CHAPTER 1

INTRODUCTION

1.1 DISTRIBUTED DATABASES

A distributed database comprises a set of nodes from a computer network and a set of local databases stored in such nodes. A centralized server for each group of nodes is present where details relating to data fragmentation and distribution are stored. The primary difference between centralized databases and distributed databases is the provision of data distribution, replication and logical correlation in the latter. Moreover, a distributed database system can be viewed as m:n client server systems where all m clients and n servers have either a homogeneous or heterogeneous database for storage. The key advantages of distributed database systems include availability, data fragmentation using horizontal and vertical fragmentation techniques, a data communication component, a set of local catalogues and a global catalogue. In a distributed database system, there are two types of database administrators (DBAs): a local DBA and a global DBA. Site autonomy is provided in distributed databases in such a way that the global DBA is practically out of sight since problems of all sorts are resolved by the local DBA. The global DBA is contacted by local DBAs only when there is a dispute among local sites beyond the purview of the local coordinators. The diverse features of distributed databases include integrity, security, query processing, concurrency control and recovery.
1.2 CLOUD DATABASES

Today, cloud computing provides continuous access to resources in the cloud with ubiquitous network access. Moreover, cloud database systems are advanced in distributed database systems effectively accepted as a promising and effective way to store and manipulate data. The system accepts not only the Structured Query Language (SQL)-based data storage but also a column-based new storage type. Cloud databases improve data utilization data and computing central resources by providing virtualization, which enhances processing speed and memory size.

Server consolidation is an effective cloud computing approach that increases computer server usage. Cloud database systems play a major role in current web-based applications owing to their storage and processing capabilities. A cloud data center has multiple servers, jointly used to store data from anywhere in the world, at any time. Moreover, these systems provide a backup for storing and retrieving data in the web and hence data can be accessed and used by clients anywhere, and at any time. Using effective storage and access methods, the cloud database system provides highly reliable and high-fidelity services to cloud database users. However, relational database technologies cannot be used as they are for storing and manipulating cloud databases, thanks to the large number of additional features provided by cloud database systems. Therefore, current storage and data manipulation techniques used in relational databases need to be enhanced with additional features, including security constraints, as there are multiple clients storing their private data in the cloud database storage. Hence, it is necessary to propose new client and server based security mechanisms based on encryption of data, trust management and temporal constraints so as to augment security.
1.3 MOBILE CLOUD DATABASES FOR HEALTHCARE MANAGEMENT

Mobile database systems have been developed by enhancing the relational model with additional features for handling mobility, including hands-off, mobile transactions and location-dependent query processing. The focal advantage of developing mobile cloud database applications is to provide a range of applications - including healthcare management systems - the features of database management systems, mobile computing and cloud computing.

In India, diabetes is a major disease calling for effective management by both patients and doctors so as to provide better healthcare services. This can be achieved effectively by transforming the practice of healthcare planning and implementation using mobile cloud databases. Diabetes is a serious disorder characterized by high levels of blood glucose in the human body and originates from defects in humans in the form of insulin production, insulin usage or both, and is a silent killer. Insulin, a hormone secreted by pancreatic beta cells, is regulated by the intake of the glucose from the blood into most cells of the human body (Alberti & Zimmer 1998). Health complications such as damage to the heart including stroke, high blood pressure, and eye problems with severe vision loss or blindness are led by the inability of the human body to generate sufficient insulin, leading to increased levels of blood glucose (Rosenthal et al. 2007, Amos et al. 1997). In such a scenario, retaining patient data using relational databases requires the user to operate the data only from specific locations. On the other hand, mobile cloud database systems provide rich data storage facilities from mobile devices, cloud-based services, and efficient wireless communication facilities from mobile cloud networks, rendering queries possible from anywhere.
A mobile healthcare application for providing healthcare services must take care of the diagnosis, treatment, and prevention of disease successfully, from anywhere and at any time. In the past, an expert system developed for medical diagnosis, including MYCIN (Elaine Rich 2001), provided effective pro-health decisions using a knowledge-based approach performing deductive inferences on symptoms. However, the large volumes of patient data and medical records at hand have made it necessary to convert knowledge-based systems into database system applications. With the onset of the internet and the web, the size of the data stored keeps increasing continuously to keep pace with patient histories. Consequently, a range of healthcare applications are moving their data into the cloud for storage. Though relational database systems provide facilities for data storage, retrieval, and transactions with concurrency and recovery, they fail to provide support for making decisions and do not permit queries from mobile devices. Also, mobile databases have limited data storage facilities. On the other hand, cloud databases provide large quantities of memory and dynamic processing power (Muthurajkumar et al. 2015). Hence, it is vital to integrate the features of database management systems, mobile computing and cloud computing with knowledge based systems to develop intelligent mobile healthcare applications.

