=1	
several definitions	1	1-0=1	
several meaning	1	1-0=1	
special meaning	1	1-0=1	
data	7	7-1=6	6-2-2-1=1
mining	3	3-1=2	2-2=0
textual	4	4-1=3	3-2-1=0
collections	1	1-0=1	1-1=0
several	2	2-0=2	2-1-1=0
formats	1	1-0=1	1-1=0
definitions	1	1-0=1	1-1=0
aimed	1	1-0=1	1-0=1
special	1	1-0=1	1-1=0
meaning	1	1-0=1	1-1=0
use	1	1-0=1	1-0=1

**Table 1. Redundant selected n-gram words**

finally, in the **fourth column** the last number of frequencies for each n-gram word is subtracted from the last obtained frequencies for each bigram word in which it also occurs. As a result of this process, the n-gram words that have null frequency in the last column are discarded from the original set of attributes candidate, because their linear combinations of frequencies have resulted in higher order n-gram words. The n-gram words that have a non zero frequency are the non redundant. The systematization of these ideas occurs in the **selection of non redundant n-gram words**, fifth step illustrated in Fig. 1, and are described as follow.

First of all, the n-gram words, with  $n = 3, 2, 1$ , must be statistically selected as candidates for domain terms, instead of the subjectively selection. This can be carried out by using a set of tests for dependence among the unigram words of the generated n-gram in the whole text collection ([12], [1]). To simplify this process, the Yule’s Q test, a robust test used in decomposed contingency tables (for details see [2]) was chosen and incorporated to the method presented here. To reach all test results, a set of frequency measures has to be taken from each n-gram word.

Let  $w_1w_2w_3$  be a trigram word and its observed frequency be  $f_{123}$ , it is also needed to obtain the observed frequency of the bigram word  $w_1w_2$  in the first and second positions of all trigram words in the collection as  $f_{12}$ , the observed frequency of  $w_2w_3$  in the second and third positions as  $f_{23}$ ; and the frequencies of  $w_1$  in the first position ( $f_{1..}$ ) and  $w_3$  in the third position ( $f_{...3}$ ), of all obtained trigram words. The frequency of all trigram words in the collection ( $f_{...}$ ) has also to be obtained.

In table 2, the frequency values for  $w_1$  in the first position ( $f_{1..}$ ) and  $w_2w_3$  in the second position ( $f_{.23}$ ) of the

trigram word  $w_1w_2w_3$  ( $f_{123}$ ) are illustrated, and, the frequency values of the n-gram words different from them in all the collection.

	$w_2w_3$	$!(w_2w_3)$	
$w_1$	$f_{123}$	$f_{1..} - f_{123}$	$f_{1..}$
$!(w_1)$	$f_{.23} - f_{123}$	$(f_{...} - f_{1..}) - (f_{.23} - f_{123})$	$f_{...} - f_{1..}$
	$f_{.23}$	$f_{...} - f_{.23}$	$f_{...}$

**Table 2.**  $w_1$  in the first and  $w_2w_3$  in the second and third positions of  $w_1w_2w_3$

To test the hypothesis of the independence between  $w_1$  in the first position and  $w_2w_3$  in the second and third positions, using the Yule's Q statistic, the following values must be obtained (for details see [2]):

$$d_1 = ((f_{...} - f_{1..}) - (f_{.23} - f_{123}))$$

$$\alpha_1 = \frac{f_{123} * d_1}{(f_{.23} - f_{123}) * (f_{1..} - f_{123})}$$

$$Q_1 = (\alpha_1 - 1) / (\alpha_1 + 1)$$

$$v_1 = \sqrt{\frac{1}{f_{123}} + \frac{1}{f_{1..} - f_{123}} + \frac{1}{f_{.23} - f_{123}} + \frac{1}{d_1}}$$

$$\sigma_{Q_1} = \frac{1}{2} * (1 - Q_1^2) * v_1$$

$$IC_{Q_1} = [Q_1 - 2 * \sigma_{Q_1}, Q_1 + 2 * \sigma_{Q_1}]$$

In table 3, the frequency values for  $w_1w_2$  in the first and second positions ( $f_{12.}$ ) and  $w_3$  in the third position ( $f_{..3}$ ) of the trigram word  $w_1w_2w_3$  ( $f_{123}$ ) are illustrated, and, the frequency values of the n-gram words different from them in all the collection.

	$w_3$	$!(w_3)$	
$w_1w_2$	$f_{123}$	$f_{12.} - f_{123}$	$f_{12.}$
$!(w_1w_2)$	$f_{..3} - f_{123}$	$(f_{...} - f_{12.}) - (f_{..3} - f_{123})$	$f_{...} - f_{12.}$
	$f_{..3}$	$f_{...} - f_{..3}$	$f_{...}$

**Table 3.**  $w_1w_2$  in the first and second and  $w_3$  in the third positions of  $w_1w_2w_3$

So, to test the hypothesis of the independence between  $w_3$  in the third position and  $w_1w_2$  in the first and second positions, using the Yule's Q statistic, the following values must be obtained (for details see [2]):

$$d_2 = ((f_{...} - f_{..3}) - (f_{12.} - f_{123}))$$

$$\alpha_2 = \frac{f_{123} * d_2}{(f_{12.} - f_{123}) * (f_{..3} - f_{123})}$$

$$Q_2 = (\alpha_2 - 1) / (\alpha_2 + 1)$$

$$v_2 = \sqrt{\frac{1}{f_{123}} + \frac{1}{f_{..3} - f_{123}} + \frac{1}{f_{12.} - f_{123}} + \frac{1}{d_2}}$$

$$\sigma_{Q_2} = \frac{1}{2} * (1 - Q_2^2) * v_2$$

$$IC_{Q_2} = [Q_2 - 2 * \sigma_{Q_2}, Q_2 + 2 * \sigma_{Q_2}]$$

If  $0 \notin IC_{Q_1}$  and  $0 \notin IC_{Q_2}$  then  $w_1w_2w_3$  is a candidate to be a domain term collocation and is also selected as a trigram word attribute candidate. All selected trigram words, as attribute candidates, are ranked by their occurrence frequency, that is, by their  $f_{123}$ . As soon as all trigram words were tested, selected and ranked, all bigram and unigram words which frequencies form the trigrams by linear combinations are eliminated with the Algorithm 1, according to their observed frequencies:  $O_{f_{ij}}$ ,  $i, j = 1, 2, 3$  for bigram words; and  $O_{f_i}$ ,  $i = 1, 2, 3$  for unigram words.

It has to be highlighted that, only the best ranked candidate trigram words can be selected as final attributes. Additionally, for each of these trigram words there are some bigram or unigram words with frequency greater or equal to their frequencies. In this way, the redundancy is a selection criteria and it is used to eliminate itself.

**Input:** *TriGramList*: list of trigram words ordered by rank value

**Output:** *FinalAttributesList*: list of final attributes with updated frequencies and selected trigram words

**for all trigram  $w_a w_b w_c$  in *TriGramList* do**  
  select next best ranked trigram word,  $w_1w_2w_3$ ;  
  search for  $w_1w_2$ , with  $O_{f_{12}} \geq O_{f_{123}}$ ;  
  search for  $w_2w_3$ , with  $O_{f_{23}} \geq O_{f_{123}}$ ;  
  search for  $w_1$ , with  $O_{f_1} \geq O_{f_{123}}$ ;  
  search for  $w_2$ , with  $O_{f_2} \geq O_{f_{123}}$ ;  
  search for  $w_3$ , with  $O_{f_3} \geq O_{f_{123}}$ ;  
  subtracted  $O_{f_{123}}$  from all above n-gram  $O_{f_{ij}}$  and  $O_{f_i}$ ;  
  add  $w_1w_2w_3$  in *FinalAttributesList*;  
**end**

**Algorithm 1:** Trigram words selection algorithm

After the trigram word selection, the same process and ideas are applied to the bigram words that were not eliminated yet. In this way, the bigram word candidates are selected using the Yule's Q test and they are ranked by their occurrence frequencies.

Let  $w_1w_2$  be a bigram word and its observed frequency be  $f_{11}$ , it is also needed to obtain the observed frequency of the unigram word  $w_1$  in the first position in all bigram words in the text collection as  $f_{1.}$  and the observed frequency of  $w_2$  in the second position as  $f_{.1}$ . The frequency of all bigram words in the collection,  $f_{..}$ , has also to be obtained; and all these values are illustrated in table 4.

To test the hypothesis of the independence between  $w_1$  in the first position and  $w_2$  in the second position, using

	$w_2$	$!(w_2)$	
$w_1$	$f_{11}$	$f_{1.} - f_{11}$	$f_{1.}$
$!(w_1)$	$f_{.1} - f_{11}$	$(f_{..} - f_{1.}) - (f_{.1} - f_{11})$	$f_{..} - f_{1.}$
	$f_{.1}$	$f_{..} - f_{.3}$	$f_{..}$

**Table 4.**  $w_1$  in the first and  $w_2$  in the second positions of  $w_1w_2$

the Yule's Q statistic, the following values must be obtained (for details see [2]):

$$\begin{aligned}
d &= ((f_{..} - f_{1.}) - (f_{.1} - f_{11})) \\
\alpha &= \frac{f_{11} * d}{(f_{1.} - f_{11}) * (f_{.1} - f_{11})} \\
Q &= (\alpha - 1) / (\alpha + 1) \\
v &= \sqrt{\frac{1}{f_{11}} + \frac{1}{f_{1.} - f_{11}} + \frac{1}{f_{.1} - f_{11}} + \frac{1}{d}} \\
\sigma_Q &= \frac{1}{2} * (1 - Q^2) * v \\
IC_Q &= [Q - 2 * \sigma_Q, Q + 2 * \sigma_Q]
\end{aligned}$$

So, if  $0 \notin IC_Q$  then  $w_1w_2$  is candidate of a domain term collocation and the bigram word is also selected as an attribute candidate. All selected bigram words, or attribute candidates, are ranked by their occurrence frequency,  $f_{11}$ . As soon as all the remaining bigram words were tested, selected and ranked, all remaining unigram word which frequencies form the bigrams by linear combinations are eliminated, according to their observed frequencies:  $O_{f_i}$ ,  $i = 1, 2$ ; and following the Algorithm 2.

**Input:** *BiGramList*: list of the remaining bigram words ordered by rank value

**Output:** *FinalAttributesList*: list of final attributes with updated frequencies and selected bigram words

```

for all bigram $w_a w_b$ in BiGramList do
 select next best ranked bigram, $w_1 w_2$ with
 remaining frequency $O_{f_{11}}$;
 search for w_1 , with $O_{f_1} \geq O_{f_{11}}$;
 search for w_2 , with $O_{f_2} \geq O_{f_{11}}$;
 subtracted $O_{f_{11}}$ from all above unigram O_{f_1} and
 O_{f_2} ;
 add $w_1 w_2$ in FinalAttributesList;
end

```

**Algorithm 2:** Bigram words selection algorithm

It has to be highlighted that, only the best ranked bigram word candidates, for which there are some unigram words with frequency greater or equal their frequency, can be selected as final attributes. Note that, the redundancy is a selection criteria and it is used to eliminate itself.

Again, it has to be highlighted that, only the best ranked candidate bigram words can be selected as final attributes. Additionally, for each of these bigrams words there are some unigram words with frequency greater or equal to their frequencies. In this way, the redundancy is a selection criteria and it is used to eliminate itself.

Moreover, it is easy to see that the two algorithms are almost the same indicating that the process can be expanded to any number n of n-gram words, if desired.

Besides, the evaluation of the method is available through an implementation of some prototypes (developed in C) to test, rank and select the candidate attribute sets and the final attribute set. In this way it is possible to obtain the candidate ranked attribute lists and the final selected attribute sets, in order to compare them. Although the best comparison should be a subjective evaluation carried out with information specialists or domain specialists, because of the data and the attribute nature, it is not an easy task over big text collections. Even a small text collection frequently results in a huge attribute set for human judgment.

To perform a non subjective evaluation of the pre-processing results and obtain a validation of the generated attribute set, some labeled document collections are used within a supervised learning process and evaluation. In this way, all the described steps are applied to a labeled text collection, but considering it as non labeled. Two attribute sets are constructed to each text collection:

1. **ranked attributes:** these are the attributes selected only by ranking the generated bigram and trigram words, using first the independence test explained and then the filter to the generated unigram words.
2. **non redundant selected attributes:** applying all the steps explained before to all generated attribute candidates.

The independence test applied to rank and select the n-gram words is the same in the two attribute sets, because the goal is to compare the attribute sets only at the presence or absence of redundancy, avoiding any other possible effect.

Then an attribute-value matrix is constructed for each attribute set, where, each row vector represents a document and each column an attribute; the cells correspond to the occurrence frequency of the attribute in the document; and finally the last column corresponds to the codified label. So, to validate the results, for each obtained attribute set, two classification models are constructed using two largely known classification algorithms: C4.5 decision trees and Support Vector Machines (SVM). Both of them were chosen due to the fact that they can face well sparse attribute matrices, besides SVM is able to generalize even in the presence of many redundant features [8]. Additionally, to estimate the classifiers accuracy rate, the 10-fold cross validation process is used. Finally, the accuracy of the ob-

tained classifiers at each attribute set are compared using the t-student test.

### 3. Some Experimental Results

Three data sets from different domains and languages were selected. The first is a collection of complete articles from the Instituto Fábrica do Milênio (IFM) [7]; that is a Brazilian organization whose actions are focused on the search for manufacturing solutions for the needs of industry. This data set is composed of 614 documents in the Portuguese language, divided into 4 classes, with 291 documents in the majority class. The second document collection is the Case Based Reasoning- Inductive Logic Programming - Information Retrieval - Sonification (CBR-ILP-IR-SON) data set<sup>1</sup> composed of 681 document abstracts in the English language, classified according to 4 classes with 276 documents in the majority class; it is a scientific production collection; and, in this paper that base is referred as **CIIS**. The third text collection is a set of complete articles, in English, divided in five sub topics of Artificial Intelligence: Agents & Multiagents, Fuzzy Logic, Machine Learning, Planning & Scheduling e Robotics<sup>2</sup>; in this paper that data base is referred as **IA**.

A random selection process was used to choose balanced samples for the three text collections, obtaining approximately fifty documents per class in each collection. In this way, effects due to the unbalanced classes are eliminated. In table 5 there are the number of documents, number of generated unigram words for each collection, the number of filtered unigram words for each collection. The chosen filter was the term contribution, TC [10], taking only 10% of the ranked unigram word attributes due to some previous experiments. In these experiments it was observed that filters based on the term contribution to the explained variance select the best attribute sets with smaller sets than the other techniques [10] and some other experiments suggested that only 10% of the ranked attributes produces results as reliable as the bigger sets [16]. In the same table, the last two columns refer to the number of generated bigram and trigram words after the unigram word filtering and the collection stopwords selection.

Taking the attribute candidates corresponding to the three last columns of table 5, the fifth step of the selection of non redundant n-gram word attributes is applied. The implemented prototypes were adapted to produce the non redundant attribute sets and the only ranked attribute sets. In this way, two attribute sets were generated for each text collection; the number of non redundant (NR) selected attributes, generated in the fifth process step, is in Table 6 and

Text Collection	#docs	#gen. 1gram	#filt. 1gram	#gen. 2gram	#gen. 3gram
IA	250	42753	4276	87297	30326
CIIS	200	10810	1081	17940	17944
IFM	193	20795	2080	82325	84200

**Table 5. Results from first,second, third and fourth process steps**

the number of ranked (R) attributes, that is with the redundancies, are in Table 7.

Text collection	#NR 1gram	#NR 2gram	#NR 3gram
IA	4264	26270	24303
CIIS	1065	1379	1831
IFM	1984	7043	42741

**Table 6. Non redundant n-gram words**

Text collection	#R 1gram	#R 2gram	#R 3gram
IA	4276	53595	30060
CIIS	1081	4284	2811
IFM	2080	48775	82354

**Table 7. Ranked n-gram words**

To each of the attribute subsets of each text collection, the two classifiers, C4.5 and SVM, were applied and their accuracy rate was estimated using 10-fold cross validation. It is important to highlight that the classifiers are not being compared; the focus is on comparing how the different attribute sets react under a same evaluation process.

Text collection	Ranked $acc \pm se$	NonRed $acc \pm se$	t-test $p - value$
IA	$90 \pm 5.42$	$89.6 \pm 7.11$	0.3838
CIIS	$87 \pm 9.19$	$87 \pm 9.19$	0.3880
IFM	$70.45 \pm 8.28$	$75.63 \pm 7.88$	0.1385

**Table 8. Classifier C4.5 -  $H_0 : \overline{acc_r} = \overline{acc_{nr}}$**

Text collection	Ranked $acc \pm se$	NonRed $acc \pm se$	t-test $p - value$
IA	$79.6 \pm 8,1$	$82 \pm 7.12$	0.2968
CIIS	$90 \pm 8.82$	$90.5 \pm 6.85$	0.3837
IFM	$59.97 \pm 10.24$	$68.84 \pm 11.74$	0.0835

**Table 9. Classifier SVM -  $H_0 : \overline{acc_r} = \overline{acc_{nr}}$**

In tables 8 and table 9, it is possible to see the evaluation results along with the estimated accuracies and their standard deviation for each classifier and for each attribute set. The attribute sets are ranked (Ranked) and non redundant (NonRed). At the last columns there are the results for the

<sup>1</sup><http://infoserver.lcad.icmc.usp.br/infosvis2/PEX>

<sup>2</sup><http://labic.icmc.usp.br/projects>

t-student test for the hypothesis of the equality of the accuracy means ( $H_0 : \overline{acc_r} = \overline{acc_{nr}}$ ), where:  $\overline{acc_r}$  and  $(se)_r$ , the accuracy mean and the standard deviation for the ranked attribute set, and  $\overline{acc_{nr}}$  and  $(se)_{nr}$  the accuracy mean and standard deviation for non redundant attribute set. It was calculated using equation 1 to obtain a  $t_{value}$  and then calculating the cumulative probability of this value, under a t-student distribution. If the cumulative probability value is the interval of  $[0.025, 0.975]$ , the hypothesis of the statistical equality of the accuracies is accepted. This interval corresponds to the values of the t-student table for nine degrees of freedom, because the accuracies have been calculated from a 10-fold cross experiment.

$$t_{value} = \frac{\overline{acc_r} - \overline{acc_{nr}}}{\sqrt{\frac{(se)_r^2}{10} + \frac{(se)_{nr}^2}{10}}} \quad (1)$$

At a first glance, focusing only on the performance, for each classifier we have the feeling that the feature redundancy removal was better, especially because of the range of the standard deviations. But after applying the t-test, the equality hypothesis was accepted for all data sets and classifiers. In this way, we can observe that the redundancy removal did not sacrifice the categorization effectiveness. Besides, as the accuracy estimates were good, we believe the attribute selection was also good.

## 4. Conclusions

A new complete proposal to solve the problem of automatically selecting good and non redundant n-gram words as attributes for textual data is proposed. The method is based on the assumption that attribute redundancy is always present when using n-gram words, for  $n \geq 2$ , due to the stochastic nature of them.

The proposed method was shown to reach the main goals: selecting good n-gram words and eliminating only the redundancies. Firstly, the accuracy estimates of the classifiers were good, what is an indication of a good attribute selection. Moreover, the categorization effectiveness has been not decreased after the redundancy removing, which indicates that only redundant attributes had been eliminated.

Since the method does not depend on class labels, it can be used in unsupervised or supervised learning. We believe that for supervised learning it is better to change the filtering process used in the second step, because an appropriated filter could produce a better unigram word selection.

The proposed method is completely dependent only on statistical tests over a bag of n-gram words instead of contiguous character n-grams, as on the method of Dejun and Maosong. Besides, Dejun and Maosong believed their method would be language independent; and the proposed method is language independent after the stemming process, as we could observe in the experiments with text

collections in two different languages (Portuguese and English).

Moreover, this method uses the collection stopwords, as defined here, to remove undesirable words among the goal n-gram words; in this way, there is no need of estimating a window length between the words, as Zhang and Zhu had proposed. The proposed method can be classified closer to the method of Zhang and Zhu, because we believe that the removal of the collection stopwords also causes a semantic information improvement similar to the window length estimation.

As future work, we intend to realize subjective analysis over the generated set of attributes, along with domain specialists, to evaluate the semantic improvement. Besides, it is necessary to expand the algorithms, and the implemented prototype, to deal with any number of n in the n-gram word model.

## References

- [1] S. Banerjee and T. Pedersen. The design, implementation, and use of the ngram statistics package. In *CICLing*, pages 370–381, 2003.
- [2] Y. Bishop, S. E. Fienberg, and P. H. Holland. *Discrete Multivariate Analysis*. MIT Press, 1984.
- [3] X. Dejun and S. Maosong. Raising high-degree overlapped character bigrams into trigrams for dimensionality reduction in chinese text categorization. In A. F. Gelbukh, editor, *CICLing*, volume 2945 of *Lecture Notes in Computer Science*, pages 584–595. Springer, 2004.
- [4] I. Dhillon, J. Kogan, and C. Nicholas. Feature selection and document clustering. In M. W. Berry, editor, *Survey of Text Mining*, pages 73–100. Springer, 2003.
- [5] N. F. F. Ebecken, M. C. S. Lopes, and M. C. d. Aragão. Mineração de textos. In S. O. Rezende, editor, *Sistemas Inteligentes: Fundamentos e Aplicações*, chapter 13, pages 337–364. Manole, 1 edition, 2003.
- [6] J. F. Gantz, D. Reinsel, C. Chute, W. Schlichting, J. McArthur, S. Minton, I. Xheneti, A. Toncheva, and A. Manfrediz. The expanding digital universe: A forecast of worldwide information growth through 2010. *External Publication of IDC (Analyse the Future) Information and Data*, pages 1–21, March 2007.
- [7] IFM. IFM - Instituto Fábrica do Milênio, 2007. Disponível em: <http://www.ifm.org.br/> [29/08/2008].
- [8] T. Joachims. Text categorization with support vector machines: Learning with many relevant features. In C. Nedellec and C. Rouveirol, editors, *ECML*, volume 1398 of *Lecture Notes in Computer Science*, pages 137–142. Springer, 1998.
- [9] L. Liu, J. Kang, J. Yu, and Z. Wang. A comparative study on unsupervised feature selection methods for text clustering. *Natural Language Processing and Knowledge Engineering, 2005. IEEE NLP-KE '05. Proceedings of 2005 IEEE International Conference on*, pages 597–601, 30 Oct.-1 Nov. 2005.

- [10] T. Liu, S. Liu, Z. Chen, and W.-Y. Ma. An evaluation on feature selection for text clustering. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML)*, pages 488–495. AAAI Press, 2003.
- [11] H. P. Luhn. The automatic creation of literature abstracts. *IBM Journal os Research and Development*, 2(2):159–165, 1958.
- [12] C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.
- [13] E. T. Matsubara, C. A. Martins, and M. C. Monard. Pre-text: uma ferramenta para pré-processamento de textos utilizando a abordagem bag-of-words. Technical Report 209, Instituto de Ciências Matemáticas e de Computação – USP – São Carlos, 2003.
- [14] Y. Miao, V. Keselj, and E. E. Milios. Document clustering using character n-grams: a comparative evaluation with term-based and word-based clustering. In O. Herzog, H.-J. Schek, N. Fuhr, A. Chowdhury, and W. Teiken, editors, *CIKM*, pages 357–358. ACM, 2005.
- [15] M. F. Moura and S. O. Rezende. Choosing a hierarchical cluster labelling method for a specific domain document collection. In *EPIA- Encontro Portugues de Inteligência Artificial, 2007, Guimarães, Portugal. New Trends in Artificial Intelligence. Lisboa, Portugal: APPIA - Associação Portuguesa para Inteligência Artificial*, pages 812–823, 2007.
- [16] B. M. Nogueira, M. F. Moura, M. S. Conrado, and S. O. Rezende. Avaliação de métodos não-supervisionados de seleção de atributos para mineração de textos. In *Proceedings of the First Workshop on Web and Text Intelligence (WTI) - Nineteenth Brazilian Symposium on Artificial Intelligence (SBIA)*, page To appear, 2008.
- [17] M. Porter. An algorithm for suffixing stripping. *Program*, 14(3):130–137, July 1980.
- [18] S. O. Rezende, J. B. Pugliesi, E. A. Melanda, and M. F. Paula. Mineração de dados. In S. O. Rezende, editor, *Sistemas Inteligentes: Fundamentos e Aplicações*, chapter 12, pages 307–335. Manole, 1 edition, 2003.
- [19] G. Salton and C. S. Yang. On the specification of term values in automatic indexing. 1973.
- [20] S. M. Weiss, N. Indurkha, T. Zhang, and F. J. Damerau. *Text Mining - Predictive Methods for Analyzing Unstructured Information*. Springer Science+Business Media, Inc., March 2005. ISBN 0-387-95433-3.
- [21] X. Zhang and X. Zhu. A new type of feature - loose n-gram feature in text categorization. In J. Martí, J.-M. Benedí, A. M. Mendonça, and J. Serrat, editors, *IbPRIA (I)*, volume 4477 of *Lecture Notes in Computer Science*, pages 378–385. Springer, 2007.



