

393
19.3.13

HRM with Optimized Functional Implementation of National Data Center Reducing Time and Spelling Error Complexity

A project paper is submitted in partial fulfillment of the requirement for the degree of **Bachelor of Science in Computer Science & Engineering**



Course Code: **CIT 402**
Course Title: Project & Thesis

Supervised By:
Md. Delowar Hossain
Assistant Professor & Chairman
Department of Computer Science & Information Technology
Faculty of Computer Science & Engineering

Submitted By:

Md. Palash Uddin
Student ID: 0802006
B. Sc. in **CSE**
Faculty of CSE

Ashfaque Ahmed
Student ID: 0802033
B. Sc. in **CSE**
Faculty of CSE

Submitted To:
Department of
Computer Science & Information Technology
Hajee Mohammad Danesh Science & Technology University,
Dinajpur, Bangladesh

January, 2013



DEDICATED TO OUR BELOVED PARENTS

Acknowledgement

The successful completion of any project depends on the successful starting and proper motion on the way of development. It also requires ceaseless cooperation, proper guidelines, encouragement, inspiration and constructive suggestions of the people who are experienced enough. The proper motion of project for the National Data Center, Bangladesh from the present manual system is accelerated by our project supervisor, **Md. Delowar Hossain**, Assistant Professor and Chairman, Department of Computer Science and Information Technology, and project co-supervisor, **Md. Arshad Ali**, Lecturer, Department of Computer Engineering. Also we want to thank **Md. Nadim**, Lecturer, Department of Computer Science and Information Technology. They are very helpful in the needs for the preparation of this project. We would like to take an opportunity to thank all other teachers and friends of our faculty who help and give suggestions during the project and preparation of the project paper.

Finally, we remember the Almighty, as without his kindness it was not possible to prepare for this project.

Date: 27-01-2013

Authors



Abstract

Data Management in a fast and efficient way is one of the demanding technologies of today's aspect. The **"HRM with optimized functional implementation of National Database reducing time and spelling error complexity"** is a system which makes the working procedures of present manual system of storing and retrieving citizen's information easier and increases its efficiency to a high degree. The implemented search methodology is efficient enough for high speed data retrieval. Though there may be spelling error in the input keywords used for searching a citizen, it is able to find out the required data about the citizen.

The search methodology accomplishes its goal due to its special design in the data structure. The main concern in this project is minimizing the total searching time for a given keyword. This can be done if we can pre-establish the idea of getting the data belonging to the keyword. The primary and secondary key-code generated by Double Metaphone Algorithm for each word is used to establish that idea about the word. This algorithm is used for creating the map of the original database, through which the keyword is matched against the data.

The whole web-based National Data Center (NDC) is accessed from different terminals of a network. This system can be used to keep records, modification of record, search for citizens, admins and find out a particular record etc.

In this report we describe the implementation procedure of NDC. We used the Entity-Relationship model to design database that will store and organize the national data. We have stored the data in MySQL Database Server and populated it with some sample data. The system can keep track of admin info, admin accounts, records of citizen etc. Using HTML PHP, JavaScript, CSS we have created an Internet-based graphical user interface that allows users to access the system from any location.

Table of Contents

		Page No.
	<i>Dedication</i>	<i>i</i>
	<i>Certification</i>	<i>ii</i>
	<i>Acknowledgement</i>	<i>iii</i>
	<i>Abstract</i>	<i>iv</i>
	<i>Table of Contents</i>	<i>v</i>
	<i>List of Tables</i>	<i>viii</i>
	<i>List of Figures</i>	<i>ix</i>
	<i>References</i>	47
	<i>Books, External Links, Web-sites</i>	47
Chapter-1:	Introduction	1
	1.1 Introduction	2
	1.2 Historical background of national database	2
	1.3 Drawbacks of existing system	3
	1.4 Considerations for solving the drawbacks of the existing system	4
	1.5 Proposed System, National Data Center of Bangladesh	5
	1.5.1 Aim and objective of proposed system	5
	1.6 Organization of the project	6
Chapter-2:	Proposed System	7
	2.1 Introduction	8
	2.2 Advantages of Proposed System	8
	2.3. Proposed Structure	8
	2.4 The information flow of the present system	9

List of Tables

Table's No.	Table's Name	Page No.
Table 3.1	Data Dictionary of division	21
Table 3.2	Data Dictionary of district	21
Table 3.3	Data Dictionary of upazilla	21
Table 3.4	Data Dictionary of union_parishad	21
Table 3.5	Data Dictionary of admin	22
Table 3.6	Data Dictionary of citizen	22
Table 3.7	Data Dictionary of admin_info	23
Table 3.8	Data Dictionary of logininfo	23
Table 3.9	Data Dictionary of ndcindex	23
Table 3.10	Data Dictionary of searchresult	23
Table 3.11	Data Dictionary of storedresult	24
Table 3.12	Data Dictionary of backup	24
Table 4.1	Comparison of different techniques	33
Table 4.2	Defining informative tables in Search Engine	34
Table 5.1	Hardware and Software requirements	39

Introduction

1.1 Introduction

Bangladesh is a developing country. In our country most of the government organizations are operating manually. As we know that the requirement of the digital world is to replace the paper work with digital components. Because, from the ancient time human always tries to ease their tasks through applying some sort of technical attempts. If the manual systems are being automated, the development of a country raises. So as technology advances we need to move on. From this thinking we have decided to digitize the records of our citizens of Bangladesh by implementing this project which is entitled by **HRM with optimized functional implementation of National Data Center reducing time and spelling error complexity.**

1.2 Historical background of national database

At present the information of the citizens of Bangladesh are not stored in a web-based central database. But partially the government organizations and offices do so in a scattered manner. For this reason the People's Republic of Bangladesh has started a Project on National Population Register. The Bangladesh Bureau of Statistics (BBS) has undertaken a pilot project to study the feasibility of creating the national database- the National Population Register [1]. As it is just the beginning of the project so what could be the actual system and what will be functionalities it is unknown. To know about the present system related with storing and retrieving the information of the citizens of Bangladesh we visited some Union Parishad and Pourasava in Dinajpur, Bangladesh. Among the available tools for gathering information about the present system, we have taken interview of different employees associated with the existing system. We preferred interview because of its multidimensional purposes. The collected information leads us to design the structure of the present system of national database. Also, we realize that the stored information of the citizens is not consistent and also the information is not centralized.

The existing structure of national database is shown in Fig. 1.1:

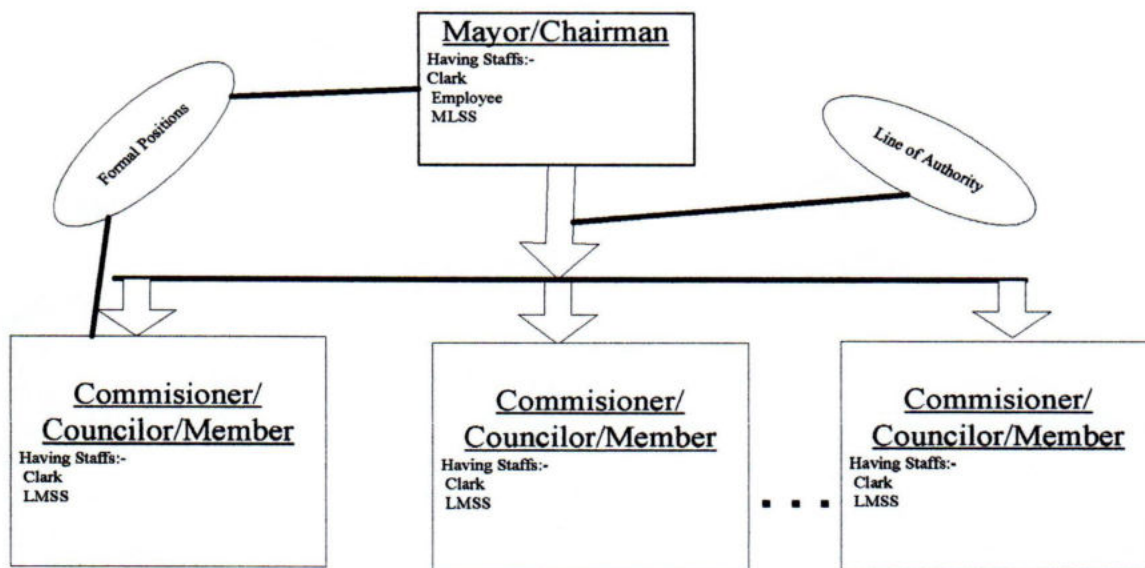


Fig. 1.1: Existing Structure

The above diagram of Fig. 1.1 shows that each formal position has certain officers and workers. The various sections of the existing structure of national database are described below:

Working Procedure: It performs all activities as shown in the above diagram manually using cataloging system.

Data Maintenance: For maintaining data it uses catalog book which has to be edited manually.

User Interaction: For maintaining/creating new record about the citizen one needs to fill a form which is provided by the organization or office and submit it to the staff member. The authorized people justify the given information and approve it. Then this information is stored in large book manually.

1.3 Drawbacks of the existing system

The present system is a manual system where staff members have to process everything manually. Present system faces many problems for its manual activities. Some of the problems are as follows:

Problems in creating central database: It is very difficult to create a central database for the system as information is not available in database.

Problems in retrieving data: Retrieving specific data from a catalog is very important for the user in time of need. But the present manual system fails to show satisfactory performance about the information. And as a result it becomes the source of hinders in searching.

Fail to maintain unique data: The records are stored may have the possibility of duplication i.e. one person can have names or other details. Further one person can entry his/her record several times by simply filling the form.

1.4 Considerations for solving the drawbacks of the existing system

The effectiveness of an NDC is greatly dependent on the order of maintenance. A computerized NDC can perform a great role in maintaining the system efficiently. It can make the whole activity user-friendly and ensure a great speed of information flow. Following topics are considered to solve the problems with our present manual system.

- **System Functionalities:** The NDC system must provide the following functionalities efficiently.

Searching: The searching of the systems should be efficient and fast enough and there is a need for designing an advanced search methodology that searches the whole database and returns a complete record.

Report/statistic generation: It should be able to generate statistical figures and any kind of reports at user's request.

Import/Export: It should be able to import or export records. It must provide data accuracy and data integrity belonging to the system.

Administration: Administrative parameters to manage and maintain any system setting, security etc.

- **Flexibility/Compatibility:** The system should be able to support expansion /upgrading, user customization, and modification. It should be compatible with other computer equipments and networks.
- **Running Environment:** The system must be easy to learn, easy to maintain, and easy to fix. It should deal everything in a user friendly manner.