Despite the availability of relational database systems for health record maintenance, a web-based solution is essential for providing e-health services and to make cost-effective treatment methods. Cloud database systems provide facilities for large-scale, real-time data storage and retrieval features, also providing distributed query processing facilities using map-reduce procedures. However, access control methods provided by cloud database systems are insufficient to guarantee data security. Further, trigger features must be enhanced with rules to provide accurate database services. In addition, existing database systems require the provision of queries from
nodes available in fixed networks. Hence, it is pivotal to propose a cloud data model that includes cloud storage, security, and mobility, as well as rules to provide effective cloud-based services.

1.4 TEMPORAL DATA STORAGE AND RETRIEVAL

Temporal databases provide facilities for storing and manipulating past, present and future data (Snodgrass & Ahn 1986). However, they require vast quanta of storage to maintain the entire history, an impossibility where relational database systems are concerned. On the other hand, cloud data bases provide cost effective storage and retrieval services by providing the required number of processors and memory. Even though, storing temporal data in cloud data bases increases storage requirements, their virtualization and data center features make it possible to store the history in its entirety. Cloud databases require an additional query language and new query processing techniques for handling temporal data and, given that these naturally temporal services are accessed through the World Wide Web (WWW), there is a need for support from cloud data centers, considering their requirements for colossal amounts of storage space.

An indubitable advantage of the cloud database system is that the cloud computing platform absolves local database administrations of the responsibility of handling such complex tasks as storage and processing requirements, by providing elasticity and virtualization. In such a case, users can only depend on the cloud data base manager to receive data services. The advantages of cloud temporal databases include maintenance of past, present and future data with reduced costs and extra storage. In addition, it motivates research focus on core query processing algorithms (Wen-ChihPeng & Ming-Syan Chen 2005) instead of infrastructure management. However, existing storage structures are scarcely adequate enough to store temporal data obtained from applications in the cloud. Therefore, it is necessary to propose a
new temporal storage structure, along with effective retrieval methods based on the temporal query processing algorithm, to resourcefully store and retrieve temporal data. In addition, security techniques such as key management, access control, intrusion detection and prevention, trust management, encryption and decryption techniques used for securing data stored in relational databases must be enhanced to suit cloud databases.

1.5 PARALLEL QUERY PROCESSING

In centralized database systems, the cost of a query is measured in terms of the number of input / output (I/O) operations, owing to the fact that I/O operations are much slower than central processing unit (CPU) operations. Therefore, query processing techniques developed for relational database management systems encourage the use of indexing and hashing techniques. However, in a distributed database system, a distributed access plan is necessary to access data present in a computer network’s multiple sites to enhance performance, achieved by writing an access plan which reads and searches for the required data in parallel. Therefore, parallel query processing techniques are more powerful than searches with indexing and hashing. In cloud computing, parallel searches can be performed using Map Reduce functions. In a distributed database system, the cost of a query is measured using I/O and network costs. In a mobile database, the cost of a query is measured using I/O, network and mobility costs. In a cloud database system, the cost of a query can be measured using the amount of parallelism achieved using Map Reduce techniques. In a mobile cloud, location specific queries are easily addressed since it does not depend on mobility. On the other hand, location dependent query processing is a serious issue to be addressed in mobile cloud database systems.
1.6 LOCATION DEPENDENT QUERY PROCESSING

