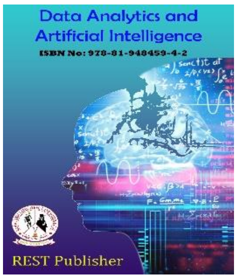
Data Analytics and Artificial Intelligence

Vol: 3(2), March 2023

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/

DOI: https://doi.org/10.46632/daai/3/2/33


Data Analytics and Artificial Intelligence
ISBN No: 978-81-948459-4-2
REST Publisher

A Multi Cluster Secure System for Out Sourced Encrypted Data Bases

[1]Kalaiselvi. K, [2]Aswini. G, [3]Jayanthi. P

[1,2,3] *St. Joseph's college of arts and science for women, Hosur,*

*Corresponding Author Email: Singhroopendra99@gmail.com

Abstract. *Recently, database users have begun to use cloud database services to outsource their databases. The reason for this is the high computation speed and the huge storage capacity that cloud owners provide at low prices. However, despite the attractiveness of the cloud computing environment to database users, privacy issues remain a cause for concern for database owners since data access is out of their control. Encryption is the only way of assuaging users' fears surrounding data privacy, but executing Structured Query Language (SQL)queries over encrypted data is a challenging task, especially if the data a re-encrypted by a randomized encryption algorithm. The existing system, which is an indexing scheme to encode the original table's tuples into bit vectors (BVs) prior to the encryption. The resulting index is then used to narrow the range of retrieved encrypted records from the cloud to a small set of records that are candidates for the user's query. Based on the indexing scheme, the designed system is used to execute SQL queries over the encrypted data. The data are encrypted by a single randomized encryption algorithm, namely the Advanced Encryption Standard-Cipher-Block Chaining (AES-CBC). In the existing scheme, the index values (BVs) are stored at user's side, and which is supported most of relational algebra operators, such as select, join, etc. In the existing system, the cloud data server is able to extract the ratio of two underlying data records, and key management server can decrypt the cipher texts of owners' data records, which break the privacy security.*

Keywords: *Modules, Module Description.*

1. Introduction

In the contemporary electronic era, both individuals and organizations need scalable data storage and high-performance computing units to process and store their data. Historically, only large organizations/companies have been able to own such units, as they were not affordable for most individuals and small companies. With the rise of cloud computing, however, this problem has been solved, as users can now rent storage and computational units as needed at an affordable price. Most cloud providers provide data bases as a service, which allow individual users and companies to outsource their data and access them at any time, from any location. According to the report in [1], the compound annual growth rate (CAGR) of cloud database market is anticipated to be46.78% in 2023, Figure.1. However, given that privacy breaches are one of the most common threats in the cloud computing environment, many people have expressed concerns about privacy when outsourcing sensitive data. For instance, untrustworthy cloud service providers might steal personal customer information, such as email addresses, mailing addresses, and phone numbers, and sell that information to third parties, who can then use it to send irritating advertisements to users via email, mail, and telephone. More importantly, attackers who target a cloud provider can gain access to customers' sensitive personal information, such as social security numbers (SSNs). This has serious consequences, as criminals can use these data to impersonate customers in situations such as financial transactions (e.g., telephone banking). Thus, sensitive data are restricted from being processed or sold to a third party. Therefore, significant evolution in the cloud computing environment could make such services unattractive to consumers if changes occur without also providing appropriate solutions for privacy breach issue. Such an issue must be tackled if cloud providers are to gain the trust of users and organizations so that they will outsource sensitive data without worrying about data leakages.