# gSVMT: Aggregating SVMs over a Dynamic Grid Learned from Data

Shaoning Pang, Tao Ban, Youki Kadobayashi, and Nik Kasabov

**Abstract**—Addressing the problem of adaptively modelling a classifier as a modular system, a new type of SVM aggregating method termed gridding SVM Tree (gSVMT) is proposed in this paper. The proposed gSVMT achieves to discover data subregions with principal discriminant knowledge through a recursive SVM-supervised data partitioning procedure. For each subregion, an individual SVM is allocated to extract the subregion knowledge. A set of such SVMs are aggregated in a specific order, resulting in a globally reliable decision rule to predict new coming samples. Experiments on a synthetic Gaussian data set and 13 benchmark machine learning data sets, have highlighted the usability of the gSVMT on its competitive classification capability. In particular, the proposed gSVMT is found to have better generalization performance than SVM classifiers for data sets with high sparseness and/or class-imbalance. Its performance has been further demonstrated with the successful real application on a face membership authentication system.

## I. INTRODUCTION

When faced with a new classification problem, one often considers, for the first instance, to employ one of the available discriminative classifiers such as the well known  $k$ -nearest neighbor classifier, feed-forward neural networks, or support vector machines (SVMs) to solve the problem swiftly in a single-stage mode. But more often, when complex applications are confronted, just as in most real-world applications, the problem could come into a scale too large and a nature too complicated for a single-stage model to be solved efficiently in terms of the convergence time in the training phase and the response time in the testing phase. In such a case, one may fall back on data preprocessing techniques, e.g. feature extraction and feature selection, to reducing the difficulty of the problem, or alternatively turn to multi-stage approaches. The idea to observe, measure, and break the learning problem into a collection of simpler subproblems and then summarize solutions of the subproblems to yield a global decision, is generally known as the ‘divide-and-conquer’ methodology. In this paper, to make difference between a single-step classifier, we call a classification method which makes use of the divide-and-conquer strategy as a multi-stage algorithm. Representative multi-stage algorithms include decision tree classifiers [19] and modular classifiers [20], e.g. SVM ensemble [10], and SVM classification tree (SVMT) [15].

Shaoning Pang and Nikola Kasabov are with the Knowledge Engineering & Discovery Research Institute, Auckland University of Technology, Private Bag 92006, Auckland 1020, New Zealand (email: {spang, nkasabov}@aut.ac.nz). Tao Ban and Youki Kadobayashi are with Information Security Research Center, National Institute of Information and Communications Technology, Tokyo, 184-8795 Japan (email: bantao@nict.go.jp, youki-k@is.aist-nara.ac.jp)

Profiting from the divide-and-conquer strategy, the advantages of a multi-stage classifier over a single-stage classifier include quick convergence, parallel training, and better generalization. In the literature, multi-stage classifiers have been successfully applied in a diversity of areas including character recognition, remote sensing, medical diagnosis, expert systems, and speech recognition [4], [6]. However, given a new data set, there is no existing guideline on how to choose an appropriate multi-stage classifier that can well suit the specific problem with a preferable performance guaranteed. Thus, one often confronts the embarrassment to implement and evaluate multiple methods on the data set. This procedure is usually frustrating because it is not only time and labor demanding, but sometimes could be resultless as well, considering the case of a dataset with a large scale that rendering some of the methods noneffective.

In this paper, we focus our discussion on the usability and construction of a novel multi-stage classifier. The contribution of the paper is twofold. First, we introduce two measurements characterizing the distribution of the data set, namely the sparseness and the class-imbalance. They help to identify whether the dataset can be fitted well by a multi-stage approach so that a generalization performance better than a state-of-the-art single-stage classifier can be expected. Hence, via a preceding exploratory analysis step, we can learn whether experiments on multi-stage classifiers are necessary. Second, we present an SVM based modular classifier termed gridding SVM Tree (gSVMT). The name of gSVMT comes from the fact that the component SVM classifiers are aggregated over a grid learned from data. gSVMT implements a ‘divide-and-conquer’ strategy by first performing data partitioning at the sparse region of the data distribution, and then conducting decision-making through aggregation of the collection of component SVM classifiers. Since each component SVM classifier not only explores the discriminant knowledge but imposes a balancing requirement in the subregion as well, gSVMT demonstrates a strong capability on sparse and/or class-imbalanced data sets.

The organization of the paper is as follows. Section II gives a brief review on related researches on multi-stage classifiers. Section III presents our pattern observation measurements: sparseness and class-imbalance. Section IV describes the detail of the proposed gSVMT methodology and its training and testing algorithms. In Section V, experimental results on 13 UCI data sets and an application of the proposed method to a face membership authentication problem are reported. Section VI concludes the paper.

## II. RELATED RESEARCHES

In this section, we first review some properties of the multi-stage approaches for classification and then discuss a certain group of modular classifiers which has enlightened our idea.

### A. Problem of Classifiers

A multi-stage classifier is generally defined as a learning model where the computation performed by the classifier is decomposed into more than one module that operates on inputs with/without communicating with each other [20], [6], [9]. Different from a single-stage classifier where input data are directly fed to the learning model, a multi-stage classifier normally performs a first step of data observation and analysis. Further steps of data partitioning and decision making are conducted only when a certain criterion is satisfied. This criterion reflects the design philosophy of the employed multi-stage approach.

Despite the mentioned successful applications in various areas, multi-stage classifiers also have confronted some failures: Huang and Ling [7] reported that decision trees performs obviously no better than single-stage method Naive Bayes and SVM on nonuniform data classification in terms of the Receiver Operating Characteristics (ROC). Drucker et. al [3] had reported that an empowered decision tree, called boosting decision tree, is not as successful as a single SVM for email spam categorization. Also, Andrew Senior [21] had found that a standard decision tree can not achieve a satisfied performance on fingerprint classification, unless the decision tree is combined properly with a hidden Markov model (HMM). This raises a problem that how a multi-stage system can be better constructed over different case studies from real world applications.

Given the learning task presented in a set of objects associated with their labels, there are three possible sources of generalization errors for a multi-stage method: 1° selection of learning models based on domain knowledge or a preceding analysis of the input data; 2° the decomposition of the problem and the aggregation of the solutions of the subproblems; 3° decision rules obtained by individual component classifiers. For errors from 2°, most multi-stage classifiers impose cost functions to guide the decomposition procedure and use statistical analysis to reduce the aggregation error as much as possible. For errors from 3°, one may employ a powerful classifier, such as an SVM or a Neural Network, as the component classifier to build the local learners. However, errors from 1° have been overlooked by researchers in most previous works. In fact, analysis of input data and selection of learning models are the basis of all subsequent data processing, especially for multi-stage classifiers which are sensitive to data distribution.

To address the errors from 1°, this paper focuses on two data characteristics that are common to see in real applications, namely the sparseness and class-imbalance. The sparseness of a data set indicates the capability of the data to be decomposed, without losing too much discriminant information presented in the data. The class-imbalance measures the property of a data

set that the samples are not evenly distributed into the two classes.

In practice, the class-imbalance problem has led to degenerated classification performance for many single-stage approaches including SVM. This is because that most classifiers have taken the assumption of a balanced prior probability, which is not true for real applications. This also gives us a very challenging difficulty to model a robust multi-stage learner in the case of imbalanced class distribution. To solve this problem, during module aggregating, we take a straightforward method to allocate more individual modules on the skewed class.

### B. Related Works

Our work can be viewed as a new attempt under the framework of SVM aggregating [14]. As an SVM based modular approach, SVM aggregating aims to create SVM-based multi-core cooperative computing techniques towards optimized problem-solving for computational intelligence. SVM ensemble learning [10] is a type of SVM aggregation in the principle of deriving a mixture of SVM experts to improve the generalization ability of single SVM decision making. SVM ensemble learning assumes that the number of SVMs in aggregation should be known in advance as prior knowledge. In fact, this number is often difficult to determine in real applications. The SVM classification tree (SVMT) was first proposed in [15], in the context of face membership authentication application [16]. SVMT overcomes the difficulty of SVM ensemble learning by determining the number of SVMs automatically during the learning.

## III. CHARACTERIZING THE DATA

In this section, we introduce two measurements to characterize the distribution of a data set. The data are assumed to be presented in a vectorial form.

### A. Sparseness

We often conduct a divide-and-conquer on a data set without caring about whether the data is dividable, i.e. be divided without losing information for classification. What if an improper data partition is performed? Due to this concern, it is interesting to see if we can define a measurement on the feasibility of a data set to be partitioned.

The sparseness is defined to describe the uniformity of a data distribution. From the view point of divide-and-conquer, it also identifies the position where the data could be partitioned, without much loss of discriminant information. Given a data set  $X = \{\mathbf{x}_i\}, i = 1, \dots, N$ , with the attribute set  $\{a_1, \dots, a_M\}$ , we analyze the data distribution of  $X$  along each attribute  $a_i$ . By partitioning the range  $[\max(a_i), \min(a_i)]$  equally into  $K$  unit intervals, a histogram  $s_l (l = 1, \dots, K)$  is calculated by counting the number of points dropping into each interval. Further, the sparseness can be calculated as the ratio of the geometric mean to the arithmetic mean of the list

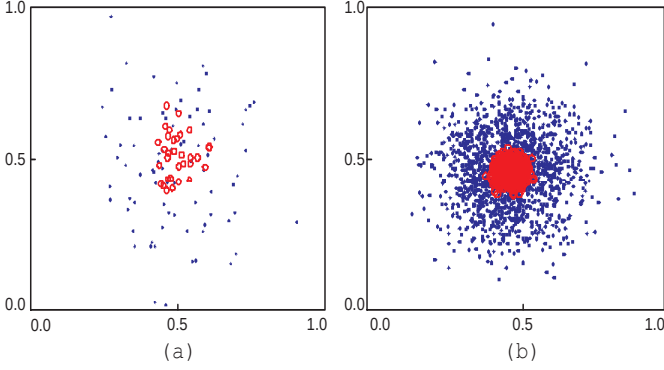


Fig. 1. The illustration of data sparseness definition. Sparsenesses of the two data sets are: (a)  $\mathcal{S}_a = 0.2045$ , and (b)  $\mathcal{S}_b = 0.8923$ .

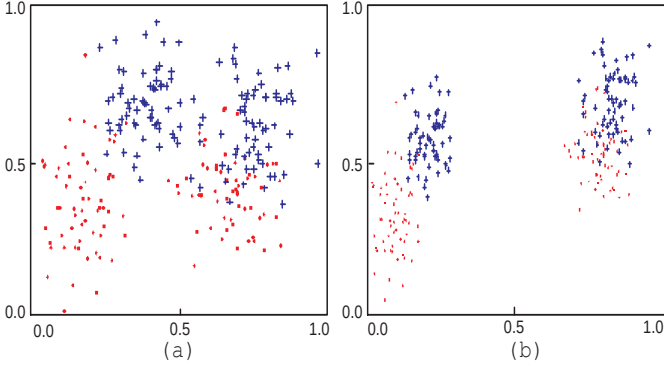


Fig. 2. Comparison of sparseness over two data sets for binary classification. (a)  $\mathcal{S}_a = 0.6545$ , and (b)  $\mathcal{S}_b = 0.6923$ .

associated with the histogram:

$$\mathcal{S}(X) = 1 - \frac{1}{\log 2} \left( \prod_{i=1}^M \frac{1}{K} \sum_{l=1}^K \log(|s_{il}| + 1) \right)^{\frac{1}{M}}. \quad (1)$$

Note that the above sparseness measures an averaged uniformity of the global data distribution, but it does not recognize any sparseness caused by data clusters. Fig. 2 gives a comparison on the sparseness of two synthetic data sets with multiple gaussian distribution. Although Fig. 2 (a) and (b) have a comparable sparseness value, obviously the capability of the two data sets for data partitioning is different. With respect to labelled data sets for classification, it is suggested that the proposed sparseness be applied to each of the classes. The obtained class-sparseness reflects the suitability of this data set for the divide-and-conquer based modelling.

### B. Class Imbalance

Class-imbalance is another problem which leads to degeneration of the generalization performance for single-stage classifiers including SVM. When class-imbalance occurs in a data set, the number of samples from one class is much larger than those of the other classes. In the case of binary classification, the class with the smaller number of samples

is called ‘skewed’ class and the other class is called ‘over-weighted’ class). In this paper, we measure the class-imbalance of a data set by computing the following class-imbalance ratio (CIR),

$$\mathcal{C}(\omega_1, \omega_2) = 1 - \frac{\min(|\omega_1|, |\omega_2|)}{\max(|\omega_1|, |\omega_2|)}. \quad (2)$$

where  $|\omega_i|$  is the cardinal number of class  $\omega_i$ . Obviously, a larger  $\mathcal{C}$  value shows a more imbalanced class distribution.

Single-stage classifiers including SVM often fail to classify the skewed class correctly [24], [23], [12], [22], since most classification algorithms assume a balanced prior probability [18]. To solve the problem caused by class-imbalance, data sampling method is a straightforward solution. Popular approaches include undersampling, which reduces the number of samples in the over-weighted class closer to the smaller class; oversampling, which repeatedly takes equal numbers of random subsamples from the two classes [1]; and resampling, which is a mixture of undersampling and oversampling [5], [25]. Algorithmic methods, e.g. adjusting decision class probability, imposing different penalty for different classes, [12], [22] modifying learning function to adjust the class boundary [24], [23], are also commonly used alternatives. Since class imbalance often occurs along with class-overlapping in many applications, the above approaches still have the following problem: improvement on the recognition rate of the skewed-class will generally leads to too much degeneration of the recognition rate on over-weighted class, especially when serious class-overlapping is involved.

Instead of a single-stage SVM with a complicated classification hyperplane, in this paper, the problem is modelled as an aggregation of a set of component SVM classifiers, e.g. linear SVMs or RBF kernel SVMs. Hence the class-imbalance problem can be treated easily by allocating more computing power, i.e. individual SVMs, for the classification of the skewed-class.

## IV. THE PROPOSED GSVMT AGGREGATION METHOD

Towards a modular SVM aggregation classifier with high generalization ability, we address three aspects of problem-solving: 1<sup>o</sup> the convergence of an SVM classification tree (SVMT), 2<sup>o</sup> problem decomposition with the minimum discriminant knowledge loss, and 3<sup>o</sup> the creation of inter-module communication amongst individual SVMs for aggregation.

### A. The Convergence of SVMT

Mathematically, a modular SVM can be formulated as a composite structural model as follows: Given a 2-class data set  $X$  for classification, and a predefined data partitioning model  $\mathcal{P}$  on  $X$ , the data set  $X$  can be partitioned via an optimized  $\mathcal{P}^*$  into  $L$  partitions  $\{g_1, g_2, \dots, g_L\}$ .

Then, an SVMT for binary classification is modelled as,

$$\mathcal{T}_{\text{SVMT}}(\mathcal{P}, f_{\text{SVM}_i^{<2>}}), i = 1, \dots, L, \quad (3)$$

where  $f_{\text{SVM}_i^{<2>}}$  is a regional 2-class decision maker modelled as a standard alone SVM classifier. Parameter  $L$  is determined

during the SVMT modelling process. Then we can have the decision function of SVMT  $\hat{f}$  as,

$$\hat{f}(\mathbf{x}) = f_{\text{SVM}_i^{<2>}} \text{ if } \mathcal{P}(\{\mathbf{x}\}) \in g_i. \quad (4)$$

Obviously, errors may occur when  $\hat{f}$  is used to predict the class label of new coming samples, because of the difference between  $\hat{f}$  and the true classification function,  $f$ . We use a real-valued loss function  $\mathcal{L} = \mathcal{L}(\hat{f}, f)$  to capture the extent of this error. The expectation of  $\mathcal{L}$  is therefore data dependent:

$$\hat{\mathcal{L}} = \int |\hat{f}(\mathbf{x}) - f(\mathbf{x})| p(\mathbf{x}) d\mathbf{x} = \int |\hat{f}(\mathbf{x}) - y| p(\mathbf{x}) d\mathbf{x}, \quad (5)$$

where  $y$  is the true class labels offered by the training set, and  $p(\mathbf{x})$  is the density function.

Given a data partitioning  $\mathcal{P}(\{X\}) = \{g_1, \dots, g_L\}$ , the minimal  $\hat{\mathcal{L}}$  can be obtained by training individual SVMs over all data partitions  $X_i$  (see proof in Appendix).

### B. Data Partitioning

Partitioning methods, such as K-means clustering, partition data normally on the basis of similarity. Yet in the spirit of searching for better individual partitions, we are aiming at partition data, which is to decompose data into partitions of different sizes, taking both input similarities and discriminant ability into consideration.

Assume the data are presented as vectors of attributes,  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ . Let  $A = \{a_1, \dots, a_M\}$  be the attributes set, and  $Y = \{y_1, \dots, y_N\}$  the class labels. The recursive data partitioning function  $\mathcal{P}$  over  $G_0 = \{X\}$  is noted as

$$\mathcal{P}(G_0) = \mathcal{P}^J(G_{J-1}, A_J, E_J) = \{g_1^J, \dots, g_{L_J}^J\}, \quad (6)$$

where  $G_j$  is the partition learned at iteration  $j$  and  $A_j$  the selected set of attributes on which the partitioning is performed. The number of partitions,  $L_J$ , and the number of iterations,  $J$ , are determined during the learning.  $E_J$  is the cost function ensuring an optimal data partitioning, which will be made clear later.

The iteration process is formulated as follows.

**Step 1.** For each partition  $g_i^j$ , an optimal attribute  $a_i^j$  can be picked from attribute set  $A$ , embodying a data partitioning with the minimal discriminant information loss for current subset  $g_i^j$ . In our implementation, we select  $a_i^j$  as

$$a_i^j = \arg \min_{a_i \in A} \frac{|u_{i1} - u_{i2}|}{\sqrt{\frac{s_{i1}}{n_{i1}} + \frac{s_{i2}}{n_{i2}}}}, \quad (7)$$

where  $u_{ik}$  and  $s_{ik}$  are the mean vector and variance of the samples in class  $k$ , respectively.

Along the selected attribute  $a_i^j$ , the data partitioning described in Eq.(6), noted as a data separation  $T(a_i^j)$ , is implemented by a crisp median-point split, or more properly by fuzzy histogram thresholding [2].

**Step 2.** For each subset  $g_i^j$  obtained from the above grid data partitioning, a cost function  $E_i^j$  is defined as

$$E(g_i^j) = \sum_{\mathbf{x}_r \in g_i^j} (y_r - \sum_s (y_s \alpha_s^* K(\mathbf{x}_r, \mathbf{x}_s) + b^*)), \quad (8)$$

where  $\sum_s y_s \lambda_s^* K(\mathbf{x}, \mathbf{x}_s) + b^*$  is the decision function of SVM and  $K(\cdot, \cdot)$  is a user defined kernel function. Equation (8) is applied to estimate the applicability of a given data partition  $g_i^j$  to act as a component classifier in SVMT. As discussed in Section II.A, this guarantees that the loss function of SVMT converges in the end.

**Step 3.** A threshold parameter  $\xi$  is put on  $E_i^j$  to see whether  $g_i^j$  bears a class discriminability large enough to be treated as a component classifier. Hence, every data partition obtained with its  $E_i^j$  less than  $\xi$  are extracted out, and the remaining data are forwarded to the next iteration of data partitioning.

### C. SVMs Aggregating over Grid

From Eq. (6), we can obtain a collection of subsets via  $J$  iterations of recursive partitioning. For regional decision making, a stand alone SVM is trained over each subset to build a 2-class classifier  $f_{\text{SVM}_i^{<2>}}$ . Towards final global decision making, these local SVM classifiers are aggregated onto a grid, whose columns and rows correspond to the iteration of data partitioning and selected attributes by Eq. (7), respectively.

Specifically, given that  $q_j$  attributes are selected by Eq. (7) at iteration  $j$ , and totally  $Q(Q \leq M)$  attributes are selected for  $J$  iterations. Then all data partitions extracted by Eq.(6), together with the corresponding partitioning functions or regional SVM classifiers, composes a composite model  $\Omega$  as

$$c = \begin{pmatrix} \{\mathcal{V}_{11}\} \\ \{\mathcal{V}_{21}\}, \{\mathcal{V}_{22}\} \\ \{\mathcal{V}_{31}\} \dots \{\mathcal{V}_{3q_3}\} \\ \vdots \\ \{\mathcal{V}_{ji}\} \dots \dots \dots \{\mathcal{V}_{jq_j}\} \\ \vdots \\ \{\mathcal{V}_{J1}\} \dots \dots \dots \{\mathcal{V}_{Jq_J}\} \end{pmatrix} \quad (9)$$

where  $\{\mathcal{V}_{ji}\}$  denotes a group of functions that use attribute  $a_i$  for data partitioning at the  $j$ th iteration. The functions could be either regional SVM classifiers  $f_{\text{SVM}_{ji}^{<2>}}$  or regional data partitioning functions  $T(a_i^j)$ .  $\Omega$  models a  $Q$ -dimensional grid which partitions the data into  $L_J$  grids, with the discriminant information in the dataset well preserved in the subsets.

Thus, over the grid learned from data, an SVM tree is constructed by building connections between every two lines of composite model  $\Omega$ .

### D. Coping with Class-Imbalance

During the aggregation, the class-imbalance is addressed in the following two parallel ways.

The first approach implements a recursive data zooming-in schema. Data size with comparatively large sample sizes, especially for the subsets from the over-weighted class, are extracted out in the first few rounds. As a result, the class-imbalance is alleviated for the remaining data. Data with class-overlapping/class-mixture are left for the next step partitioning. Hence, as we drilled down in SVMT, decisions are made only at an appropriate scale where the decision is reachable.

In the second approach, the skewed-class is protected by SVMT via turning SVM loss function  $E$  in Eq. (8) to the skewed class as,

$$L_{\text{SVM}} = \frac{1}{n} \sum_{r=1}^n (y_r - \sum_i (y_s \alpha_s^* K(\mathbf{x}, \mathbf{x}_s) + b^*)). \quad (10)$$

where  $n$  is the sample size of the *regional skewed class*, and  $\mathbf{x}$  a sample from the class.

Note that the *regional skewed class* represents the skewed-class within a data partition. In other words, a *regional skewed class* also could be a globally over-weighted class, if its class distributes are skewed regionally. Hence, Eq. (10) does no overbalanced adjustment as most data resampling methods do to favor the skewed class by using the skewed-class data repeatedly. In fact, the *regional skewed class* loss function Eq. (10) protects equally both the skewed class and the over-weighted class from being underestimated in the regional decision functions and thus helps to overcome the class-imbalance. In practice, Eq. (10) eventually favors the skewed class, because the probability for the globally skewed class to be regionally skewed is often much higher than that of the globally over-weighted class.

#### E. The Proposed gSVMT Algorithm

To build a gSVMT, Eq. (6) is implemented recursively in two steps. First, the grid partitioning is employed to split the data recursively. During this procedure, data partitions with a cost function  $E_i^j$  smaller than threshold  $\xi$  are extracted out. For each of these partitions, an SVM classifier is built and associated with a leaf node of the gSVMT structure.

To test a constructed gSVMT, an input sample  $\mathbf{x}$  is first judged by the test function  $\mathcal{T}(\mathbf{x})$  at the root node of the gSVMT. Depending on the decision made by the root node,  $\mathbf{x}$  will be redirected to one of the children of the root node. This procedure is repeated until a leaf node is reached. The final prediction  $\hat{f}(\mathbf{x})$  is made by the SVM classifier associated with the leaf node. Algorithm 2 presents the testing procedure of a gSVMT.