- **Maintenance:** The maintenance of the system considers the Labor charges for installation, Cost of parts replacement, Cost of system upgrade, Cost of phone/email services etc.
- **Performance:** The performance of the system will be determined by Time saving, Manpower saving, System response time, Online service, i.e. online renewal, online checking of circulation records
- **Others:** Other features are required based on the demand of the system. For instance:
 - Display functions.
 - Feature to support reading activities the users.

1.5 Proposed System, National Data Center of Bangladesh

The National Data Center (NDC) is built based as per international standards. An NDC provides on-demand data access and submission via a secure web-based interface. The main purpose of this project is to store the information of citizens of Bangladesh in a web-based central database and then to retrieve the information of the citizens from a huge database ignoring spelling error and reducing searching time using a phonetic algorithm namely **Double Metaphone Algorithm**.

The NDC is a web-based integrated application. It is designed to meet the needs of digitization of national database. The system will help the top level management to make decisions regarding development of the country. It will provide necessary information to administrative personnel. These services have been designed to help users to get their jobs done faster and more efficiently ignoring spelling error complexity. The graphical interface brings new power and flexibility to their users. The most outstanding part of this web-based system is that the search technique is enhanced greatly.

1.5.1 Aim and objective of proposed system

The automation of the system will enable the user to interact with the system in an easy way. They can search records of the citizens with ease and high speed. This will provide a fast delivery of services from the system to the user. The system will be useful also to administrative personnel to take necessary decisions regarding the development of the country.

The purposes of this application are as follows:

- Admin can enter details related to a particular citizen.
- Super Admin can provide membership to the Admins of NDC.
- Admin Can read and write information about any citizen, and can update, create, and delete the record of a particular citizen as per requirement and implementation plans.
- Admin and Super Admin can view their status and make search for citizen.
- Providing high security

1.6 Organization of the project

Among the available models for developing a project we have used **SDLC**, System Development Life Cycle model. Here we organized our project in the following pattern:

Chapter 2 includes instruction to the proposed system, advantages, proposed structure, proposed information flow, and functional components about the **Proposed System**.

Chapter 3 traces **Structured Analysis** with the tools such as Use case diagram, Data flow diagram or Bubble chart, Class diagram, E-R diagram, Relational diagram, Data dictionary, Activity diagram, Swimlane diagram, State diagram, Architecture diagram and so forth.

Chapter 4 covers the **Search Methodology** that is used to implement the project.

Chapter 5 includes **Implementation and Testing** with some snapshots of the proposed system.

Finally, **Chapter 6** describes **Conclusion** and gives suggestion for further expanding of the proposed system.

CHAPTER **2**

PROPOSED SYSTEM

Proposed System

2.1 Introduction

Analyzing the present system we came to a decision that it has certain drawbacks. Some of the major drawbacks are listed below:

- It is a manual system
- It has failed to centralize the database
- Information may be duplicated by present system
- It takes long time to find a particular information etc.

To overcome these drawbacks we need to implement the proposed system, NDC.

2.2 Advantages of Proposed System

Our proposed system will have several advantages including fast and easy way to store information about citizens, retrieval of records with less time from a huge database etc. Further benefits are explained here:

- **Central Database:** The principal advantage of our propose system is that it will carry all information of the citizens of Bangladesh in a central database of NDC.
- **Automatic and real-time update of database:** NDC is more than three times faster for quick database update.
- **Lower manpower utilization in the system:** Huge amount of task is done with the less utilization of manpower through implementing the proposed system.
- **Easy and convenient to use:** NDC is very user friendly and easy to use. It facilitates faster retrieval of records.
- **Insert through files:** Another advantage of the proposed system is that new records can be inserted to the database through uploading files such **.xls** and **.txt**. This will certainly be a great benefit and reduce time of insertion.

2.3. Proposed Structure

Analyzing the present system we can come up to a decision that the structure of the proposed system will be like as follows:-

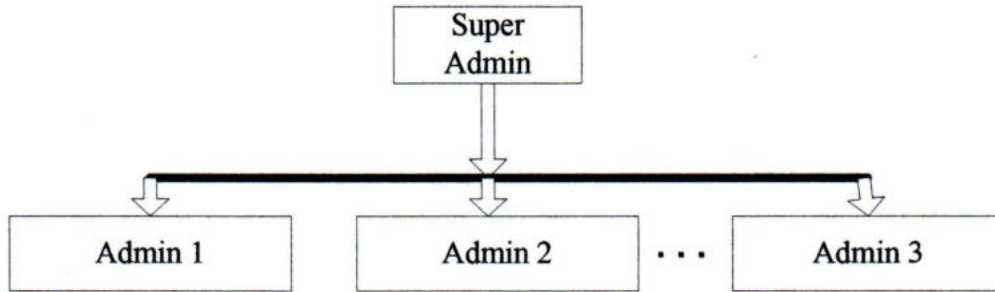


Fig. 2.1: Proposed Structure of NDC

2.4 The information flow of present system

The information flow of the existing system can be expressed by the following diagram of figure 2.2.

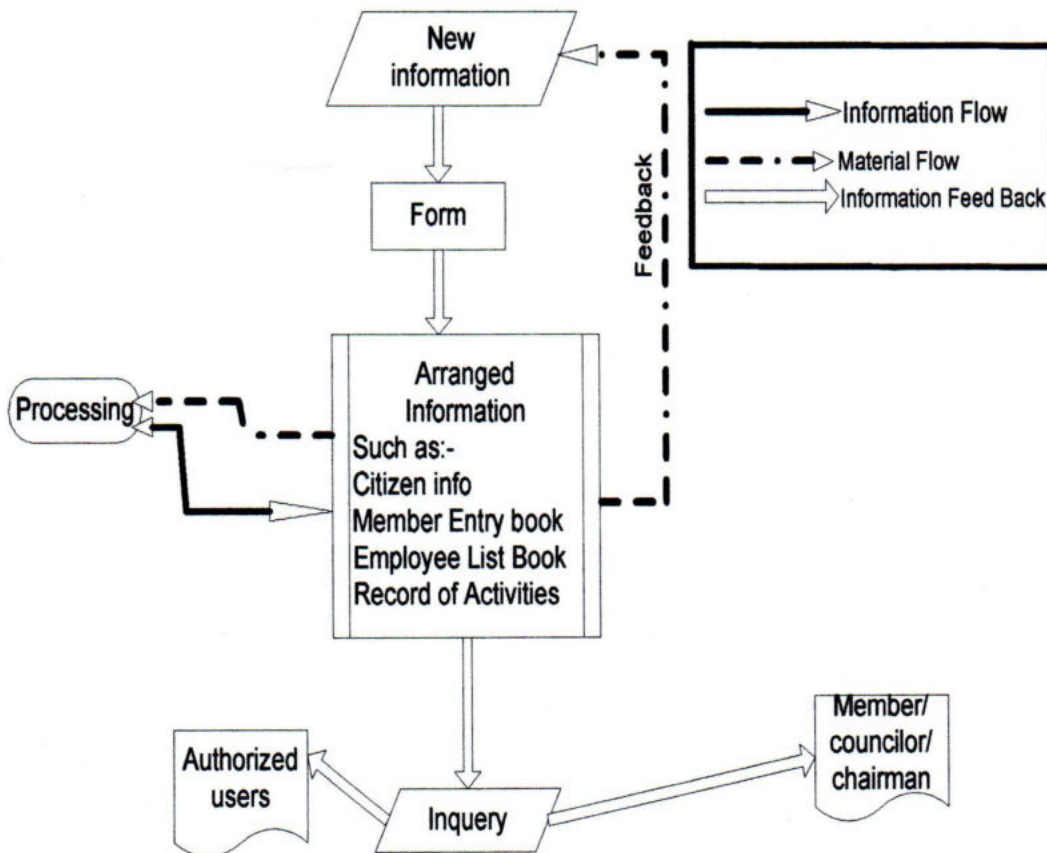


Fig. 2.2: Current Information Flow

2.5 Proposed Information Flow

Comparing the information flow of the present system of figure 2.3 the information flow of the proposed system will be like as follows:

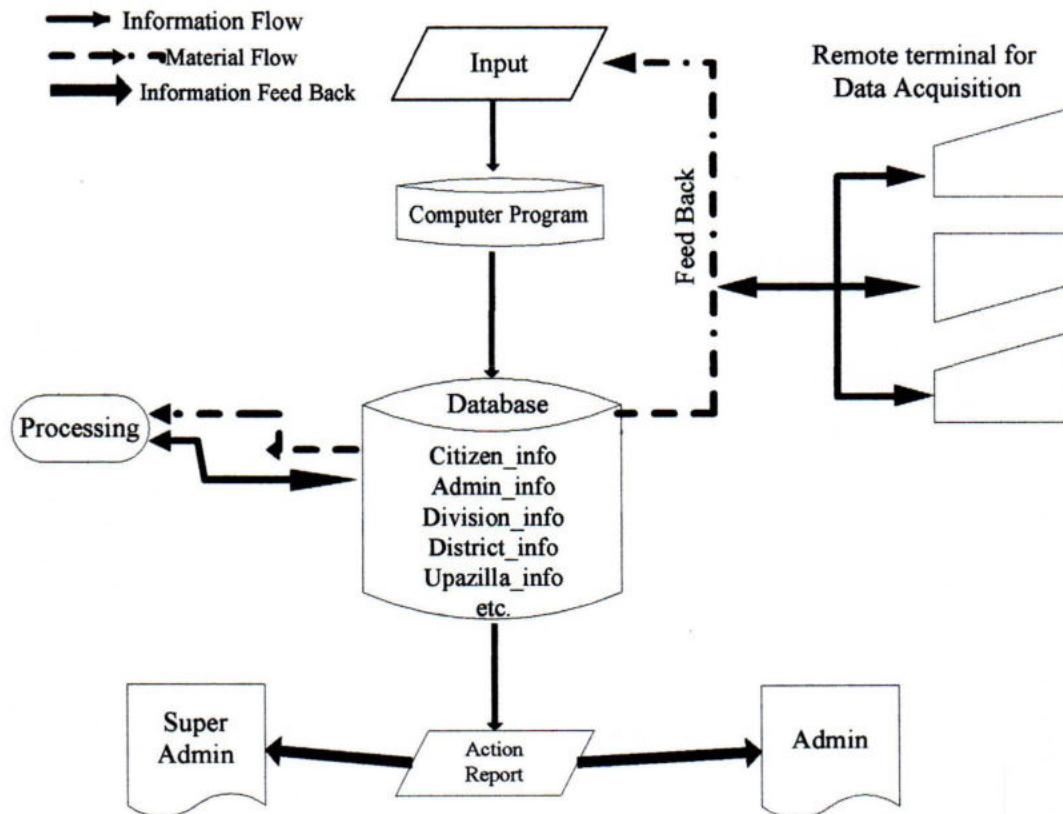


Fig. 2.3: Proposed Information Flow of NDC

2.6 Functional Components

Our project is to implement National Data Center, which provides necessary information related to the citizens of Bangladesh to its users. A super admin is responsible for maintaining the overall system functionalities. The super admin can have the authority to create more admins for the maintenance of the system. Our candidate system provides various types of functional components to the users. These functionalities have different types of scope for different types of user agents. Following is a list of functional components of the candidate system:

2.7 Common facilities

- Login to the system through the first page of the application
- He/She can search for particular information about the citizen i.e. name, division, district, upazilla, occupation etc.
- Insert new information.
- Insert new information through uploading file.
- Delete records from the database.
- Create Backup of the database
- Update information for particular record.

