A HEALTH DECISION SUPPORT SYSTEM FOR RURAL INDIA

A THESIS SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology
In
Computer Science and Engineering

By

Sumeet Priyadarshee Dash
Roll No: 10306022
&
Supreet Padhi
Roll No: 10306008

Department of Computer Science and Engineering
National Institute of Technology, Rourkela

May, 2007

A HEALTH DECISION SUPPORT SYSTEM FOR RURAL INDIA

A THESIS SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology

In

Computer Science and Engineering

By

Sumeet Priyadarshee Dash

Roll No: 10306022

&

Supreet Padhi

Roll No: 10306008

Under the guidance of

Prof. B. D. Sahoo

Department of Computer Science and Engineering

National Institute of Technology, Rourkela

May, 2007

National Institute of Technology
Rourkela

CERTIFICATE

This is to certify that the thesis entitled, "A HEALTH DECISION SUPPORT SYSTEM FOR RURAL INDIA" submitted by Sumeet Priyadarshee Dash and Supreet Padhi in partial fulfillment of the requirements for the award of Bachelor of Technology Degree in Computer Science and Engineering at the National Institute of Technology, Rourkela (Deemed University) is an authentic work carried out by them under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other university / institute for the award of any Degree or Diploma.

Date:

Prof B. D. Sahoo
Dept. of Computer Science and Engineering
National Institute of Technology, Rourkela
Rourkela - 769008

ACKNOWLEDGEMENT

We are sincerely thankful to **Prof B.D. Sahoo,** department of Computer Science and Engineering, NIT Rourkela for giving us the opportunity to work under him and lending every support possible at every stage of this project work. The level of flexibility offered by him in implementing the project work is highly applaudable.

We would also like to convey our sincerest gratitude and indebtedness to all other faculty members and staff of the Department of Computer Science and Engineering, NIT Rourkela, who bestowed their great effort and guidance at appropriate times without which it would have been very difficult on our part to finish the project work.

Date: Sumeet Priyadarshee Dash

Date: Supreet Padhi

Abstract

In India and other developing nations, the bulk of the morbidity & mortality is due to commonly occurring communicable diseases & parasitic diseases coupled with malnutrition especially in rural areas. The effective decision making at the top level of any health services mostly depends on the availability of various resources such as human expertise, equipments, and medicine. People die from infectious and/or chronic diseases are the leading causes of death, especially in rural areas. By analyzing mortality, morbidity, and behavioral data, one can attempt to quantify health problems and the behavioral risk factors that contribute to them Hence in a country like India an effective multi disease surveillance system is essential for the General Health Care System to detect an outbreak, monitor the trend, prevent an epidemic & decrease the morbidity & mortality rate of India The proposed DSS is targeting to assist the top management of the State health service which will provide a practical, relatively inexpensive and replicable model of disease surveillance.

The proposed system consists of application and management software that support clinical and operational data. The software is designed for multi-site use in individual medical facilities and health workers in remote villages. The disease surveillance data is collected and updated periodically by the health workers to the central database through SMS. This disease surveillance system through SMS will provide real time data and extract the statistical and customized information and even facilitate the prediction of the outbreak of epidemics and report emergencies. It also provides an automatic response messaging through SMS to people regarding basic preventive measures and cures. A smart phone application is built using J2ME which make data transmission error free and secured. The use of SMS as the mode of data transmission will help reduce bureaucratic delays and will automate the task of disease surveillance by providing an inexpensive replacement to the existing trend.

Contents

List of Abbreviations

DSS	Decision Support System
SMS	Short Messaging Service
HQ	Head Quarter
GSM	Global Standard for Mobile Communication
GPRS	General Packet Radio Service
NGO	Non-Government Organization
PHC	Primary Health Center
CHC	Community Health Center
SDH	Sub-divisional Hospital
DHH	District Head Hospital
OHMIS	Orissa Health Management Information System
FCT	Fixed Cellular Terminal
PDU	Portable Data Unit
J2ME	Java 2 Micro Edition
LBS	Location Based Service
BTS	Base Transceiver Station
KB	Knowledge Base
HW	Health Worker

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 INTRODUCTION

In India and other developing nations the bulk of the morbidity & mortality is due to commonly occurring communicable diseases & parasitic diseases coupled with malnutrition. The spectrum of emerging and reemerging communicable diseases along with anti- microbial resistance poses a major threat to mankind. Incidence of non-communicable diseases like DM, Hypertension, IHD, Cancer, Mental Illness and accidents is also a major cause of concern. The emerging picture is a mixture of communicable & non-communicable diseases. The figure 1.1 compares India with some other countries in terms of certain health indicators. Hence in a country like India an effective Multi disease surveillance system is essential for the General health Care System to detect an outbreak, monitor the trend, prevent an epidemic & decrease the morbidity & mortality. The surveillance system needs to change to respond to the changing health situation in India. Surveillance is the foundation of an effective disease prevention and control program. It involves timely collection, analysis, and interpretation of health data, followed by dissemination of the findings to take action for prevention and control of diseases. Conversely, there are costs to not having an effective disease surveillance program in place. For example, the failure to detect and control the epidemic of plague in 1994 at its early stages resulted in tragic loss of life as well as disruption of trade, transportation, and tourism. Estimates of the total cost of the plague are as high as USD1.5 billion. Another recent example is the outbreak of SARS in Asia.

India in comparison with other countries

Indicator	India	China	USA	Sri Lanka	Thailand
IMR/1000 live-births	68	<30	2	8	15
Under-5 mortality /1000 live-births	87	37	8	15	26
Fully Immunized (%)	67	84	93	99	94
Births by skilled attendents	43	97	99	97	99
Health expenditure as % of GDP	4.8	5.8	14.6	3.7	4.4
Government share of Total Expenditure (%)	21.3	33.7	44.9	48.7	69.7
Government health spending to total governmentspending (%)	4.4	10	23.1	6	17.1
Percapita spending in international dollars	96	261	5274	131	321

Source: WHR, 2005

Fig 1.1

The proposed DSS is targeting to assist the top management of the State health service. The effective decision making at the top level of any health services mostly depends on the availability of various resources such as human expertise, equipments, and medicine. Information on disease and deaths are of vital importance for planning and evaluating community health promotion. People die from infectious and/or chronic diseases are the leading causes of death, especially in rural areas. By analyzing mortality, morbidity, and behavioral data, one can attempt to quantify health problems and the behavioral risk factors that contribute to them. As a result of this, Disease surveillance ought to be an important component of public health program in every country. It has two essential purposes. One is to monitor the progress of ongoing interventions for disease reduction. For example, the childhood immunization program is incomplete without surveillance for vaccine preventable diseases. Second, disease surveillance is essential for early detection of outbreaks in order to initiate investigations and control measures. Our model will provide a practical, relatively inexpensive and replicable model of disease surveillance.

One of the significant barriers to health outcomes in developing countries is the lack of access to medical expertise, a problem exacerbated by brain drain and made even more apparent in the shortage of specialists. The idea is to establish a central clinical knowledge base accessible to the general people which they can use to avail health care information at any time of need.

The prime intension of resolving an emergency situation is to mobilize necessary resources in a minimum time frame. In a case where a victim is not aware of the location of the incident, the emergency controller has no way but to have wide area scannings in search of the place of occurring. Our model is targeting to provide a location tracking system which will help the emergency controller to have information about the location of the emergency thereby facilitating it to fasten task of resource mobilization.

1.2 OBJECTIVES

Thus the objectives of our project include:

- To strengthen coordination at the state and district levels by using SMS as the provision of information from the rural hospitals and health workers to the district HQ.
- To provide an automatic response messaging through SMS to people regarding basic preventive measures and cures.
- To facilitate the auto-detection of the area concerned in case of an emergency through a GSM/GPRS location tracking module so that resources can be mobilized in a minimum time frame.
- To facilitate criteria based report generation to ease the task of decision making.
- To provide services to NGOs and private organizations so that they will be encouraged to report to district surveillance units, which in turn will provide them with important feedback, motivating them further to be part of the plan.

Chapter 2

CASE STUDY OF THE EXISTING MODEL: A COMPARATIVE ANALYSIS

2.1 EXISTING APPROACH

The need for a *disease surveillance system* that fulfills the crucial criterion of surveillance - "information for action", is indispensable for every state to monitor the events in the community, take corrective measures & minimize the suffering of the people.

We hereby provide a case study of the health care system in the state of Orissa to facilitate understanding of existing measures. The entire state is covered by Health infrastructure like Health Sub- centers, Primary Health Centers, Referral hospitals, District Head Quarter hospitals. The existing surveillance approach uses predominantly rural based general health care infrastructure.

Data on morbidity and mortality is collected under two age groups (< 5 & >= 5 Yrs). For morbidity only the new cases are taken in to consideration. Simple and uniform format is used for data collection at all level. Data is collected from the community by the health worker while discharging their routine field activity during the week. Data from OP and IP registers of the Sector PHC, block PHC, SDH & DHH are collected weekly. The data is compiled at 4 levels - Sector PHC, Block PHC, District HQ, and State HQ. Up to District HQ level all the work is done manually. At the District HQ the data are entered in OHMIS & transmitted through email or Fax to Disease Surveillance Cell at State HQ. Analysis of the data at sector & block level is under taken by the concerned Medical officer.

The data is neither complete nor is the transmission regular and hence the information is largely for archival purpose only and not suited for action.

2.2 OUR APPROACH

The proposed system consists of application and management software that support clinical and operational data. The software is designed for multi-site use in individual medical facilities and health workers in remote villages. At the district level it will facilitate services like Event Driven Reporting, Patient Care Encounter, Patient Activity by Clinic, Caseload Profile by Clinic, Workload by clinic, Diagnosis ranked by frequency, Patient Activity by location and Remote medicine supply etc.