In the above algorithms,  $\mathbf{K}$  specifies the type of SVM used in gSVMT construction.  $\xi$  is the permitted minimal loss function value, which gives a criterion to select partitions which can be well modelled by a stand alone SVM classifier. The default value for  $\xi$  is zero. However, in practice, a small positive value near zero should be given. Obviously, the smaller  $\xi$ , the more time cost to train the gSVMT. Fig. 3 gives an example of gSVMT. As can be seen, gSVMT is a binary tree with ‘P0’ as the root node. Every ellipse here represents an SVM node, and every internal node is a testing function for data partitioning. The size of an ellipse identifies the size of training data set associated with the node.

## V. EXPERIMENTS AND DISCUSSION

In this section, the experimental results on UCI Machine Learning benchmark data sets and a facial membership authentication application are reported.

#### Function: gSVMTtraining

##### input:

$\mathcal{X}_{\text{train}}$ ; /\* training data set \*/  
 $\xi$ ; /\* the threshold of SVM loss function \*/  
 $\mathbf{K}$ ; /\* the used type of SVM Kernel \*/

##### output:

$\mathcal{T}$ ; /\* generated gSVMT \*/

##### variables:

$\mathbf{x}$ ; /\* a training instance \*/  
 $\mathcal{T}_p$ ; /\* an SVM tree node \*/  
 $\mathcal{X}_p$ ; /\* data partition on an SVM node \*/  
 $\mathcal{X}_{pt}$ ; /\* merged data partition for SVM node training \*/

##### function Calls:

SVMnode\_train( $\mathcal{T}_p$ ,  $\mathbf{K}$ ); /\* train an SVM classifier \*/  
Partition( $\mathcal{X}$ ); /\* data partitioning function \*/  
UpdateTree( $\mathcal{T}$ ,  $\mathcal{T}_p$ ); /\* adding a new node \*/

##### begin

```

1 $\mathcal{T} \leftarrow 0$; /* initialize gSVMT as a root node */
2 $d \leftarrow \text{variable_select}(\mathcal{X}_{\text{train}})$; /* initialize variable */
3 if $\mathcal{X}_{\text{train}}$ is empty
4 return $\mathcal{T}$; /* Iteration stops, and gSVMT generated */
5 $\mathcal{P} \leftarrow \text{Partition}(\mathcal{X}, d)$;
6 for each $\mathcal{X}_p \in \mathcal{P}$ {
7 if $\sum_{\mathbf{x} \in \mathcal{X}_p} |Svm^{<m>}(\mathbf{x}) - f(\mathbf{x})| > \xi$ {
8 UpdateTree($\mathcal{T}$, $\mathcal{T}_p$);
9 $\mathcal{X}_{\text{train}} \leftarrow \mathcal{X}_{\text{train}} - \mathcal{X}_p$;
10 }
12 gSVMTtraining($\mathcal{X}_{\text{train}}$, ξ , $\mathbf{K}$); /* zooming in */
13 for $\mathcal{T}_p \in \mathcal{T}$ {
14 $\mathcal{X}_{pt} = \text{SVMnode_Merge}(\mathcal{T}_p, \mathcal{T})$;
15 SVMnode_train($\mathcal{X}_{pt}$, $\mathbf{K}$);
end

```

Algorithm 1. gSVMT Training

#### A. UCI Data sets

We test the proposed gSVMT algorithm on 13 UCI data sets. Table 1 reports the error rates obtained by the proposed gSVMT method with a comparison against previously reported results on SVM classifiers [13]. We divide the 13 data sets into two categories according to the sparseness value. Six data sets with the sparseness greater than 0.3 are taken as sparse data sets, and the remaining 7 data sets are counted as non-sparse data sets.

As shown in the table, the proposed gSVMT matches SVMs on 5 out of 6 sparse data sets when proper RBF kernels are employed. It is surprising that gSVMTs wins SVM over 15% on the classification accuracy of 2 non-sparse data sets, Flare-solar and Titanic. As we exploring the distribution characteristics of these two data sets in their feature space, it is found that both of them have some discrete attributes. In other words, both two data sets are a type of discrete data set, which is a special case of sparseness covered by our sparseness definition. The extraordinary performance on these two data sets suggests that the proposed gSVMT is very efficient on the classification of discrete data sets. The proposed gSVMT performs not as well as single SVMs on those non-sparse data set, this possibly because the grid partitioning of gSVMT has destroyed too much information useful for classification.

**Function:** gSVMTtesting

**input:**

$T$ ; /\* a constructed gSVMT \*/  
 $x$ ; /\* a testing instance \*/

**output:**

$C$ ; /\*output classifier \*/

**variables:**

$Current$ ; /\*current node number of  $T$  \*/  
 $Next$ ; /\* the node next to  $Current$  node \*/  
 $T_i$ ; /\*The test function on node  $i$  \*/

**Function Calls:**

$SVMTest(x, K)$ ; /\* trained SVM classifier \*/  
 $SgnTest(x)$ ; /\* trained one-class classifier \*/  
 $SearchNode(T, i)$ ; /\* search the node next to  $i$  in  $T$  \*/

**begin**

```

1 $Current \leftarrow 0$;
2 while($SearchNode(T, Current) == Null$){
3 $Next \leftarrow SearchNode(T, Current)$;
4 $Current \leftarrow Next$;
5 if $T_{Current}$ is 2-class
6 $C \leftarrow SVMTest(x, K, Current)$;
7 else
8 $C \leftarrow SgnTest(x, Current)$;
9 return C ;
end
```

Algorithm 2. gSVMT Testing

TABLE I

COMPARISON BETWEEN SVM AND gSVMT ON 13 UCI DATA SETS.

Data sets	Sparseness	SVM	gSVMT	$\xi$
Breast Cancer	0.1039	$0.26 \pm 0.05$	$0.28 \pm 0.05$	0.1
German	0.0777	$0.24 \pm 0.02$	$0.37 \pm 0.03$	0.05
Heart	0.2637	$0.16 \pm 0.03$	$0.16 \pm 0.03$	0.02
Image	0.2806	$0.03 \pm 0.006$	$0.06 \pm 0.01$	0.5
Splice	0.0318	$0.11 \pm 0.01$	$0.16 \pm 0.02$	0.01
Flare-solar *	0.0265	$0.32 \pm 0.02$	$0.05 \pm 0.02$	0.01
Titanic *	0.0942	$0.24 \pm 0.01$	$0.06 \pm 0.04$	0.7
Waveform *	0.7136	$0.10 \pm 0.01$	$0.13 \pm 0.01$	0.01
ringnorm *	0.6575	$0.02 \pm 0.001$	$0.02 \pm 0.01$	0.01
Diabetes *	0.3822	$0.24 \pm 0.02$	$0.24 \pm 0.02$	0.0001
Banana	0.7727	$0.12 \pm 0.01$	$0.14 \pm 0.01$	4
Thyroid *	0.7452	$0.05 \pm 0.02$	$0.04 \pm 0.02$	1
Twonorm *	0.7043	$0.03 \pm 0.002$	$0.02 \pm 0.00$	0.002

### B. Face Membership Authentication Application

To test gSVMTs on problems where dynamic class-imbalance and class-overlapping are prevalent, we have also applied the algorithm to the face membership authentication (FMA) problem [17], [16], [15]. The objective of an FMA system is to distinguish the membership class from the non-membership class based on their facial features presented as images.

FMA is counted as a problem involving sparse data, because the within-class similarity is rather low: Facial appearance of one person may look very different from another, despite the fact that they both belong to the same membership group. In other words, face samples for one person can be far away

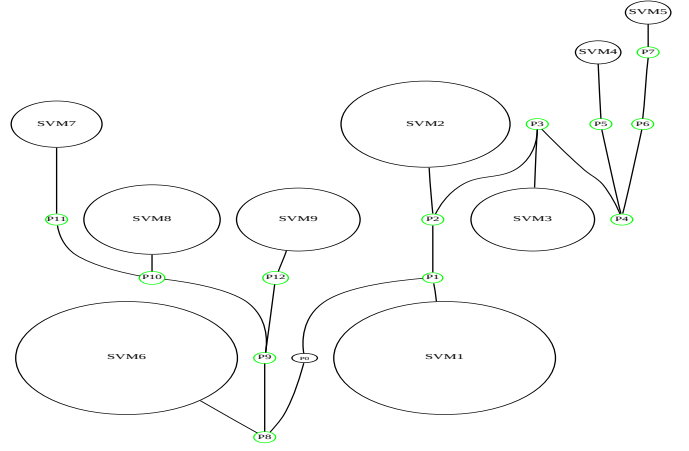


Fig. 3. An example of gSVMT

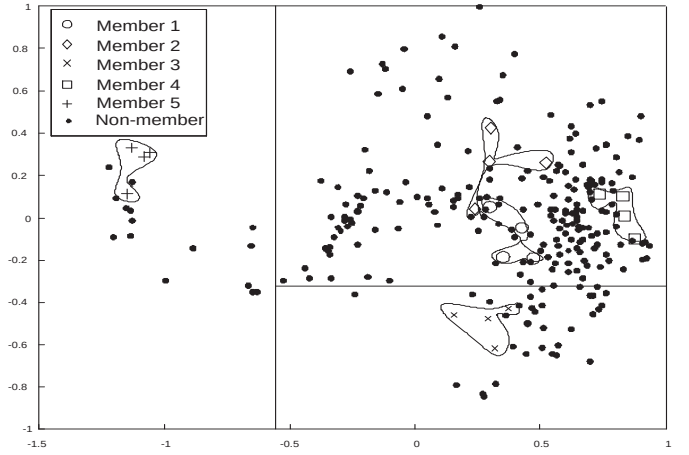


Fig. 4. The illustration of FMA of 5 members against 100 non-members

from samples for another person in the feature space. Another challenging property of a FMA problem is severe class-imbalance: The size of membership group is generally much smaller than that of non-membership, and the class-imbalance may vary dynamically over time in real applications.

In the experiment, we work on a database composed of facial images from 271 individuals, with 5 different images for each. Facial features are extracted using principle component analysis (PCA), and the eigenvector associated with top 100 eigenvalues are used in our experiment. The sparseness of the data set is 0.6605. Although this value does not really reflect the sparseness of the data set, the capability of the data set for data partitioning is obvious. Fig. 4 presents an FMA problem of 5 members against 100 non-members in the feature space defined by the first two eigenvectors. As we can imagine, to distinguish these 5 members from the 100 non-members, a complicated decision function is required. However, such classification difficulty can be reduced greatly, as shown in Fig. 4, if 5 members can be individually authenticated by separated simpler decision rules. This indicates that the FMA problem

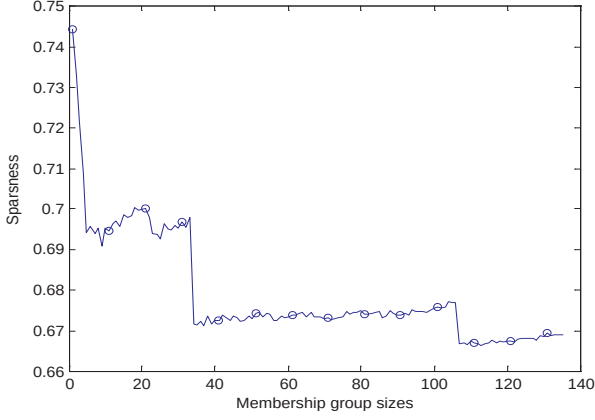


Fig. 5. The sparseness for the membership class data, with membership group size increasing from 10 to 135.

can be well fitted by a divide-and-conquer model.

To test the robustness of gSVMT to class imbalance, we formulate a sequence of FMA problems with the membership group size increasing from 10 ( $C_1 = 0.962$ ) to 120 ( $C_1 = 0.205$ ), with a step size of 10 persons. Here, as an extreme case of class-imbalance, results for the member group size of 135 ( $C_1 = 0.007$ ) are also reported. Note that the sparseness for the whole data set does not change over time, however, that of the membership class data is decreasing as shown in Fig. 5.

In particular, when the size of the membership group is 10, it indicates that 10 members scatter randomly in the feature space among the remaining 261 non-members. In this case, although the sparseness for the whole data set is the same as before, the sparseness within membership group reaches the maximum because the within-class similarity of the 10 members is less than cases with more than 10 members. Hence, membership group data distribution shows a large sparseness value.

Figure 6 gives a comparison of the member authentication rate (the classification rate of the skewed class) versus class-imbalance. Four curves have recorded how the four classifiers perform under the variation of class-imbalance. As shown in the figure, SVMs get large variations on the classification of the skewed class (the membership group), while the membership group size is growing from a small size, 10, to the maximum, 135. On the other hand, the proposed gSVMT shows obvious robustness to the class-imbalance variation, keeping the member authentication rate around 80% under all class-imbalance situations.

Especially for cases where the sparseness is large, it is difficult for ordinary SVMs to find a complicated but continuous decision rule to distinguish a small number of members accurately from a large number of non-members. On the contrary, the proposed gSVMT has shown a very good performance on these cases. The sparsity of the data set can be seen as rather an advantage for gSVMT, than a disadvantage for ordinary SVMs.

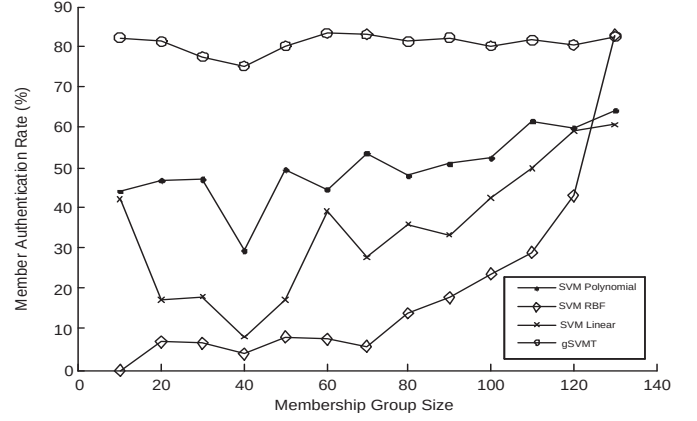


Fig. 6. The proposed gSVMT is compared with 3 single SVMs on FMA

## VI. CONCLUSION AND FUTURES WORK

Addressing the problem caused by the sparsity and class-imbalance of data distribution, this paper proposes a new type of SVM based modular methods. The advantages of the proposed gSVMT are as follows. First, gSVMT keeps every iteration of data partitioning to be performed only at the sparse region of the data, minimizing the loss of discriminant information during the SVMT construction. Second, gSVMT addresses the class-imbalance problem by adjusting the partitioning cost function on the skewed-class, so that the skewed-class is emphasized by every SVM node in gSVMT. The proposed gSVMT is experimented on 13 UCI data sets and a real Facial Membership Authentication application which is characterized by the dynamic class-imbalance and class-sparseness property. The experimental results demonstrated that, compared with stand alone SVMs, the proposed gSVMT is able to treat with the sparseness and class-imbalance of the input data and delivers competitive classification performance.

From the viewpoint of SVM aggregating intelligence [14], the proposed gSVMT follows the same route of recursive data partitioning plus aggregating as the previous SVM tree methods. The gSVMT is different from all previous SVMT aggregating methods [15] because 1<sup>o</sup> grid partitioning is used for problem decomposition; 2<sup>o</sup> every internal node of the gSVMT is a data partitioning function, and every leaf node is an SVM node. Note that gSVMT has no one-class node, and decisions are made by the leaf SVM node.

The proposed gSVMT is merely sustainable for 2-class problems. It can be simply extended to  $m$ -class problems by learning  $m$  binary classifiers for each class versus the others, or by learning  $m(m-1)$  binary classifiers for each class versus another class. This topic is generally discussed under the framework of SVM decision forests [8]. Alternatively, in the principle of SVM aggregating, a multi-class gSVMT (m-gSVMT) can be constructed by decomposing a  $m$ -class task into a certain number of 2 to  $m$  classes subtasks as

$$\mathcal{T}_{m\text{-gSVMT}}(\mathcal{P}^n, f_{\text{SVM}^{<2>}}, \dots, f_{\text{SVM}^{<m>}}, \mathbf{x}), \quad (11)$$



where  $\mathcal{P}^n$  represents a partitioning function which is able to split a set into multiple subsets. Similar to 2-gSVMT,  $f_{\text{SVM}^{<i>}}$  is an  $i$ -class ( $2 \leq i \leq m$ ) local SVM classifier. Accordingly, the classification function of a  $m$ -gSVMT can be written as,

$$\hat{f}(\mathbf{x}) = f_{\text{SVM}^{<i>}} \text{ if } \mathcal{P}^n(\mathbf{x}) \in g_j. \quad (12)$$

As a measurement of the sparsity of a data set, sparseness is an important concept for divide-and-conquer data modelling. However, it is also a concept difficult to give an accurate and consistent evaluation, just as has been reflected in our work. Still, sparsity does exist for most life data, thus for better generalization ability of gSVMT, it will be an interesting future work to derive a new sparseness measurement consistent for most data sets.

## REFERENCES

- [1] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research (JAIR)*, no. 16, pp. 321–357, 2002.
- [2] H.D. Chen and Y. H. Chen "Fuzzy partition of two-dimensional histogram and its application to thresholding," *Pattern Recognition*, vol. 32, no. 5, pp. 825–843, May 1999.
- [3] H. Drucker, Donghui Wu, and V. N. Vapnik, "Support vector machines for spam categorization," *IEEE Transactions on Neural Network*, vol. 10, no. 5, pp. 1048–1054, Sept. 1999.
- [4] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*, Wiley & Sons, 2 edition, 2001.
- [5] J. Estabrooks, and Japkowicz, "A multiple resampling method for learning from imbalanced data sets. *Computational Intelligence*," no. 20, pp. 18–36, 2004.
- [6] S. Haykin. *Neural Networks - A Comprehensive Foundation*. Macmillan Colledge Publishing Company, New York, 1994.
- [7] J. Huang, and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Trans. on Knowledge and Data Engineering*, vol. 17, no. 3, pp. 299–310, Mar. 2005.
- [8] Tin Kam Ho, "The random subspace method for constructing decision forests Tin Kam Ho," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), pp. 832–844, Aug. 1998.
- [9] R. A. Jacobs and M. I. Jordan, "Learning piecewise control strategies in a modular neural network architecture," *IEEE Trans. on System, Man and Cybernetics*, vol 23, pp. 337–345, 1993.
- [10] Hyun-Chul Kim, Shaoning Pang, Hong-Mo Je, Daijin Kim, and Sung Yang Bang, "Constructing support vector machine ensemble," *Pattern Recognition*, vol. 36, no. 12, pp. 2757–2767, 2003.
- [11] Sun Kim, Soo-Jin Kim, Byoung-Tak Zhang, "Evolving hypernetwork classifiers for microRNA expression profile analysis," *Proc. of IEEE Congress on Evolutionary Computation*, pp. 313–319, Sept. 2007.
- [12] Y. Lin, Y. Lee, and G. Wahba, "Support Vector Machines for Classification in Nonstandard Situations," *Machine Learning*, vol. 46, pp. 191–201, 2002.
- [13] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K. R. Müller, "Fisher discriminant analysis with kernels," *Neural Networks for Signal Processing IX*, pages 41–48, 1999.
- [14] Shaoning Pang, "SVM Aggregation: SVM, SVM Ensemble, SVM Classification Tree," *IEEE SMC eNewsletter* Dec. 2005. Available at <http://www.ieeesmc.org/Newsletter/Dec2005/R11Pang.php>
- [15] Shaoning Pang, D. Kim, and S. Y. Bang, "Face Membership Authentication Using SVM Classification Tree Generated by Membership-based LLE Data Partition," *IEEE Trans. on Neural Network* vol. 16, no. 2, pp. 436–446, Mar. 2005. C
- [16] Shoning Pang, D. Kim and S. Y. Bang "Membership authentication in the dynamic group by face classification using SVM ensemble," *Pattern Recognition Letters*, vol. 24, pp. 215–225, 2003.
- [17] Shoning Pang, Seiichi Ozawa and Nikola Kasabov "One-Pass Incremental Membership Authentication by Face Classification," *ICBA 2004, LNCS 3072*, Springer pp. 155–161, July 2004.
- [18] F. Provost, "Machine Learning from Imbalanced Data Sets 101," *Working Notes AAAI'00 Workshop Learning from Imbalanced Data Sets*, pp. 1–3, 2000.
- [19] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 660–641, Aug. 2002.
- [20] M. Sarkar, "Modular pattern classifiers: a brief survey," *Proc. of IEEE International Conference on Systems, Man, and Cybernetics*, vol. 4, pp. 2878–2883, Oct. 2000.
- [21] A. Senior, "A combination fingerprint classifier," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1165–1174, Oct. 2001.
- [22] K. Veropoulos, C. Campbell, and N. Cristianini, "Controlling the Sensitivity of Support Vector Machine," *Proc. Int'l Joint Conf. Artificial Intelligence*, pp. 55–60, 1999.
- [23] G. Wu and E. Chang, "Adaptive Feature-Space Conformal Transformation for Imbalanced Data Learning," *Proc. 20th Int'l Conf. Machine Learning*, pp. 816–823, 2003.
- [24] G. Wu and E. Chang, "KBA: Kernel Bounday Alignment Considering Imbalance Data Distribution," *IEEE Trans. Knowledge and Data Engineering*, vol. 17, no. 6 pp. 786–795, 2005.
- [25] Z. H. Zhou and X. Y. Liu, "Training cost-sensitive neural networks with methods addressing the class imbalance problem," *IEEE Trans. on knowledge and data engineering*, vol. 19, no. 1, pp. 63–77, 2006.



## **Runtime Thread Rescheduling: An Extended Scheduling Algorithm to Enhance the Performance of the Gridbus Broker**

Altaf Hussain, Abu Awal Md. Shoeb, Md. Abu Naser Bikas and Mohammad Khalad Hasan

*Department of Computer Science & Engineering*

*Shahjalal University of Science & Technology*

*Sylhet, Bangladesh*

*{altaf\_sust\_82, johnshoeb, bikasbd}@yahoo.com, khalad-cse@sust.edu*

### **Abstract**

*Grid computing is becoming a requirement for the processing of large amount of data now-a-days. The Gridbus broker schedules jobs depending on data and compute resources. Current scheduling process does not reassign a job from lower compute resource to higher compute resource if higher compute resource is available. In this paper, we have proposed a technique to reassign a thread to higher grade executor by preempting the thread in lower grade executor by using the data restoration technique which track the information of the thread so far ran on a lower rate compute resource. It is done only if there is an idle higher computer resource is available. The performance as well as the reliability of the Grid has been improved by this approach in a considerable extent.*

### **1. Introduction**

Grid computing is very useful in the large scientific projects. A grate number of researcher organization and communities using grid computing to carry out their sheared data resource out come. However the using of the power of Grid Computing remains to be a challenging problem for the users due to the complexity involved in the creation of the application and deployment on distributed computational resources [1]. The Gridbus broker is an advanced service-oriented meta-scheduler for data Grids which deals with computation on very large amount of data where data host may geographically distributed. It supports different Grid middleware and services. It takes care of many functions that Grid applications require such as resource selection, job scheduling and management, data access to any distributed Grid

resources and handling faults that may occur during execution, collecting result and presenting them. The Gridbus Broker schedules the job depending on some criteria like data and computes resources, job completion history etc. Here we present a modified scheduling algorithm, where we can use any idle higher power compute resource extensively and thus can improve the performance.

This paper is organized as follows: section 2 discusses the related works and section 3 shortly discusses Gridbus broker architecture, section 4 describes the problem scope of Gridbus broker architecture and scheduling, section 5 elaborates our proposed algorithm to counteract the problems of Gridbus broker scheduling, section 6 illustrates the experimental evaluation and performance analysis of algorithm with some examples. We conclude in section 7.

### **2. Related works**

As the scheduling of independent jobs on distributed heterogeneous resources [2] is a well-studied problem, so here we discussing only a few major work which deal with Grid. In the context of Data Grid, the Storage Resource Broker (SRB) [3] from San Diego Supercomputing Centre (SDSC) provides middleware for storing datasets over a network and accessing them; although it does not deal with applications executing directly.

The Gridbus broker extends the Nimrod-G [4] computational Grid resource broker model to distributed data-oriented grids. Nimrod-G specializes in parameter-sweep computation. However, the scheduling approach within Nimrod-G aims at optimizing user-supplied parameters such as deadline and budget [5] for computational jobs only. Like

Nimrod-G, the AppLeS PST [6] supports deployment of parameter-sweep applications on computational grids, but its adaptive scheduling algorithm emphasizes on data-reuse.