2.7.1 Serving module for Admin

It provides following functionalities for the member:

- When the admin login to the system through the first page of the application, the admin views his/her current information that was provided during the create admin phase.
- The admin can update his/her account such as change password or other information through the first page after login.

2.7.2. Serving Module for Super Admin

It provides functionalities for the administrative tasks. The main functions of this module are

- It provides all functionalities of the system.
- The super admin can create admin.
- Can view admin details.
- Also he/she can delete admin.
- It provides the facility to monitor the whole system.

CHAPTER 3

STRUCTURED ANALYSIS

Structured Analysis

3.1 Introduction

Structured analysis is a set of techniques and graphical tools that allow a system analyst to develop a new kind of system specifications that are easily understandable to the user. The most common tools used in structure analysis are:

- Use case diagram
- Data flow diagram or Bubble chart
- Class diagram
- E-R diagram
- Relational diagram or schema diagram
- Data dictionary
- Activity diagram
- Swimlane diagram
- State diagram
- Architecture diagram

The system is analyzed and the movement of data between different components of the system is designed so that the performance of the system can be improved.

3.2 Use case diagram

A use case diagram at its simplest is a graphical representation of a user's interaction with the system and depicting the specifications of a use case. Here, it covers the whole system of National Data Center and how it works. It makes easier to communicate between the client and the system developers.

- It is a set of scenarios that describing an interaction between a user and the system.
- It displays the relationship among actors and use cases.

The two main components of a use case diagram are actors and use cases as specified in the following diagram.

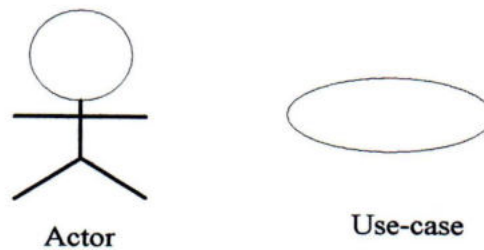


Fig. 3.1: Symbols of Actor and use-case

An actor represents a user or another system that will interact with the system you are modeling. A use case is an external view of the system that represents some action the user might perform in order to complete a task.

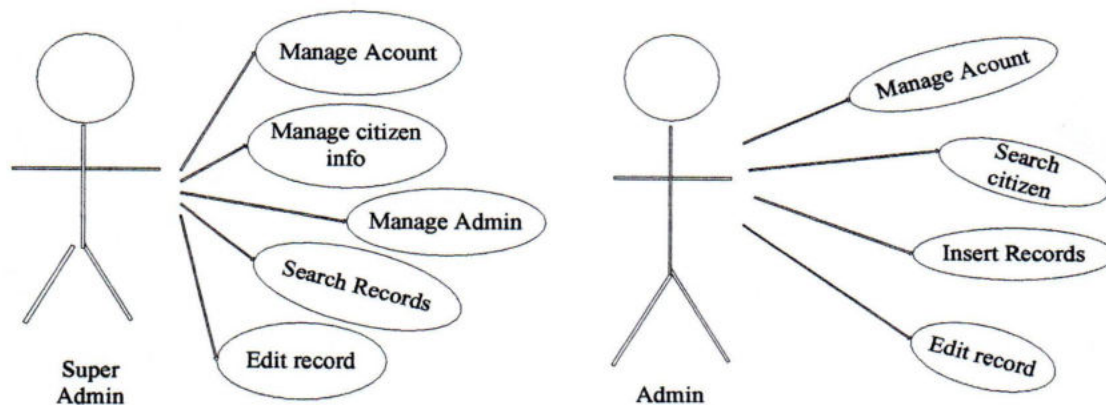


Fig. 3.2: Use-case for Super Admin and Admin

3.3 Data Flow Diagram

A Data Flow Diagram (DFD) shows what kinds of information will be input to and output from the system, where the data will come from and go to, and where the data will be stored. Again, a Data Flow Diagram is a logical model of the system; the model does not depend on hardware, software, and data structure or file organization. It is a graphical picture of the logical system. It is an excellent tool for communication with user to explain the system for him.

Data Flow diagrams have a notation for each of the components of a system. They also have a notation for representing different level in a hierarchy of detail used to describe the system. This notation makes it possible to represent an overall view of a large complex system and a detailed

view of a part of a system using the same notation. Parts of a system can be isolated into independent subsystems. These work together as a unit to perform a set of processes that must be done together at one time. There is a standard process that an analyst can follow to make sure that all components of a system are identified.

Data Flow Diagrams are constructed from four basic building blocks:

- Terminators
- Data stores ,
- Data Flows, and
- Process.

The different levels of data flow of the candidate system can be described by the following diagrams:-

3.3.1 Level-0 Data Flow Diagram

Level-0 Data Flow Diagram shows the systems major processes, data flows and data stores at a high level of abstraction. In level-0 DFD input going into a process is different from outputs leaving the process. Data stores are first shown at this level. Actually the basic structure of the system is shown in this level.

Level 0 DFD For NDC:

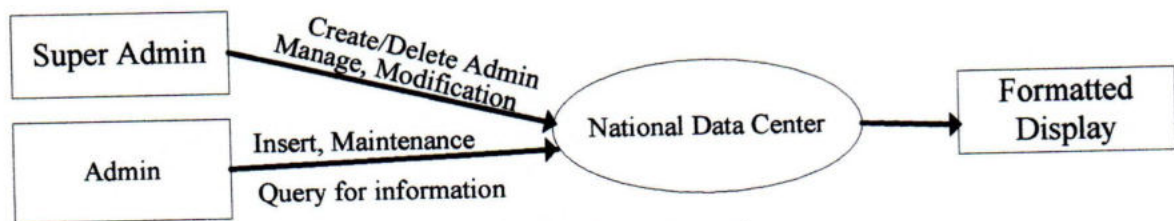


Fig. 3.3: level 0 Data Flow Diagram

3.3.2 Level-1 Data Flow Diagram

Level-1 DFD shows the systems major processes, data flows and data stores at a second high level of abstraction. Level -1 DFD must balance with the Level-0 it describes. Input going into a process is different from outputs leaving the process.

Level 1 DFD Expanding NDC

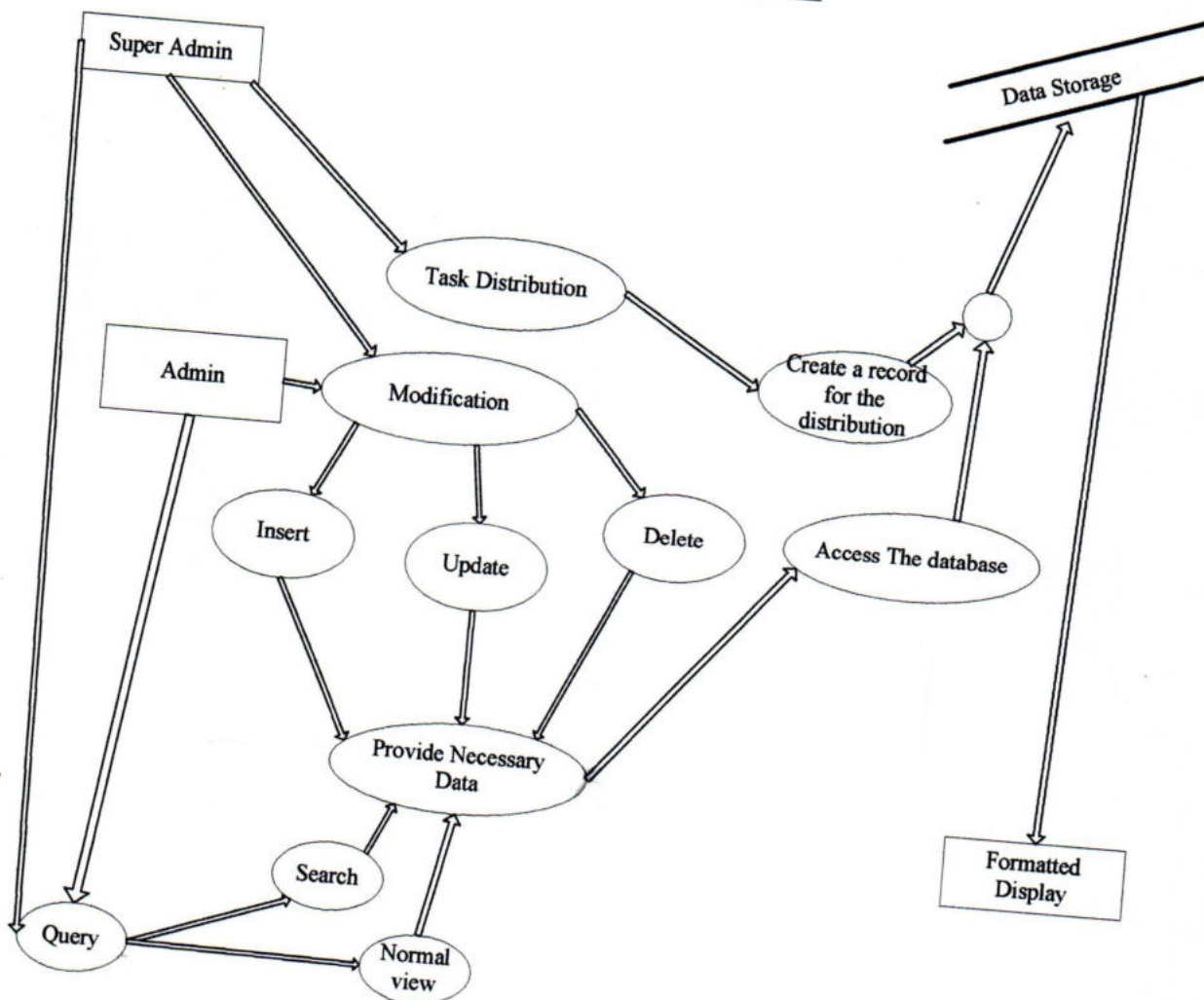


Fig. 3.4: level 1 Data Flow Diagram

3.3.3 Level-2 Data Flow Diagram

Level-2 DFD shows a process broken down into greater detail. Level-2 diagrams are only necessary where the level-1 process is more complex, and where the particular process is relevant to the analysis.