1. R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan [2012] done research on "CryptDB: Processing queries on an encrypted database" [1]. This research analysed about the recent years; encrypted databases have emerged as a promising direction that provides data confidentiality without sacrificing functionality: queries are executed on encrypted data. However, many practical proposals rely on a set of weak encryption schemes that have been shown to leak sensitive data. In this paper, we propose Arx, apracticalandfunctionallyrichdatabasesystemthatencryptsthedataonlywithsemantically secure encryption schemes. We show that Arx supports real applications such as Share La Te X with a modest performance overhead.
2. R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu [2004] done research on "Order preserving encryption for numeric data" [2]. This research analyzed about the Encryption is a well-established technology for protecting sensitive data. However, once encrypted, data can no longer be easily queried aside from exact matches. Authors present an

order-preserving encryption scheme for numeric data that allows any comparison operation to be directly applied on encrypted data. Query results produced are sound (no false hits) and complete (no false drops). Proposed scheme handles updates gracefully and new values can be added without requiring changes in the encryption of other values. It allows standard database indexes to be built over encrypted tables and can easily be integrated with existing database systems. The proposed scheme has been designed to be deployed in application environments in which the intruder can get access to the encrypted database, but does not have prior domain information such as the distribution of values and cannot encrypt or decrypt arbitrary values of his choice. The encryption is robust against estimation of the true value in such environment.

3. 3.Y.-D. Jang and J.-H. Kim [2016] done research on “A comparison of the query execution algorithms in secure database system” [3]. This research analyzed about the accordance with the database management, DAS (database as service) model is one solution for outsourcing. However, we need some data protection mechanisms in order to maintain the database security The most effective algorithm to secure databases from the security threat of third-party attackers is to encrypt the sensitive data within the database. However, once we encrypt the sensitive data, we have difficulties in queries execution on the encrypted database. In this research, we focus on the search process on the encrypted database. We proposed the selective tuple encryption method using Bloom Filter which could tell us the existence of the data.
4. 4.H. Hacigümü, B. Iyer, C. Li, and S. Mehrotra [2002] done research on “Executing SQL over encrypted data in the database-service-provider model” [4]. This research analyzed about the Rapid advances in networking and Internet technologies have fueled the emergence of the "software as a service" model for enterprise computing. Successful examples of commercially viable software services include rent-a-spreadsheet, electronic mail services, general storage services, disaster protection services."Database as a Service" model provides users power to create, store, modify, and retrieve data from anywhere in the world, as long as they have access to the Internet. It introduces several challenges, an important issue being data privacy. It is in this context that we specifically address the issue of data privacy. There are two main privacy issues. First, the owner of the data needs to be assured that the data stored on the service-provider site is protected against data thefts from outsiders. Second, data needs to be protected even from the service providers, if the providers themselves cannot be trusted. In this paper, we focus on the second challenge. Specifically, we explore techniques to execute SQL queries over encrypted data. Our strategy is to process as much of the query as possible at the service providers' site, without having to decrypt the data. Decryption and the remainder of the query processing are performed at the client site. The research explores an algebraic framework to split the query to minimize the computation at the client site.
5. 5. O. M. B. Omran and B. Panda [2015] done research on “Efficiently managing encrypted data in cloud databases” [5]. This research analysed about the Cloud computing has brought many advantages to organizations and computer users. It allows different service providers to distribute many applications as services in an economical way. Therefore, many users and companies have begun using cloud computing. However, they are concerned about their data when they store it on a third-party server, the cloud. The private data of individual users and companies is stored and managed by the service providers on the cloud, which offers services on the other side of the Internet in terms of its users, and consequently results in privacy concerns. In this research, a technique has been explored to encrypt the data on the cloud and to execute and run SQL queries on the cloud over encrypted data. The strategy is to process the query at the service providers' site without having to decrypt the data. Also, to achieve efficiency, no more than the exact set of requested data is returned to the client. Data decryption is performed at the clients it’s to prevent any leakage at the cloud or during transmission. Two techniques have been provided to effectively store the encrypted data.

2. Modules

1. Data set pre-processing
2. Data set encryption using Advanced Encryption Standard-Cipher-Block Chaining (AES-CBC).
3. Implementation of indexing scheme for tuples
4. Implementation of Query Execution System
5. Implementation of Multi Clustered Indexing Scheme for tuples
6. Performance Comparison

3. Module Description

1. Dataset Pre-Processing

In this module, Data owner having their individual registration and login forms. After the login, data owner uploads the adult dataset file collected from the UCI bench mark resources (https://archive.ics.uci.edu/ml/datasets/adult) and storing the min to database server.