Harold, Shazia, Liao, Krishna, Buyya developed a Grid Thread Programming Environment (GTPE) [7] for the Gridbus broker, a software grid resource broker that provide transparent access to grid services. However, it does not provide any way to save the output of thread at any point of its runtime.

Venugopal, Buyya, Winton [1] focuses on a resource scheduling strategy within a Data Grid but concentrated on adaptive scheduling algorithms and brokering for heterogeneous resources that are shared by multiple user jobs.

Similar to [1], our work focuses on a resource scheduling strategy within a Data Grid but we concentrate on runtime thread scheduling algorithms and brokering for heterogeneous resources that are shared by multiple user jobs. However our work focuses on using the higher rate compute resources more extensively. In addition, the scheduling strategy is based on our 'File based Grid Thread Implementation' that provides a way for saving the output of a thread and creating restoration files.

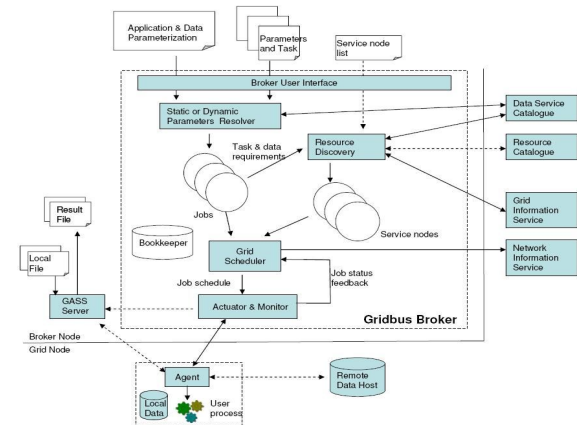
### 3. Gridbus broker overview

The architecture of the Gridbus broker is shown in Figure 1[1]. The Gridbus broker inputs are considered as task and the parameter values associated with the task. The parameters can be static or dynamic. Among the other component the Grid Index Information service (GIIS), gathers information from remote information services i.e. availability and the properties of compute resources. The Data Service Catalogue organizes the data files as Logical File Names (LFNs). Network weather service (NWS) obtains the information about the current available bandwidth between the data sources and the compute resources. Global Access to Secondary Storage (GASS) obtains the inputs or parameter file from the secondary storage if required. Logical and Physical File Names (LFNs), each LFN maps to one or many Physical File Names (PFNs) somewhere on the Grid and the Job Monitor keeps track of job status, whether the jobs are queued, executing, finished successfully or failed.

A task is further considered as sequence of commands that focuses user requirements. The task along with its associated parameters decomposed in to jobs. A job is unit of work that is send to a Grid node. The set of jobs along with the set of service nodes are

an input to the scheduler. For jobs requiring remote data, it interacts with a NWS to obtain the information about current available bandwidth between the data sources and the compute resources. It then uses this information to schedule jobs by optimizing on the amount of data transfer involved.

In the current implementation, the jobs are dispatched to the remote node through the Actuator component. The Actuator submits the job to the remote node using the functionality provided by the middleware running on it. The Actuator has been designed to operate with different Grid middleware frameworks and toolkits such as Globus 2.4 that primarily runs on Unix-class machines and Alchemi [8], which is a .NET based grid computing platform for Microsoft Windows-enabled computers.



**Figure 1: Gridbus broker architecture and its interaction with other Grid entities**

The Gridbus broker scheduling algorithm schedules a job depending compute resources and data host associated with a job that is to be deployed to an executor. Firstly it gathers the information about the compute resources and data hosts using GIIS and Data Service Catalogue and determines the job completion rate for every compute resources for its history. For a job to be deployed, identifies the data hosts associated with that job and for every pair of data and compute resources, it finds the pair with earliest job completion rate and if such a pair exist the scheduler assigns the job to that pair. After the job has successfully finished at the remote node (i.e. executor), its returns any results to the Broker. If any job becomes failed, it is marked as failed and reset for rescheduling.

#### 4. Problem scope of Gridbus Broker architecture and scheduling

In the present scheduling algorithm, the Gridbus Broker schedules the jobs depending on data and compute resources available at the time of the deployment of jobs. An executor (i.e. remote node) with a higher history of job completion gets the higher priority for executing new job i.e. if there is one thread and two executors, one with relatively rich job completion history but lower compute resource and other with poor job completion history but higher compute resource then the first one will get the job. Then the scheduler takes decision depending on the compute resource. Also in the case when a compute resource with a higher job completion rate comes after the scheduling and deployment all jobs, the present scheduling algorithm does not take advantage of that compute resource. Also if there are two executors working for an application one with much higher computing power than the other and the slower one executes the last thread then the completion time of the application depends on the completion of the last thread while the faster one remains idle after completing its last job long before the slower one. In these contexts, the completion time of a job as well as for an application gets much higher as a result of threads running on lower grade compute resource. But in this case if we can use the idle time of the higher compute resource then it can reduce the computation time by a considerable factor.

#### 5. Proposed solution

For the solution of the problem stated in the previous section, if any job running on lower grade compute resource and there is a much higher grade compute resource (let it can finish the job in half of the time than current one) available after the task assignment or gets connected to grid after the first deployment of the job. Then we can preempt the job from the current executor with lower resources and assign the job to the newly available compute resource depending on the data resource for the job. Though this approach gives better performance than existing one however as the new executor starts the job from the beginning the margin of performance improvement is low depending on the time of preemption. The closer the preemption happens to the closer to the start time of the thread the higher the degree of performance improvement. However, there are cases can be found where this approach gives much

worse result if the preemption time is very closer to the finished time of the job.

So, we further move to an approach where we can save the output of a thread in any point of runtime. As we can make restoration point of a thread then we can save the computation time of the job has been executed before preemption. By this approach, whenever a compute resource with higher computation power gets idle or available in the Grid, we can preempt the job running on a lower grade computation resource and reschedule the job to execute on that new executor. At the time of reassigning the job to the new compute resource we can make the restoration file of the job available to the new executor so that it can start executing from the saved point of the job. Suppose the previous node compute 40% of the job before preemption, when the job is reassigned to new node with generated output file. This node will restart the job from the saved point (i.e. from the 40%, not from the beginning). As in this approach the new executor does not need to start from the beginning of the job, as a result the degree of performance improvement gets much higher.

In our application (File Based GRID Thread Implementation), we created a random access file associated with each thread and writing the results to that file and thus creating restoration point. The file only contains the last two computed output of a thread. The file size depends on the type of application but in our application we managed it to be quite small. This generated file is stored in the Manager Node. The file of a thread is deleted automatically after the successful completion of the thread. Threads received from the User are placed in a pool and scheduled to be executed on the various available Executors by the Manager. Manager can reassign that thread to any executor node with the same thread ID from the thread pool. When the Manager distributes the thread to any Executor node, the node first checks if there is any file available with its own thread ID. If any file is available for that thread, the node then recovers the resultant value of that thread from the file created and starts execution from the recovered value. If no file is found for that thread then the thread will start execution from the beginning and also create a file in the Manager Node using Thread ID (i.e. filename) as Thread ID remains same throughout the lifetime of an application.

In the global grid environment, these generated restoration files can be stored in the remote data host assigned with that for that Grid node or can be stored in some different host depending on the bandwidth

and cost of data transfer etc. For the jobs requiring remote data, the scheduler interacts with the Network Weather Service (NWS) [9] to get the current bandwidth condition. We can create a new instance for the restoration file in the Logical File Names (LFNs) which maps to Physical File Name of the restoration file stored in the appropriate location or data host. This instance can be updated and can use for the selection of the restoration file host. The Gridbus broker can resolve the restoration files by querying the Data Service Catalogue. When the scheduler reschedules that job for execution the restoration file can be assigned with that job so that the new executor can retrieve the file from the data host before execution and start execution from that successful executed point.

Depending on this approach, we modified the existing scheduling algorithm of the Grid bus broker. The modified Gridbus broker scheduling algorithm is stated in *Figure 2*. We added a new module (module 6) in the existing algorithm for rescheduling the jobs running on lower grade compute resource while there available any idle higher grade compute resource and no job in unassigned-job-list.

The scheduler schedules the job considering the earliest job completion time between two compute resources depending on the data host associated with that job. If the higher grade compute resource has the earliest job completion time then the scheduler preempt the job and reassign it to the new executor with the restoration files. The new executor collect the restoration files and start the job from the last saved point.

#### **Initialization**

1. Identify characteristics, configuration, capability, and suitability of compute resources using the Grid information services (GIS).

2. From the task definition, obtain the data query parameters (if present), such as the logical file name.

a. Resolve the data query parameter to obtain the list of Logical Data Files (LDFs) from the Data Catalog.

b. For each LDF, get the data sources or Data Hosts that store that file by querying the Data Catalog.

#### **Scheduling loop**

Repeat while there exists unprocessed jobs. [This step is triggered for each scheduling event. The event period is a function of job processing time, rescheduling overhead, resource share variation, etc.]:

3. For each compute resource, predict and establish the job consumption rate or the available

resource share through the measure and extrapolation strategy taking into account the time taken to process previous jobs. Use this estimate along with its current commitment to determine expected job completion time.

4. If any of the compute resource has jobs that are yet to be dispatched for execution and there is variation in resource availability in the Grid, then move such jobs to the unassigned-Jobs-List.

5. Repeat until all unassigned jobs are scheduled or all compute resources have reached their maximum job limit.

a. Select the next job from the Unassigned-Jobs-List.

b. Identify all Data Hosts that contain the LDF associated with the job.

c. Create a Data-ComputeResource-List for the selected job:

- For each data host, identify a compute resource that can complete the job earliest given its current commitment, job completion rate, and data transfer time using current available bandwidth estimates.

d. Select a data host and compute resource pair with the earliest job completion time from the Data-ComputeResource-List.

e. If there exists such a resource pair, then assign the job to the compute resource and remove it from the Unassigned-Jobs List.

6. Repeat while there is no unassigned job and there exists any idle higher grade compute resource then any node in the Grid [This step is triggered for each scheduling event. The event period is a function of rescheduling overhead, restoration file recovering over head, etc.]:

a. Select the lowest grade compute resource executing any job.

b. Select a data host and compute resource with the earliest job completion time between the two compute resources depending Data-ComputeResource-List.

c. If the idle compute resource can complete the job in earliest time then preempt the job and

i. Select the restoration file host (data host) that contain LDF associated with that job.

ii. Reassign the job to the higher grade compute resource.

7. End of scheduling loop.

**Figure 2: Modified adaptive scheduling algorithm for Data Grid**

## **6. Experimental result and performance analysis**

We setup the experiment environment with heterogeneous executors, one with lower grade computation power and the other with much higher

grade computation power. This higher grade compute resource can perform almost twice as faster than the other lower grade compute resource. The configuration of the two executors is given in *Figure 3* and *Figure 4*.



Processor: Inter Celeron® CPU 2.66 GHz  
 Ram: 512 MB  
 OS: Microsoft Windows XP Professional V2002 Service Pack 2

**Figure 3: Executor 1 (lower grade compute resource)**



Processor: Inter Core2Duo 1.86 GHz  
 Ram: 1024 MB  
 OS: Microsoft Windows XP Professional V2002 Service Pack 2

**Figure 4: Executor 2 (higher grade compute resource)**

Two different application programs have been used to evaluate the performance comparison. In our Implementation, both applications take an integer as input and give the number of factors as output. In case of first application program, we developed with existing algorithm that does not use any technique to save the output of any thread. As there is a non-zero probability of thread failure so, if any thread became failed, we can not get the partial result so far computed by that thread. In contrast, in our ‘File based Grid Thread implementation’, we created a random access file associated with each thread and writing the results to that file and thus creating restoration point. The file only contains the last two computed output of a thread. Two status characters for two results are also stored to indicate which of the results are stored successfully or not. The size of each result is 10 byte. Since the file contains two results, the total size of the file is 20 byte.

In both applications, same input is used to compare between them, both give the same results as well. Both application give the total number of factors of a given

input value (e.g.  $10^{10}$ ). We will show the performance comparisons in the context of a single thread and also in the context of an application comprises by a number of thread. These cases will give us an overview of how these approaches effect in case of a thread and in the application context. Here we have shown different performance comparisons where our modified approach gives better performance than the existing approach in almost every case. The performance of modified approach depends on how long a thread is executed before it is being preempted. The much shorter a thread executed before it is being preempted the much better performance is obtained by our modified file based implementation.

### 6.1. In the context of a single thread

This case gives an overview how the modified run thread scheduling using preemption approach improves performance in case of a single thread. Here we give the comparisons between the performance of thread scheduling technique using the existing implementation (where preempted thread starts form the beginning at the reassigned executor) and the modified file based implementation (where preempted thread starts form the last saved point at the reassigned executor).

**6.1.1. Test case 1 (without preemption):** In this test module we showed how much time a single thread takes to finish by *Executor 1* and by the *Executor 2* separately. In all the test cases here threads are executed without any preemption (i.e. finished successfully). This table also shows the comparisons between the times taken by the existing implementation verses our modified file based implementation.

**6.1.2. Test case 2 (with preemption):** In this test module we show the comparison between the performance of the existing implementation and the modified implementation by preempting a thread at several times in its life time. We took the thread same as the thread describe in *Test Module 1*. The thread using existing implementation, first executed in the *Executor 1*, which takes an average of 1:21.9 min to execute normally. So we preempt the thread for  $10^{\text{th}}$  to  $70^{\text{th}}$  second with 10 sec interval after starting execution at the *Executor 1*. Then we make the higher compute resource (*Executor 2*) available to the grid and the thread starts execution form the saved point using restoration point. This technique finishes the job

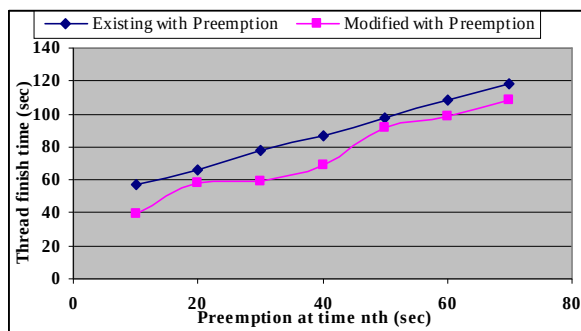
in much lesser time in almost every case though the second executor starts the job from the beginning.

**Table 1: Comparisons between the execution time of Executor 1 and Executor 2 with a single thread**

Test case	# of thread	Existing implementation		Modified file based implementation	
		Executor 1	Executor 2	Executor 1	Executor 2
1	1	0:1:21.87500	0:0:36.390625	0:1:21.828125	0:0:37.687500
2	1	0:1:21.78125	0:0:36.437500	0:1:21.109375	0:0:36.421875
3	1	0:1:22.12500	0:0:37.390625	0:1:21.890625	0:0:36.390625

**Table 2: Comparison between two implementation with job preemption**

Test Case	Preempt time (n <sup>th</sup> sec)	Finish time with preemption (min : sec)	
		Existing Implementation	Modified Implementation
1	10	0:57.265625	0:39.421875
2	20	1:06.171875	0:57.984375
3	30	1:17.546875	0:59.015625
4	40	1:26.859375	1:09.359375
5	50	1:37.187500	1:32.078125
6	60	1:48.515625	1:38.953125
7	70	1:58.156250	1:48.609375



**Figure 5: Comparison between two implementation with job pre-emption**

By using the modified file based implementation we use the same technique but the second executor does not need to start the job from the beginning. As the *Executor 2* starts the job from last saved point it

has executed in the first executor thus we can save more time. Thus this approach gives a greater degree of performance improvement, which can be clearly seen from the *Figure 5* and *Table 2*.

## 6.2. In the context of an application (multiple threads)

In this case we show how the run time thread scheduling with preemption effect on completion time of an application comprises of several number of threads. We take an application with five threads. For this test case last thread is executed on the lower grade compute resource. Here we preempt the last thread in several point of its life time.

**6.2.1. Test case 1 (without preemption):** This test module gives an overview of how much time the executors takes (separately or both executors together) to finish an application with five threads. This table also compares between the existing and modified implementation.

**6.2.2. Test case 2 (with preemption):** In this test module we show the comparison between the performance of the existing implementation and the modified implementation by preempting the last thread of an application which runs at the lower grade compute resource (*Executor 1*) by the existing Gridbus Broker scheduling algorithm. Firstly, we run the application with existing implementation and preempt the last thread at *Executor 1*. We took the similar approach described in the *Test Module 2* to do this section experiment. We preempted the job with 10

**Table 3: Comparison of execution time of an application with five threads between existing and modified implementation**

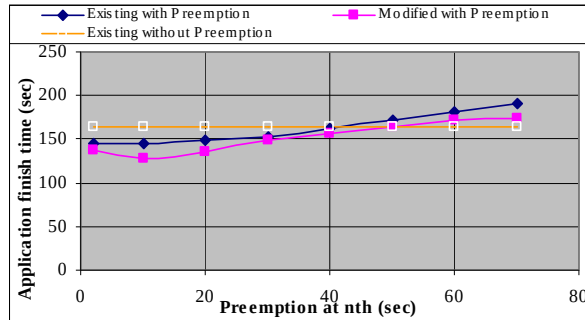
Modified implementation (min : sec)			Existing implementation (min : sec)		
Executor 1	Executor 2	Together	Executor 1	Executor 2	Together
6:45.7344	3:01.4063	3:00.1094	6:47.2031	3:01.8069	2:43.2812
6:44.9817	3:00.4063	2:54.8594	6:48.1709	3:01.4687	2:43.0570
6:44.1986	3:00.1563	2:54.8906	6:47.5803	3:00.8390	2:44.0196

seconds interval and let the *Executor 2* to execute the job which can do the job much faster than the

*Executor 1*. This technique finishes the application in much lesser time in almost every case as we remove the bottleneck created by the lower grade executor though the second executor starts the job from the beginning. By using the modified file based implementation we use the same technique but as the second executor does not need to start the job from the beginning. Since the *Executor 2* starts the job from last saved point it has executed by the first executor, it reduces the execution time greatly. Thus this approach gives a greater degree of performance improvement, which we can clearly see in the *Figure 6* and *Table 4*.

**Table 4: Performance comparison between two implementations with job preemption**

Test Case	Preempt time (n <sup>th</sup> sec)	Application finish time (min : sec)	
		Existing Implementation	Modified Implementation
1	2	2:25.406250	2:16.578125
2	10	2:25.312500	2:07.171875
3	20	2:27.921875	2:16.250000
4	30	2:32.500000	2:28.812500
5	40	2:41.937500	2:35.843750
6	50	2:51.859375	2:44.875000
7	60	3:01.156250	2:51.468750
8	70	3:10.734375	2:54.281250



**Figure 6: Performance comparison between two implementations with job preemption**

## 7. Conclusion

In this paper, a modified algorithm has been presented for the scheduling of the Gridbus broker. By using this modified file based implementation and the described approach to reschedule the threads at run time; it is clear that algorithm gives better time improvement with compare to the existing thread

implementation. It is known that, Grid thread is the last unit of the job that is executed in the executor node. As we get better performance by using discussed approach so we can propose that if this mechanism can be added to the Gridbus Broker it can push the Grid computing one step further.

## 8. References

- [1] A.B. Smith, C.D. Jones, and E.F. Roberts, "Article Title", *Journal*, Publisher, Location, Date, pp. 1-10.
- [2] Jones, C.D., A.B. Smith, and E.F. Roberts, *Book Title*, Publisher, Location, Date.
- [1] S. Venugopal, R. Buyya and L. Winton, "A Grid Service Broker for Scheduling Distributed Data-Oriented Applications on Global Grids", *Technical Report*, GRIDS-TR-2004-1, Grid Computing and Distributed Systems Laboratory, University of Melbourne, Australia, February 2004.
- [2] M. Maheswaran, S. Ali, H. J. Siegel, D. Hensgen and R. F.Freund, Dynamic Mapping of a Class of Independent Tasks onto Heterogeneous Computing Systems, *Journal of Parallel and Distributed Computing*, Volume 59, Issue 2, November 1999, Pages 107-131.
- [3] C. Baru, R. Moore, A. Rajasekar, and M. Wan, The SDSC Storage Resource Broker, in *Proceedings of the 1998 conference of the IBM Centre for Advanced Studies on Collaborative research (CASCAN'98)*, Toronto, Canada, October 1998, ACM Press, New York, USA.
- [4] A. Samar, H. Stockinger. Grid Data Management Pilot (GDMP): A Tool for Wide Area Replication, *IASTED International Conference on Applied Informatics (AI2001)*, Innsbruck, Austria, February 2001, ACTA Press, Calgary, AB, Canada.
- [5] A. Chervenak, E. Deelman, I. Foster, L. Guy, W. Hoschek, A. Iamnitchi, C. Kesselman, P. Kunst, M. Ripeanu, B. Schwartzkopf, H. Stockinger, K. Stockinger, B. Tierney. Giggie: A Framework for Constructing Scalable Replica Location Services, *Proceedings of Supercomputing 2002 (SC2002)*, November 2002, IEEE Computer Society Press, Los Alamitos, CA, USA.
- [6] W. Hoschek, J. Jaen-Martinez, A. Samar, H. Stockinger, K. Stockinger, Data Management in an International Data Grid Project, *Proceedings of the 1st International Workshop on Grid Computing (Grid 2000, Bangalore, India)*, Springer-Verlag, Berlin, Germany, 2000.
- [7] Harold Soh, Shazia Haque, Weili Liao, Krishna Nadiminti and Rajkumar Buyya, "GTPE: A Thread Programming Environment for the Grid".

[8] A. Luther, R. Buyya, R. Ranjan, and S. Venugopal, Alchemi: A .NET-based Grid Computing Framework and its Integration into Global Grids, *Technical Report*, GRIDS-TR-2003-8, Grid Computing and Distributed Systems Laboratory, University of Melbourne, Australia, December 2003.

[9] R. Wolski, N. Spring, and J. Hayes, “The Network Weather Service: A Distributed Resource Performance Forecasting Service for Metacomputing”, *Journal of Future Generation Computing Systems*, Volume 15, Numbers 5-6, pp. 757-768, Elsevier Science, Amsterdam, Netherlands.



## An Efficient Clustering based Texture Feature Extraction for Medical Image

Marghny.H.Mohamed  
Faculty of computers and  
Information  
Assiut University, Assiut, Egypt,  
marghny@aun.edu.eg

M.M.AbdelSamea  
Faculty of Science  
Assiut University, Assiut,  
Egypt,  
mm\_abdelsamea@yahoo.com

### Abstract

*In some medical applications where a tissue of interest covers a large fraction of the image or a prior knowledge on the region of interest is available, extracting features by fixed blocs in the image is sufficient. However in the general case, one would like to identify features for each tissue in the image. This would require prior image segmentation. Medical image segmentation is one of the most challenging problems in medical image analysis and a very active research topic. Therefore, there is no algorithm available in the general case for isolating medical image regions [1].*

*This paper presents an accurate method for extracting texture features from medical image for classification. It is based on bloc wise clustering of medical images. The proposed technique extracts accurate and general set of textural features. Experimental result showed the high accuracy of the extracted textural features. Experiments held on Mammographic Image Analysis Society (MIAS) dataset.*

### 1. Introduction

Image analysis techniques have played an important role in several medical applications. In general, the applications involve the automatic extraction of features from the image which is then used for a variety of classification tasks, such as distinguishing normal tissue from abnormal tissue. Chabat [2] used 13 texture parameters, derived from the histogram, co-occurrence matrix and run-length matrix categories, to differentiate between a variety of obstructive lung diseases in thin-section CT images. Kovalev [3] used texture parameters derived from gradient vectors and from generalized co-occurrence matrices for the characterization of texture of some MR-T2 brain images. Herlidou [4] used texture parameters based on the histogram, co-occurrence matrix, gradient and run-length matrix for the characterization of healthy and pathological human brain tissues (white matter, grey matter, cerebrospinal fluid, tumours and oedema). Mahmoud [5] used the texture analysis approach based on a three-dimensional co-occurrence matrix in order to improve brain tumour characterization. Du-Yih Tsai [6] used four texture features derived from the co-occurrence matrix was used for classification of the heart disease. H.S. Sheshadri [7] used Six textural features for mammogram images derived from the histogram categories was used as a part of developing a computer aided decision system for early detection of breast cancer. Maria-Luiza [8] used texture parameters based on the histogram for tumour classification in mammograms.

