International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 3 Mar 2022, pp: 1365-1368 www.ijaem.net ISSN: 2395-5252

Review on Encrypted Vector Using RSA for Multi-Keyword Search Over Encrypted Data in Hybrid Clouds

Miss Gauri Banore, Prof. A.A. Chinchamalatpure
ME Final year CSE, Dr. Sau. K.G.I.E.T. Amravati, Maharashtra
Prof, Dr. Sau. K.G.I.E.T. Amravati, Maharashtra

Submitted: 25-03-2022 Revised: 31-03-2022 Accepted: 03-04-2022

ABSTRACT — Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. In this paper, we present a secure multi keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. Specifically, the vector space model and the widely-used TF _ IDF model are combined in the index construction and query generation. We construct a special tree-based index structure and propose a “Greedy Depth-first Search” algorithm to provide efficient multi-keyword ranked search. The secure kNN algorithm is utilized to encrypt the index and query vectors, and meanwhile ensure accurate relevance score calculation between encrypted index and query vectors. In order to resist statistical attacks, phantom terms are added to the index vector for blinding search results.

Due to the use of our special tree-based index structure, the proposed scheme can achieve sub-linear search time and deal with the deletion and insertion of documents flexibly. Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.

I. INTRODUCTION

Cloud computing has been considered as a new model of enterprise IT infrastructure, which can organize huge resource of computing, storage and applications, and enable users to enjoy ubiquitous, convenient and on demand network access to a shared pool of configurable computing resources with great efficiency and minimal economic overhead [1]. Attracted by these appealing features, both individuals and enterprises are motivated to outsource their data to the cloud, instead of purchasing software and hardware to manage the data themselves.

Despite of the various advantages of cloud services, outsourcing sensitive information (such as e-mails, personal health records, company finance data, government documents, etc.) to remote servers brings privacy concerns.

The cloud service providers (CSPs) that keep the data for users may access users’ sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing [2]. However, this will cause a huge cost in terms of data usability. For example, the existing techniques on keyword-based information retrieval, which are widely used on the plaintext data, cannot be directly applied on the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical.

In order to address the above problem, researchers have designed some general-purpose solutions with fully-homomorphic encryption [3] or oblivious RAMs [4]. However, these methods are not practical due to their high computational overhead for both the cloud sever and user. On the contrary, more practical special-purpose solutions, such as searchable encryption (SE) schemes have made specific contributions in terms of efficiency, functionality and security. Searchable encryption schemes enable the client to store the encrypted data to the cloud and execute keyword search over ciphertext domain. So far, abundant works have been proposed under different threat models to achieve various search functionality, such as single keyword search, similarity search, multi-keyword Boolean search, ranked search, multi-keyword ranked search, etc. Among them, multi-keyword ranked search achieves more and more attention for its practical applicability. Recently, some dynamic schemes have been proposed to support inserting and deleting operations on document collection. These are significant works as it is highly possible that the data owners need to update their data on the cloud server. But few of the dynamic schemes support efficient multi-keyword ranked search.

II. PROPOSED OBJECTIVE

To enable secure, efficient, accurate and dynamic multi-keyword ranked search over outsourced encrypted cloud data under the above models, our system has the following design goals. Dynamic. The proposed scheme is designed to provide not only multi-keyword query and accurate result ranking, but also dynamic update on document collections. Search efficiency. The scheme aims to achieve sublinear search efficiency by exploring a special tree-based index and an efficient search algorithm. Privacy-preserving. The scheme is designed to prevent the cloud server from learning additional information about the document collection,[5] [6] the index tree, and the query. The specific privacy requirements are summarized as follows,

1) Index confidentiality and query confidentiality. The underlying plaintext information, including keywords in the index and query, TF values of keywords stored in the index, and IDF values of query keywords, should be protected from cloud server;

2) Trapdoor unlink ability. The cloud server should not be able to determine whether two encrypted queries (trapdoors) are generated from the same search request;

3) Keyword privacy. The cloud server could not identify the specific keyword in query, index or document collection by analyzing the statistical information like term frequency. Note that our proposed scheme is not designed to protect access pattern, i.e., the sequence of returned documents.