The disease surveillance data is collected and updated periodically by the health workers at PHC level to the central database located at district HQ through SMS. This disease surveillance system through SMS will provide real time data and extract the statistical information and even facilitate the prediction of the outbreak of epidemics and report emergencies.

The use of SMS as the mode of data transmission will help reduce bureaucratic delays and will automate the task of disease surveillance by providing an inexpensive replacement to the existing trend.

With the use of SMS messages we obtain the following benefits.

- SMS Messages can be sent and read at any time.
- SMS Messages can be sent to an offline mobile phone.
- SMS Messages are confirmed on delivery.
- SMS Messaging is less disturbing while you can still stay in touch.
- SMS Messages are supported by 100% GSM Mobile Phones and they can be exchanged between different wireless carriers.

In order to facilitate general people to avail help regarding the preventive measures of some class of diseases the system will maintain a central clinical knowledge base which will be made accessible to the general people through the use of mobile phones. The users can send their queries through SMS and the system will reply them with the necessary diagnosis measures. The system maintains the knowledge base in the form of a decision tree where non-terminal nodes represent tests/decisions on one or more attributes and terminal nodes reflect decision outcomes. Maintaining a tree structure helps reduce the effective searching time in the knowledge base.

The proposed system contains a GSM-based location tracking module which enables the auto-detection of the area under concern in the case of an emergency so that resources can be mobilized in a minimum time frame. This helps the emergency controller in efficient resolving of emergency situations.

Chapter 3

METHODOLOGIES

We now provide a list of methodologies we have come up with in order to fulfill our requirements.

3.1 USE OF GSM FCT DEVICE FOR MESSAGING

A GSM FCT router is used to build an SMS application that facilitates communication between a cellular device and a computer. The GSM FCT router is a “fixed cellular terminal” device that provides an option of connecting to a computer and use it for data applications. The device is a GSM modem which can be configured to send and receive SMS messages from a mobile device to computer via RS-232 serial port. The SMS communication is depicted in Figure 3.1

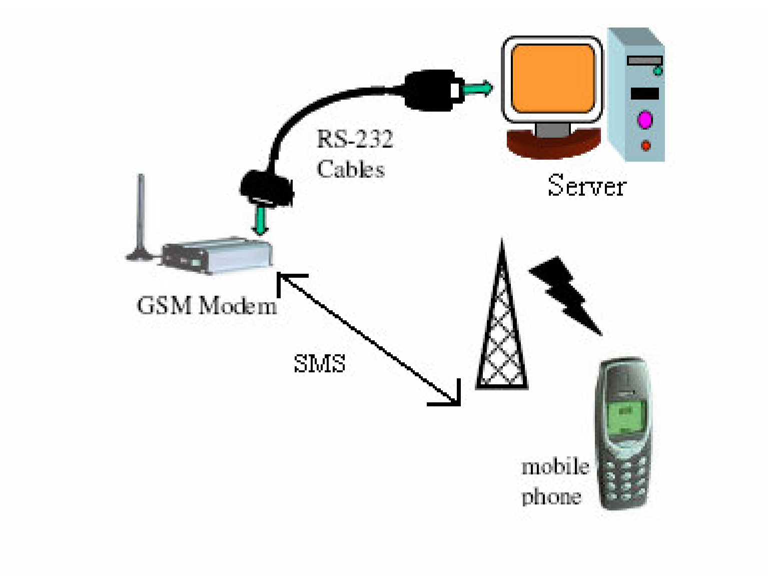


Fig 3.1

3.2 USE OF GSM ATTENSION CONTROL COMMANDS

The operations of GSM modem are controlled by a set of control commands called AT commands. There are various groups of AT commands. We use a class of AT commands that enables us to send and receive SMS on the server machine. For example the command AT+CMGS is used to send SMS and AT+CMGR allows the application to receive SMS. SMS messages are received encoded in PDU mode and before they can be read they have to be converted in to text mode. A PDU decoder is used to achieve this conversion.

3.3 USE OF SMS TEMPLATES FOR INFORMATION TRANSFER

Appropriate SMS templates are used to make the task of data interpretation efficient and error-free. These templates are used when a rural hospital/health worker sends messages to the district HQ either for submitting report or requesting help. Figure 3.2 depicts the SMS template used for submitting disease surveillance report. The <DISCODE> field represents a unique code corresponding to a certain disease. The subsequent fields contain the collected surveillance data. Figure 3.3 is the template which is used when a rural hospital places a request for certain essential drugs. The <MEDCODE> field refers to a particular medicine and the <QUANTITY> field indicates the corresponding "medicine" quantity per box. Figure 3.4 is the template when a person wants to avail diagnosis help and sends a request of the form where the field <SYMPTOMCODE> indicates a particular symptom belonging to a class of symptoms and the next filed indicates whether he is showing the symptom.

<DISCODE> <NEWCASES<5>
<NEWCASES>5> <DEATHS<5>
<DEATHS>5>

Fig: 3.2

<MEDCODE> <QUANTITY>

Fig: 3.3

<SYMPTOMCODE> <YES/NO>

Fig: 3.4

3.4 USE OF J2ME FOR BUILDING SMART PHONE APPLICATION

A smart phone application is developed using J2ME to provide communication ability at the user end and for sending surveillance data. The health worker enters the data to the application running on his mobile phone and the application solely has the responsibility of constructing messages following the formats of some predefined templates and sending them to the server. This, in a way, ensures error free transmission and unauthorized access as only the worker having the application deployed on its phone can send messages in appropriate format. Only authorized workers are provided with the application software which they can use for submitting surveillance report. An additional layer of security can be provided by having each worker supply his unique user id and password with each message.

3.5 USE OF DECISION TREE FOR DISEASE DIAGNOSIS

A decision tree is a classification scheme which generates a tree and a set of rules, representing the model of different classes, from a given data set. With the help of decision trees we are able to generate understandable rules. The decision tree provides a clear indication of which fields are most important for prediction and classification.

3.5.1 SPLITTING ATTRIBUTE

With every node of the decision tree, there is an associated attribute whose values determine the partitioning of the data set when the node is expanded.

3.5.2 SPLITTING CRITERION

The qualifying condition on the splitting attribute for data set splitting at a node is called the splitting criterion at the node. For a numeric attribute, the criterion can be an equation or inequality. For a categorical attribute, it is a membership condition on a subset of values. A decision tree construction method is based on the principle of recursively partitioning the data set till homogeneity is achieved.

3.5.3 BEST SPLIT

There are several alternatives to choose from for the splitting attribute and the splitting criterion. But in order to build an optimal decision tree, it is necessary to select those corresponding to the best possible split. The main operations during the tree building are

1. evaluation of splits for each attribute and the selection of the best split; determination of the splitting attribute,
2. determination of the splitting condition on the selected splitting attribute, and
3. partitioning the data set using best split.

The desirable feature of splitting is that it should do the best job of splitting at the given stage. The first task is to decide which of the independent attributes makes the best splitter. The best split is defined as the one that does the best job of separating records into groups, where a single class predominates. To choose the best splitter at a node, we consider each independent attribute in turn. Assuming that an attribute takes on multiple values, we sort it and then, using some evaluation function as the measure of goodness, evaluate each split. We compare the effectiveness of the split provided by the best splitter from each attribute. The winner is chosen as the splitter of the root node.

3.5.4 SPLITTING INDICES

We now present a method for determining the goodness of a split. We make use of an index derived from economics as a measure of diversity. This is called the *gini index.*

3.5.4.1 GINI INDEX

If a data set *T* contains examples from *n* classes, gini index, *gini*(*T*) is defined as

$$gini(T) = 1 - \sum_{j=1}^{n} p_j^{\,2}$$

where *pj* is the relative frequency of class *j* in *T*.

If a data set *T* is split into two subsets *T1* and *T2* with sizes *N1* and *N2* respectively, the *gini* index *gini*(*T*) is defined as

$$gini_{split}(T) = \frac{N_1}{N} gini(T_1) + \frac{N_2}{N} gini(T_2)$$

The attribute provides the smallest *ginisplit*(*T*) is chosen to split the node (*need to enumerate all possible splitting points for each attribute*).

3.5.5 SUPERVISED VS. UNSUPERVISED LEARNING

Supervised Learning builds a learner model, or concept definitions, using data instances of known origin and uses the model to determine the outcome new instances of unknown origin. Decision tree construction mechanisms belong to this learning paradigm. Unsupervised Learning is a data mining method that builds models from data without predefined classes usually for classification/clustering.

Let us consider the Symptoms/Diagnosis dataset for a supervised classification.

Patient ID#	Sore Throat	Fever	Swollen Glands	Congestion	Headache	Diagnosis
1	Yes	Yes	Yes	Yes	Yes	Strep throat
2	No	No	No	Yes	Yes	Allergy
3	Yes	Yes	No	Yes	No	Cold
4	Yes	No	Yes	No	No	Strep throat
5	No	Yes	No	Yes	No	Cold
6	No	No	No	Yes	No	Allergy
7	No	No	Yes	No	No	Strep throat
8	Yes	No	No	Yes	Yes	Allergy
9	No	Yes	No	Yes	Yes	Cold
10	Yes	Yes	No	Yes	Yes	Cold

Table: 3.1

In order to determine the best splitting attribute to serve as the root node of the tree we compute the gini indices for each the symptoms which are obtained as follows.

Symptom	Gini Index
Sore Throat	0.64
Fever	0.40
Swollen Glands	0.34
Congestion	0.47
headache	0.64

Table: 3.2

In the above case "Swollen Glands" provides the least *ginisplit*(T) value and hence is chosen as the split node. We now divide the data set depending on each of the available splitting criterion on the best split and enumerate all possible splitting points. In our case splitting the data set upon "Swollen Gland" we obtain two datasets as follows.