Feature extraction [9] is a vital component of the Computer Aided Diagnosis (CAD) System that can discriminate between medical tissues to serve as a second reader to aid radiologists. The feature extraction unit is used to prepare

data in a form that is easy for a decision support system or a classification unit to use. Compared to the input, the output data from the feature extraction unit is usually of a much lower dimension as well as in a much easier form to classify. Medical images possess a vast amount of texture information relevant to clinical practice [10]. Hence texture is the most promising feature to work on. Texture analysis gives information about the arrangement and spatial properties of fundamental image elements [11]. Coggins [12] has compiled a catalogue of texture definitions.

One of the most commonly used texture parameters come from Co-occurrence matrix as a statistical approach [10] which represents texture in an image using properties governing the distribution and relationships of grey-level values in the image methods normally achieve higher discrimination indexes than the structural or transform methods.

This paper introduces a bloc wise clustering as a region of interest selection (ROI) approach which provides a more accurate extraction of textural features. We use four texture features measured from a gray-level co-occurrence matrix generated from the breast images for classification of the images. A statistical discrimination method (fisherfaces algorithm) [13, 14] for feature selection algorithm is also used for extracting discriminative information from extracted feature of medical images to be used as inputs to our classification system.

### 2. Data collection

The data collection which has been used in our experiments was taken from the MIAS [15]. The same collection has been used in other studies of automatic mammography classification. Its corpus consists of 322 images, which belong to three big categories: normal, benign and malign. There are 208 normal images, 63 benign and 51 malign, which are considered abnormal. In addition, the abnormal cases are further divided in six categories: microcalcification, circumscribed masses, speculated masses, ill-defined masses, architectural distortion and asymmetry.

### 3. Pre-processing phase

Mammograms are images difficult to interpret, and a preprocessing phase of the images is necessary to improve the quality of the images and make the feature extraction phase more reliable. We applied to the images two

techniques: a cropping operation and an image enhancement one. The first one was employed in order to cut the black parts of the image as well as the existing. Image enhancement helps in qualitative improvement of the image with respect to a specific application. In order to diminish the effect of over brightness or over darkness in the images and accentuate the image features, we applied a widely used technique in image enhancement to improve visual appearance of images known as Histogram Equalization. This process equalizes the illumination of the image and accentuates the features to be extracted [8].

## 4. Feature extraction

A major component in analyzing images involves data reduction which is accomplished by intelligently modifying the image from the lowest level of pixel data into higher level representations. From these higher level representations we can gather useful information; a process called feature extraction [9]. Extracting features by fixed blocs in the image has been considered to be sufficient as an ROI selection method in some medical applications where a large fraction of the image is covered by tissue of interest.

### 4.1. Measurements of texture features

The gray-level co-occurrence matrix (GLCM) is a matrix used to express the correlation of spatial location and gray-level distribution of an image [16]. From it, the local variation of gray levels on an image or sub-image can be statistically investigated and in turn, enable us to know the manner of change in gray level as a whole. In the current application, we used the following conditions to generate gray-level co-occurrence matrices.

- Direction: In general the gray-level co-occurrence matrices from 0, 45, 90, and 135 directions are used. Only the direction of 0 was used in the study.
- Distance: The length of 1-pixel was used.

Of the 14 original statistics developed by Haralick et al. (1973) for generating texture features based on co-occurrence probabilities, we chose the four most commonly used features for our evaluation.

Dissimilarity, uniformity, entropy, and contrast are often used in practice. The following four statistics will be used exclusively in this paper:

1. Dissimilarity =  $\sum_{i,j=1}^G C_{ij} |i-j|$ .
2. Uniformity =  $\sum_{i,j=1}^G C_{ij}^2$ .
3. Entropy =  $-\sum_{i,j=1}^G C_{ij} \lg C_{ij}$ .

$$4. \text{ Contrast} = \sum_{i,j=1}^G C_{ij} (i-j)^2.$$

Where,  $C_{ij}$  represents co-occurring probabilities stored inside GLCM.  $G$  represents number of grey level available.

Now, we introduce and study two ROI selection methods that give us better texture feature extraction when co-occurrence matrix method is used and hence a more accurate texture feature (see section 7).

#### 4.1.1. Extraction based on pixel wise segmentation approach.

For a more accurate feature extraction, apply extraction based pixel wise partitioning method as follows (see Figure.1):

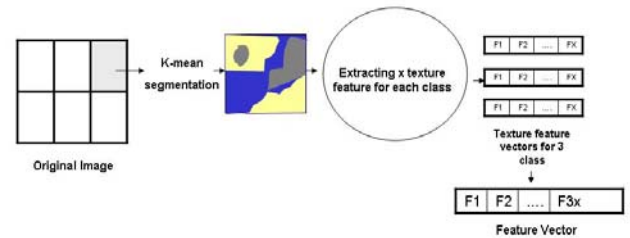
**Step1:** divide the entire image into SN non-overlapping sub-images  $SI = \{I_1, I_2, \dots, I_{SN}\}$ .

**Step2:** use the k-means algorithm [17] to cluster the sub-images ( $SI$ ) into several classes based on pixel intensity for each  $I_i$ ,  $i=1,2,\dots,SN$  independently.

**Step3:** for each cluster in  $I_i$ ,  $i=1,2,\dots,SN$ , construct a sub-image representing set of texture feature vectors  $F_K = \{f_1, f_2, \dots, f_X\}$ ,  $k=1,2,\dots,L$ ; where  $L$  is the number of classes each of which contains  $X$  texture features.

**Step4:** build the final set of texture features representing the overall image in the form of a single transaction of the final dataset (set of images),  $T = \{t_1, t_2, \dots, t_c\}$ , where  $c$  is the number of images,  $t_i$  is a vector of the size  $(SN \times L \times X)$ ,  $i=1, 2, \dots, c$ .

**Step 5:** for each  $t_i$ ,  $i=1,2,\dots,c$  add the class label of its image.



**Figure.1** Illustrate simple example describes feature extraction phase: where  $SN=6$  and  $L=3$ .

#### 4.1.2. Extraction method based on proposed ROI selection approach.

For a more accurate feature extraction and a further investigation of the localization, apply extraction based bloc wise partitioning method as follows (see Figure.2):

**Step1:** divide the entire image into SN non-overlapping sub-images  $SI = \{I_1, I_2, \dots, I_{SN}\}$ .

**Step2:** split each of these SN sub-image into other M blocs  $I_j = \{B_1, B_2, \dots, B_M\}$ ,  $j=1,2,\dots,SN$ .

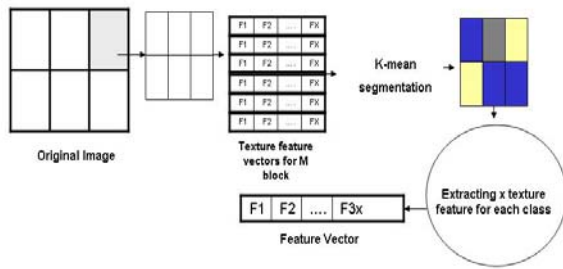
**Step3:** for each bloc  $B_i$ ,  $i=1,2,\dots,M$ , construct a bloc representing set of texture feature vectors.

**Step4:** use the k-means algorithm to cluster the feature vectors into several classes for each sub-image  $I$  independently.

**Step5:** for each cluster in  $I_i, i=1,2,\dots,SN$ , construct a sub-image representing set of texture feature vectors  $F_K = \{f_1, f_2, \dots, f_X\}, k=1,2,\dots,L$ ; where  $L$  is the number of classes each of which contains  $X$  texture features.

**Step6:** build the final set of texture features representing the overall image in the form of a single transaction of the final dataset (set of images),  $T=\{t_1, t_2, \dots, t_c\}$ , where  $c$  is the number of images,  $t_i$  is a vector of the size  $(SN \times L \times X), i=1, 2, \dots, c$ .

**Step7:** for each  $t_i, i=1,2,\dots,c$  add the class label of its image.



**Figure.2** Illustrate simple example for describe feature extraction phase: where  $SN=6, M=6$  and  $L=3$ .

## 5. Feature selection

The statistical discrimination methods are suitable not only for classification but also for characterization of differences between a reference group of patterns and the population under investigation.

For the image classification case, without class labels, Principle Component Analysis (PCA) can be applied, when some class labels are available, linear discriminant analysis (LDA) is applied to transform the features into the most discriminating Feature space [18-21].

Let us consider a set of  $N$  sample images  $\{x_1, x_2, \dots, x_N\}$  taking values in an  $n$ -dimensional image space, and assume that each image belongs to one of  $c$  classes. Let us also consider a linear transformation mapping the original  $n$ -dimensional image space into an  $m$ -dimensional feature space, where  $m < n$ . The new feature vectors  $Y$  are defined by the following linear transformation:

$$Y_k = W^T x_k, \quad k=1, N. \quad (1)$$

### 5.1. Fisher's Linear Discriminant Analysis (FLD)

In this method,  $W$  in (1) can be selected in such a way that the ratio of the between-class scatter and the within-class scatter is maximized. Let the between-class scatter matrix be defined as

$$S_b = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T,$$

and the within-class scatter matrix be defined as

$$S_w = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T,$$

where  $\mu_i$  is the mean image of class  $X_i$ , and  $N_i$  is the number of samples in class  $X_i$ . If  $S_w$  is nonsingular, the optimal projection  $W_{opt}$  is chosen as the matrix with orthonormal columns which maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples, i.e.,

$$W_{opt} = \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|}. \quad (2)$$

In this paper, we can not apply (FLD) directly to solve the recognition problem since the dimension of the sample space is typically larger than the number of samples in the training set. As a consequence,  $S_w$  is singular in this case.

**5.1.1. Fisherfaces.** Swets and Weng [22] proposed a two stage PCA+LDA method, also known as the Fisherfaces method, in which PCA is first used for dimension reduction so as to make  $S_w$  nonsingular before the application of LDA.

In this method an alternative to the criterion in (2) is applied. This method avoids this problem by projecting the image data set to a lower dimensional space so that the resulting within-class scatter matrix  $S_w$  is nonsingular. This is achieved by using PCA to reduce the dimension of the feature space to  $N-c$ , and then applying the standard FLD defined by (2) to reduce the dimension to  $(c-1)$  [20]. More formally,  $W_{opt}$  is given by

$$W_{opt}^T = W_{fld}^T W_{pca}^T,$$

where

$$W_{pca} = \arg \max_W |W^T S_T W|,$$

$$W_{fld} = \arg \max_W \frac{|W^T W_{pca}^T S_b W_{pca} W|}{|W^T W_{pca}^T S_w W_{pca} W|},$$

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T.$$

Note that the optimization for  $W_{opt}$  is performed over  $n \times (N-c)$  matrices with orthonormal columns, while the optimization for  $W_{fld}$  is performed over  $(N-c) \times m$  matrices

with orthonormal columns. In computing  $W_{opt}$ , we have thrown away only the smallest (c-1) principal components.

## 6. Measures for performance evaluation

We evaluated the performance of the proposed methods in terms of sensitivity, specificity and overall accuracy [6], Such that k-fold cross validation process is used.

K-fold cross validation is used for model testing and evaluation to determine how accurately a learning algorithm will be able to predict data that it was not trained on. In this method it is not important how the data is divided. Every data point appears in a test set exactly once, and appears in a training set k - 1 times. We can independently choose the size of the each test and the number of trials (Kohavi & Provost, 1998).

## 7. Experimental results

In our experiment a sample of 22% from the MIAS data set is selected randomly for model testing and evaluation. The sample is distributed among classes as follows:

Normal class (n=30), Microcalcification class (n=10), Circumscribed Masses class (n=7), Spiculated Masses class (=8), Ill-Defined Masses class (n=5), Architectural Distortion class (n=7) and Asymmetry class (n=4).

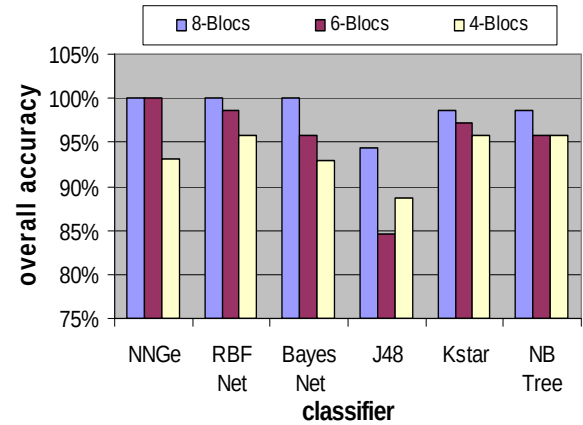
We evaluated the performance of the proposed methods described in sections 4.1.1 and 4.1.2 with standard classifiers. Using weka experimenter [23], Such that 10-fold cross-validation process is used.

Table.1 present the overall accuracy of textural feature set extracted by the extraction method based on bloc wise segmentation as a ROI selection approach (see section 4.1.2) for SN=6 and L=3. Several values of M were checked for the possibility of improving the results and increasing features localization. See table. 1.

**Table.1** Overall accuracy for the proposed method with M=4, 6 and 8.

Classifier	Overall accuracy		
	4-blocs	6-blocs	8-blocs
NNGe	93.10%	100%	100%
RBF net	95.77%	98.59%	100%
Bayes net	92.95%	95.77%	100%
J48	88.73%	84.5%	94.37%
Kstar	95.77%	97.18%	98.59%
NB tree	95.77%	95.77%	98.59%

From results in the above table, we can notice that, using a certain feature selection algorithm (Fisherfaces), classifiers accuracy is dependent on number of blocs regardless of classification algorithms used. The overall accuracy with different bloc size is presented in Figure.3.



**Figure.3** Success rates of classifiers with decreasing of bloc size.

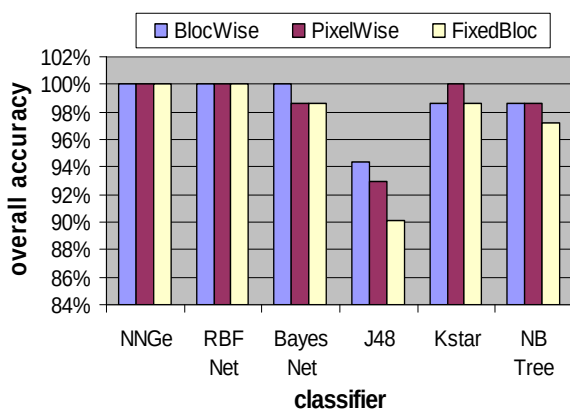
Table.2 illustrates comparison between the following textural feature sets based on different ROI selection methods:

1. Textural feature sets based on fixed bloc segmentation (simply divide the whole image into 6 non-overlapping sub-images then split each of these 6 sub-images into other 8 non-overlapping blocs).
2. Textural feature sets based on pixel intensity segmentation described in section 4.1.1 with SN=6 and L=3.
3. Textural feature sets based on our bloc wise segmentation described in section 4.1.2 with SN=6, M=8 and L=3.

**Table.2** The classification rate comparison.

Classifier	ROI	Sensitivity	Specificity	Overall accuracy
NNGe	Fixed bloc	100%	100%	100%
RBF net		100%	100%	100%
Bayes net		96.77%	100%	98.59%
J48		85.71%	95.35%	90.14%
Kstar		96.77%	100%	98.59%
NB tree		93.75%	100%	97.18%
NNGe	Pixel wise	100%	100%	100%
RBF net		100%	100%	100%
Bayes net		96.77%	100%	98.59%
J48		88.24%	97.62%	92.96%
Kstar		100%	100%	100%
NB tree		96.77%	100%	98.59%
NNGe	Bloc wise	100%	100%	100%
RBF net		100%	100%	100%
Bayes net		100%	100%	100%
J48		90.9%	97.62%	94.37%
Kstar		96.77%	100%	98.59%
NB tree		96.77%	100%	98.59%

From results shown in the table.2, we can notice that, using Fisherfaces feature selection algorithm, our proposed extraction methods gave better accuracy than extraction method based on fixed bloc partitioning as a ROI selection method, which means a more accurate extraction of textural feature and hence a more efficient image representation. The classifier NNGe and RBF-Network gave the highest performance when compared to the others which implies that NNGe and RBF-Network are much more suitable for such application. Figure.4. presents a comparison between classification results.



**Figure.4** Success rates of classifiers with our proposed methods.

## 8. Conclusions

In this work, four texture features derived from the co-occurrence matrix was used as a part of developing CAD system for early detection of breast cancer.

This paper present textural extraction method based on three different ROI selection methods for obtaining efficient image representation. We proposed an ROI selection method based on bloc wise segmentation designed to obtain more accurate extraction of textural feature from medical images. It was shown that this approach improves the localization of texture features and produce general texture features. A fisherfaces feature selection algorithm is used for extracting discriminative information from extracted feature of medical images to be used as inputs to our classification system.

Our experiments confirm the importance of our proposed ROI selection approaches as a primary step of feature extraction in the final decision of the CAD.

## 9. Reference

[1] Tristan Glatard, Johan Montagnat, and Isabelle E.Magnin, "Texture based medical image indexing and retrieval: application to cardiac imaging", Proceedings of the 6th ACM SIGMM

international workshop on Multimedia information retrieval, October 15-16, 2004, New York, NY, USA.

[2] Chabat F, Yang G-Z, and Hansell DM, "Obstructive lung diseases: texture classification for differentiation at CT", *Radiology*2003;228:871-7.

[3] Kovalev VA, Petrou M, and Bondar YS, "Texture anisotropy in 3D images", *IEEE Trans Image Process* 1999; 8:346-60.

[4] Herlidou-Me'me S, Constans JM, and Carsin B, "MRI texture analysis on texture test objects, normal brain and intracranial tumours", *Magn Reson Imaging* 2003;21:989-93.

[5] Mahmoud-Ghoneim D, Toussaint G, Constans JM, and De Certaines JD, "Three dimensional texture analysis in MRI: apreliminary evaluation in gliomas", *Magn Reson Imaging* 2003; 21: 983-7.

[6] Du-Yih Tsai and Katsuyuki Kojima, "Measurements of texture features of medical images and its application to computer-aided diagnosis in cardiomyopathy", *Measurement* 37 (2005) 284–292.

[7] H.S.Sheshadri and A. Kandaswamy, "Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms", *Computerized Medical Imaging and Graphics* 31 (2007) 46-48.

[8] Maria-Luiza Antonie, Osmar R. Zaiane, and Alexandru Coman, "Application of data mining techniques for medical image classification", In *Proc. of Second Intl. Workshop on Multimedia Data Mining (MDM/KDD'2001)* in conjunction with Seventh ACM SIGKDD, pages 94-101, San Francisco, USA, 2001.

[9] V.S.Kodogiannisa, M.Boulougourab, E.Wadgea and J.N.Lygyourasc, "The usage of soft-computing methodologies in interpreting capsule endoscopy", *Engineering Applications of Artificial Intelligence* 20 (2007) 539-553.

[10] G.Castellano, L.Bonilha, L.M.Li, and F.Cendes, "Texture analysis of medical images", *Clinical Radiology* (2004) 59, 1061-1069.

[11] V.S.Kodogiannisa, M.Boulougourab, J.N.Lygyourasc, and I.Petrouniasd, "A neuro-fuzzy-based system for detecting abnormal patterns in wireless-capsule endoscopic images", *Neurocomputing* 70 (2007) 704-717.

[12] Coggins, J. M, "A Framework for Texture Analysis Based on Spatial Filtering", Ph.D. Thesis, Computer Science Department, Michigan State University, East Lansing, Michigan, 1982.

[13] Carlos E. Thomaz, Nelson A. O. Aguiar, Sergio H. A. Oliveira, Fabio L. S. Duran, Geraldo F. Busatto, Duncan Fyfe Gillies, and Daniel Rueckert, "Extracting Discriminative Information from Medical Images: A Multivariate Linear Approach", *SIBGRAPI* 2006: 113-120.

[14] I.Cohen, Q.Tian, X.Zhou, and T.S.Huang, "Feature Selection and Dimensionality Reduction Using Principle Feature Analysis", submitted to the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI-2000) Workshop on Fusion of Domain Knowledge with Data For Decision Support, Stanford University, Stanford, CA, June 30, 2000.

[15] <http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html>

- [16] David A.Clausi, and Yongping Zhao, "Rapid extraction of image texture by co-occurrence using ahybrid data structure", Computers and Geosciences 28 (2002) 763-774.
- [17] Jose Marroquin, and Federico Girosi, "Some extensions of the K-means algorithm for image segmentation and pattern classification", Technical report, Massachusetts Institute of Technology, January 1993.
- [18] K.Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, Boston, 1990.
- [19] C.E.Thomaz, J.P.Boardman , S.Counsell , D.L.G.Hill, J.V.Hajnal, A.D.Edwards, M.A.Rutherford, D.F.Gillies, and D. Rueckert, "A multivariate statistical analysis of the developing human brain in preterm infants", Image and Vision Computing 25 (2007) 981–994.
- [20] Peter N.Belhumeur, Joao P.Hespanha, and David J.Kriegman. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7, JULY 1997.
- [21] P.A.Devijver, and J.Kittler, "Pattern Classification: A Statistical Approach", Prentice-Hall, Englewood Cliffs, N. J., 1982.
- [22] D.L.Swets, and J.Weng, "Using discriminant eigenfeatures for image retrieval", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 831-836, August 1996.
- [23] <http://www.cs.waikato.ac.nz/ml/weka/>.

## Ant Colony Clustering by Expert Ants

Zahra Sadeghi<sup>1</sup>, Mohammad Teshnehlab<sup>2</sup>

<sup>1</sup> Computer Engineering Department, Science and Research Branch, Islamic Azad University - Member of Young Researchers Club

<sup>2</sup> Faculty member, Electronic Engineering Department, K.N.Toosi University, Tehran, Iran

<sup>1</sup>z.sadeghi@seiau.ir, <sup>2</sup>teshnehlab@eetd.kntu.ac.ir

### Abstract

*In this article a new ant clustering algorithm based on case based reasoning (CBR) is presented. Every ant has a case base which is updated iteratively by the process of CBR. The ant which is successful in dropping an item becomes an expert and can use its knowledge for future picked up items. Also expert ants are capable of cooperating to share their knowledge for even better clustering. Our simulation results demonstrated better performance than previous approaches.*

**Keywords:** ant based clustering, ant colony optimization

### 1. Introduction

Machine learning is divided into two primary sub-fields: supervised learning and unsupervised learning. Within the category of unsupervised learning, one of the primary tools is clustering [1]. Clustering means the act of partitioning an unlabeled dataset into groups of similar objects. Each group, called a 'cluster', consists of objects that are similar between themselves and dissimilar to objects of other groups [2]. Swarm Intelligence emerged out of social insect collective behavior shows many interesting properties such as flexibility, robustness, decentralization and self organization [3]. The swarm intelligence clustering models and algorithms have advantages in many aspects, such as no need of priori information, self-organization, flexibility, robustness and decentralization [4]. Ant-based clustering is a heuristic clustering method that draws its inspiration from the behavior of ants in nature [5]. Some of the proposed ant based clustering algorithm can be found in [6-11].

In the present article, we have introduced a clustering algorithm based on expert ants. After each successful dropping, the value and the coordination of

the object being dropped is saved in the ant's case base. In the beginning, the case bases of all ants are empty. We have introduced two types of ants: normal ants, which have empty case base and have not dropped an object yet, and expert ants, who were able to drop at least one object. Whenever an expert ant picks up a new load, it searches its case base for the best match and if such a case is found, it will be retrieved and then the ant immediately drops down its load in the nearest neighborhood of the grid coordination of the stored case. Otherwise it uses the drop down function. The case base of all expert ants updates each time they drop down an object. Also expert ants are able to cooperate with each other to find the best place for dropping their load. The results of our approach demonstrated that it can speed up the process of clustering and prevents from producing too many small clusters which is one of the main drawbacks of ant based clustering.