In a mobile cloud database system, since a user is allowed to move from one location to another, spatial query processing features and temporal query processing techniques cannot be applied directly in mobile cloud databases for querying. As a result, the location-dependent query processing techniques proposed for mobile databases can be enhanced and used in processing location-dependent queries in mobile cloud database systems. Consider the example of finding the restaurant nearest to Anna University in Chennai. Since the location of the university is static, data stored in relational databases with location information are sufficient for processing the query. However, if spatial queries are used answers come in fast. This requires the use of spatial data structures for storing data in the database. On the other hand, if new restaurants crop up frequently, temporal constraints are necessary, in addition to spatial database features. In processing a query such as finding a user the nearest restaurant, factors such as location information, mobility speed, the direction of movement, spatial constraints, temporal constraints and storage structures are required. A query of this sort can be answered by a mobile cloud database system more effectively than a mobile database system in which the size of the cache is miniscule, and the main memory and secondary memory have limited capabilities as well. Therefore, processing location-dependent queries requires support in the form of continuous power supply, spatial features, temporal features and effective storage structures. On the other hand, cloud database systems provide large memory and elastic processing power. Therefore, suitable methods must be developed to process location-dependent queries in mobile cloud database systems.
1.7 DIFFERENT TYPES OF QUERIES IN MOBILE CLOUD DATABASES

Mobility in a mobile cloud database environment can be classified into three categories of entities: “mobile cloud client that submits a query,” “mobile cloud server that processes a query,” and “a moving object which represents data to be provided by a query.” Mobile database queries are classified into four categories-location-dependent, moving object database, spatio-temporal and continuous based on the entities mentioned above (Hocine Grine et al. 2000). All types of queries can be used in cloud databases with their mobility to perform effective data storage and retrieval.

1.7.1 Location Dependent Queries

There are three types of geographical location-dependent queries, according to Ayse et al. (2001). The first type is termed a non location related query in which all the predicates and attributes used in the query are non-location related. For example, the query “Select all restaurants with Indian specialties” is a location-independent query. The second type is a Location Aware Query in which a query has at least one location-related simple predicate or location-related attribute. For example, consider the query ‘What’s the weather like in Tamil Nadu?’ where the location is plainly specified in the query. Finally, if the results of a query depend on the location of a user who submitted the query and is on the move such a query is called a location dependent query. For example the query, ‘Find me the nearest hotel within 2 kilometers of my current position,’ contains implicit location information where location values change with respect to the user’s mobility.
1.7.2 Moving Object Database Queries

All queries issued from a mobile client where the data are stored either in the form of objects and classes or in the form of spatial objects are called moving object database queries, which are themselves on the move. For example, a query of this type could be ‘Find all cars within a 10-meter distance from mine’ (Prasad Sistla et al. 1998).

1.7.3 Spatio-Temporal Queries

In a mobile cloud environment, responses to user queries may vary with respect to a change in locations i.e., query results may depend on a query’s spatial properties (Mohamed et al. 2004). Therefore, query results must be such that they are both relevant to the query and valid for the bounded location for a location-bounded query. On the other hand, a spatio-temporal query type includes all queries combining location information with temporal information in the form of instants or intervals. There are two types of spatio-temporal queries in the literature. The first type considers trajectories of moving objects describing a temporal history of the object’s movement, and the second type focuses on the current position of a moving object and possibly its future position using predictions.

1.7.4 Continuous Queries for Mobile Cloud Databases

Continuous queries for mobile cloud databases can be adapted from mobile databases with extended temporal constraints. Moreover, continuous queries allow users to receive new results when they become available during movement for a time interval. For example, a driver may ask for petrol bunks located within 10 kilometers from his current position. The result of the query is a set of petrol bunks that varies continuously with the movement of the driver (Prasad Sistla et al. 1998).
1.7.5 Location Dependent Continuous Queries

Porkaew (2000) termed continuous queries depending on a change of location (with respect to time, movement, direction and space) location-dependent continuous queries (LDCQ). In this query, future temporal logic is used to discover the future location of a moving object before providing a query answer. The user provides a time interval or a spatial location as a condition for the query to be terminated. Up to such time, the query processor present in mobile cloud databases searches of the services available in the current location and provides answers continuously, based on a time slot, to the user. The query is terminated when the spatio-temporal constraints are satisfied.

1.8 FEATURE SELECTION

Feature selection plays a major role in a decision making system since it is used to perform attribute reduction. Moreover, feature selection algorithms help improve the performance as well as prediction accuracy with a considerable decrease in training time. Feature selection (or preprocessing) comprises finding the relevant attributes from a data set and discarding irrelevant ones. The central goal of feature selection is to obtain a subset of features that describes the given problem properly with minimal performance degradation. It has several advantages (Guyon et al. 2006), including performance improvement in the system. It also, helps in data understanding, gaining a knowledge of the process and helping visualize it. The major advantages of preprocessing are data reduction, limiting storage requirements and cost reduction.