Patient ID#	Sore Throat	Fever	Swollen Glands	Congestion	Headache	Diagnosis
1	Yes	Yes	Yes	Yes	Yes	Strep throat
4	Yes	No	Yes	No	No	Strep throat
7	No	No	Yes	No	No	Strep throat

Table 3.3

Patient ID#	Sore Throat	Fever	Swollen Glands	Congestion	Headache	Diagnosis
2	No	No	No	Yes	Yes	Allergy
3	Yes	Yes	No	Yes	No	Cold
5	No	Yes	No	Yes	No	Cold
6	No	No	No	Yes	No	Allergy
8	Yes	No	No	Yes	Yes	Allergy
9	No	Yes	No	Yes	Yes	Cold
10	Yes	Yes	No	Yes	Yes	Cold

Table 3.4

The process proceeds recursively for each stage until homogeneity is achieved and the construction of decision tree is complete at this stage. The final decision tree generated is depicted in Figure 3.5

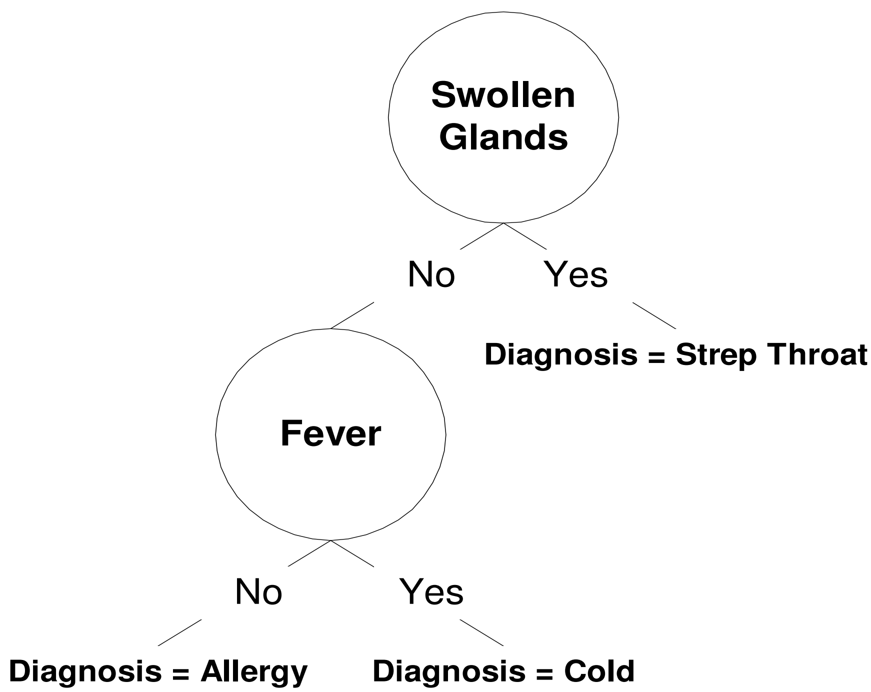


Fig 3.5

We can summarize the Decision Tree by listing the decisions (production rules) along each path from the starting node to each terminal node:

1. IF *Swollen Glands = Yes* THEN *Diagnosis = Strep Throat*
2. IF *Swollen Glands = No & Fever = Yes* THEN *Diagnosis = Cold*
3. IF *Swollen Glands = No & Fever = No* THEN *Diagnosis = Allergy*

By determining the best split at each stage, the Decision Tree discards the unnecessary attributes. We can use the Decision Tree for future diagnoses, (or prediction of diagnosis).

3.6 GSM LOCATION TRACKING FOR EMERGENCY CONTROL

Location-based services answer three questions: Where am I? What's around me? How do I get there? They determine the location of the user by using one of several technologies for determining position, and then use the location and other information to provide personalized applications and services. Location-Based Services (LBS) provide personalized services to the user on request, based on their current position. Many LBS applications establish the location of the user by simply being informed by the cellular network which base station the user is currently connected to - a technique known as Cell-ID. This basic form of location tracking is supported by all GSM handsets. The subscriber's location can be tracked to an area of between 200 meters and 1 kilometer. The current cell ID can be used to identify the Base Transceiver Station (BTS) that the device is communicating with and the location of that BTS. The Cell ID is then converted into a geographic location using the knowledge of operator network. Clearly, the accuracy of this method depends on the size of the cell. We are using the above technique to avail information of the location to resolve emergency situations.

The location information of a mobile phone in GSM network includes 4 parameters as depicted in Fig 3.5. All the fields are in hexadecimal format.

For instance, location parameters for a mobile phone user at NIT Rourkela, Orissa are:

- MCC-404 for India,
- MNC-76 for BSNL in Orissa,
- LAC-118 for Rourkela,
- CELLID-3F35 for Jagda

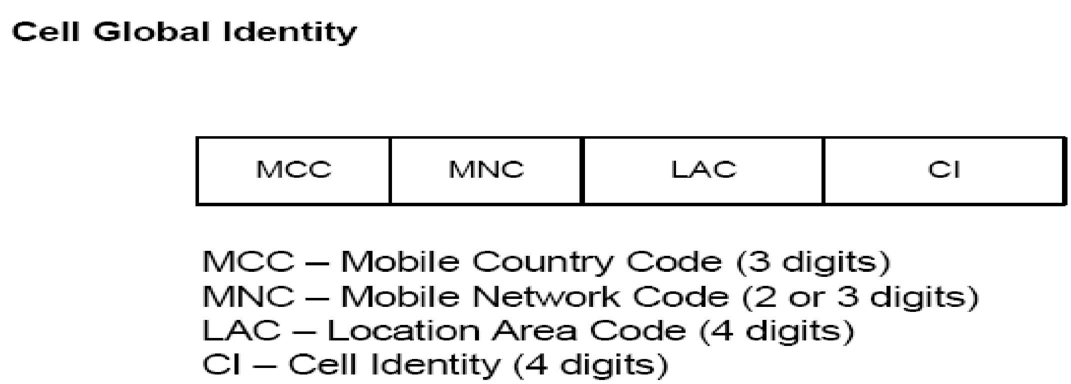


Fig: 3.6

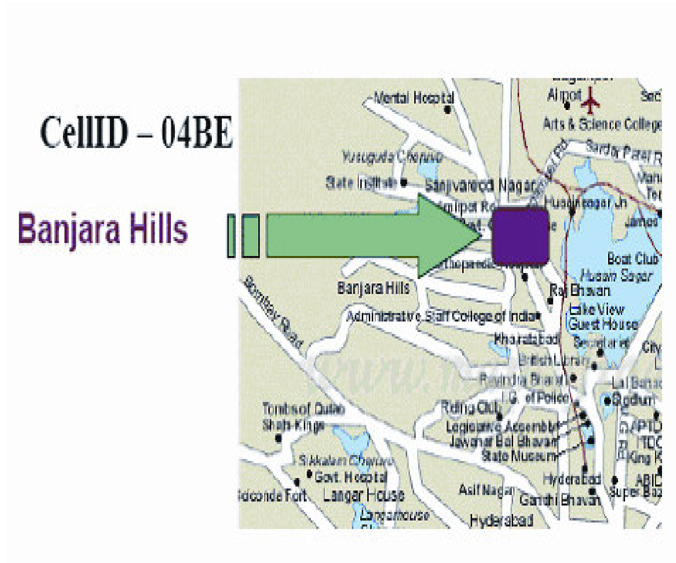


Fig: 3.7

Figure 3.6 shows a sample CELLID representing the location, Banjara Hills in Hyderabad. The Location API for J2ME (JSR-179) specification enables developers to write wireless location-based applications and services for resource-limited devices like mobile phones, and can be implemented with any common location method. The compact and generic J2ME location APIs provide mobile applications with information about the device's present physical location and orientation (compass direction), and support the creation and use of databases of known landmarks, stored in the device.

3.7 USE OF QUERY MANAGER FOR REPORT GENERATION

A Query Manager is used to generate dynamic health reports according to user query type. It performs the following operations:

- Generation of appropriate views: When a user's query is provided, the required data(s) are stored in various tables in the database. Thus in order to access data from multiple tables, the query manger creates view on the fly and provides the required result to the user.
- Direct queries to the appropriate table(s): It may sometimes happen that a user' query is constrained to only one table, in that case queries are directly executed.

Architecture of the query manager is depicted in figure 3.6

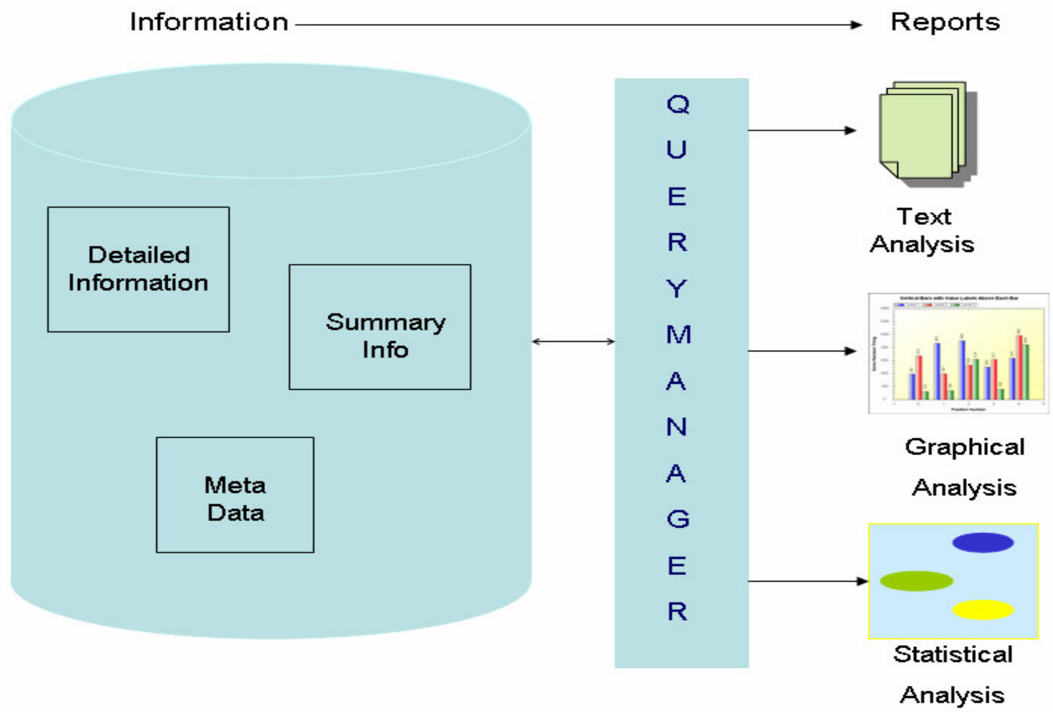


Fig 3.8

The query manager takes input as the information (tables, views, stored procedures) from the database and provides output in from of customized reports which performs *text analysis*, *graphical analysis* and *statistical analysis*

Text analysis involves generating combinational data depending on user query type and presenting them in a tabular format. Some of the reports generated by our system are listed below.