Level 2 DFD Expanding only "Modification"

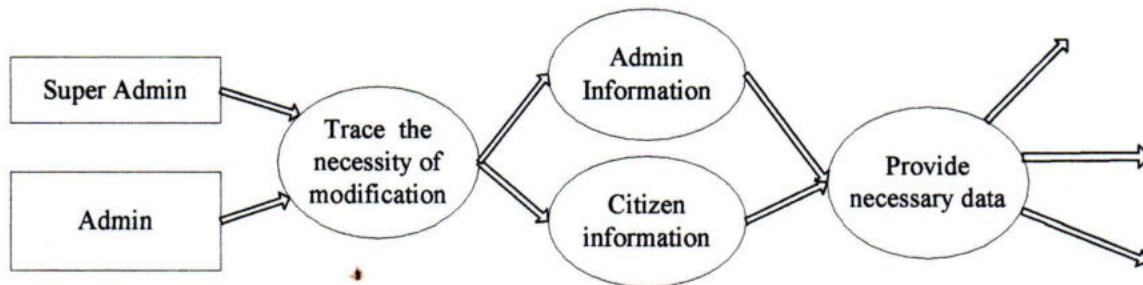


Fig. 3.5: level 2 Data Flow Diagram

3.4 Class Diagram

A class diagram is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. In this conceptual model we describe the basic concepts of the system and its relationships with each other. In this point we do not pay much attention to the system behavior. For determining the conceptual model we analyze the administrative, membership, quarry generation, modification tasks of the system. Different functionalities of the system can be accomplished by different classes. The classes used in designing the system are shown in Fig. 3.6.

3.5 E-R Diagram

The physical objects of the system; the admins, citizens, and locations – correspond to entities in the Entity-Relationship model, and the operations to be done on those entities – holds, checkouts, and so on – correspond to relationships. After some considerations, we have decided on the following design, which minimizes the redundancy in the stored information:

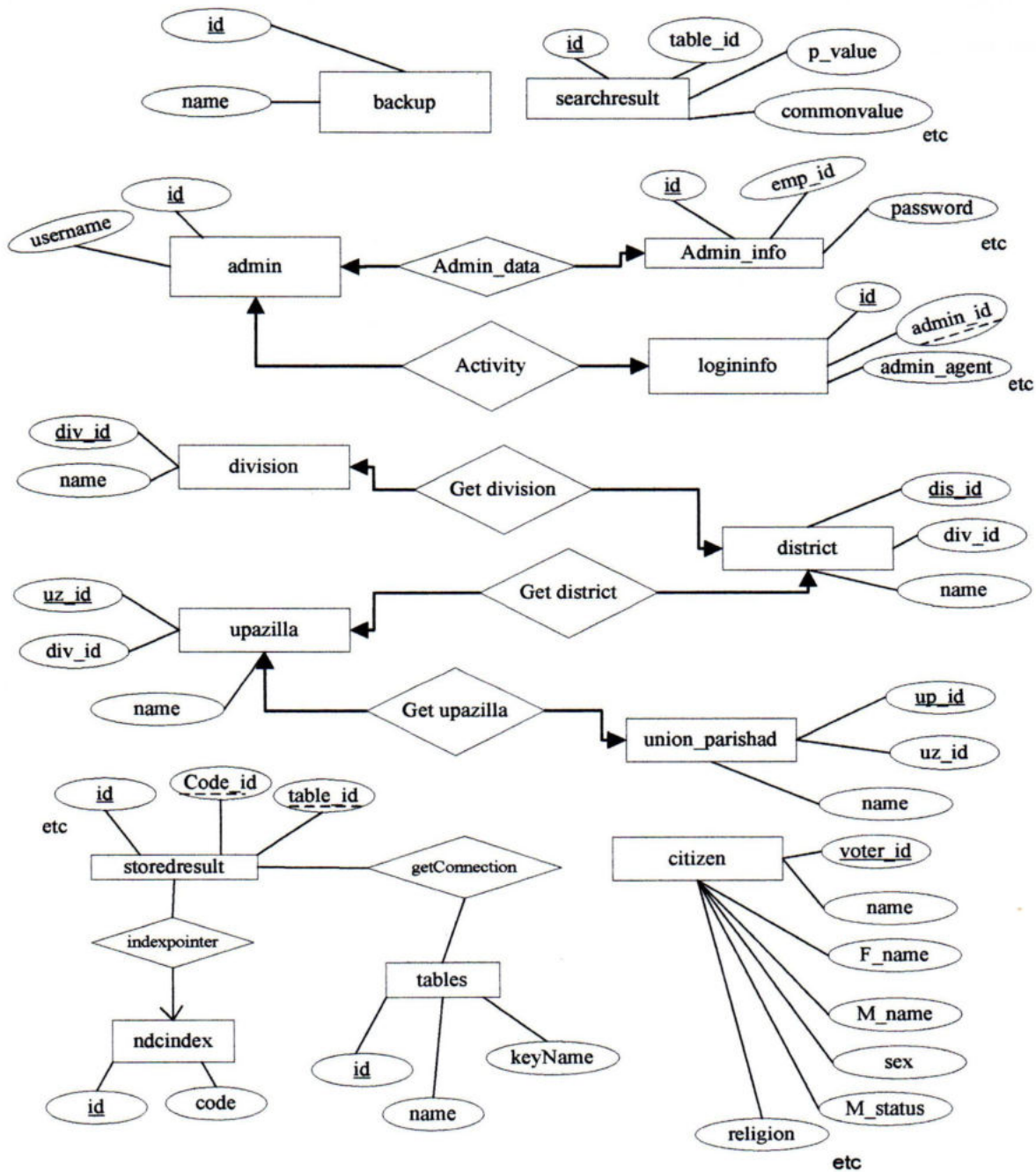


Fig. 3.7: ER Diagram of NDC

3.7 Data Dictionary

The data dictionary is a structured repository of data about data. It is a set of rigorous definitions of all DFD data elements and data structures. The National Data Center of Bangladesh will have the following Data Dictionary for storing the data belonging to the system.

Field	Type	Null	Default	Comments
div_id	int(2) unsigned	No		Identity of division (primary key)
div_name	varchar(10)	No		Name of the division

Table 3.1: **division**

Field	Type	Null	Default	Comments
dis_id	int(2) unsigned	No		Identity of district (primary key)
div_id	int(2) unsigned	No		Division ID to which the district belongs
dis_name	varchar(13)	No		Name of the district

Table 3.2: **district**

Field	Type	Null	Default	Comments
uz_id	int(3) unsigned	No		Identity of upazilla (primary key)
dis_id	int(2) unsigned	No		District ID to which the upazilla belongs
uz_name	varchar(16)	No		Name of the upazilla

Table 3.3: **upazilla**

Field	Type	Null	Default	Comments
up_id	int(3) unsigned	No		Identity of union_parishad (primary key)
uz_id	int(3) unsigned	No		District ID to which the union_parishad belongs
uz_name	varchar(16)	No		Name of the union_parishad

Table 3.4: **union_parishad**

Field	Type	Null	Default	Comments
Id	int(20) unsigned	No		Identity of admin (primary key)
userid	varchar(32)	Yes	Null	Name of the admin
md5_pw	varchar(32)	Yes	Null	Password of the admin
login	tinyint(1)	Yes	Null	Is login or not?
status	tinyint(1)	Yes	Null	Is active or not?
category	int(1)	Yes	Null	Super admin/admin

Table 3.5: **admin**

Field	Type	Null	Default	Comments
voter_id	bigint(20) unsigned	No		voter_id is Primary key
name	varchar(30)	yes	Null	Name of the citizen
f_name	varchar(30)	yes	Null	Father name of the citizen
m_name	varchar(30)	yes	Null	Mother name of the citizen
dob	date	yes	Null	Date of birth
sex	varchar(8)	yes	Null	Sexual definition
religion	varchar(10)	yes	Null	Religion of the citizen
marital	varchar(16)	yes	Null	Marital status of the citizen
s_name	varchar(30)	yes	Null	Spouse name of the citizen
occupation	varchar(15)	yes	Null	Occupation of the citizen
division	varchar (16)	Yes	Null	Current living division
district	varchar (16)	yes	Null	Current living district
upazilla	varchar (16)	yes	Null	Current living upazilla
union_parishad	varchar (16)	yes	Null	Current living union_parishad
ward	int(2)	yes	Null	Current living ward
village	varchar(20)	yes	Null	Current living village
contact	varchar(50)	yes	Null	Contact information
image	varchar(500)	yes	noimage	Picture of the citizen

Table 3.6: **citizen**

Field	Type	Null	Default	Comments
id	bigint(20) unsigned	No		Id number is the primary key
codeId	int(20) unsigned	No		Identity of Metaphone Code.
tableId	int(5)	yes	NULL	Identity of the data stored table
p_value	int(20)	yes	NULL	Value of the primary key.
commonvalue	varchar(500)	yes	NULL	The value for which matching is occurred.

Table 3.11: storedresult

Field	Type	Null	Default	Comments
Id	Tinyint(20) auto_increment	No		Id number is the primary key
Date	Date	Yes	Null	Backup creation date
name	varchar(50)	No		Backup database file name

Table 3.12: backup

3.8 Activity Diagram

Activity diagram is basically a flow chart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The activity at different level of the system can be shown in the following diagram. The facts of the system explained by the following diagram are:

1. Every user of the system requires logging in providing valid user name and password.
2. The number of tries (in case of wrong user name or password) is limited.
3. After log in the system provides appropriate facilities for super admin or admin.
4. If the idle time of any user crosses the limit the system logout automatically.