There are two key models that deal with feature selection: filter methods and wrapper methods (Kohavi & John 1997). While wrapper models involve optimizing a predictor as part of the selection process, filter models
rely on the general characteristics of the training data to select features independent of a predictor. Wrapper models tend to give better results and are more precise than filter models.

Feature selection can be carried out using soft computing techniques. Genetic algorithm is a soft computing technique which is empirical in nature and hence they are capable of providing optimal number of features. Genetic algorithms, reduces the search space by applying suitable fitness function. Hence, genetic algorithm based feature selection is more appropriate for medical diagnosis systems. In addition, classified results using genetic algorithm in a medical diagnosis system cannot be quantified easily in order to provide a treatment plan. On the other hand, fuzzy logic allows for a qualitative analysis in medical diagnosis based on the gradation of truth values. Such a fuzzy logic integrated with temporal constraints can provide more appropriate medical decisions when an optimal feature is used for classification. Therefore, a new feature selection algorithm based on genetic algorithms and a new classification algorithm using fuzzy temporal constraints has been proposed in this work to perform effective decision making.

In this work, a new genetic algorithm based feature selection algorithm is proposed that selects a suitable set of optimal number of features for TYPE 1 diabetics and a new and different set of optimal number of features for TYPE 2 diabetics. The chief advantages of these proposed algorithms are attribute selection, tuple reduction and data cleaning.

1.9 CLASSIFICATION TECHNIQUES

Classification is used to learn a model from a set of labeled data instances and, thereafter, to classify a test instance into one of the classes using the learned model (Duda et al. 2000, Tan et al. 2005). Classification techniques function in a two-phase manner. The training phase learns a
classifier using the available labeled training data. The testing phase classifies a test instance as normal or abnormal, using the classifier. This kind of classification can be done either by a multi-class or one-class classifier. Multi-class classification techniques assume that the training data contains labeled instances belonging to multiple normal classes (Stefano et al. 2000). Such classifiers distinguish between each normal class and the rest of the classes. A test instance is considered abnormal if it is not classified as normal by any classifier. Certain techniques in this sub category associate a confidence score with the prediction made by the classifier. If none of the classifiers are confident in classifying the test instance as normal, it is declared to be abnormal. One-class classification-based anomaly detection techniques assume that all training instances have only one class label. Such techniques learn a discriminative boundary around normal instances using a one-class classification algorithm. A test instance that does not fall within the learned boundary is declared a disease-affected record.

1.9.1 Decision Trees

Each non-terminal node represents a test or decision, the outcome of which is based on the choice of a certain branch, on the considered data item in decision trees. To classify a particular data item, decision tree algorithms start at the root node and follow the assertions down until they reach a leaf node. A decision is made when the leaf node is reached. Decision trees are a special form of a rule set, characterized by their hierarchical organization of rules.

1.9.2 Neural Networks

Neural Networks (NN) are classification technique modeled to simulate the working of a human brain (Ahmad Ghadiri & Nasser Ghadiri 2011). Just as the human brain consists of millions of neurons interconnected
by synapses, a neural network is, likewise, a set of connected input or output units in which each connection has a weight associated with it. In the learning phase, the network learns by adjusting the weights so as to predict the correct class label of the input. An artificial neural network consists of a connected set of processing units with weights assigned to each connection determining how one unit affects the other. The subsets of such units act as input nodes output nodes, while the remaining nodes constitute the hidden layer. A neural network performs a functional mapping from input values to output values by assigning an activation function to each input node and allowing it to propagate through nodes in the hidden layer to the output nodes.

1.9.3 Naive Bayesian Classifiers

Naive Bayesian classifiers utilize Bayes’ theorem to classify new instances of data samples (Srilatha Chebrolu et al. 2005). Each instance is a set of attribute values described by a vector, $X = (x_1, x_2, \ldots, x_n)$. Considering $m$ classes, the sample $X$ is assigned to the class $C_i$ iff $P(X|C_i)P(C_i) > P(X|C_j)P(C_j)$ for all $j$ in $(1, m)$. The sample belongs to the class with the maximum posterior probability for the sample. The ratio of the frequency of value $X_k$ for attribute $A_k$ and the total number of samples in the training set are evaluated for categorical data $P(X_k|C_i)$. And for continuous valued attributes, the Gaussian distribution can be assumed without loss of generality. The basic assumption in Bayesian approach is that it considers attributes to be conditionally independent, due to which, naïve Bayesian classifiers give satisfactory results in identifying classes for instances, not exact probabilities. Applications that use naïve Bayesian classifiers are spam mail, text classification, to name only a few. Ideally, Bayesian classifiers are least prone to errors and their limitation is the requirement of prior probabilities. The amount of probability information required is exponential in terms of the number of attributes, the number of classes and the maximum
cardinality of attributes. With an increase in the number of classes or attributes, space and computational complexity of Bayesian classifiers increase exponentially.