1. Disease surveillance data -district wise
2. Disease surveillance data-subdivision wise
3. Disease surveillance data-block wise
4. Disease surveillance data –hospital wise
5. Disease surveillance data –date wise (from date – to date) (for district, subdivision, block, hospital)
6. Weekly disease surveillance data (for district, subdivision, block, hospital)
7. Monthly disease surveillance data (for district, subdivision, block, hospital)
8. Yearly disease surveillance data (for district, subdivision, block, hospital)
9. Disease surveillance data according to number of new cases, deaths registered for particular diseases (in a district, subdivision, block, hospital)

10. Frequency of disease surveillance requests received for a particular disease from various areas (districts, subdivision, block, hospital)

Graphical analysis involves presenting the surveillance data in the form of graphs such as

- Bar Graph
- Pie Chart
- Line Graph

to facilitate understanding and interpretation

Statistical analysis is used to segregate different locations into different groups on the basis of incidence of diseases. The different locations are classified into *highly critical*, *critical* and *lowly critical* regions. In order to classify a new data as to be a critical data we perform a variance operation over the previously received data with respect to the newly arrived data. In probability theory and statistics, the **variance** of a random variable (or somewhat more precisely, of a probability distribution) is a measure of its statistical dispersion, indicating how its possible values are spread around the expected value. If this value exceeds certain variance threshold we conclude that the data is of critical importance and requires attention. For instance some district reports (new cases>5) for the disease of Jaundice per week as 1, 3, 2, 5, 3, 4 … (somewhere in the range of 1-5) on a weekly basis. If the newly arrived data is 4, then variance of the available data trend with respect to the newly reported data would be 0.76. On the other hand if the newly received data is 25, the variance would be 20.40. This indicates that the new data is deviated form the general trend. If the threshold value is set at 15, we conclude that the receiving of 25 is critical and requires attention.

Chapter 4

SYSTEM ARCHITECTURE AND UTILITY

4.1 SYSTEM ARCHITECTURE

The working of the proposed system an be categorized into three functional units as

- Surveillance unit
- Diagnosis unit
- Emergency controller

Figure 4.1 depicts the functioning of the proposed model. The message received at the server end is interpreted in two phases. Parsing and classification. In the parsing phase the message is checked for correctness of syntax. If the received message is out of order, then it is discarded. Otherwise the message is classified either as a surveillance report or diagnosis query or location information. Depending on the type of message the system behaves accordingly.

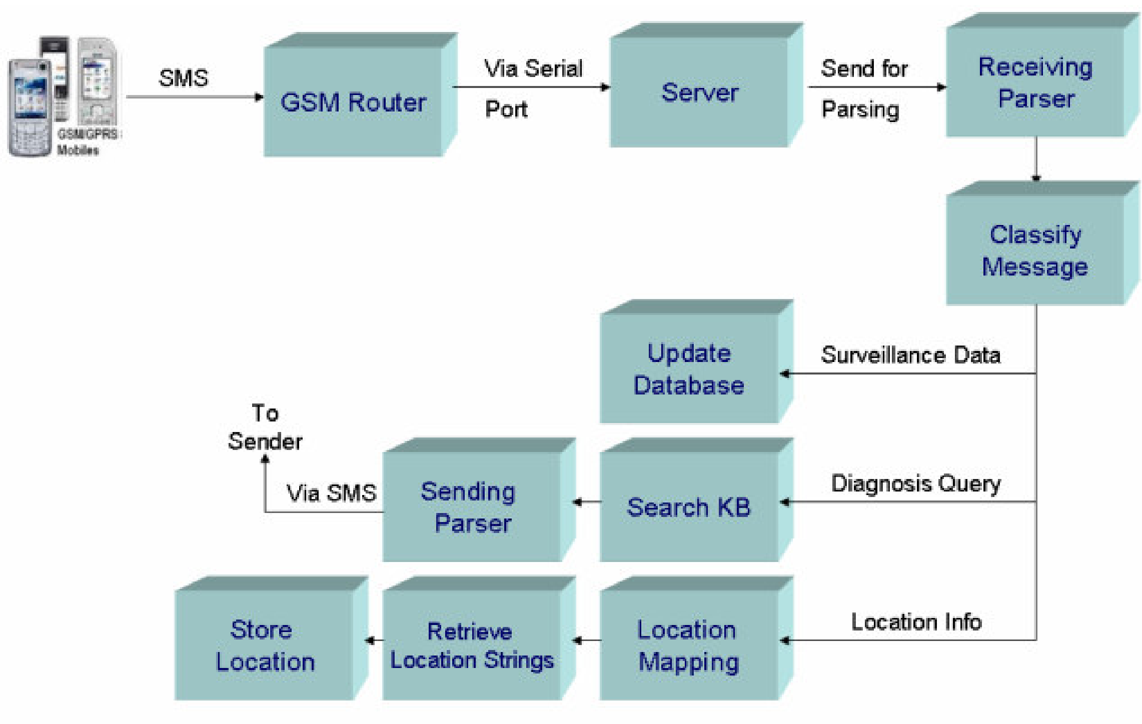


Fig 4.1

If the received message is a surveillance report the system updates its database accordingly; on the other hand if it's a diagnosis query the system searches its KB to find out a possible diagnosis which is then sent to the sender in the form of a reply message through the sending parser. If the message is found to be a location information string then the system stores the string and alerts the emergency controller of the received location.

Figure 4.2 depicts the use of the proposed model in the organizational hierarchy.

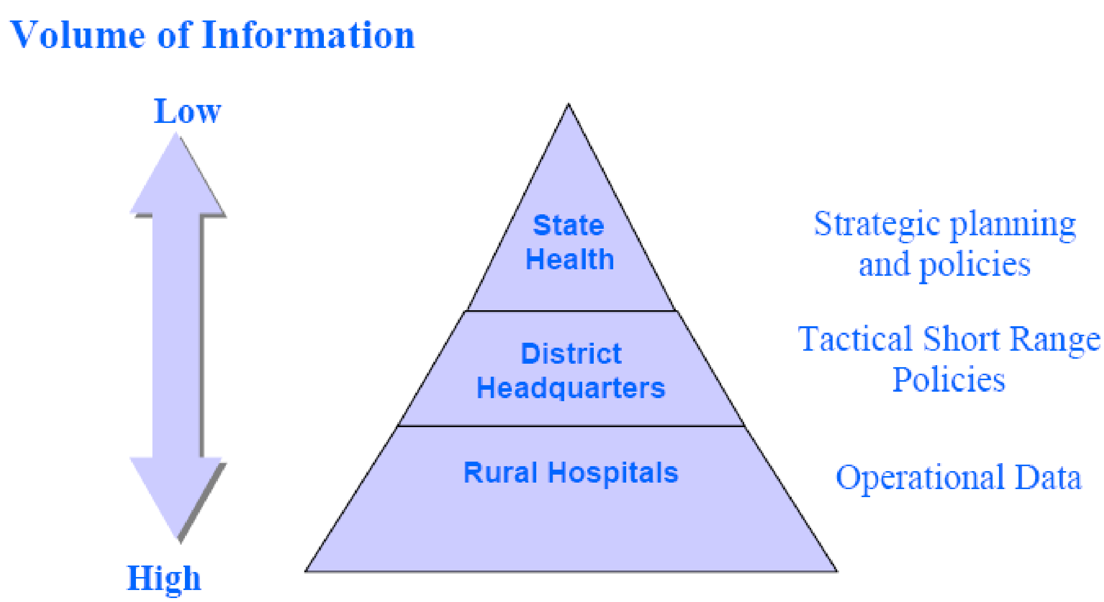


Fig 4.2

Detailed data for Operational decisions include disease surveillance data. Summarized data for Tactical short range policies are the organizing of health camps, awareness campaigns etc...Condensed data for Strategic planning and policies include long-term health policies. Our model is compatible with the above structural hierarchy.

Figure 4.3 illustrates the flow of information in the proposed model. The health workers at the PHC and CHC level are assigned the job of collecting and submitting the periodic surveillance data to the district HQ through SMS messaging.