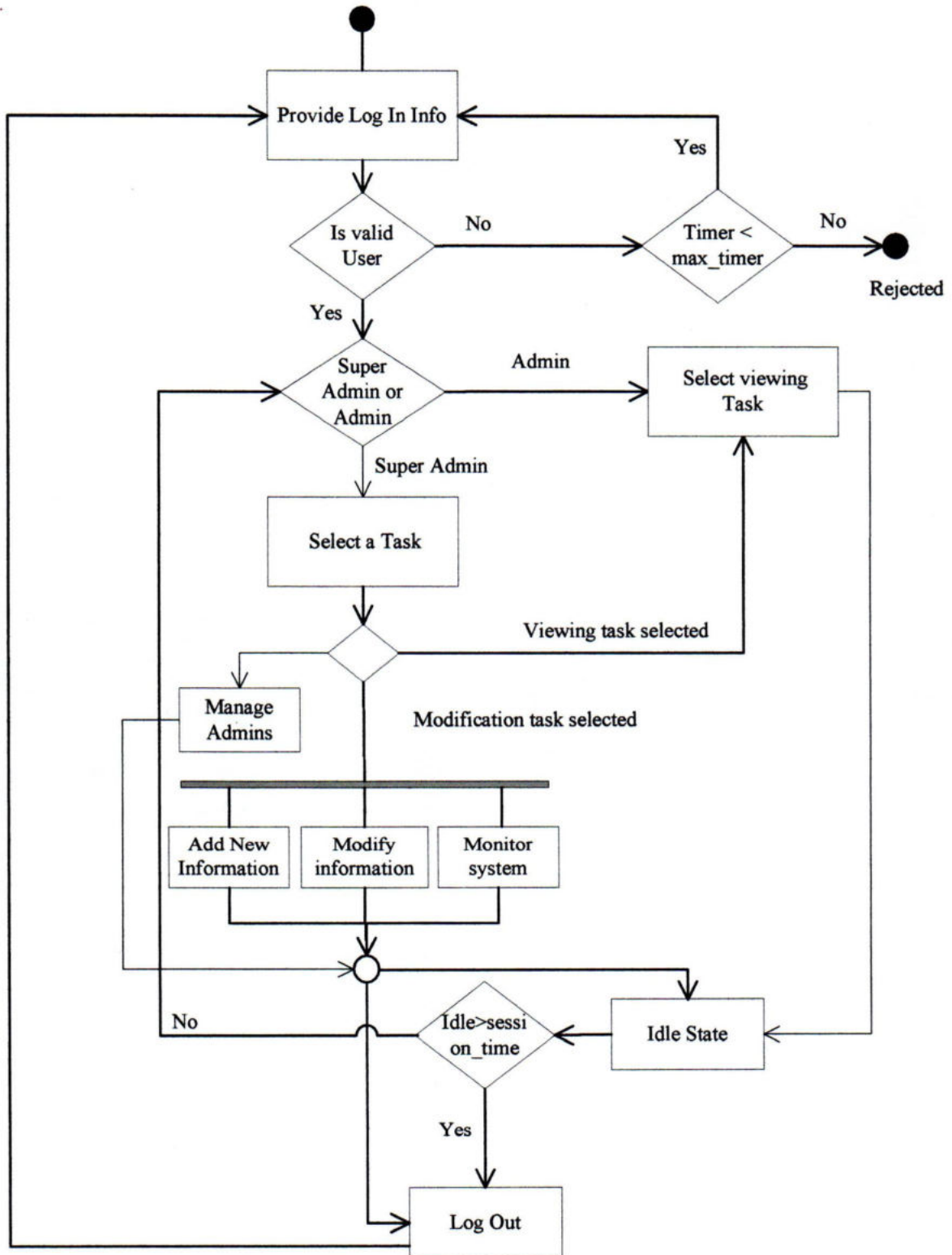


Fig. 3.9: Activity Diagram of NDC

3.9 Swimlane Diagram

A swimlane is a visual element used in process flow diagrams or flowcharts that visually distinguishes responsibilities for sub-processes of a business process. It is a useful variation of the activity diagram and allows the modeler to represent the flow of activities described by the use-case and the same time it describes which actor (if there are more than one actors) is responsible for a action described by the activity diagram.

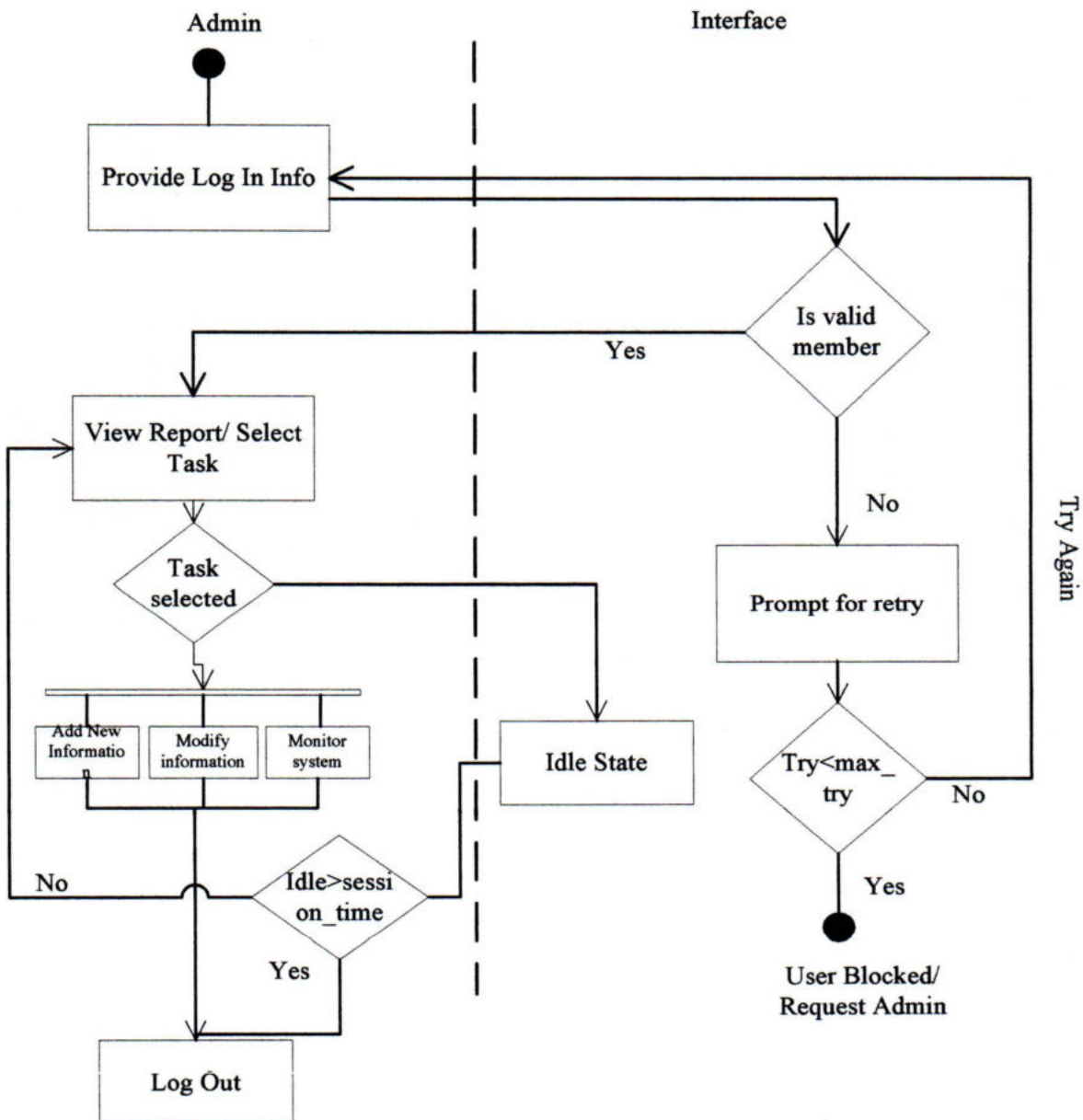


Fig. 3.10: Swimlane Diagram of **admin Vs interface**

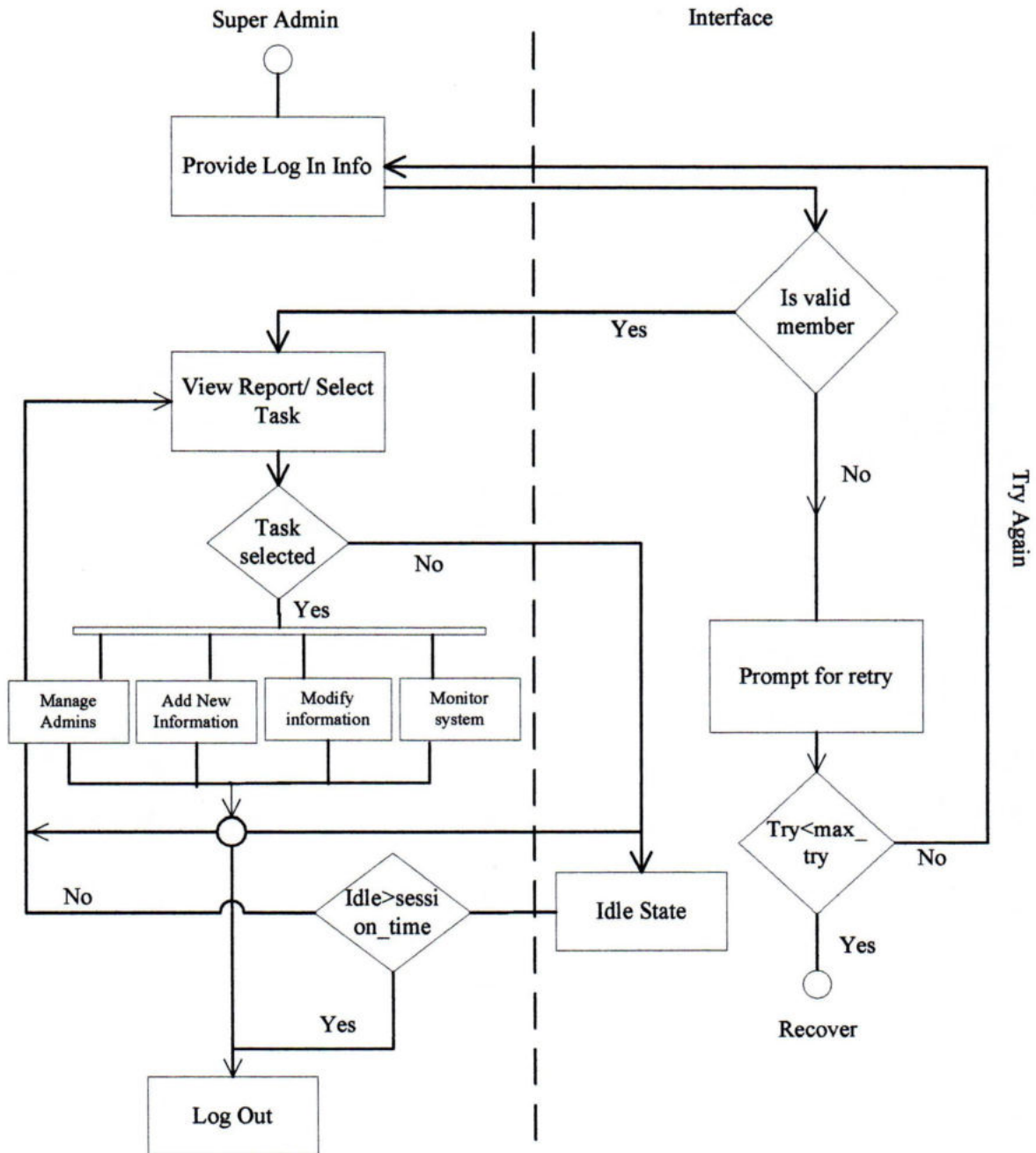


Fig.3.11: Swimlane Diagram of super admin Vs interface

3.10 State Diagram

State diagram requires that the system described is composed of a finite number of states; sometimes, this is indeed the case, while at other times this is a reasonable abstraction. It shows how the object transactions from an initial state to other states, when certain events occur or when certain conditions are satisfied. It represents dynamic models of objects which changes their states in response to events. It is a model of the states of an object and the events that cause the object to change from one state to another. Different states can be represented by a rectangle with rounded corners. In figure we have shown different states of citizen info object such as user entrance, insertion form, database accessing, searching result and getting results.