### 2. The general principle of ant clustering

Ant based clustering is a distributed process. The pioneers of this work are Deneubourg *et al.* [12]. Their model is known as basic model (BM) in which ants are simulated by simple agents that randomly move in an environment which is a square grid. Initially, each data object that represents a multi-dimensional pattern is randomly distributed over the 2-D space. Data items that are scattered within this environment can be picked up, transported and dropped by the agents in a probabilistic way. The picking and dropping operations are influenced by the similarity and density of the data items within the ant's local neighborhood. Deneubourg *et al.* proposed a computational model for spatial sorting. Their model is based on the following probability of picking and dropping functions:



$$Pp = \left(\frac{Kp}{Kp + f}\right)^2. \quad (1)$$

$$Pd = \left(\frac{f}{Kd + f}\right)^2. \quad (2)$$

Where  $Pp$  is the probability of picking,  $Pd$  is the probability of dropping, and  $Kp$  and  $Kd$  are constants.  $f$  is the perceived fraction of items in the neighborhood of the ant. The probability of picking ( $Pp$ ) is increased if a data object is surrounded by dissimilar data, or when there's no data in its neighborhood. Also, ants tend to drop data in the vicinity of similar ones, so the probability of dropping ( $Pd$ ) is increased if ants are surrounded with similar data. Deneubourg *et al.*'s model was later extended by Lumer *et al.* [13] to allow its application to exploratory data analysis. They modified the dropping probability functions as the following:

$$Pd(i) = \begin{cases} 2f(i) & \text{if } f(i) < Kd \\ 1 & \text{if } f(i) \geq Kd. \end{cases} \quad (3)$$

Where  $Pd$  is the same as before, i.e., the probability of dropping and picking. In addition, they introduced the density function as the following:

$$f(i) = \max(0, \frac{1}{\sigma^2} \sum_{j \in L} (1 - \frac{\delta(i, j)}{\alpha})). \quad (4)$$

Where  $\delta(i, j) \in [0, 1]$  is the dissimilarity function between two data object  $i$  and  $j$ .  $\alpha \in [0, 1]$  is the scaling parameter that permits further adjustments of resulting values.  $\sigma^2$  is the size of local neighborhood  $L$  around the ant's current position. The basic ant clustering algorithm is like the one summarized as bellow [14]:

Let  $R \in [0, 1]$  be a random number drawn for each use, the basic algorithm is then given by:

Randomly scatter data items on a square grid.

For a number of steps or until a criterion is met, repeat for each ant:

- If the ant is unladen and placed on location  $l$  occupied by object  $o$ , the object is picked up if

$$P_{pick\_up}(o) \geq R.$$

- If the ant is carrying object  $o$  and placed on an empty location  $l$ , the object is dropped if

$$P_{drop\_down}(o) \geq R.$$

- Move the ant to a new randomly selected (neighboring) location.

### 3. Case Based Reasoning (CBR)

CBR is a familiar process that is a normal, intuitive method of decision-making for humans in everyday life. It is a process of considering past cases and arriving at decisions on comparison between the current situation and the old cases [15].

In CBR terminology, a case usually denotes a problem situation. A previously experienced situation, which has been captured and learned in a way that it can be reused in the solving of future problems, is referred to as a past case, previous case, stored case, or retained case. Correspondingly, a new case or unsolved case is the description of a new problem to be solved [16].

CBR is done by considering the similarity between the current situation and old seen situations. The basic idea of CBR for solving a new problem is as follows:

- Remembering a previous similar situation.
- Comparing the new problem to the old solved problems
- Reusing information and knowledge of that situation.

At the highest level of generality, a general CBR cycle may be described by the following four processes [16]:

1. RETRIEVE the most similar case or cases.

The goal of this step is to retrieve old cases stored in the case library [17]. Retrieving is the selection of the most similar cases (source cases) to the new problem specifications (target case) from the library of cases.

2. REUSE the information and knowledge in that case to solve the problem.

3. REVISE the proposed solution.

Because a new case may not exactly match the old one, the old knowledge may often need to be fixed to fit the new one [17].

4. RETAIN the parts of this experience likely to be useful for future problem solving.

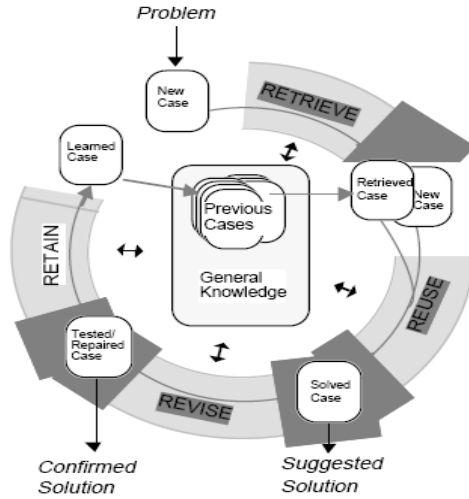
In "Fig. 1", this cycle is illustrated [16].

### 4. The Framework of the Algorithm

Our ant clustering algorithm uses the same probability and density function as the one proposed by Lumer *et al.* [13], but instead of consuming a long time for finding an item to pick, we have kept the location of all available items on the grid in an array. After an ant drops its load, it immediately picks up one item from the list of available items on the grid [18] and the ant jumps to the location of the picked up item.



We have considered a grid with  $s \times s$  cells. The number of cells must be greater than the number of all objects. All ants move with different speeds on the grid. We have considered a maximum speed for all ants which is set as  $\text{max\_speed} = s/3$ . The step of movement of each ant is generated randomly between 1 and  $\text{max\_speed}$  before each movement. We decrease the  $\text{max\_speed}$  by one every 200 iterations. The details of our proposed method are discussed in the following subsections.



**Figure 1.** The CBR cycle [16].

#### 4.1 Expert Ants

In the beginning of clustering, all ants are identical, but only with different loads. The memory of all is empty and no one has any kind of information about the search space and the location of items in the grid. We call these ants as normal ants. They move randomly in grid and try to pick up loads using equation (3) and drop their load according to equation (4). An ant becomes an expert by memorizing the location of at least one dropped data. The method of reasoning of these expert ants is based on case based reasoning. Each normal laden ant when for the first time wants to drop an item, needs to use the probability of dropping. It obtains a new experience with each dropping and it keeps the value and location of that data in a repository called case base. After dropping at least one data, it is called an expert ant. When an expert ant is laden, it first refers to its case base, looking for a similar case. If it found such a case, it decides to put its load somewhere next to the location of most similar data it has dropped before. In this way

there is no need to compute the probability of dropping and it causes the gathering of similar data in a single area.

#### 4.2 Case and Case Base of Expert Ants

Each ant has a private case base for keeping the different cases it has seen. A case consists of value of attributes and the place they have been dropped. The place of dropping, is a two dimensional matrix (because the search space is a grid). The value of attributes of a data is a multidimensional matrix, in which the number of dimensions depends on the dimensions of data. There is also a variable for each ant which specifies the number of the existing cases in the case base.

#### 4.3 Retrieve and Revise

Whenever an expert ant wants to put its load, it searches for a case in its case base, and compares the value of all cases with its current load to find the most similar case. For this, each ant first collects all the cases having a distance less than a threshold from its load in an array, then it sorts the array ascendingly, and finally selects the first element of the array which has the lowest distance value. But, before retrieving this case, it first checks whether the case is a valid one, or not. Because ants pick up and drop down the data iteratively, it is so probable that the value of case bases become invalid during clustering process. A case is valid for an ant, only if its current location on the grid is the same as the one recorded in the case base of that ant. If the first element of the array was invalid, the second element is checked. This process repeats until a case is retrieved or no other case remains in case base. If a case was retrieved it can be used. It is possible that whenever an ant is looking for a case, it faces a case with identical attribute values with its current load. Under this condition, the case must be revised and then updated. For this purpose, the dropping probability must be computed. When a suitable location for dropping that item is found, the old stored location in case base is replaced with a new one.

#### 4.4 Cooperation between Ants

According to the number of successful droppings, each ant gains some kind of expertise so that the expertise of each ant is different from that of others. In this way we have a multi expert- multi agent system. Whenever a laden ant fails in dropping its load, and an

unladen ant fails for picking up an item, they can sense their neighborhood for getting help from other ants. So, a laden ant can put its load on an unladen ant, and an unladen ant can get load from a laden ant. If both ants were laden, they only exchange their loads. Nothing will happen if both were unladen. If there were more than one neighbor ant, one is chosen randomly. All the process of getting load from other ants and exchanging of loads happen with probability of 0.5. This kind of cooperation can prevent ants from moving aimlessly and it can also increase the speed of convergence.

#### 4.5 The algorithm of clustering by expert ants

The details of our algorithm are as follows:

//Initialization part

1. Spread the data and ants randomly on the grid and give all ants a random load.

//main part

1. **While** iter < max\_iter
2. **For** i=1 to #ants
3. a = Choose a random ant without replacement
4. step = Generate a random speed from interval (1,max\_speed)
5. new\_location=Move (a, step)
6. **if** ant a is unladen and there is an object o in location l then pick\_done=try\_pick(o)
7. **if** pick\_done=0 then try to find a laden expert ant in the neighborhood of ant a and take the load of it with probability of .5
8. **end if**
9. **end if**
10. **if** ant a is laden then search its case base for a case which has the similarity of more than .9 with the current load,
11. **if** such a case was found then
12. // retrieve  
Drop the load in a neighborhood of 2 from the place of that case
13. **end if**
14. **if** a case was found with the same value with the current load then
15. // revise  
Find a place for dropping it and change the place part of the case with the new found location
16. **end if**
17. **if** no similar case was found and the current location of it is empty then
18. drop\_done=try\_drop(l)
19. **end if**

20. **if** drop\_done=1
21. // retain  
Make a new case and save it in case base
22. **end if**
23. **if** drop\_done=0
24. find an unladen expert ant in the neighborhood of 2 and give the load of ant a to it with probability of .5
25. **if** no unladen expert ant was found, find a laden expert ant in the neighborhood of 2 and exchange the load of ant a with the load of the neighboring ant
26. **end if**
27. **End For**
28. **End while**

**try\_pick(o)**

1. r= a random number from interval (0,1)
2. p=compute probability function (1)
3. **if** p>r then
4. pick up the object o
5. pick\_done=1
6. **else**
7. pick\_done=0
8. **end if**

**try\_drop(l)**

1. r= a random number from interval (0,1)
2. d=compute probability function (3)
3. **if** d>r then
4. drop down load l
5. drop\_done=1
6. **else**
7. drop\_done=0
8. **end if**

#### 5. Experimental Results and Comparison

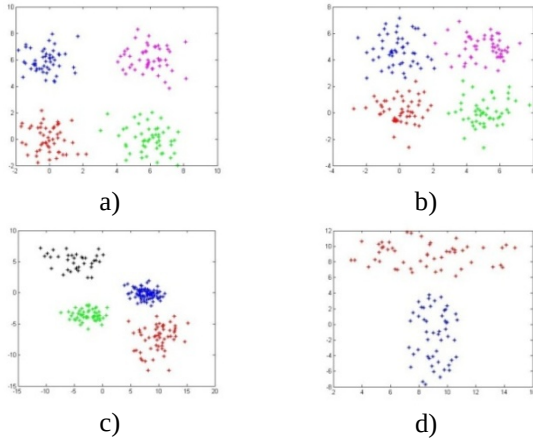
The algorithm is applied to six datasets. The datasets are summarized in Table 1 and all except for Iris and Wine datasets are shown visually in “Fig. 2”. Iris and Wine datasets are not illustrated because their dimensionality is more than 3.

We have compared our algorithm with k-means algorithm, which works as follows: Starting from a partitioning of the data into k randomly generated clusters, the k-means algorithm repeatedly (i) computes the current cluster centers (i.e., the average vector of each cluster in data-space) and (ii) reassigns each data item to the cluster whose center is closest to it. It terminates when no more reassignments take place. As k-means can sometimes generate empty clusters, these are identified in each iteration and are randomly

reinitialized. This enforcement of the correct number of clusters can prevent convergence, and we therefore set the maximum number of iterations to 100. To reduce suboptimal solutions k-means is run repeatedly (20 times) using random initialization, and only the best result in terms of intra-cluster variance is returned. [7].

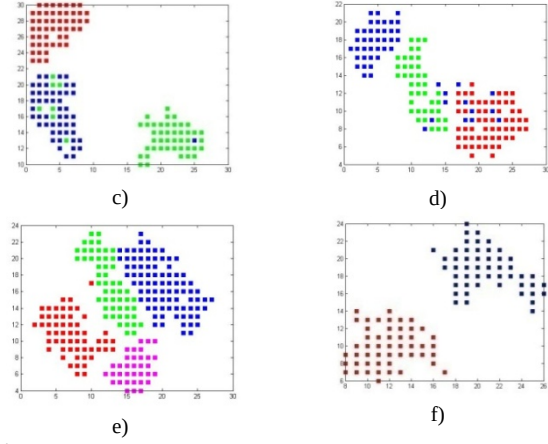
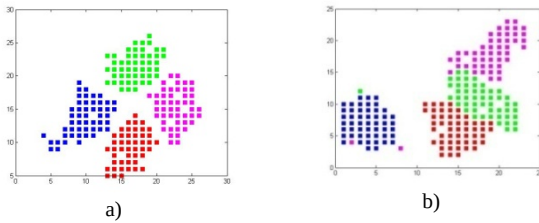
**Table 1.** Summary of the used data sets. k gives the number of clusters, size shows the number of members of a cluster, and dim is the dimensionality,.

Name-k-size	dim	Source
Square6- 4*50	2	N([0,0],[2,2]),([0,6],[2,2]),([6,0],[2,2]),([6,6],[2,2])
Square5- 4*50	2	N([0,0],[2,2]),([0,5],[2,2]),([5,0],[2,2]),([5,5],[2,2])
Iris- 3*50	4	UCI
Wine-3-(59,71, 48)	13	UCI
D2C2-2*50	2	Artificial data
D2C4-4-(35,57,84,52)	2	Artificial data



**Figure 2.** Benchmark data. a) Square6, b) Square5, c) D2C4, d) D2C2

The results of applying k-means algorithm and ant clustering by normal ants are compared to the results of our algorithm in Table 2 and the best result out of a few runs is selected and is illustrated in “Fig. 3”.



**Figure 3.** Best visual results of benchmark data. a) Square6, b) Square5, c) Iris, d) Wine, e) D2C4, f) D2C2

A crucial matter in ant based clustering is the requirement for spatial separation between clusters. Fig.3 shows that there is a good boundary between clusters that are generated by our algorithm, which makes the cluster retrieval process to be done with less ambiguity. The results of Table 2 shows the number of clusters that our algorithm can find is nearer to the true number of clusters than normal ants, the average error of clustering which is computed based on the number of objects that were assigned to incorrect clusters is decreased in our approach and the number of iteration needed for our proposed method is much less than that of normal ants.

**Table 2.** Test results on benchmark data for K-means, normal ants and expert ants.

Square6	K-means	Normal Ants	Expert Ants
Min Error	0	0	0
Max Error	10	7	6
Average Error	2.9	2.8	2.77
Average Number of Detected Clusters	4	4.4	4.05
Maximum Iterations	100	50000	10000
Square5	K-means	Normal Ants	Expert Ants
Min Error	0	2	2
Max Error	12	15	16
Average Error	3.9	6.7	5.3
Average Number of Detected Clusters	4	4.37	4.31
Maximum Iterations	100	50000	10000
Iris	K-means	Normal Ants	Expert Ants
Min Error	4	8	8
Max Error	17	16	15

Average Error	8.8	9.85	8.72
Average Number of Detected Clusters	3	3.12	2.91
Maximum Iterations	100	500000	100000
<b>Wine</b>	<b>K-means</b>	<b>Normal Ants</b>	<b>Expert Ants</b>
Min Error	4	6	4
Max Error	8	10	9
Average Error	8.13	8.21	7.75
Average Number of Detected Clusters	3	3.28	3.12
Maximum Iterations	100	500000	100000
<b>D2C2</b>	<b>K-means</b>	<b>Normal Ants</b>	<b>Expert Ants</b>
Min Error	0	0	0
Max Error	9	4	3
Average Error	2.7	1.02	0.8
Average Number of Detected Clusters	2	2.1	2.21
Maximum Iterations	100	50000	10000
<b>D2C4</b>	<b>K-means</b>	<b>Normal Ants</b>	<b>Expert Ants</b>
Min Error	0	0	0
Max Error	9	8	7
Average Error	6.8	2.7	2.5
Average Number of Detected Clusters	4	5.85	4.17
Maximum Iterations	100	50000	10000

## 6. Conclusion

In this article we presented a novel method for ant clustering which uses case based reasoning. We have introduced “expert ants” who are able to decide about the best place for dropping their load using their case bases. The case base of each expert ant contains different information about promising places for dropping loads. We have added a mechanism for taking advantage of the knowledge of other ants. Those ants that were unsuccessful in dropping their loads can give their loads to other unladen expert ants in their neighborhood, and those that were unsuccessful in picking up an object can take the loads of other laden ants in their neighborhood. If two neighboring ants both were laden they can exchange their loads. Our proposed method has shown better results in terms of speed, correctness and compactness of clusters.

## 7. References

- [1] G. Fung, “A Comprehensive Overview of Basic Clustering Algorithms”, June 22, 2001
- [2] A. Abraham, S. Das, S. Roy, “Swarm Intelligence Algorithms for Data Clustering”, In *Soft Computing for Knowledge Discovery and Data Mining*. Springer Science, 2007.
- [3] Q. Li, Z. Shi, J. Shi, and Z. Shi, “Swarm Intelligence Clustering Algorithm Based on Attractor”, Springer Berlin / Heidelberg, 2005, pp.496-504
- [4] W. bin, S. Zhongzhi, “A clustering algorithm based on swarm intelligence”, *Proc. of the Int. Conf. on Info-tech and Info-net*, Beijing, China, 2001, pp. 58-66
- [5] E. Bonabeau, M. Dorigo, and G. Theraulaz. “Swarm Intelligence – From Natural to Artificial Systems”, Oxford University Press, New York, NY, 1999.
- [6] J. Handl, B. Meyer, “Improved ant-based clustering and sorting in a document retrieval interface”, In J. J. Merelo, P. Adamidis, & H.-G. Beyer (Eds.), *Lecture notes in computer science: Vol. 2439*, 2002.
- [7] J. Handl, J. Knowles, M. Dorigo, “Ant-based clustering and topographic mapping”, *Artificial Life*, 12(1), 35–61, 2006.
- [8] A. L. Vizine, L. N. de Castro, R. R. Gudwin, “Text document classification using swarm intelligence”, In *International conference on the integration of knowledge intensive multi-agent systems*, Piscataway: IEEE Press. problem solving from nature—PPSN VII, 2005, pp. 134–139.
- [9] A. L. Vizine, L. N. de Castro, E. R. Hruschka, R. R. Gudwin, “Towards improving clustering ants: an adaptive ant clustering algorithm” *Informatica*, 29, 2005, pp. 143–154.
- [10] V. Ramos, J. Merelo, “Self-organized stigmergic document maps: environments as a mechanism for context learning”, In *Proc. of the first Spanish conf. on evolutionary and bio-inspired algorithms*, 2002, pp. 284–293.
- [11] S. Schockaert, M. D. Cock, C. Cornelis, E. E. Kerre, “Efficient clustering with fuzzy ants”, In D. Ruan, P. D’hondt, M. D. Cock, M. Nachtegaele, & E. E. Kerre (Eds.), *Applied computational intelligence, proceedings of the 6th international FLINS conference*, 2004, pp. 195–200.
- [12] J.-L. Deneubourg, S. Goss, N. Franks, C. Detrain, and L. Chretien “The Dynamics of Collective Sorting: Robot-Like Ant and Ant-Like Robot”, In J.-A. Meyer & S. Wilson (Eds.), *from animals to animats: proc. of the first int. conf. on simulation of adaptive behavior*, Cambridge: MIT Press, 1991, pp. 356-365.
- [13] E. Lumer, B. Faieta, “Diversity and Adaptation in Populations of Clustering Ants”, Cambridge: MIT Press, In D. Cliff, P. Husbands, J.-A. Meyer, & S. W. Wilson (Eds.), *from animals to animats: proceedings of the third international conference on simulation of adaptive behavior*, 1994, pp. 501-508.
- [14] E. Magnus, H. Pedersen, “Ant Colony Clustering & Sorting”, 2003.

- [15] J. Dahlgren, A. Khattak, P. McDonough, I. Banerjee, Ph. Orrick, A. Sharafsaleh, "Developing Case-Based Reasoning and Expert Systems and Incorporating New Material", 2004.
- [16] A. Aamodt, E. Plaza, "Case-Based Reasoning: Foundational Issues", Methodological Variations Methodological Variations and System Approaches. AI Communications, Vol. 17(1), 1994.
- [17] Mu-Jung Huang, Mu-Yen Chen, Show-Chin Lee, "Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis", 2007.
- [18] J. Handl, B. Meyer, "Improved Ant-Based Clustering and Sorting in a Document Retrieval Interface", In J.J. Merelo, J.L.F. Villacañas, H.G.Beyer, P. Adamis Eds. Proceedings of the PPSN VII – 7th Int. Conf. on Parallel Problem Solving from Nature, Granada, Spain, Lecture Notes in Computer Science 2439, Springer-Verlag, Berlin, 2002, pp. 913-923.

## Detection of faulty products using data mining

M. A. Karim<sup>a\*</sup>, G. Russ<sup>b</sup>, and A. Islam<sup>c</sup>

<sup>a</sup>School of Engineering System  
Queensland University of technology, QLD 4001

<sup>b</sup>Faculty of Informatics

University of Magdeburg, Germany

<sup>c</sup>Build and Integration Engineering, Nokia Devices

Hoimotie 19, 00380 Helsinki, Finland

### Abstract

*The manufacturing process is complex due to the large number of processes, diverse equipment set and nonlinear process flows. Manufacturers constantly face yield and quality problems as they constantly redesign their processes for the rapid introduction of new products and adoption of new process technologies. Solving product yield and quality problems in a manufacturing process is becoming increasingly difficult. There are various types of failures and their causes have complex multi-factor interrelationships. High innovation speed forced today's manufacturers to find failure causes quickly by examining the historical manufacturing data. Data mining offers tools for quick discovery of relationships, patterns, and knowledge in large databases. This has been applied to many fields such as biological technology, financial analysis, medical information, etc. Application of data mining to manufacturing is relatively limited mainly because of complexity of manufacturing data. Growing self-organizing map (GSOM) algorithm has been proven to be an efficient algorithm to analyze unsupervised DNA data. However, it produced unsatisfactory clustering when used on some manufacturing data. Moreover, there was no benchmark to monitor improvement in clustering. In this study a method has been proposed to evaluate quality of the clusters produced by GSOM and to remove insignificant variables from the dataset. With the proposed modifications, significant improvement in unsupervised clustering was achieved with complex manufacturing data. Results show that*

*the proposed method is able to effectively differentiate good and faulty products.*

### 1. Introduction

Manufacturing databases usually comprise of process control, process step, and quality control data. In many applications data are automatically generated by sensors and therefore the number of inputs can be very large. This large volume of data coupled with quicker time to market expectations is making finding and resolving problems quickly an overwhelming task. The analysis of such a large volume of data, interpreting results, and implementing design improvements is computationally intensive and time-consuming. A high priority goal for today's manufacturing is finding the most probable causative factor(s) that discriminate between low yield and high yield products by quickly examining the historical manufacturing data.

Engineers often perform a series of experiments to identify and resolve potential problems with the design. Traditional techniques of dealing with quality problems are the use of statistical process control (SPC) and design of experiments (DOE). However, these techniques fail to extract underlying features from complex data [1]. Moreover, these methods are highly time-consuming. In order to decrease design cycles and the time-to-market new products, it is important to have a method for analyzing manufacturing data quickly and efficiently, predicting the effects of design changes and determining the best design parameters. Effective product data management has a significant influence on product quality improvement [2, 3]. However, traditional techniques

---

\* Corresponding Author: (617) 3138 1516 , Email: [azharul.karim@qut.edu.au](mailto:azharul.karim@qut.edu.au)

are incapable of dealing with large volumes of manufacturing data [4].

Advances in the data mining techniques enhanced automatic knowledge discovery as well as extraction of useful information from large volumes of data. Literature shows that the dynamic GSOM is a powerful tool for unsupervised data analysis [5]. In this study, a modified GSOM is used to detect faulty products from manufacturing data. Several modifications are proposed in order to reduce noisy data and to improve clustering quality.

## 2. Unsupervised Clustering Method

Cluster analysis is an unsupervised data mining technique that processes data that do not have known class labels, or where the analyst opts not to use them. Data within a cluster are highly similar to one another, but are very dissimilar to data in other clusters. Each cluster can be treated as a class of data, on which class or concept characterisation and discrimination can be performed.