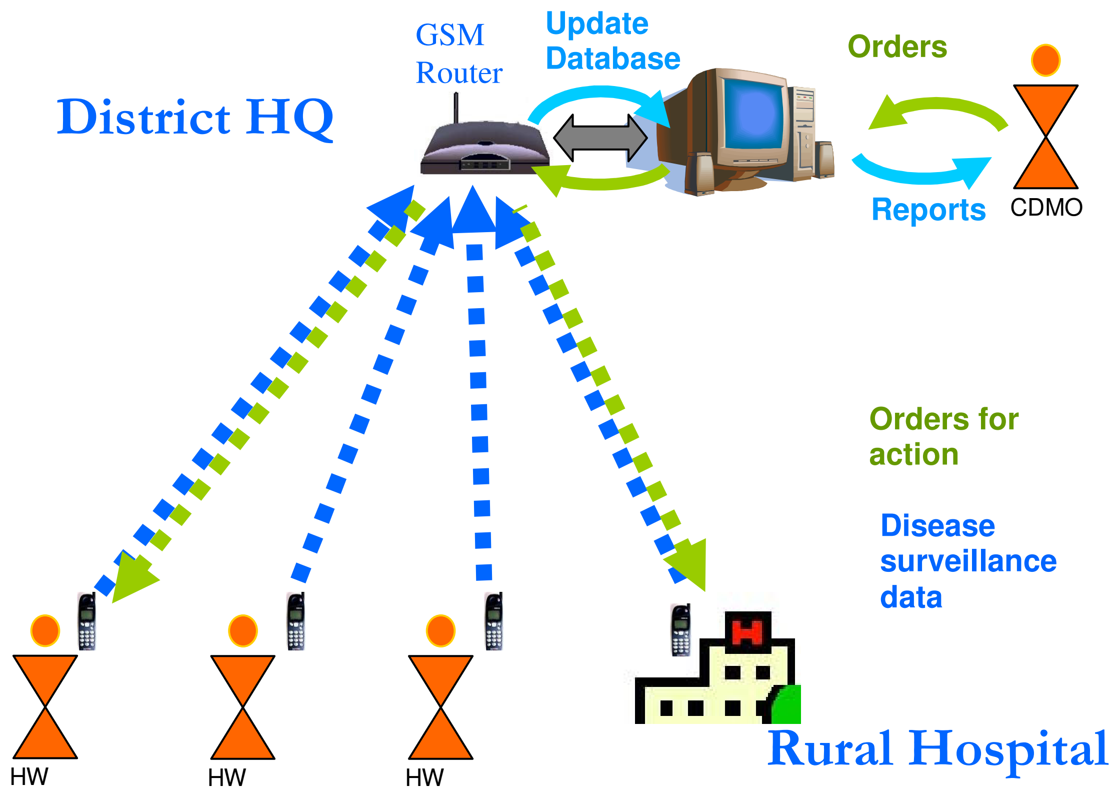


Fig: 4.3

4.2 SYSTEM UTILITY

- **Pertaining to the State Health Department**

- Assist the top management of the State Health Department to make strategic decisions.

- Provide real time information about availability of various resources and the health status of the state.

- **Pertaining to the District Level**

- Acquire required information from the rural hospitals and field workers.

- Provide an effective and instantaneous automated system to detect an outbreak and monitor the trend to prevent an epidemic and decrease the mortality and morbidity.

- Handle emergencies and mobilize the available resources optimally with in a given time frame.

- Organizing health camps at needy areas depending on the data at hand.

- **Pertaining to Rural Hospitals**

- Allow the hospitals to act in a more organized manner with constant hierarchical supervision.

- To inform the higher authorities quickly, especially in case of an emergency.

- Uninterrupted quality data flow and improving the response system during epidemics is of paramount importance.

- **Pertaining to Health Workers**

- Makes their job easier by providing them with an easy to fill template.

- The collected data is immediately updated to the central database through SMS.

- **Pertaining to NGO(s)**

- Get the required information through forums about the needs of different demographic locations, which will help them to provide better health service.

Chapter 5

SYSTEM IMPLEMENTATION AND RESULTS

The proposed system consists of 5 major modules:

1. SMS Module for sending and receiving SMS between the server and mobile phone.
2. Database Design Module for creating and maintaining the database.
3. SMS parser for validation and authentication of received and sent SMS messages.
4. Disease Surveillance Module for the collection, storage and interpretation of surveillance data.
5. Query Manager for generation of customized reports.
6. Diagnosis Module for resolving user diagnosis query.
7. Location Tracking Module for detection of concerned location.

5.1 THE SMS MODULE

The SMS module is built using the tool Visual Studio.NET 2005. It is further divided into two parts, one being the SMS receiving module and another being the SMS sending module. The SMS receiving module periodically checks for messages on the serial port connected to the GSM Modem. If data is available, it reads the messages using the "AT+CMGR" command in PDU mode. After reading all the messages available on the serial port, it converts the SMS messages from PDU to text mode using a PDU decoder. The PDU decoder extracts necessary data (message, date, mobile number) from the SMS message and sends to the parser for further processing. The SMS sending module periodically checks for pending messages in the database. If messages are available, it sends the data to the required destination using the "AT+CMGS" command through the GSM modem until there are no pending messages. Before executing the "AT+CMGS" command, the transmission on the serial port is converted from PDU to text mode using "AT+CMGF" command.

5.2 THE DESIGNING OF THE DATABASE

The ER-Model of our database and the integrity relationship between the entities are depicted in figures from 5.1 to 5.5. SQL Server – 2000 is used to implement our database.