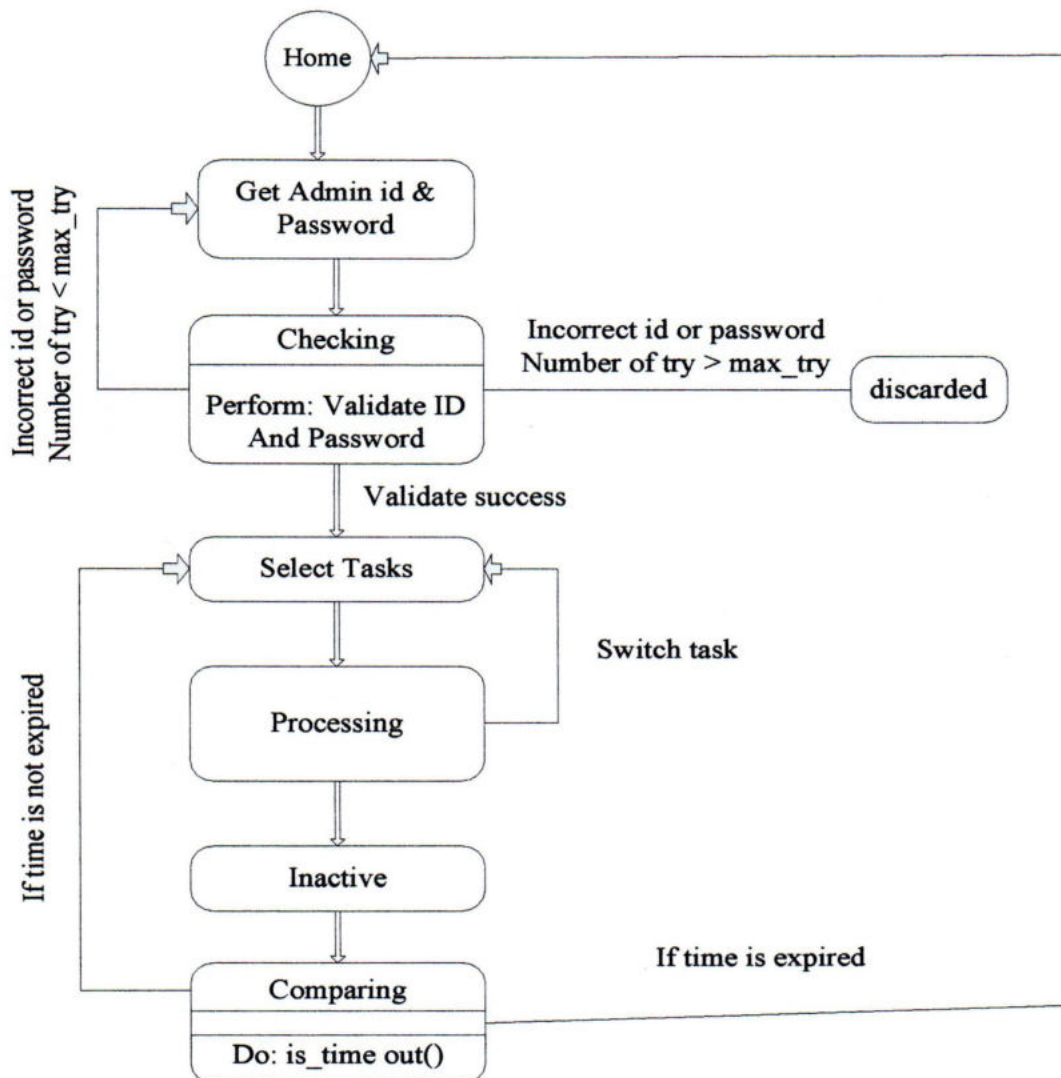


Fig. 3.12: State Diagram of NDC

3.11 Architecture Diagram

Architecture Diagram of NDC Designed for Bangladesh shows the different components of the system. We have populated the central database of the system by MySQL database and populated it in the MySQL Essential-5.0.67. The web server used is Apache 2.2.10. Our system designed by HTML, PHP, JavaScript, and CSS provides an interface with the user using both the servers. This interaction requires a network among the computer from which the personnel will provide their requirements.

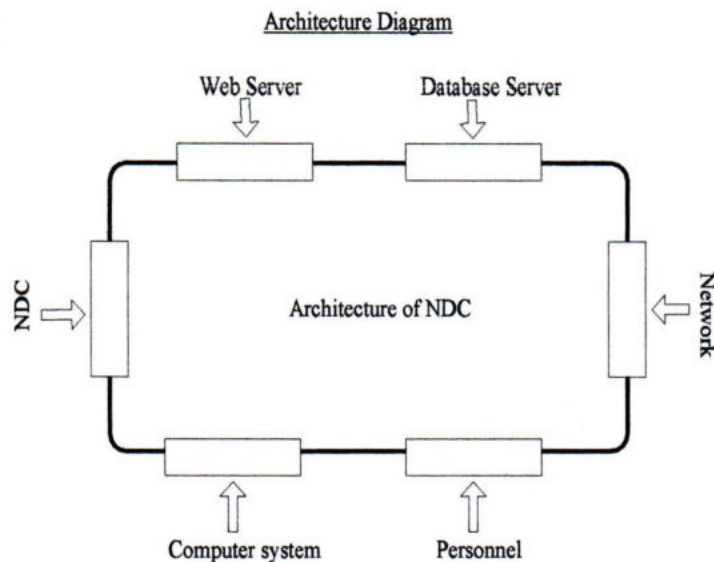


Fig. 3.13: Architecture Diagram of NDC

3.12 Feasibility Considerations

The main objective of Feasibility Study is documenting each of the potential solutions of a system to identify the likelihood of one or more solutions meeting the stated requirement. In its simplest terms, the two criteria to judge feasibility are cost required and value to be attained. In feasibility study phase of National Data Center we have undergone through different Level of Analysis and designed a web-based System. Our designed system shows following feasibility considerations:-

3.12.1 Technical Feasibility

The technologies used in developing the candidate system are well-proven technology and there will be no difficulty on implementation, because the required technology, resources, manpower, programmers, testers & debugger, software and hardware all are available.

3.12.2 Economic Feasibility

Economic Feasibility means the project benefit of the proposed system outweigh. For any system if the expected benefits equal or exceed the expected costs, the system can be judged to be economically feasible. Economic analysis is used for evaluating the effectiveness of the proposed system. The implementation of our system is economically feasible as it provides:

- greater reduction on operational cost
- saving in salaries
- better service time
- Improved services

3.12.3 Operational Feasibility

A system that has operational feasibility is one that will be used effectively after it has been developed. Our Candidate System is Operationally Feasible because there would be no operational difficulties with adequate training and it will be accepted by the user.

In feasibility study phase we had undergone through various steps which are describe as under:

1. Identify the origin of the information at different level.
2. Identify the expectation of user from computerized system.
3. Analyze the drawback of existing system (manual) system.

4.1 Objective

To fulfill the requirement of the implementation of NDC we have used a search methodology which is based on **Double Metaphone Algorithm**. The algorithm is very useful for this purpose because it eliminates the problem during search due to **misspelling** of given keyword and increasing the speed of search. The Double Metaphone phonetic encoding algorithm is the second generation of the Metaphone algorithm. Its implementation was described in the June 2000 issue of C/C++ Users Journal by Lawrence Philips. It makes a number of fundamental design improvements over the original Metaphone algorithm. It is called "Double" because it can return both a primary and a secondary code for a string; this accounts for some ambiguous cases as well as for multiple variants of surnames with common ancestry. For example, encoding the name "Smith" yields a primary code of SM0 and a secondary code of XMT, while the name "Schmidt" yields a primary code of XMT and a secondary code of SMT--both have XMT in common [5].

4.2 Previous Work on Double Metaphone Algorithm

Double Metaphone algorithm is mainly designed for finding the phonetic similarity between two words. It is implemented in several applications such as:

- Naushad UzZaman and Mumit Khan Designed **A Bangla Phonetic Encoding for Better Spelling Suggestions** using this algorithm. [2]
- Chakkrit Snae and Michael Brückner Designed **Novel Phonetic Name Matching Algorithm with a Statistical Ontology for Analyzing Names Given in Accordance with Thai Astrology**. [3]
- A.K. Mandal, M.D. Hossain and M. Nadim designed **Developing an efficient search suggestion generator, ignoring spelling error for high speed data retrieval using Double Metaphone Algorithm**[4]

4.3 Why Double Metaphone Algorithm?

For minimizing the searching comparisons fuzzy indexing is required, and Double Metaphone algorithm is chosen for this purpose. The Available choices are None, Stemmer Algorithm, Soundex Algorithm, Metaphone Algorithm, Double Metaphone Algorithm etc.

4.5 Interfaces between search result and the database

The interface between search result and the database is maintained by introducing data pointers. Data pointer is a mean by which a particular row of a table can be identified. For example if we know value of the primary key of a particular table then we can easily get the value of the specific row. Thus during the searching phase it not needed to originally search the whole database, just getting the appropriate data pointer and retrieve the original data located by the pointer. The process of the creation of the data pointer is described bellow:-

1. The entire tables whose information may be required to search are given unique id.
2. Each Data Pointer comprises with table id and primary key value of the row.