The neural network approach to data mining has gained popularity due to the facts that most real world data have complex decision boundaries and contain a substantial amount of noise and the neural networks are robust against noise. A popular and well-accepted self-organising method of neural network analysis is self-organising maps (SOMs) [6]. A SOM with two-dimensional constraints can provide a visualisation of high-dimensional data by projecting it into a two dimensional network. However, the two-dimensional SOMs are rectangular grids of neurons that have predefined sizes and are specified by widths and heights. This can potentially hinder the real representation of input space topology [7].

Consequently, several modifications to SOM have been proposed to overcome the problem of predefined grids. One of the most recent variants, called the growing self-organising map (GSOM) was proposed to let the algorithm determine the size of the feature map [7-9]. A significant difference between these two algorithms is that SOM is not able to grow but GSOM grows according to its own growing criterion. Hsu et al. [9], proposed a hierarchical clustering algorithm using multiple layers of GSOM that automatically suggests an appropriate number of clusters that are present in the data.

### 2.1. GSOM algorithm

GSOM is an unsupervised learning algorithm and the basic idea is to:

1. Represent high-dimensional data in a low-dimensional form without losing any of the important features of the data.
2. Organise data on the basis of similarity by grouping objects that are geometrically close to each other.

This makes GSOM popular for visualizing high-dimensional data. Though computationally more expensive than other unsupervised methods, it has proved to be extremely useful with high-dimensional and complex input data.

A GSOM consists of neurons that are usually arranged in a two-dimensional grid with rectangular or hexagonal topology. Each neuron is represented by an n-dimensional weight vector  $m_i = [m_{i1}, m_{i2}, \dots, m_{in}]$ , where, n is equal to the respective dimension of the input vectors. The neurons are connected to adjacent neurons by means of neighbourhood relations.

The GSOM utilizes a competitive learning method [9]. When a training sample is given to the network, its distance to all weight vectors is computed. The best-matching unit (BMU) or winning neuron is the unit whose weight vector possesses greatest similarity with respect to the input sample  $\mathbf{x}$ . The distance measure used to define the similarity is typically Euclidean distance. Thus for each input vector  $\mathbf{x}$ , there exists a winning neuron  $m_b$  with minimum Euclidean distance to input  $\mathbf{x}$ , such that

$$\|\mathbf{x} - m_b\| = \min_i \|\mathbf{x} - m_i\| \quad (1)$$

This is referred to as the competitive learning part of the learning process, since all neurons compete to be the winner. Having found the BMU, the weight vectors of the GSOM are updated. The weight vectors of the BMU and its topological neighbours are moved closer to the input vector from the input space. This phenomenon is illustrated in Figure 1. The magnitude of the change decreases with time and is smaller for neurons physically far away from the BMU.

The update rule for changing the respective weight vectors of unit  $i$  is:

$$m_i(t+1) = m_i(t) + \alpha(t)h_{bi}(t)[x(t) - m_i(t)] \quad (2)$$

where,  $t$  denotes the time step,  $x(t)$  is the input vector chosen from the input data set at time  $t$ ,  $\alpha(t)$  is a monotonically decreasing learning rate, and  $h_{bi}(t)$  the

neighbourhood kernel around the winner unit at time  $t$ . The neighbourhood kernel defines the region of influence that the input sample has on the GSOM. This is called cooperative learning.

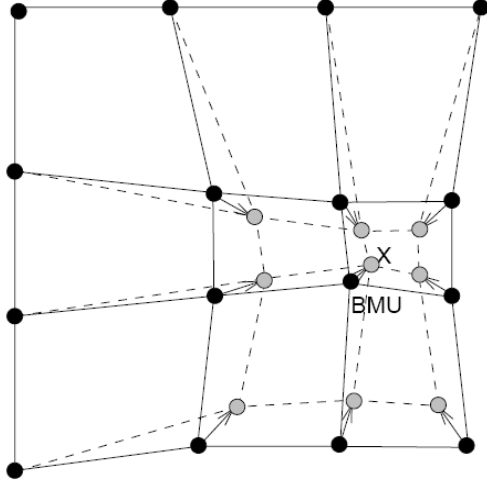


Figure 1: Updating the BMU and its neighbors towards the input sample [10]

There are many possible types of neighbourhood functions, the simplest and most commonly used are the following [10]:

- Bubble: In this simplest form it is one for all neurons close enough to BMU and zero for others

- Gaussian:  $e^{-\frac{\|r_b - r_i\|^2}{2\sigma^2(t)}}$

where,  $r_b$  and  $r_i$  are positions of the winning neuron  $b$  and  $i$  respectively, and  $\sigma(t)$  is the neighborhood radius. The learning rule makes the weights of the winning neuron and its neighbours more similar to the input  $x$ . Regardless of the functional form, the neighbourhood function shrinks with time. For each input vector this process is repeated for large number of cycles. The final network associates the output nodes with groups or patterns in the input data set. During the training process a map is built and the neural network organises itself using a competitive process. As the weights of the whole neighborhood are moved in the same direction during the training phase, similar items tend to excite adjacent neurons. Therefore, GSOM forms a semantic map where similar samples are mapped close together and dissimilar apart. An analyst can then evaluate the map with the knowledge of input dataset.

The characteristic feature distinguishing neural maps from other neural network paradigms and regular vector quantisers is the preservation of neighbourhoods. This desirable feature obviously depends on the choice of the output topology. The proper dimensionality of the output space is usually not known a priori, but has to be specified prior to learning in the GSOM algorithm.

GSOM starts with a minimal number of nodes (usually four) and grows new nodes on the boundary based on a heuristic. By using the value called the spread factor (SF), the data analyst has the ability to control the growth of the GSOM. A parameter of growth known as, the growth threshold (GT) is defined as [9]:

$$GT = -D \times \ln(SF) \quad (3)$$

where,  $D$  is the dimensionality of data and SF is the user defined spread factor that takes values between 0 and 1. An SF value close to 0 represents minimum growth and close to 1 represents maximum growth. Note that the outer boundaries of the spread factor (0 and 1) cannot be included since  $\ln(0) = \text{undefined}$ , and  $\ln(1) = 0$  renders the growth threshold meaningless.

The algorithm first identifies the winning node. Then an accumulated error  $E$  (difference between input vector and weight) of the winning node is updated by the following rule:

$$E(t+1) = E(t) + |x - m_b| \quad (4)$$

where,  $x$  is the input vector and  $m_b$  is the weight vector of the winning node. If the winning node is the boundary node and  $E$  exceeds GT, growing is initiated on that node to fill the surrounding unoccupied spaces of the lattice. If  $E$  of the winning node exceeds GT but the winning node is not a boundary node, then  $E$  is propagated outwards to other neighboring nodes. Weights of the new nodes will be initialized according to the following equation [3]:

$$m_{new} = 2m_{winner} - m_{opposite} \quad (5)$$

where  $m_{opposite}$  represents the weight of the node topologically opposite to the new node. If there are no topologically opposite nodes, weights of the new nodes will be calculated according to the following equation:

$$m_{new} = m_{winner} + m_{other1} - m_{other2} \quad (6)$$

where,  $m_{other1}$  and  $m_{other2}$  are weights of the nodes nearest to the new node. For hexagonal topology,



equation (6) is applicable as there will always be a neighbor of the winning node that is topologically on the opposite side of the new node. Growing phenomenon in GSOM is illustrated in Figure 2. New nodes are inserted around the high ‘error node’ (node ‘Error’ in Figure 2), with accumulated error exceeding the predefined growth threshold.

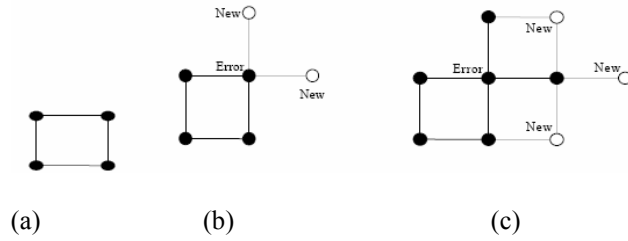


Figure 2: Growth in GSOM: (a) initial state (b) two new nodes and (c) three new nodes

## 2.2. Clustering quality improvement

Main objective of GSOM is to transform high-dimensional data in a low-dimensional cluster maps. However, with complex datasets, it may not be straightforward to obtain a good clustering. It is more applicable for manufacturing datasets, which are considered as complex due to the large number of processes, diverse equipment set, and nonlinear process flows. Manufacturing datasets often comprise of hundreds of process control, process step and quality control data. It might be necessary to change different variables and monitor the change of ‘quality’ of the cluster. This study proposes a clustering quality (CQ) measure, as a benchmark to evaluate the changes in clustering quality with different parameter changes. Mathematically, clustering quality can be expressed as:

$$CQ = \sum_i \max \left\{ \frac{\frac{g_i - G}{n_i - N} * \frac{n_i}{N}, 0 \right\} + \sum_i \max \left\{ \frac{\frac{b_i - B}{n_i - N} * \frac{n_i}{N}, 0 \right\} \quad (7)$$

where,  $B$  is total number of faulty products,  $G$  is total number of good products,  $N$  is total number products ( $B+G$ ),  $b_i$  is number of faulty products in neuron  $i$ ,  $g_i$  is number of good products in neuron  $i$ , and  $n_i =$  number of products in neuron  $i$  ( $b_i + g_i$ ).

A CQ of 1 would mean a perfect separation of good and faulty products and a 0 would mean no separation

at all. This proposed CQ takes the complete cluster map into account and can be generated automatically as an objective quantifier of the programme to separate good and faulty products from the data provided

Many datasets (manufacturing databases for example) are large and complex and may have substantial amounts of noise. Noise reduction from the data is one of the primary conditions for obtaining good clustering. Some datasets may contain many categorical variables, which are comprised of letters or a combination of numbers and letters. These categorical data should be transformed into numerical data to allow their use in data mining programmes. One of the ways to transform categorical data to numerical data is to use the binary convention.

Investigation of sample manufacturing datasets has revealed that there are many categorical variables that do not have any affect on separating good and faulty products. For example, a categorical variable may have an equal distribution among good and faulty products. These variables, when expanded, only add noise to the dataset. To filter out these unnecessary categorical variables, a method is proposed to remove them from the dataset. Mathematically the constraint can be expressed as:

$$FI = \left( \frac{X_{ig}}{N_g} - \frac{X_{if}}{N_f} \right) \leq \alpha \quad (8)$$

where,  $FI$  = Filtration index,  $X_{ig}$  = number of good products having a particular categorical variable,  $X_{if}$  = number of faulty products having a particular categorical variable,  $N_g$  = number of good products in the sample,  $N_f$  = number of faulty products in the sample,  $\alpha$  = user defined constraint limit.

If the Filtration Index for a variable is close to zero, it can be considered that the variable is equally distributed among good and faulty products and hence, should not be considered for analysis. For example, if a dataset has 50 good products and 25 faulty products and among them 10 good products and 5 faulty products have the same variable, then  $FI$  will be  $[(10/50)-(5/25)] = 0$ . As the variable is similarly distributed among the good and faulty products, it will have no effect on clustering. However, it is not expected that a variable would have  $FI$  value exactly zero. In practice, the user has to define the value (or range of values) for  $\alpha$  depending on the characteristics of the dataset. Although the  $FI$  has been proposed for categorical variables, it can similarly be used for numerical variables.

An improved GSOM process taking clustering quality and  $FI$  into consideration is described below:

1. Initialize the weight vectors of the starting nodes (usually four) with random numbers between 0 and 1.
2. Calculate the growth threshold (*GT*) using equation (3)
3. Determine the winning neuron using equation (1).
4. Update the weight vectors of the neighbourhood using equation (2).
5. Calculate error values and if the error of node *i* is greater than *GT*, grow nodes if *i* is a boundary node. Distribute weights to neighbors if *i* is a non-boundary node.
6. Initialize the new node weight vectors to match the neighboring node weights.
7. Repeat steps 3 – 6 until all inputs have been presented and node growth is reduced to a minimum level.
8. Reduce learning rate and fix a small starting neighborhood.
9. Find winner and adapt the weights of the winner and neighbors in the same way as in growing phase.

The GSOM process with clustering quality improvement described above can be summarized as follows. Primary or basic data is first pre-processed. Then the dataset is filtered to remove insensitive variables. If there are some categorical data in the dataset, these are converted to numerical values. Data are then normalized to feed into the GSOM. The quality of the cluster maps are then evaluated using equation (7). If the clustering quality is not acceptable, the dataset is further refined by removing insensitive data. Finally the cluster maps are analyzed and interpreted with the knowledge of the dataset used.

### 3. Results and Discussions

The method proposed has been applied to a complex manufacturing datasets. The results demonstrate the effectiveness of the methodology. Details of the simulation results are described in the following sections.

Recorded manufacturing process data can be used to analyze the performance of a process. However, it is difficult to find the causes of any abnormal output or the factors resulting in lower yield rates [11]. To determine whether GSOM could deal with a complex manufacturing dataset, a simulation has been run with a large dataset obtained from Motorola USA. The

quality problem of the Motorola wafer fabrication process is described in reference [1].

In the context of analysis of manufacturing quality problems, the focus involves two main aspects - separation of good and faulty products and identifying the reason for yield failure. The present study focuses on the first aspect. The challenge is not solely in clustering, but also to obtain a meaningful and adequate number of clusters. With meaningful clusters, grouped in appropriate numbers, identification of the reasons that contribute significantly to the differentiation of clusters should become a simpler task.

The dataset is the historical wafer data collected for 2500 wafers over a 2-month period. The input database measured 133 parameters by 16,381 entries organized into an Excel file. The data consisted of wafer probe data of 39 wafer probe functional tests, process control data (59 numerical electrical PC measurements probed at 8 sites per wafer) and process step data such as material vendor/lot, wafer boat position, etc. A sample of the dataset is shown in Table 1. The second column in the table is the product ID, the 3<sup>rd</sup> column is the product reference number, and columns C1 to X133 are measured parameters. In the original dataset, there are 59 'C' columns (C1-C59), 38 'K' columns (K60-K98), and 34 'X' columns (X99-X133). The reference number in the 3<sup>rd</sup> column identifies which product is good and which product is faulty. Reference numbers above 8750 indicate good products and below 8750 indicate faulty product.

As shown in Table 1, there are many categorical variables especially in the X-columns. This dataset must be converted into a suitable format to use in GSOM. Pre-processing of data includes removal of outliers, transformation of categorical data into numerical values, and normalization. It was found that there are some excessively large values in the dataset and if these are not removed, most of the values will become zero after normalization. After removing the outliers, the categorical data are transformed into binary. Considering the complexity of the dataset, about one quarter of the entries (4000) was considered for simulation purposes.

No.	NAME	REF	C1	C2	...	K60	K61	...	X132	X133
1	J546040_12_1	9628	1.00E-09	2.21E-02	...	33	18	...	R2793	Dec-18-95
2	J546040_12_2	9628	-1.33E-09	2.39E-02	...	33	18	...	R2793	Dec-18-95
3	J546040_12_3	9628	8.33E-10	2.31E-02	...	33	18	...	R2793	Dec-18-95
4	J546040_12_4	9628	-4.77E-10	2.36E-02	...	33	18	...	R2793	Dec-18-95
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
16381	F606617_21_5	9584	-2.69E-08	2.23E-02	...	19	18	...	R285	Feb-26-96

**Table 1.** Motorola Wafer Manufacturing Data showing dimensionality and layout)

SF(IF $\leq$ 0)	<b>0.1</b>	0.5	0.7	0.8	0.9	<b>0.95</b>
CQ	<b>0.25</b>	0.40	0.45	0.49	0.51	<b>0.52</b>
IF (SF=0.95)	$\leq$ <b>0.05</b>	$\leq$ 0.10	$\leq$ 0.12	$\leq$ 0.13	$\leq$ 0.14	$\leq$ <b>0.15</b>
CQ	<b>0.71</b>	0.76	0.78	0.79	0.80	<b>0.81</b>

**Table 2.** Improvement of Clustering Quality Changing SF and IF

Simulation was run at SF 0.1 first but a very poor (0.25) CQ was obtained. Clustering quality was improved with higher SF but the maximum CQ was only 0.52, obtained at an SF of 0.95. It is thought that the large number of attributes imposed unnecessary noise in the dataset. It may be necessary and of great insight to study the effect of removing several attributes from the input dataset and to check the impact on the generated maps.

The categorical data from the dataset were reduced using equation (8). First using the constraint  $IF \leq 0.05$ , 163 categorical variables were deleted and the simulation was rerun. A significantly higher CQ of 0.71 was achieved. To test further, more categorical variables were deleted. Using  $IF \leq 0.15$  a CQ of 0.81 was achieved. As further deletion of variables did not produce any more significant improvement, no further deletion was done. Moreover, if many variables are deleted, some important features may be lost. A summary of the results is presented in

Table 2 and a cluster map of SF=0.95 and  $IF \leq 0.15$  is presented in Figure 3.

Although higher clustering quality is desired, it may not be possible to obtain high CQ with complex manufacturing data as 100% separation of all good and faulty products is not possible. For example, consider that the desired dimension of a product is 10mm. If the dimension of the finished product is 11mm it will be considered a failure and if the dimension of final product is 15mm it is also a failure but certainly different to the previous failure. It cannot be expected that all products with a dimension 10mm (plus

tolerance) be clustered in one group and the rest will be grouped in another cluster. There will certainly be some mixing. In practice, some clusters with 100% good and faulty products and the rest of the clusters with majority either good or faulty products are expected. In this study some clusters with 100% good and faulty products were obtained and other clusters either contained majority of good products or faulty products.

#### 4. Conclusions

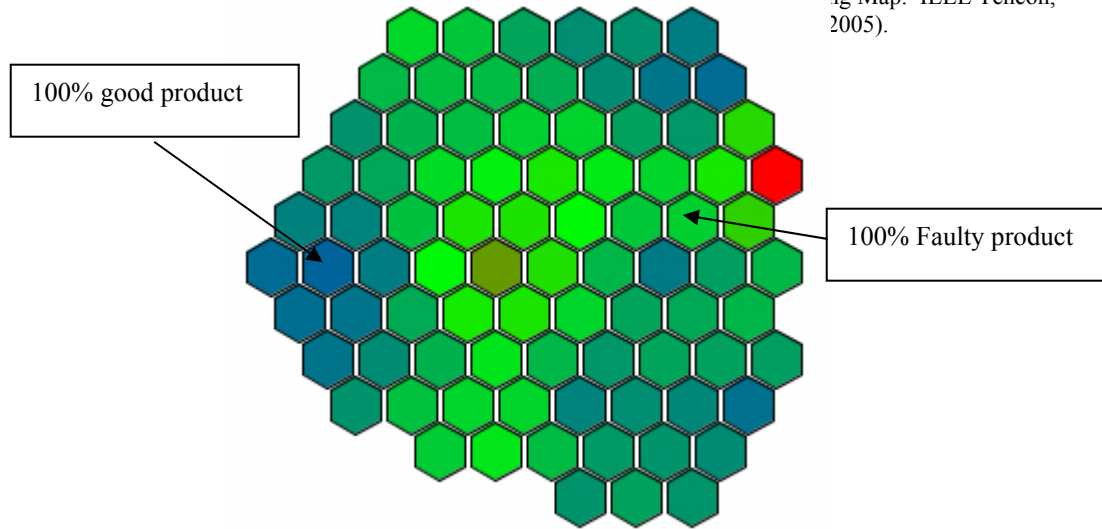
The proposed methodology is a significant contribution

to the data mining techniques available, especially in dealing with complex manufacturing datasets. The GSOM algorithm has been proven to be an efficient algorithm to analyze unsupervised DNA data. However, it produced unsatisfactory clustering when used on manufacturing data. Moreover, there was no benchmark to monitor improvement in clustering. The present study has proposed methods to evaluate quality of the clusters produced by GSOM and to remove insignificant variables from the dataset. With the proposed modifications, significant improvement in unsupervised clustering was achieved with simple as well as complex manufacturing data. Possible further improvement in cluster analysis may be achieved by applying evolutionary optimization algorithms.

For manufacturing data, the objective is not limited to finding the separation of good and faulty

products. The main objective is to find the underlying reason for poor yield. To discover this, it is necessary

4. Russ, G., Karim, M. A., Islam, A., Hsu, A. L., Halgamuge, S. K., Smith, A. J. R., Kruse, R.: Detection of Faulty Semiconductor Wafers using Dynamic Self-Organising Map. IEEE Tencon, 2005).



**Figure 3.** Cluster Map of Motorola Wafer Data (SF: 0.95 and IF= $\leq 0.15$ )

to create clusters of good and faulty products, as has been done in this study. The technique is being extended to model the failure causes of the lower yielding products.

**Acknowledgements:** The authors acknowledge the significant contribution of Professor Saman Halgamuge and Dr Alan Smith of University of Melbourne to this research work. The authors would also like to thank Robert M. Gardner, Director Intelligent Systems Research Lab, Motorola USA, for kindly providing the valuable semiconductor wafer data.

## References

1. Gardner, M., Bieker, J., Elwell, S., Thalman, R., Rivera, E.: Solving Tough Semiconductor Manufacturing Problems using Data Mining. IEEE/SEMI Advanced Semiconductor Manufacturing Conference, Boston, MA (2000) 46-55.
2. M. A. Karim, A. J. Smith, M. Islam and S. Halgamuge, A comparative study of manufacturing practices and performance variables, *International Journal of Production Economics (ISI impact factor-1.183)*, **112** (2008), 841–859
3. M. A. Karim, A. J. Smith and S. Halgamuge, Empirical relationships between some manufacturing practices and performance, *International Journal of Production Research*, Vol. 46, No. 13, 1 July 2008, 3583–3613.
5. Hsu, A. L. (2005). Interactive data mining with dynamic self-organising maps, Ph.D. Thesis, The University of Melbourne, Melbourne.
6. Kohonen, T.: Self-Organising Maps. Springer , Berlin (1997).
7. Alahakoon, D., Halgamuge, S. K., Srinivasan, B.: Dynamic Self-Organising Maps with Controlled Growth for Knowledge Discovery. *IEEE Transactions on Neural Networks*. 11 (2000) 601-614.
8. Alahakoon, D.: Controlling the Spread of Dynamic Self Organising Maps. *Journal Neural Computing and Applications*, 13(2), (2004) 168-174.
9. Hsu, A. L., Tang, S.-L., Halgamuge, S. K.: An Unsupervised Hierarchical Dynamic Self-Organizing Approach to Cancer Class Discovery and Marker Gene Identification in Microarray Data. *Bioinformatics* 19 (2003) 2131–2140.
10. Vesanto, J. (2000). Using SOM in Data Mining. *Masters thesis*, Helsinki University of Technology.
11. Hou, T.-H., Liu, W.-L., and Lin, L.: Intelligent Remote Monitoring and Diagnosis of Manufacturing Processes Using an Integrated Approach of Neural Networks and Rough Sets. *Journal of Intelligent Manufacturing* 14 (2003) 239-253.

## Significant Cancer Risk Factor Extraction: An Association Rule Discovery Approach

Jesmin Nahar and Kevin S Tickle

School of Computing Science, Faculty of Business and Informatics  
Central Queensland University, Rockhampton, QLD-4702, Australia  
E-mail: {j.nahar,k.tickle@cqu.edu.au}

### Abstract

*Cancer is the top most death threat for human life all over the world. Current research in the cancer area is still struggling to provide better support to a cancer patient. In this research our aim is to identify the significant risk factors for particular types of cancer. First, we constructed a risk factor data set through an extensive literature review of bladder, breast, cervical, lung, prostate and skin cancer. We further employed three association rule mining algorithms, Apriori, Predictive apriori and Tertius algorithm in order to discover most significant risk factors for particular types of cancer. Discovery risk factor has been identified to shows highest confidence values. We concluded that apriori indicates to be the best association rule-mining algorithm for significant risk factor discovery.*

### 1. Introduction

Cancer is the one of the most common cause of death all over the world. Furthermore, in Australia, United States and the United Kingdom, breast cancer is the second leading cause of death [1]. Research shows that most of the cancers are not detected in the early stage-resulting in high death rates. Medical practitioner diagnostic systems (mammography, biopsies, ultrasound, etc) can takes a long time to inform the patient about the cancer status, thus early cancer prevention could save human life. The aim of this research is to find out the different cancer risk factors using modern machine learning techniques. We consider bladder, breast, cervical, lung, prostrate and skin cancer in this research. Association Rule Mining (ARM), a particular method of machine learning technique could prove to have great potential within the medical domain identifying responsible factors

for a specific disease. The ARM technique has been considered as a tool for our research.