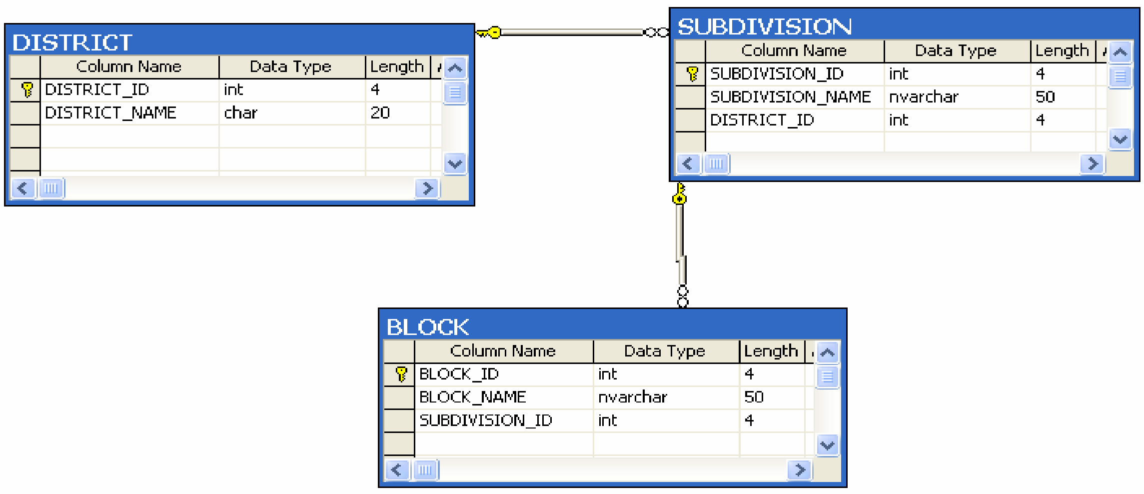


Fig 5.1

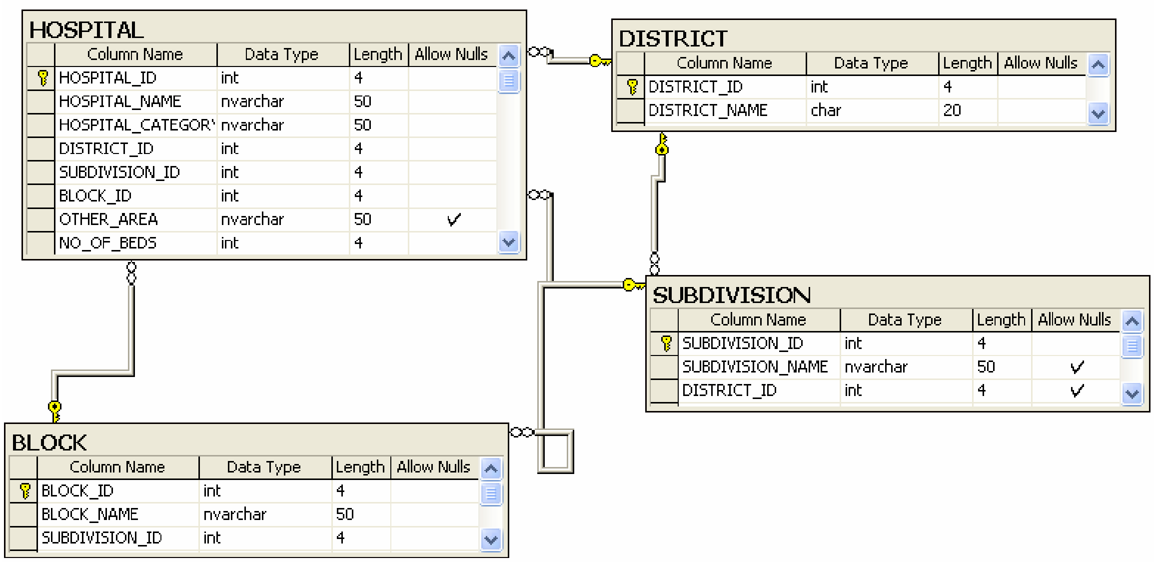


Fig 5.2

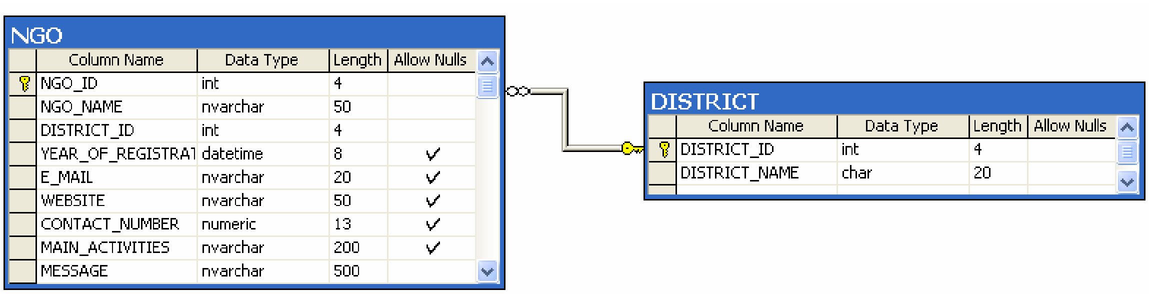


Fig 5.3

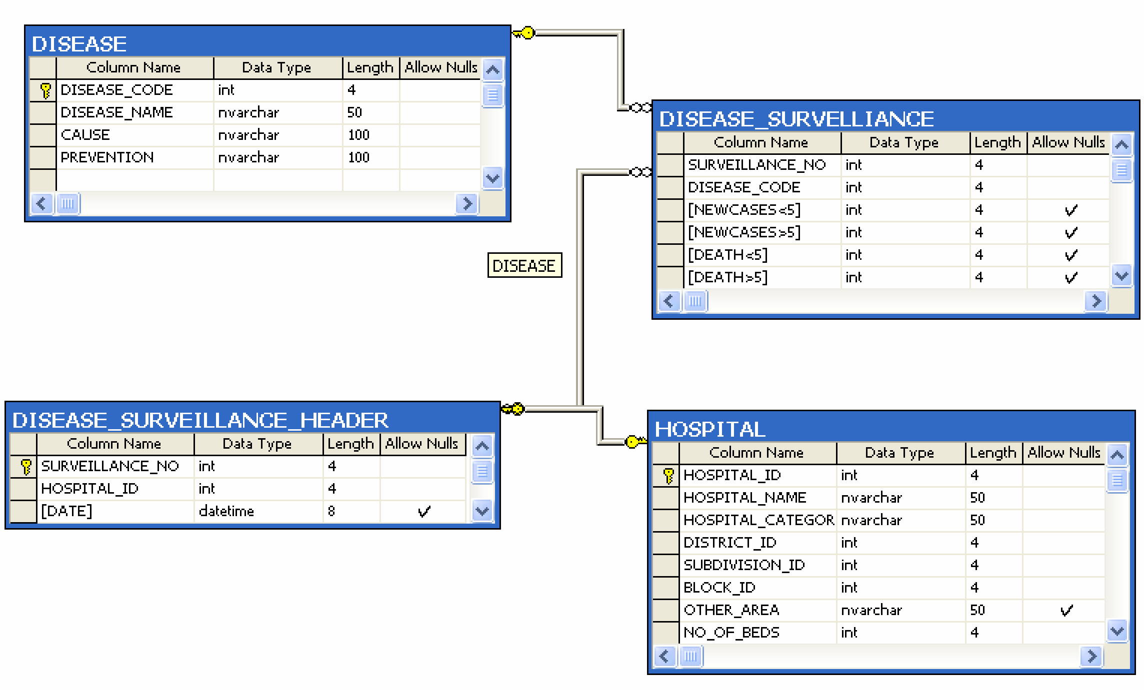

DISEASE
Column Name | Data Type | Length | Allow Nulls
DISEASE_CODE int 4
DISEASE_NAME nvarchar 50
CAUSE nvarchar 100
PREVENTION nvarchar 100
DISEASE_SURVELLIANCE
Column Name | Data Type | Length | Allow Nulls
SURVEILLANCE_NO int 4
DISEASE_CODE int 4
[NEWCASES<5] int 4 ✓
[NEWCASES>5] int 4 ✓
[DEATH<5] int 4 ✓
[DEATH>5] int 4 ✓
DISEASE
DISEASE_SURVEILLANCE_HEADER
Column Name | Data Type | Length | Allow Nulls
SURVEILLANCE_NO int 4
HOSPITAL_ID int 4
[DATE] datetime 8 ✓
HOSPITAL
Column Name | Data Type | Length | Allow Nulls
HOSPITAL_ID int 4
HOSPITAL_NAME nvarchar 50
HOSPITAL_CATEGOR nvarchar 50
DISTRICT_ID int 4
SUBDIVISION_ID int 4
BLOCK_ID int 4
OTHER_AREA nvarchar 50 ✓
NO_OF_BEDS int 4

Fig 5.4

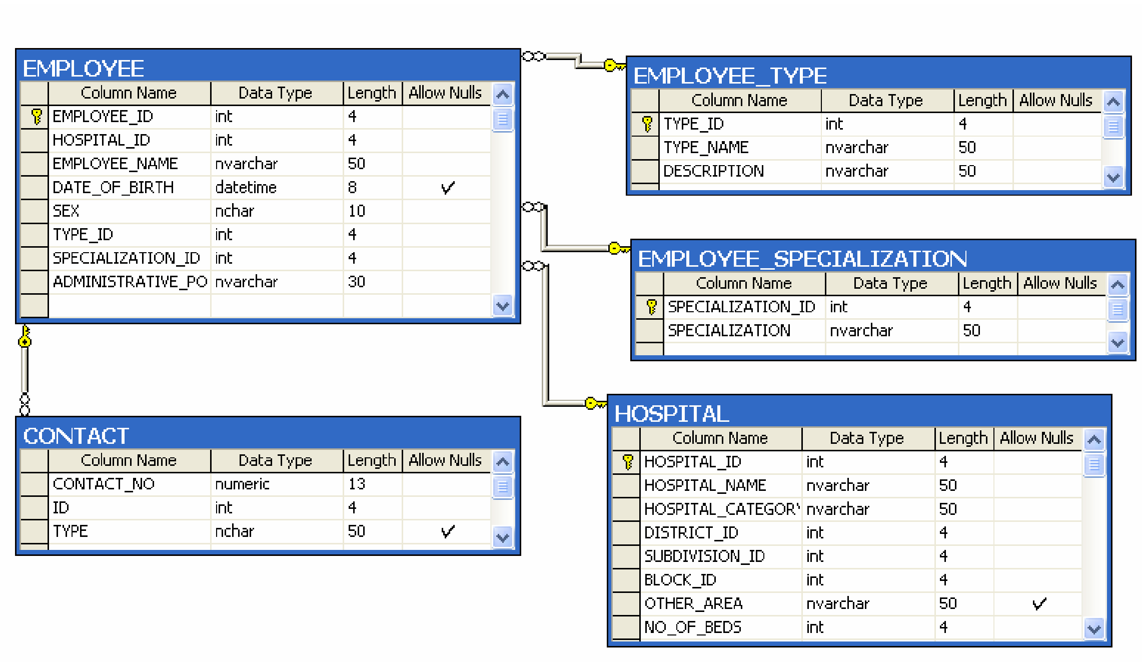

EMPLOYEE
Column Name | Data Type | Length | Allow Nulls
EMPLOYEE_ID int 4
HOSPITAL_ID int 4
EMPLOYEE_NAME nvarchar 50
DATE_OF_BIRTH datetime 8 ✓
SEX nchar 10
TYPE_ID int 4
SPECIALIZATION_ID int 4
ADMINISTRATIVE_PO nvarchar 30
EMPLOYEE_TYPE
Column Name | Data Type | Length | Allow Nulls
TYPE_ID int 4
TYPE_NAME nvarchar 50
DESCRIPTION nvarchar 50
EMPLOYEE_SPECIALIZATION
Column Name | Data Type | Length | Allow Nulls
SPECIALIZATION_ID int 4
SPECIALIZATION nvarchar 50
CONTACT
Column Name | Data Type | Length | Allow Nulls
CONTACT_NO numeric 13
ID int 4
TYPE nchar 50 ✓
HOSPITAL
Column Name | Data Type | Length | Allow Nulls
HOSPITAL_ID int 4
HOSPITAL_NAME nvarchar 50
HOSPITAL_CATEGOR' nvarchar 50
DISTRICT_ID int 4
SUBDIVISION_ID int 4
BLOCK_ID int 4
OTHER_AREA nvarchar 50 ✓
NO_OF_BEDS int 4

Fig 5.5

5.3 THE SMS PARSER

The SMS parser is also built using Visual Studio.NET 2005. It performs two major functions, *validation* and *authentication* of the messages received from the SMS module. Once the message is parsed, it is then passed to the SMS classifier to determine the type of message.

5.4 THE DISEASE SURVEILLANCE MODULE

The main job of the disease surveillance module is to store the parsed surveillance report in the database and to generate reports using the query manager. A smart phone application developed using J2ME is installed on the mobile phone on the user's side. This facilitates the orderly collection and error free transmission of the collected data.

5.5 THE QUERY MANAGER

As discussed earlier, the query manager performs three types of analysis namely text, graphical and statistical. The text analysis is done using the *DataGridView* component of VB.NET 2005. A graphical package called *Zedgraph,* for Visual Studio.NET 2005 is used to generate graphical reports. The results of statistical analysis are portrayed upon geographical maps of various block(s), subdivision(s) and district(s) for better understanding.

5.6 THE DIAGNOSIS MODULE

Its function is to resolve the diagnosis user query received from the SMS parser by searching the knowledge base to find the corresponding diagnosis. The decision tree is implemented using VB.NET 2005.

5.7 THE LOCATION TRACKING MODULE

This module retrieves the location string received from the parser and alerts the emergency controller of the information. A J2ME application at the user end provides the capability to send location information.