1.9.4 Fuzzy Sets

Fuzzy sets are an efficient approach for representing and processing uncertain information (Zadeh 1998, Wei Zhang et al. 2010). Uncertainty arises in the form of imprecision, non-specificity, inconsistency, vagueness, etc., in today’s databases. Fuzzy sets exploit uncertainty as a challenge to make system complexity manageable. Furthermore, they constitute a powerful approach to deal not only with incomplete, noisy or imprecise data, but are also helpful in developing uncertain models of data that provide smarter and smoother performances than traditional systems. Moreover, Song et al (2010) introduced a new neuro-fuzzy cognitive model.

In this research, a novel Fuzzy Temporal Classification Algorithm (FTCA) based on fuzzy temporal rules is proposed, applying fuzzy rules to perform the first level-classification. Groups are further refined by developing fuzzy cognitive maps constrained with temporal intervals to make the final classification. Fuzzy cognitive maps use 80% of the dataset for training and 20% for testing. The main advantage of this proposed algorithm is that it uses temporal rules for feature selection, classification and temporal analysis-based decision making. Finally, all the decisions are given to a collaborative filtering agent for providing effective recommendations.

1.10 MEDICAL RECOMMENDATION SYSTEMS

In recent years, technological improvements in networking, computer architecture, database systems, data warehousing and mining, distributed systems and grid computing have led to the development of
sophisticated medical diagnostic systems providing support to people affected by serious diseases such as diabetes, heart disease and cancer. Among these, diabetic is one of the most important, given that the patient is aware of its existence only after it has reached an advanced stage, with the immediate family of diabetic parents most prone to diabetes themselves. Therefore, an automatic diagnosis and recommendation system is necessary to assist the families of diabetic patients. Such a system needs to store very large volumes of data evolving over time, leading to a near-constant increase in storage requirements. Therefore, existing relational database management systems are unsuitable for storing medical data. Given its temporal nature.

Temporal reasoning for medical diagnosis can perform four crucial tasks: analyze patient histories, predict the effects of medicine, learn rules based on behavioral changes brought on by prolonged treatments over a given period of time, and plan alternate treatment methods. Time-based reasoning in medical diagnosis is called on to consider different granularities of time based on the nature of the disease. For example, the granularity of time for heart patients can be either seconds or milliseconds. On the other hand, granularity for diabetics can start with days and weeks and may continue for months and years. Therefore, time stamping for different applications in the domain of healthcare must be different, based on their nature. However, interval stamping of tuples can be an appropriate choice since all medical histories are analyzed over a period of time. Storing history was a challenge when relational databases were used. Hence, extended relational models including temporal databases, were proposed by researchers in the field of database management systems. In addition, since the artificial intelligence community focused largely on developing expert systems for medical diagnosis-attention was on rules rather than data to resolve the problem of storage space. However, when rule chaining was applied repeatedly in an application to obtain inferences, processing power became a challenge. Both storage and
processing powers can be enhanced by using virtualization in cloud computing, so cloud databases are the most preferred destination for most researchers in the area of medical informatics using computational intelligence. Hence, cloud databases (Sidorov & Ng 2013) which can handle big data well have been proposed in this work for storing the history of diabetic patients and to suitably analyze such data so as to provide recommendations. The major advantages of cloud computing include scalability, elasticity, virtualization and storage features.

1.11 PROPOSED WORK

This work proposes a novel secured storage and retrieval system for medical information in the cloud. Further, an intelligent medical diagnosis system which performs temporal analysis using fuzzy temporal rules is proposed for the effective diagnosis of diabetes in patients using historical data stored in cloud databases.