The tables and id used in the database search of National data Center are as follows:

Table Name	Purpose of the Table	ID
citizen	Storing the information of citizen	0

Table 4.2: Defining informative tables in Search Engine

Thus if we get a Data Pointer (0, 20) thus following SQL query will be performed for getting the data from the table:

```
select * from citizen where p_value=20;
```

4.6 Purposes of the data tables used in search methodology

1. Table ***ndcindex***, stores the Metaphone Code of each words, is indexed against keycode, for ensuring high speed retrieval of id.
2. Table ***tables*** is defined by using two dimensional array.
3. Table ***storedresult***, stores the map of original database, is indexed against tableid and codeid for ensuring high speed navigation of exact location of the searched value.
4. Table ***searchresult*** is a temporary storage of searched result, and it will be updated with every search operation.

The above tables are introduced to make the search methodology easy and efficient.

4.8 Explanation with Example

The searching process for a particular input (Here spelling error occurred and user desiring output for **Rahim Dinajpur**) will be performed as follows:

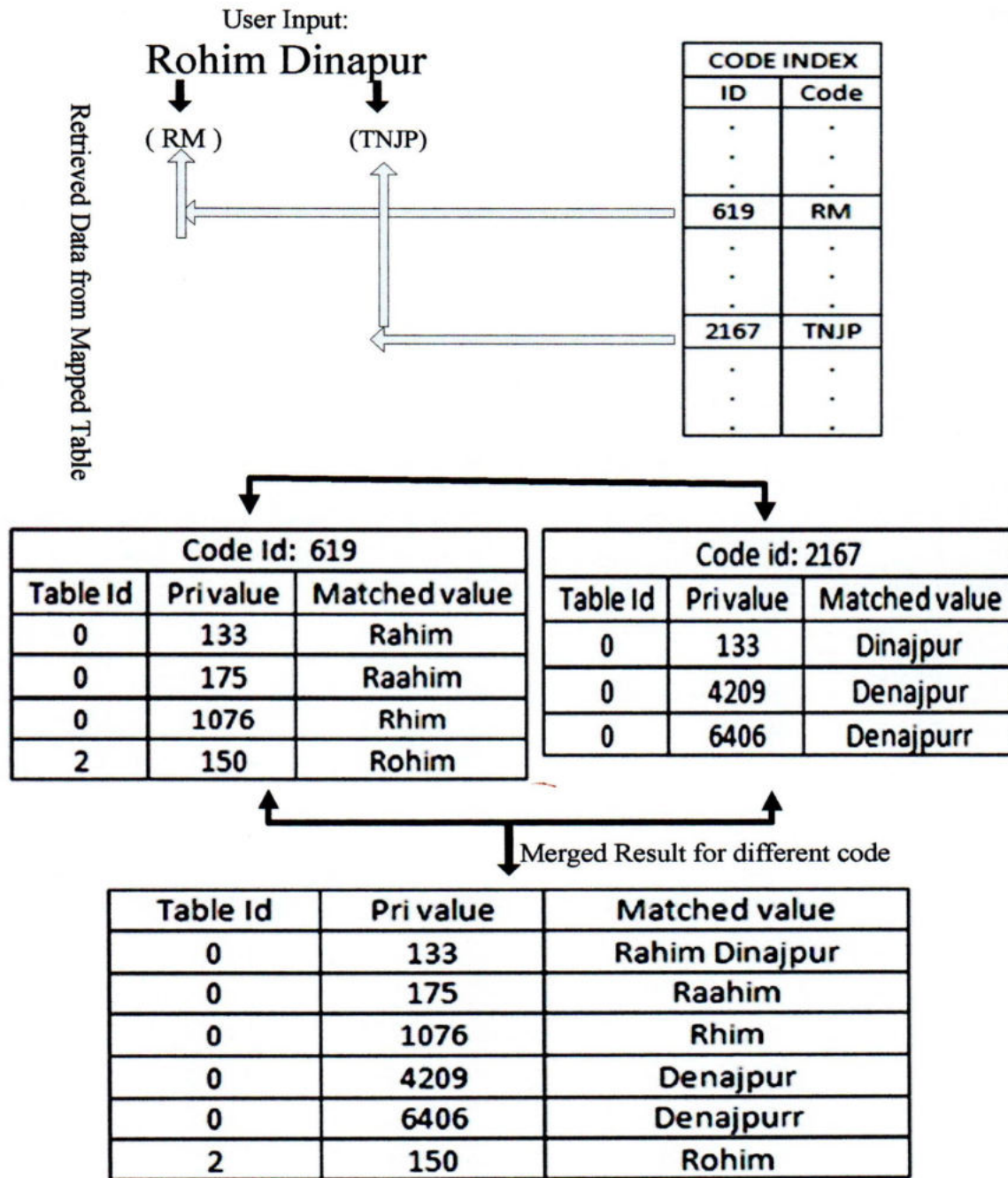


Fig. 4.2: Explanation by Example of Search Engine

4.9 Performance Analysis

The linear search would have a time complexity of $O(n)$ in the worst case. However, the time complexity of the Binary search method is $O(\log n)$, which is much more efficient than the Linear Search.

The algorithm used to implement the search engine shows excellent performance in such a situation when the input keywords are misspelled. Because the algorithm not only implements binary search in such situation but also reduces the search domain depending on the input keywords.

We have performed an experiment by taking random citizen information, where increase of comparison is expressed with the increase of number of citizen.

4.10 Load analysis

In the phase of load analysis we examined several parameters including number of inserted data, time required to search a specific record from database, time required to create a backup and time required to import record from files such as .xls and .txt.

The amount of data versus time required to search the database is plotted here:

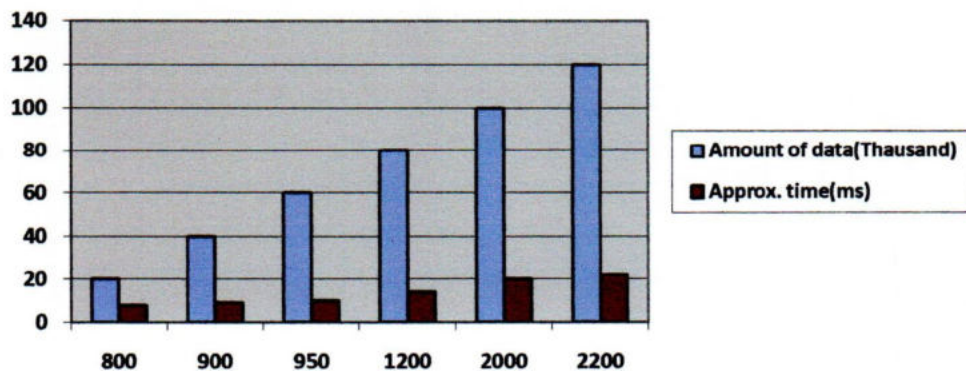


Fig. 4.3: Amount of data vs. time graph.

4.11 Decisions from the analysis

Observing the above figure we can easily conclude that, if the rate of number of data increases, the searching time increases slowly or not. Thus the speed of Metaphone searching technique is very high than the normal linear search technique

CHAPTER 5

IMPLEMENTATION AND TESTING

Implementation And Testing

5.1 Tools and Technology

The tools denote the components that are used to create the environments for the programmers. The technology describes the way to code a particular project. In the implementation of this project the following tools and technology are used:

Languages: HTML, PHP, CSS, JavaScript

Database: MYSQL,

Web server: Apache

Tools:

- Macromedia Dreamweaver 8, Notepad++
- Apache HTTP Server
- PHP 5.2.6
- MySQL Server 5.0

5.1.1 Requirement Analysis

Hardware Requirement
1. Processor- Pentium-IV
2. RAM- 512 MB
3. Disk Space- 200 MB
Software Requirement
1. Operating system- Windows XP
2. Server – Apache
3. Database – MYSQL

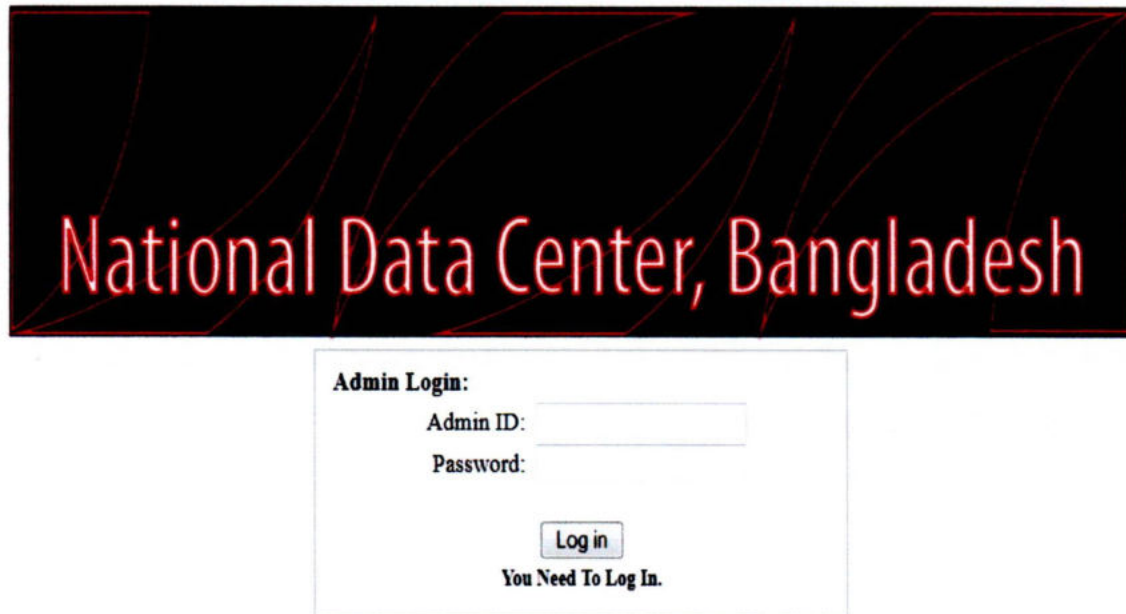
Table 5.1: Hardware and Software requirements

5.2 Testing

Testing is an operation to detect the difference between the expected result and the actual result. Working of the system is tested by configuring Apache Server along with PHP, MySQL database Server. Proper Modification is made in different level of implementation.