5.8 WORKING OF THE SYSTEM

5.8.1. SMS Module at the Server Site:

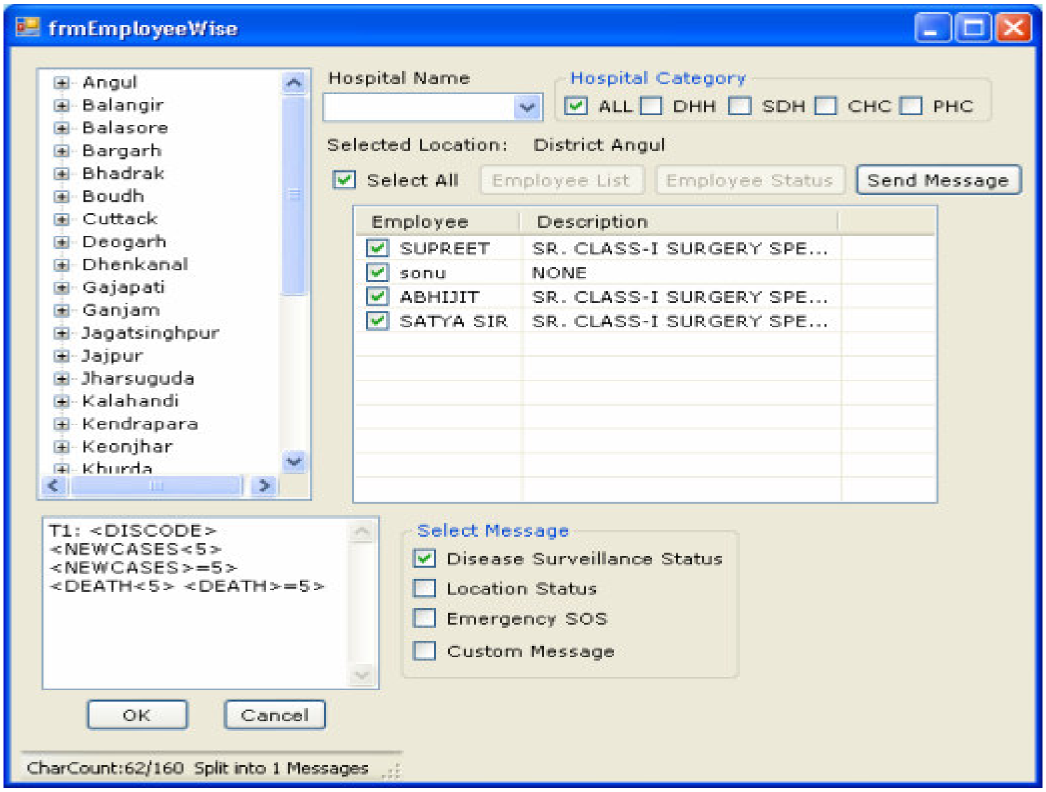


Fig: 5.6

5.8.2 Smart Phone Application at the Client Site:

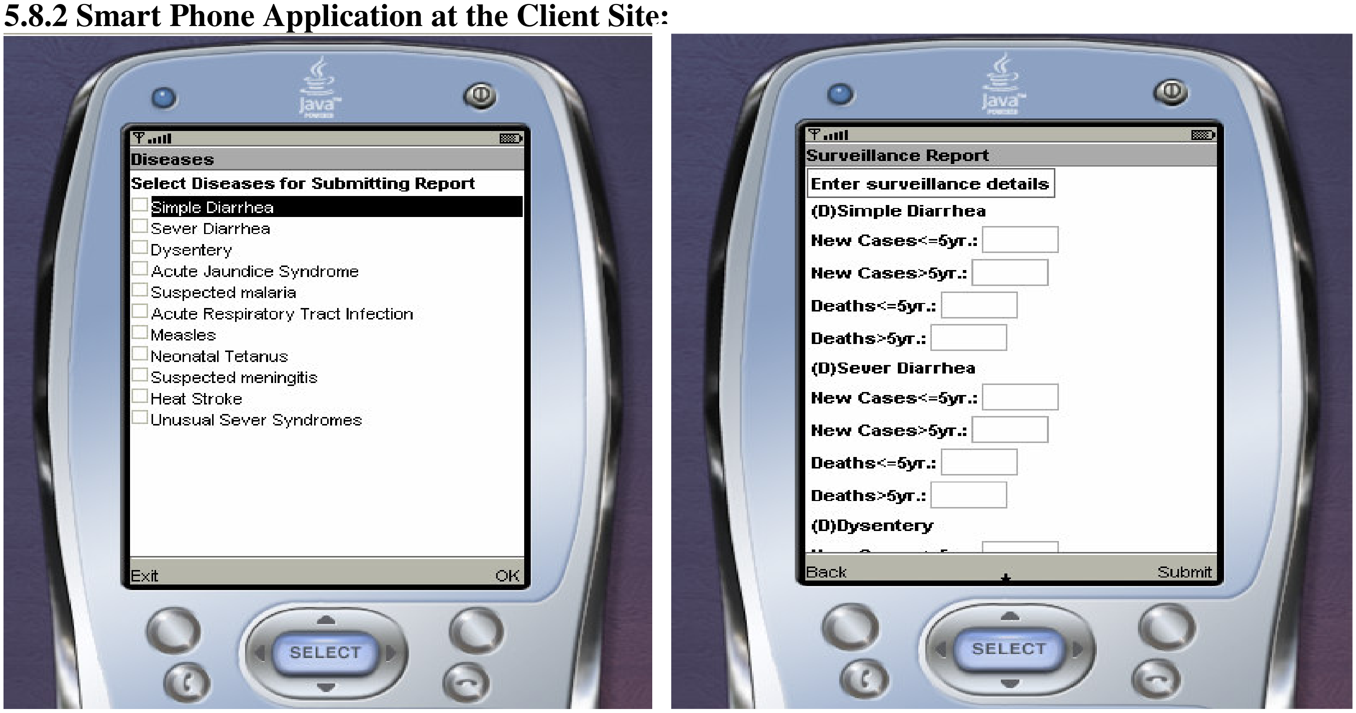


Fig: 5.7

5.8.3. Text Analysis of Collected Data:

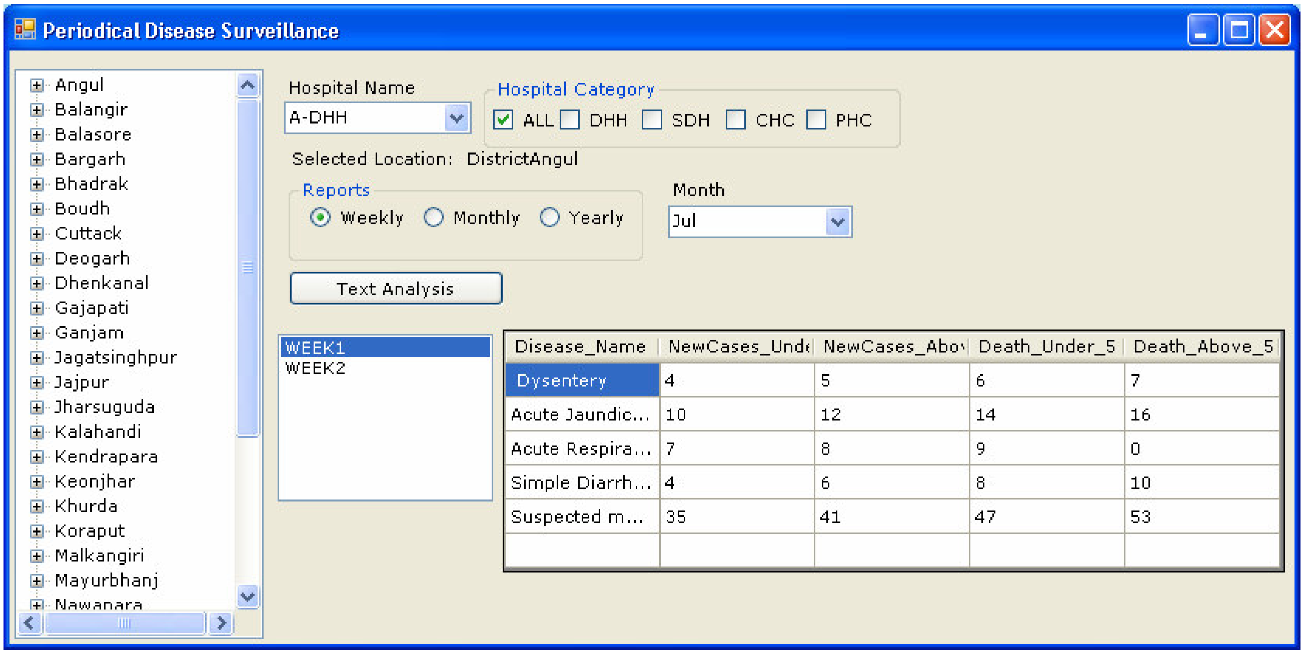
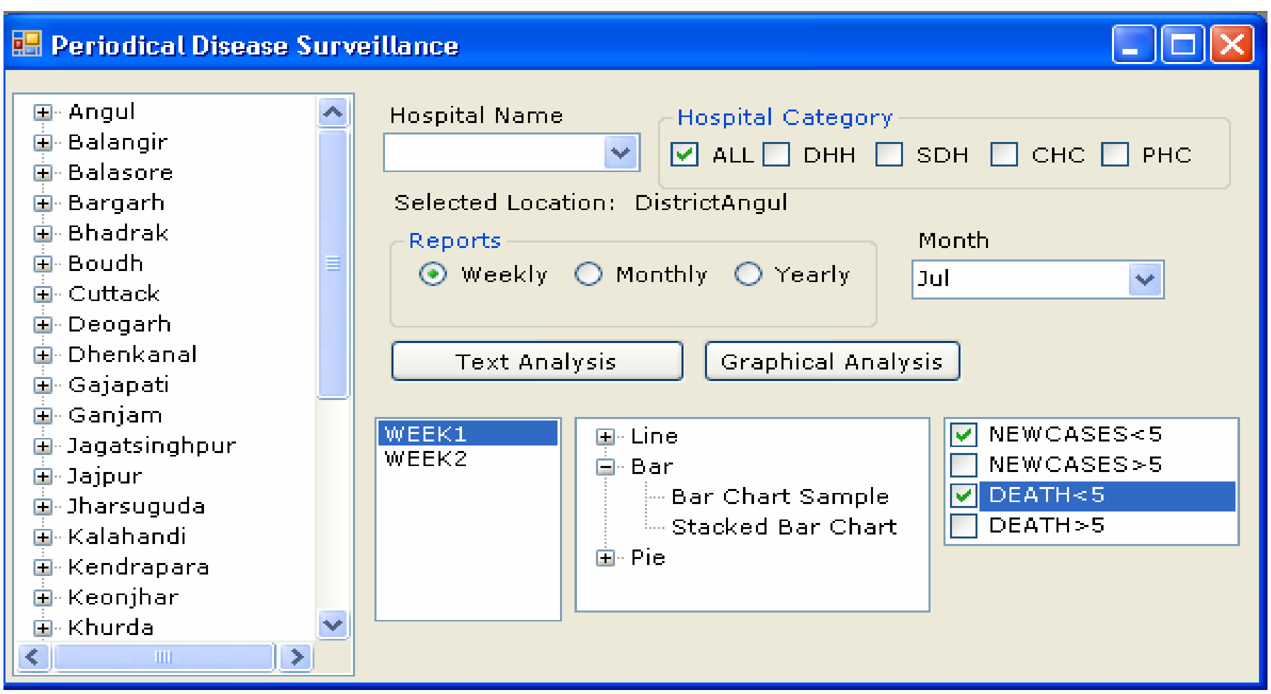