2. Encryption using Advanced Encryption Standard-Cipher-Block Chaining (AES-CBC).

In this module, data owner fixing the sensitive attribute to encrypt the records using AES algorithm and cipher-text and stores into separate table of the database server. To provide privacy and a high level of security, our approach used AES-CBC to encrypt sensitive data. Only the query manager (QM) keeps and maintains the secret keys (SKs).

3. Implementation of indexing scheme for tuples

In this module, we construct the Query Manager (QM) as a trusted server that resides in cloud server. It works as an intermediary between users and the Cloud and is responsible for processing queries and encoding Bit Vector (BV) for each table. The Partitioning Tree (PT) is the primary element of proposed system in which the query is appropriately rewritten for execution by the cloud server. The owner of the table participates in the construction of the PT by specifying which columns are sensitive and should be encrypted. The name of the table is the root of the tree, and the second-level nodes are the sensitive columns that have to be encrypted. Nodes in the third level, each of which is assigned an ID, represent the partitions of all the sensitive columns. This information is stored in cloud database and processed in Query Manager.

4. Implementation of Query Execution System

Select is an essential statement in all database applications. In this module, Query Manager (QM) will decrypt the encrypted tuples retrieved from the Cloud. If the column(s) in the query condition is sensitive, so they are stored in encrypted form in the Cloud. In this case of non sensitive, the QM directly searches and retrieves the data corresponding to the query condition from the Cloud.

5. Implementation of multi clustered indexing scheme for tuples

In this module, we initially construct Partitioning Tree (PT) using Multi Clustering algorithm. The owner of the table participates in the construction of the PT by specifying which columns are sensitive and should be encrypted. The data collected are grouped into several clusters based on the attributes by using Multi cluster index algorithm; the grouping is done by corresponding cluster centroid. Then the distance of each object to the centroids are calculated, the value is based on selection of particular attribute. Since we are not sure about the location of the centroid, we need to adjust the centroid location based on the current updated data which is stored in Partitioning Tree.

6. Performance Comparison

The System is evaluated against the following parameters namely, Total processing time, Total Processing Memory, through put record counts., etc.,

4. Conclusion

Solve the above-mentioned problem, we propose a highly secure privacy- preserving out sourced multi clustering scheme under multiple keys in cloud computing. In the proposed system, cloud computing service (CCS) and Key Management Server (KMS) jointly perform clustering over the encrypted data records without exposing data privacy. Specifically, we use encryption which has additive homomorphic property and AES encryption to double encrypt data records, where the former crypto system prevents CCS from obtaining any useful information from received cipher texts and the latter one protects data records from being decrypted by KMS. We tested the proposed system by implementing it and comparing its performance against the existing system. We evaluated it in terms of execution time and space requirements. We find that the proposed system requires both less execution time and space when compared with existing systems.

References

[1] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, ``Crypt DB: Processing queries on an encrypted database, "Commun. ACM, vol.55,no.9,pp.103-111,Sep.2012.

[2] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, ``Order preserving encryption for numeric data, "in Proc. ACMSIGMOD Int. Conf .Manage. Data (SIGMOD),2004,pp.563-574.

[3] Y.-D. Jang and J.-H. Kim, ``A comparison of the query execution algorithms in secure database system, "Int. J. Electr. Comput .Eng., vol.6,no.1,p.337,Feb.2016.

[4] H. Hacigümü³, B. Iyer, C. Li, and S. Mehrotra, ``Executing SQL over encrypted data in the database- service-provider model," in Proc. ACM SIGMOD Int. Conf. Manage. Data (SIGMOD), 2002,pp.216-227.

[5] O.M.B.OmranandB.Panda,``Efficientlymanagingencrypteddataincloudddatabases,"inProc.IEEE2ndInt.Conf.CyberSecur.CloudComput.,Nov.2015,pp.266-271.