The use of association rules on medical data records with heart perfusion measurements is explored for the first time in [2]. Discovering association rules on segmented images is studied in [3]. Another study use association rule mining algorithm to study associations of different female hormones with the occurrence of breast cancer and also bench mark the performance of artificial neural networks (ANN) and data envelopment analysis (DEA) against the standard parametric Fisher's linear discriminant analysis (FLDA) technique [4]. Ho et al presented an induction data mining approach in order to predict cervical cancer by using a combination of demographic, environmental and genetic factors. They examined the risk factors for cervical cancer using a combination of logistic regression and a decision tree algorithm, and reported the relationship among demographic, environmental and genetic factors for cervical cancer using induction technique. In addition, to the induction technique, they compared logistic regression and the decision tree algorithm, and present the practical use of rule induction for the management of cervical cancer [5]. Their emphasis was on a specific cancer with several limitations to their research. One being the limited access to data, relying on one single source subjecting themselves to bias in an otherwise potentially data rich subject group its. Sufficient validation for the generalization of their findings to other geographic areas or groups was not performed. The other limitation was the lack of input variables for patient characteristics, such as exercise behaviour, nutrition, stress, and depression. Further study limitations also included a low sensitivity for the logistic regression and Chi-squared Automatic Interaction Detection (CHAID) algorithm. An association rule mining algorithm was applied to

extract associations between the selected socioeconomic variables and the colorectal, lung, breast, and prostate cancers mortality from 1988 to 1992 cancer mortality rates in United States [6]. Geographic information system technology is used for data integration which is defined at different spatial resolutions, and to visualize and analyze the results from the association rule mining process. This study provided the emphasis on mortality rates. We failed to find a single study which was able to demonstrate on the significant risk factor for a specific cancer. We therefore, decided to contribute towards cancer risk factor issue through this research.

To the best of our knowledge, this is the first risk factor extraction research using association rule mining algorithms. Our paper is organized as follows: section two will explain the problem description. Data collection evidence is described in section three. The association rule mining algorithm is summarized in section four. In section five, we publish our experimental finding concluding in section six

## 2.0 Problem Descriptions

We initially we considered bladder, breast, cervical, lung, prostate and skin cancer oriented risk factors. This section will summarize each problem definition and its significance.

### 2.1 Bladder Cancer

Bladder cancer is a cancer where abnormal cells multiply without control in the bladder. In most of the cases bladder cancer begins in cells lining of the bladder and is called urothelial cell or transitional cell carcinoma (UCC or TCC). If the bladder cancer that is only in the lining is called *superficial bladder cancer*, and if the cancer spreads into the muscle wall of the bladder, it is called *invasive bladder cancer* [7]. The American Cancer Society estimates that in the United States there will be about 68,810 new cases of bladder cancer diagnosed in 2008 (about 51,230 men and 17,580 women). The chance of a man developing this at any time during his life is about 1 in 27 and for a woman, 1 in 85. In 2008, there will also be about 14,100 deaths from bladder cancer in the United States (about 9,950 men and 4,150 women) [8]. The cancer affected bladder is presented in Figure 1.

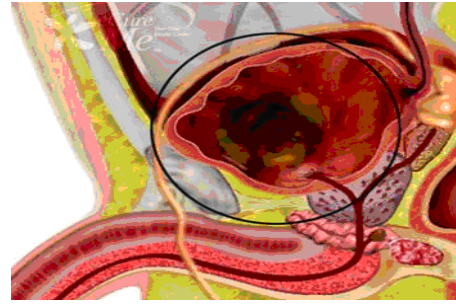


Figure 1. Affected cancer bladder shows by circle. (Source: [http://www.ecureme.com/emyhealth/data/Bladder\\_Cancer.asp](http://www.ecureme.com/emyhealth/data/Bladder_Cancer.asp).)

### 2.2 Breast cancer

Breast cancer is the combination of more than 100 different diseases. Breast cancer develops while some of the cells in the breast begin to grow out of control. Cancerous growth begins to appear anywhere inside the breast. Most breast cancers starts in the lining of the ducts though a small number can start in the milk sacs. In time if not treated properly, cancer cells from the breast can spread to other parts of the body [8]. One in eight Australian women will develop breast cancer before the age of 85. In Australia, about 12,000 women and 84 men were diagnosed with breast cancer in 2002. In 2005, breast cancer caused 502,000 deaths worldwide (7% of cancer deaths; almost 1% of all deaths). It is predicted that by 2011, the number of new diagnoses will increase to about 14,800 women and 122 men [7]. The cancer affected breast is presented in Figure 2.

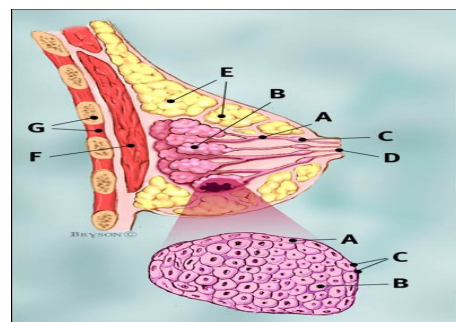


Figure 2. Normal breast with lobular carcinoma in situ (LCIS) in an enlarged cross-section of the lobule. **Breast profile:** A ducts, B lobules, C dilated section of duct to hold milk, D nipple, E fat, F pectoralis major muscle, G chest wall/rib cage. **Enlargement:** A normal duct cells, B

ductal cancer cells, **C** basement membrane, **D** lumen (center of duct). ([http://www.cancer.org/docroot/CRI/content/CRI\\_2\\_2\\_1X\\_What\\_is\\_breast\\_cancer\\_5.asp?sitearea=](http://www.cancer.org/docroot/CRI/content/CRI_2_2_1X_What_is_breast_cancer_5.asp?sitearea=))

### 2.3 Cervical

Cervical cancer is a type of cancer that starts in the cervix (part of the female reproductive systems), the lower part of the uterus (womb) that opens at the top of the vagina. Most of the time, early cervical cancer has no symptoms. Cervical cancer is divided into two stages; early or pre-invasive stage, and the late or invasive stage. Cervical cancer is a major worldwide health problem with incidence and mortality rates second only to breast cancer. It is the number one killer of young women in under developed countries. In Korean women for example, cervical cancer is the third most common form of cancer, after stomach and breast cancer [9]. The affected cervical cancer is presented in Figure 3.

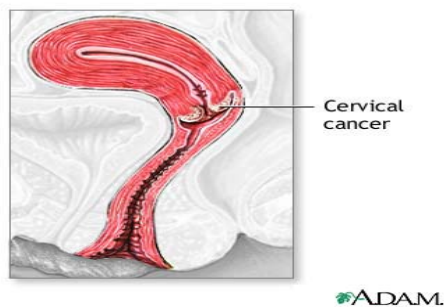


Figure 3. Cervical cancer, ( Source: [http://www.righthealth.com/Health/Cervical\\_Cancer\\_Pictures/-od-definition\\_adam\\_2%25252F9163-s](http://www.righthealth.com/Health/Cervical_Cancer_Pictures/-od-definition_adam_2%25252F9163-s))

### 2.4 Lung cancer

There are different types of lung cancer, depending on which cells are affected. The lungs are made up with three right lobes and two left lobes. As soon as we breathe, air goes through the nose, down the windpipe (trachea), and into the lungs where it spreads through tubes called bronchi. The majority of lung cancer begins in the cells that line these tubes [7]. Statistics shows that for both men and women people die from lung cancer more often than any other type of cancer. In 2004 lung cancer accounted for more deaths than breast cancer, prostate cancer, and colon cancer combined. In that year, 108,355 men and 87,897 women were diagnosed with

lung cancer, and 9,575 men and 68,431 women died from lung cancer. During 2008, there will be about 215,020 new cases of lung cancer (114,690 among men and 100,330 among women) [8]. The cancer affected lung is shown in Figure 4.

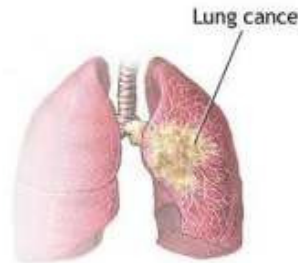


Figure 4. Lung cancer. (Source: <http://search.live.com/images/>)

### 2.5 Prostate cancer

The cancer which originates in the prostate with an uncontrolled (malignant) growth of cells in the prostate gland which is located at the base of the urinary bladder and is responsible for helping control urination as well as forming part of the semen. It is a common malignancy in men older than 50 years of old and it may cause no symptoms in the early stage. Prostate cancer may remain in the prostate gland or may spread to nearby lymph nodes to other parts of the body. It is a second leading cause of death of males in the U.S.A. In Australia one in nine males affects for prostate cancer at the age of 75 years and one on five by the age 85 years. The death rate of prostate cancer is lower which is one in 84 men under the age of 75 years and one in 22 by the age of 85 years [10]. The prostate cancer is shown in Figure 5.

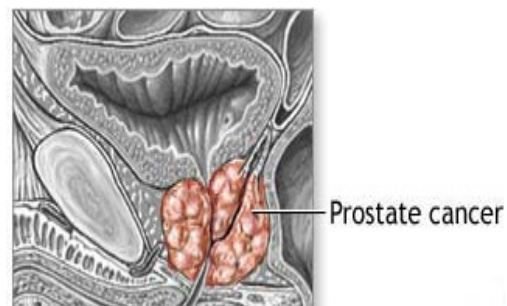




Figure 5. Prostate cancer. (Source, <http://health.allrefer.com/pictures-images/prostate-cancer.html>)

## 2.6 Skin cancer

Skin cancer is the most common form of cancer which estimated that over 1 million new cases occur annually. Skin cancer occurs in the outer layers of your skin. Each year the annual rates of all forms of skin cancer are increasing, representing a growing public concern. Research estimated that nearly half of all Americans who live to age 65 will develop skin cancer at least once ([www.cpaaindia.org](http://www.cpaaindia.org)). The affected skin cancer is shown in Figure 6.



Figure 6. Skin cancer. (Source: <http://www.taconichills.k12.ny.us/webquests/noncomdisease/skin%20cancerpic.jpg>)

## 3.0 Data Collection

We constructed a risk factor data set containing  $n$  from patient records  $M = \{m_1, m_2, \dots, m_n\}$  with categorical attributes. All attributes are uniformly treated as categorical to simplify the problem. Due to lack of space the sources of data are not available here. All the data are available upon request.

## 4.0 Association Rule Mining

In this section we introduce the standard definition of association learning [11,12]. An association rule is a simple rule, such as:

Smoking  $\cap$  family\_history\_positive  $\Rightarrow$  lung\_cancer

Where ‘Smoking  $\cap$  family\_history\_positive’ is called the body rule and lung\_cancer the head of the rule. It associates the rule body with its head.

In context of medical data, our example expresses the fact that people who are smoking and past family history have lung cancer are likely to have risk of lung cancer. This association rule makes no assertion about people who are not getting skin or breast cancer.

Let  $D$  be a database consisting of one table over  $n$  attributes  $\{a_1, a_2, \dots, a_n\}$  and the table contain  $k$  instances. In the following an attribute-value-pair will be called an item. An item set is a set of distinct attribute-value-pairs. For instance, smoking, family\_history\_positive, lung\_cancer and so on.

Let  $d$  be a database record and it satisfies an item set  $X \subseteq \{a_1, a_2, \dots, a_n\}$  if  $X \subseteq d$ . An association rule is an implication  $X \rightarrow Y$  where  $X, Y \subseteq \{a_1, a_2, \dots, a_n\}$ ,  $Y \neq \emptyset$ ; and  $X \cap Y = \emptyset$ .

The *support* ( $s$ ) and *confidence* ( $c$ ) are two measures of rule interestingness that reflect usefulness and certainty of a rule respectively. The association rules that satisfy user specified minimum support threshold (*minSup*) and minimum confidence threshold (*minCon*) are called strong association rules. The formulations of support and confidence measure are as follows:

Support ( $X \rightarrow Y$ ) =  $P(X \cap Y)$

Confidence ( $X \rightarrow Y$ ) =  $P(Y|X) = \frac{P(X \cap Y)}{P(X)}$

where  $P$  is used for probability measure in support and conditional probability in confidence. A probabilistic interpretation of support and confidence is discussed in [13]. Another rule measure metric called lift, defined as

$$\text{Lift}(X \rightarrow Y) = \frac{P(X \cap Y)}{P(X)P(Y)}$$

Lift quantifies the relationship between  $X$  and  $Y$ . Basically, a lift value greater than 1 provides strong evidence that  $X$  and  $Y$  depend on each other. A lift value below 1 state  $X$  depends on the absence of  $Y$  or vice versa.



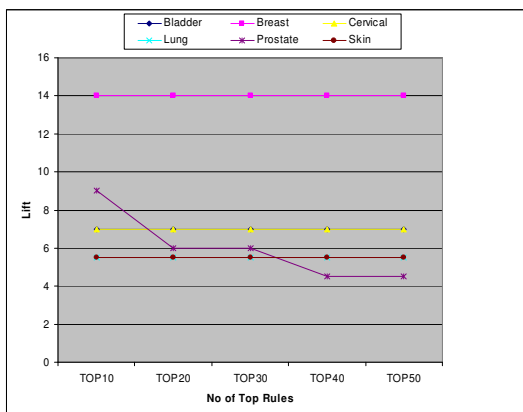
The task for association rule mining can be divided into two folds:

- Find all combinations of the risk factors for a specific cancer whose supports are greater than a user-specified minimum support (threshold).
- Use the risk factors from frequent risk factors to generate the desired rules. Generally the confidence of each rule is computed, and if it is above the confidence threshold, the retrieve the rule in the system.

Among the long list of association learning algorithms Apriori [11], Predictive Apriori [14], Tertius [15] are most popular algorithms in the machine learning community. The Apriori algorithm is a state of the art algorithm most of the association rule algorithms are somewhat variations of this algorithm [11].

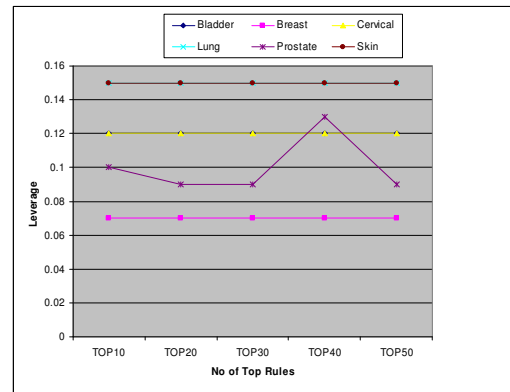
## 5.0 Experimental Results

We considered three association rule mining algorithms: Apriori, Predictive Apriori and Tertius for our experiment. We found among these three algorithms Apriori to be the number one choice to extract useful rules for our cancer data set. The reputed rules were selected based on confidence and lift measure. All the reported rules obtained confidence 1. However, in the lift measure all types of cancer discussed in this paper show a stable performance except prostate cancer as shown in Figure 7. The breast data-set shows the highest lift values, while on the other hand skin cancer showed the worst lift values.



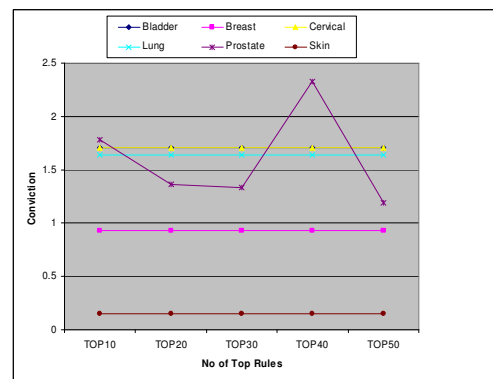
**Figure 7.** Lift performances measures with Apriori association rule mining algorithm for six types of cancer problems.

Moreover we verified Apriori performance with leverage measures as shown in Figure 8. Similar to lift measures we found all other types of cancers suggested similar steady performance, except prostate cancer. The most reliable performance was shown with the skin type cancer data-set while breast cancer data demonstrated the worst performance in the leverage measure.



**Figure 8.** Leverage performances measures with Apriori association rule mining algorithm for six types of cancer problems.

Finally we compared the apriori performance with conviction measures as shown in Figure 9. We found similar steady performances for all the cancer data set except prostate cancer, showing the highest performance. On the other hand again the breast cancer data set showed the worst performance.



**Figure 9.** Conviction performances measures with Apriori association rule mining algorithm for six cancer types' problems.

We therefore concluded a rule for each cancer dataset and summarized as below. All the rules were shown highest confidence. We found the breast and prostate cancer rules are very simple. Whereas rest of the rules contained several significant risk factors to predict an appropriate cancer.

**Relation: bladder**

**Rule:** (work\_exposure  $\cap$  work\_leather\_industry  $\cap$  work\_rubber\_industry  $\cap$  work\_cable\_industry  $\cap$  work\_printing\_industry)  $\Rightarrow$  bladder\_cancer

**Relation: breast**

**Rule:** (gender (female)  $\cap$  age(>40)  $\cap$  family\_history)  $\Rightarrow$  breast\_cancer

**Relation: cervical**

**Rule:** (poor\_hygienic\_practices  $\cap$  low\_socioeconomic\_status  $\cap$  low\_vegetables  $\cap$  sexually\_transmitted\_infections,smoking)  $\Rightarrow$  cervical\_cancer

**Relation: lung**

**Rule:** (chemical\_exposure  $\cap$  smoking  $\cap$  previous\_lung\_disease  $\cap$  urban\_residence  $\cap$  family\_history)  $\Rightarrow$  cervical\_cancer

**Relation: prostate**

**Rule:** (family\_history  $\cap$  age (>50))  $\Rightarrow$  prostate\_cancer

**Relation: skin**

**Rule:** (weak\_immune\_system  $\cap$  fragile\_skin  $\cap$  exposure\_to\_environmental\_hazards  $\cap$  previous\_skin\_cancer)  $\Rightarrow$  skin\_cancer

All the generated rules demonstrated high confidence. Thus, we conclude that these rules are highly acceptable. Comparatively breast and prostate cancer showed shorter rules comparing other cancer. This research argues to consider more risk factors to be considered in the experimental dataset for both breast and prostate cancer to extract most significant risk factors.

## 6.0 Conclusions

This study has demonstrated the use of association rule mining to identify some specific risk factors for a particular cancer problem. In this research we employed three types of association rule mining algorithms to extract significant risk factor for a specific cancer. Finally we suggested in our research to use Apriori algorithm for such type of task. All the generated rules were showing highest confidence level in our experiment.

We aim to extend this research by considering more risk factors in the raw data-set for each cancer problem, in order for association rule mining algorithm to extract more useful and significant risk factors for a particular type of cancer.

## References

- [1] Breasthealth, <http://www.breasthealth.com.au/statisticsresearch/index.html>, Accessed on 12 August, 2008.
- [2] Cooke, D., Ordonez C., Ernest V. Garcia, E. Omiecinski, Elyzabeth Krawczynska, Russell Folks, Cesar Santana, Leven de Braal, and N. Ezquerra. Data mining of large myocardial perfusion SPECT (MPS) databases to improve diagnostic decision making. *Journal of Nuclear Medicine*, 40(5), 1999.
- [3] Ordonez, C. and Omiecinski. E. Discovering association rules based on image content. In *IEEE Advances in Digital Libraries Conference (ADL'99)*, 1999, 38–49.
- [4] Pendharkar, P.C., J.A. Rodger, G.J. Yaverbaum, N. Herman and M. Benner, Association, statistical, mathematical and neural approaches for mining breast cancer patterns, *Expert Systems with Applications* **17** 1999, 223–232.
- [5] Ho, S.H., Jee, S.H., Lee, J.E. and Park, J.S. Analysis on risk factors for cervical cancer using induction technique, *Expert Systems with Applications*, 27(1). 2004, 97-105.
- [6] Vinnakota S. and Lam, N.S.N. Socioeconomic inequality of cancer mortality in the United States: a spatial data mining approach, *Int J Health Geogr.* 2006, 5- 9.
- [7] Wikipedia, [http://en.wikipedia.org/wiki/Bladder\\_cancer](http://en.wikipedia.org/wiki/Bladder_cancer), Accessed 20<sup>th</sup> August 2008.
- [8] Cancer.org, [http://www.cancer.org/docroot/CRI/content/CRI\\_2\\_4\\_1X\\_What\\_are\\_the\\_key\\_statistics\\_for\\_bladder\\_cancer\\_44.asp](http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_are_the_key_statistics_for_bladder_cancer_44.asp), Accessed, 20<sup>th</sup> August, 2008.
- [9] Herzog, T.J. New approaches for the management of cervical cancer, *Gynecologic Oncology* **90** (2003), S22–S27.

- [10] Pharmacy, [http://www.pharmacy.gov.my/self\\_care\\_guide/Urogenital/Postate%20Cancer.pdf](http://www.pharmacy.gov.my/self_care_guide/Urogenital/Postate%20Cancer.pdf), Accessed, 25<sup>th</sup> August, 2008
- [11] Agrawal, R. T. Imielinski, & A. Swami. Mining association rules between sets of items in large databases, In Proceedings of the 1993 ACM SIGMOD international conference on Management of data, 207–216.
- [12] Ordonez, C. Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine, 10(2), 2006. 334 – 343.
- [13] Hastie, T., R. Tibshirani, & J.H. Friedman. *The Elements of Statistical Learning*. Springer, New York, 1st edition, 2001.
- [14] Scheffer, T. Finding Association Rules that Trade Support Optimally Against Confidence. *Proceedings of the 5th European Conference on Principles and Practice of Knowledge Discovery in Databases(PKDD'01)*, Freiburg, Germany : Springer-Verlag, 2001. 424-435.
- [15] Flach, P.A. & Lachiche, N. Confirmation-guided discovery of first-order rules with Tertius. *Kluwer Academic Publishers*. The Netherlands, 42, 2001. 61-95.

## **Secure Multi-party Computation: Problems, Techniques and Applications for Privacy-Preserving Data Mining**

Dr. Durgesh Kumar Mishra

Professor (CSE) and Dean (R&D), Secretary IEEE MP-Subsection,  
Acropolis Institute of Technology and Research, Indore, MP, India,  
[durgeshmishra@ieee.org](mailto:durgeshmishra@ieee.org), [dkmishra@acropolis.in](mailto:dkmishra@acropolis.in)

### **ABSTRACT**

Consider a set of parties who do not have trust in each other, nor in the channel by which they communicate. Even then the parties wish to correctly compute some common function of their local inputs, while keeping their local data secure from others. This, in a nutshell, is the problem of secure multi-party computation (SMC). This problem is fundamentally in cryptography and in the study of distributed computations. It takes many different forms depending on the underlying network, on the function to be computed, and on the amount of distrust the parties have in each others and in the network.

In this tutorial, we present several aspects of secure multi-party computation. We first present the definition of this problem in various situations. Our definition is drawn from the previous idea and formulations, and incorporate aspects that were previously overlooked. We also present several problems associated with SMC.

Next we show the problem of hiding the data from trusted third party (TTP) which computes the result. We present the existing solutions of SMC along with the protocols developed by us. First, we have introduced an anonymizer between the parties and the TTP to hide the data. In this case each party is having its own anonymizer and the entire data of a party goes through a single anonymizer. Keeping in mind the anonymizer may become malicious so we have chosen a particular anonymizer randomly to transfer the data. Apart from the randomly selection of anonymizer, the party will divide the data into number of packets and then send to different anonymizer so that the entire data will not reach to a single anonymizer and the privacy of individual will be maintained.

Next we present another problem, which enables the SMC to perform the correct computation of the result as well as the authentication of computational body. We have introduced multiple TTPs instead of a single one. If there is only one TTP then its behavior can be suspicious. Using this multiple TTP concepts we have the option to choose a TTP from domain of TTPs for computation. In this method we divide the data in several packets and these packets are sent to multiple TTPs and a randomly selected master TTP will perform the computation after accumulating data from other TTPs. In effect no body knows where computation will be performed till the last moment and thus problem of security is solved. For the authentication of TTP, we have introduced the concepts of equivalence classes. With the help of this concept we remove the malicious TTP from the system for further computation.

Finally, we present the problem of dealing with adversaries in SMC and minimizing their effects. We investigate the power of adversaries in several situations. In particular, we show how to construct a secure protocol for computing any function, where the communication channels can be tapped by the adversary, and secure communication is achieved by cryptographic primitives based on the computational limitations of the adversary.