5.3 Home page of NDC

The home page of NDC designed by HTML, CSS codes is as follows:



National Data Center, Bangladesh

Admin Login:

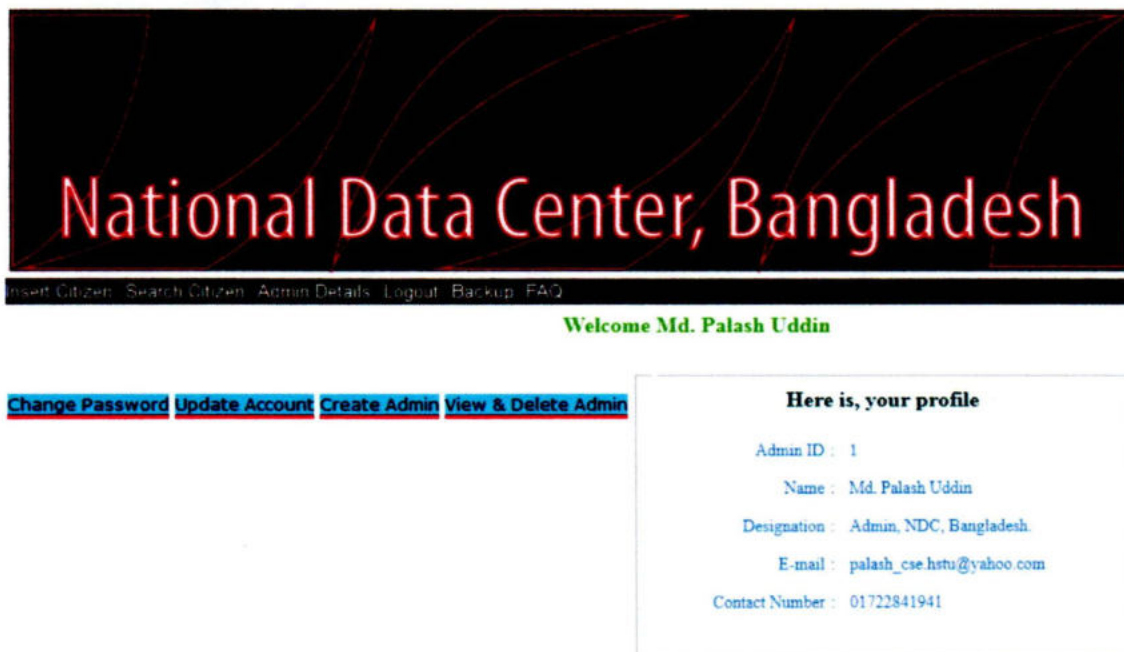
Admin ID:

Password:

You Need To Log In.

Fig. 5.1: Home page of NDC

5.4 Admin Details page of NDC



National Data Center, Bangladesh

[Insert Citizen](#) [Search Citizen](#) [Admin Details](#) [Logout](#) [Backup](#) [FAQ](#)

Welcome Md. Palash Uddin

[Change Password](#) [Update Account](#) [Create Admin](#) [View & Delete Admin](#)

Here is, your profile

Admin ID : 1

Name : Md. Palash Uddin

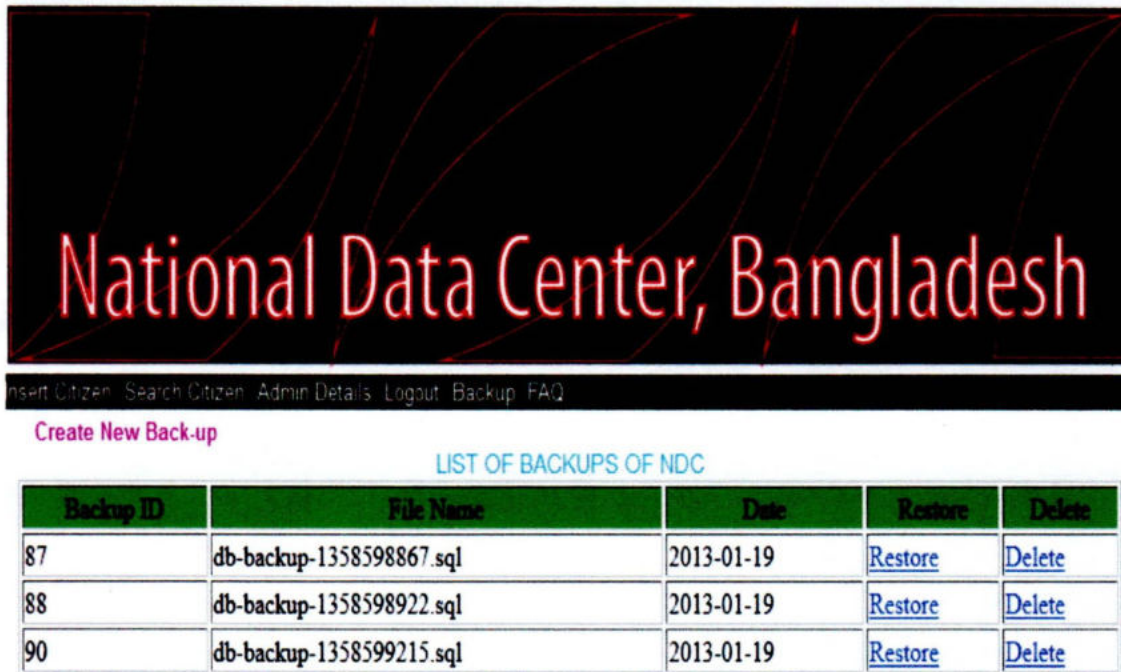
Designation : Admin, NDC, Bangladesh

E-mail : palash_cse.hstu@yahoo.com

Contact Number : 01722841941

Fig. 5.2: Admin Details page of NDC

5.8 Backup page of NDC



National Data Center, Bangladesh

Insert Citizen Search Citizen Admin Details Logout Backup FAQ

[Create New Back-up](#)

LIST OF BACKUPS OF NDC

Backup ID	File Name	Date	Restore	Delete
87	db-backup-1358598867.sql	2013-01-19	Restore	Delete
88	db-backup-1358598922.sql	2013-01-19	Restore	Delete
90	db-backup-1358599215.sql	2013-01-19	Restore	Delete

Fig. 5.6: Database Backup page of NDC

CHAPTER 6

CONCLUSION AND FUTURE PLAN

Conclusion and Future Plan

6.1 Conclusion

Since Bangladesh is a developing country its faster development totally depends on how much technologies she can use in an effective way. In other words, how fast her organizations are being digitalized. Bangladesh is still using the manual system for storing and retrieving the information of its citizens. There a lot of time and resources are being wasted. To save that time and resource we have tried to develop this web-based NDC. If this system will be used, we guarantee that the government of Bangladesh must have a big savings. Again analyzing the proposed system that was performed in the design and implementation process of NDC, it can be concluded that it is a highly efficient **GUI** (Graphical User Interface) based system. We have tried our level best to make the system most efficient, user-friendly and effective for the fulfillment of the users' requirement. When preparing the design and implementation for this system, our goal was to replace the manual system by computerized one, which implements a faster search technique that uses a phonetic algorithm called **Double Metaphone Algorithm**. The algorithm is able to trace misspelled keywords and for that the proposed system still shows relevant search results which make the system more comfortable. Using this algorithm the searching time for a particular citizen stored in the NDC is reduced much more. In our analysis we have used advanced techniques to prepare a convenient design to make the implementation phase easy. Therefore we hope that this system will appear in a good case. The new system can be integrated with other datacenter in the world to get full benefit of modern web technology.

6.2 Future plan

Most of the project and research might have certain drawbacks because it's uncompleted caused by lack of time and proper resource. The NDC itself is not free from some drawbacks. If we would have sufficient time and proper resource, we could cover the drawbacks. In future we effort to extend and integrate the following features:

Online Public Access: Online public access will make the system information updated. But in that case some rules and regulations are applied.

Biometric Information: In future some more data will be added to enhance the system for this purpose the biometric information such as finger print and retina information will provide with great impact on the system.

Linked up with Govt. Organizations: If the system is linked up with the government organizations than at the time of recruitment for govt. jobs the organization can easily trace the person and get proper information and at the same time citizen doesn't need to provide his/her biodata again.

Provide National ID: If someone lost his/her National ID card in that case he/she will have to show a GD report and after that he/she can download their ID card.

Image Search: In future, Search by image will be added to our system. This will be quite easy to find out the right person.

References:

- [1] **Bangladesh starts National Population Register Project**
<http://www.futuregov.asia/articles/2011/jul/14/bangladesh-starts-national-population-register-pro/>
- [2] **A Bangla Phonetic Encoding for Better Spelling Suggestions** By Naushad UzZaman and Mumit Khan, BRAC University, Dhaka, Bangladesh.
naushad@bracuniversity.ac.bd, mumit@bracuniversity.ac.bd
- [3] **Novel Phonetic Name Matching Algorithm with a Statistical Ontology for Analysing Names Given in Accordance with Thai Astrology** By Chakkrit Snae and Michael Brückner Faculty of Science, Naresuan, University, Phitsanulok, Thailand,
chakkrits@nu.ac.th; michaelb@nu.ac.th
- [4] **Developing an efficient search suggestion generator, ignoring spelling error for high speed data retrieval using Double Metaphone Algorithm** By A.K. Mandal, M.D. Hossain and M. Nadim, Faculty of Computer Science & Engineering, Hajee Mohammad Danesh Science & Technology University, Dinajpur, Bangladesh. Journal of Science & Technology, Vol. 8, pp 30-36, June 2010, ISSN 1994-0386
http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5723876&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D5723876
- [5] **The Double Metaphone phonetic encoding algorithm**
<http://en.wikipedia.org/wiki/Metaphone>

Books, External Links, Web-sites

Books:

1. Beginning PHP5, Apache, and MySQL® Web Development By: Elizabeth Naramore, Jason Gerner, Yann Le Scouarnec, Jeremy Stolz, Michael K. Glass
2. MySQL Cookbook By: Paul DuBois
3. Professional JavaScript™ for Web Developers. By: Nicholas C. Zakas
4. The Absolute Beginner's Guide To Starting a Web Site. By: Milana Nastetskaya
5. PHP/MySQL Tutorial by Graeme Merrill
6. Beginning CSS: Cascading Style Sheets for Web Design By: Richard York

External links

- [The Soundex Indexing System](#) (U.S. National Archives and Records Administration)
- [Complete PHP String Reference](#)
- [Text::Soundex Perl](#) module from [CPAN](#)
- [PHP soundex function](#)
- [CString Management](#) Learn how to effectively use CStrings.

Web Sites:

1. <http://www.apache.org/>
2. <http://www.php.net/>
3. <http://www.mysql.com/>
4. <http://www.techotopia.com/>
5. <http://www.tutorialized.com/>
6. <http://www.pagetutor.com/>
7. <http://www.w3schools.com/>
8. <http://www.tech-funda.com/>
9. <http://www.chipchapin.com/>
10. <http://www.accessify.com/>