Fig: 5.8

5.84. Graphical Analysis of Collected Data:



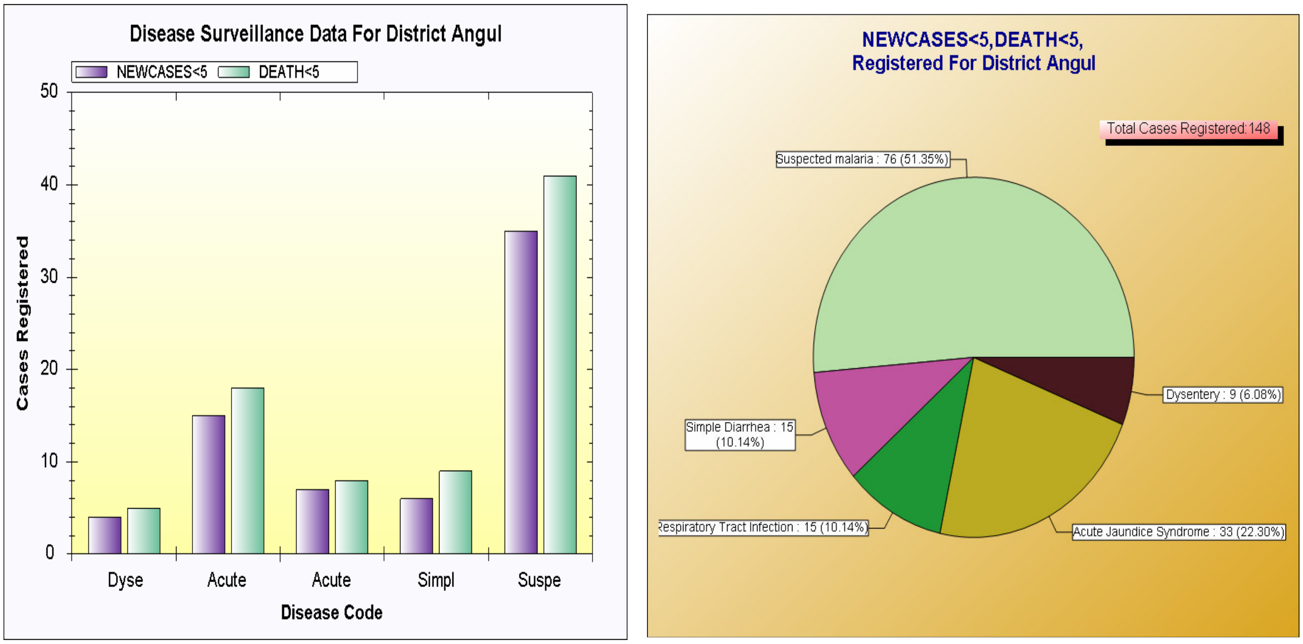


Fig: 5.9

5.8.5 Statistical Analysis of Collected Data:

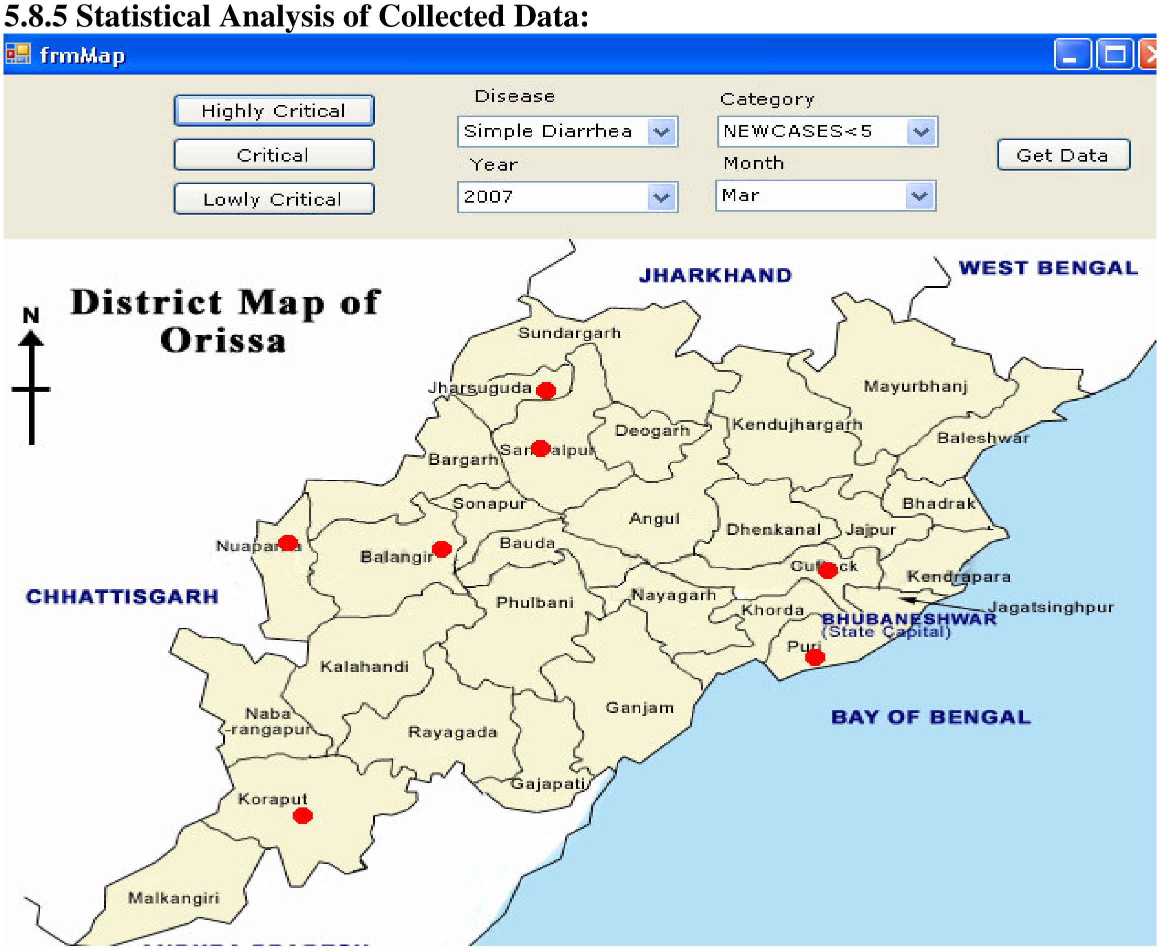


Fig: 5.10

Chapter 6

CONCLUSION AND REFERENCES

6.1 CONCLUSION

To facilitate access to high-quality evidence on what is best-practice health care, electronic decision support systems are essential. They can make a difference to the quality of health care by giving clinicians and consumers' access to relevant, evidence-based information at the point of care. Today in this new era of Mobile communication SMS is one of the cheapest and most reliable ways to send small amount of data and information. With the mobile penetration increasing and the cost of mobile handsets falling, this system will provide a cost effective, efficient and fast solution to the health department.

6.2 REFERENCES

[1] http://www.orissagov.nic.in/health

[2] http://www.gsmworld.com/technology/gsm.html, "*Today's GSM".*

[3] Wavecom, "AT Commands Interface Guide", Version.001/9.1, pp.127-128, November 2001.

[4] A.D.Sarma, P.S.Ravikanth, D.Krishna Reddy, "integration of gps and gsm for determination of cellular coverage area", 2004

[5] Melinda Wharton, "Analysis of Surveillance Data", *VPD Surveillance Manual, 3rd Edition*, 2003

[6] Tan, J. and Sheps, "S. Health Decision Support Systems" *Aspen Publishers, Inc., Gaithersburg, MD,* 1998

[7] Keeratpal Singh and Mohamod Ismail, "Location Based Services for GSM and 3G Networks", 2004

[8] http://java.sun.com

[9] James Keogh, “The Complete Reference J2ME”, *TATA McGraw HILL Publishers*, *Inc*., 2003.

[10] Arun K.Pujari, “Data Mining Techniques”, *Universities Press*, 2001

[11] Sam Anahory, Dennis Murray, “Data Warehousing in the Real World”, *Pearson Education*, 2005

[12] N.B. Szirbika, C. Pelletier a, T. Chaussaletb, “Six methodological steps to build medical data warehouses for research”, *international journal of medical informatics* 75 (2006) 683–691

[13] Abdullah A. AlAqeell, Nicholas Noyes2, John Enderle', “Custom Report Generator”, *IEEE*, 2000

[14] W. Raghupathi & J. Tan (2002) “Strategic IT Applications in Health Care” *Communications ACM*, 45(12): 56-61.