1.11.1 Secured Storage Mechanism

The proposed secured storage and retrieval system consists of five algorithms namely intelligent secured storage algorithm, data aggregation algorithm, data partitioning algorithm, data merging algorithm and location dependent query processing algorithm. The intelligent secured storage algorithm uses a combination of the Caesar and Hill ciphers with temporal rules for encrypting data. These data are used by the indexing component of the database manager, along with spatial and temporal constraints based on locations and time intervals, to develop a spatio-temporal tree based data structure towards productively storing the encrypted data. The storage and retrieval processes are supported by the proposed three new algorithms: data aggregation, data partitioning and data merging. These techniques are useful
for query processing while the temporal analysis is being undertaken by the proposed location-dependent query processing algorithm.

1.11.2 Location Dependent Temporal Query Processing

In this thesis, a new algorithm for performing location-dependent query processing is proposed for the mobile cloud. In this model, data is stored in the cloud but the user permitted to move from one place to another during query processing. The location-dependent query processing algorithm uses the spatio-temporal tree data structure created during data indexing for enhanced query processing performance. It considers the costs incurred by the I/O, network, encryption, decryption, cloud resources and mobility to perform query optimization. A major contribution of this algorithm is that it effectively takes into consideration security aspects and large storage requirements, with temporal analysis for competent decision making.

1.11.3 Sequential Pattern Analysis

Efficient algorithms for finding sequential patterns from medical datasets stored in encrypted form and pertaining to diabetics are proposed, based on association rules and intelligent agents. For this purpose, data encrypted using the Caesar and Hill ciphers are analyzed using rules based on the importance of the application domain. Moreover, a multi-agent system is employed along with the association rule mining algorithm to identify sequential patterns. The proposed system provides in-depth analysis of medical datasets stored in the cloud efficaciously and securely. In addition, communication and coordination are performed using intelligent agents in order to make accurate decisions with regard to diabetics.
1.11.4 Recommendation System

The proposed medical recommendation system uses a newly-proposed feature selection algorithm for optimal feature selection, and a new classification algorithm using temporal fuzzy rules for effective decision making on medical data. Also, it uses the results obtained from the location-dependent query processing module and sequential pattern analysis module to offer worthwhile recommendation. The recommendation system uses secret keys for decrypting the recommended information and providing it to the user in decrypted form.

1.11.4.1 Feature selection

In this work, feature selection has been carried out using the University of California, Irvine (UCI) Diabetic Repository Dataset. New and additional features have been added following consultations with domain experts on Indian conditions. For performing successful feature selection, a new genetic algorithm based feature selection algorithm is proposed that selects a suitable set of optimal number of features for TYPE 1 diabetics and a new and different set of optimal number of features for TYPE 2 diabetics.

1.11.4.2 Fuzzy temporal classification

Further, a novel Fuzzy Temporal Classification Algorithm (FTCA) based on fuzzy temporal rules is proposed. It applies fuzzy rules to perform the first-level classification. Groups are further refined by developing fuzzy cognitive maps constrained with temporal intervals to make the final classification. The fuzzy cognitive maps use 80% of the dataset for training and 20% for testing. The principal advantage of the proposed algorithm is that it uses temporal rules for feature selection, and classification and temporal
analysis based decision making. Finally, all the decisions are given to a collaborative filtering agent to provide effective recommendations.

1.12 THESIS CONTRIBUTIONS

The major contributions of this research are:

1. The introduction of a new encryption algorithm using the Hill and Caesar ciphers for securing intellectual medical data stored in cloud databases,

2. The proposal of a new data aggregation algorithm for effective data storage and improved security,

3. An intelligent data merging algorithm planned for accessing encrypted and original datasets,

4. Intelligent query processing algorithms projected for handling location-dependent queries, location-specific queries and location-dependent continuous queries,

5. A new cluster based transaction model advanced to divide transactions into clusters and maintain their consistency,

6. The application of a new medical expert system for decision making using newly proposed techniques for feature selection, classification and query processing using spatio-temporal constraints and fuzzy rules, and

7. The scheme of a classification model using association rule mining based on frequent pattern mining, to perform medical data mining.
1.13 THESIS ORGANIZATION

The reminder of the thesis is organized as follows:

**Chapter 2** discusses the related work.

**Chapter 3** depicts the architecture of the proposed system.

**Chapter 4** explains the techniques propounded for secured storage and retrieval.

**Chapter 5** elaborates on the details of location-dependent query processing in the cloud database.

**Chapter 6** explains the medical diagnosis system presented in this work.

**Chapter 7** concludes the work and suggests future advances.