4) Index Encryption : in this to achieve the level of security proposed methodology finds the required data index and perform the index encryption.

III. EXISTING SYSTEM

MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED (MRSE): Now a day's cloud computing has become essential for many utilities, where cloud customers can slightly store their data into the cloud. Its huge suppleness and financial savings are attracting both persons and enterprise to outsource their local complex data management system into the cloud. To safe guard data privacy and struggle unwanted accesses in the cloud and away from, sensitive data, for example, emails, personal health records, photo albums, videos, land documents, financial transactions, and so on, may have to be encrypted by data holder before outsourcing to the business public cloud;[7] [8] on the other hand, obsoletes the traditional data use service based on plaintext keyword search. Furthermore, apart from eradicating the local storage management, storing data into the cloud supplies no purpose except they can be simply searched and operated. Thus, discovering privacy preserving and effective search service over encrypted cloud data is one of the supreme importance. Ranked search can also gracefully remove redundant network traffic by transferring the most relevant data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking operation on the other hand, should not reveal any keyword to related information. To get better the search result exactness as well as to improve the user searching experience, it is also essential for such ranking system to support multiple keywords search, as single keyword search often give up far too common results. And each keyword in the search demand is able to help narrow down the search result further. "Coordinate matching", as many matches as possible, is an efficient resemblance measure among such multi-keyword semantics to refine the result significance, and has been widely used in the plaintext information retrieval (IR) community. Encryption is a helpful method that treats encrypted data as documents and allows a user to securely search through a single keyword and get back documents of interest. On the other hand, direct application of these approaches to the secure large scale cloud data utilization system would not be necessarily suitable, as they are developed as crypto primitives and cannot put up such high service-level needs like system usability, user searching experience, and easy information discovery. Even though some modern plans have been proposed to carry Boolean keyword search as an effort to improve the search flexibility, they are still not sufficient to provide users with satisfactory result ranking functionality. The solution for this problem is to secure ranked search over encrypted data but only for queries consisting of a single keyword. The challenging issue here is how to propose an efficient encrypted data search method that supports multi-keyword semantics without privacy violation. In this paper, we describe and solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving exact system wise privacy in the cloud computing concept. Along with various multi keyword semantics, select the efficient resemblance measure of "coordinate matching," it means that as various matches as possible, to confine the significance of data documents to the search query. Particularly, inner product similarity the numbers of query keywords show in a document, to quantitatively calculate such similarity assess of that document to the search query.[9] [10] The search query is also illustrates as a binary vector where each bit means whether corresponding keyword appears in this search request, so the resemblance could be exactly calculated by the inner product of the query vector with the data vector. we propose a basic idea for the MRSE using secure inner product computation, which is modified from a secure k-nearest neighbour (kNN) method, and then give two considerably improved MRSE method in a step-by-step way to accomplish different severe privacy needs in two risk models with enlarged attack competence.

IV. PROPOSED SYSTEM

Data owner**:** Data owner has a collection of documents F = {f1; f2; :::; fn} that he wants to outsource to the cloud server in encrypted form while still keeping the capability to search on them for effective utilization. In our scheme, the data owner firstly builds a secure searchable tree index I from document collection F, and then generates an encrypted document collection C for F. Afterwards, the data owner outsources the encrypted collection C and the secure index I to the cloud server, and securely distributes the key information of trapdoor generation (including keyword IDF values) and document decryption to the authorized data users. Besides, the data owner is responsible for the update operation of his documents stored in the cloud server. While updating, the data owner generates the update information locally and sends it to the server.

Data users: Data users are authorized ones to access the documents of data owner. With t query keywords, the authorized user can generate a trapdoor TD according to search control mechanisms to fetch k encrypted documents from cloud server. Then, the data user can decrypt the documents with the shared secret key.

Cloud server Cloud Server stores the encrypted document collection C and the encrypted searchable tree index I for data owner. Upon receiving the trapdoor TD from the data user, the cloud server executes search over the index tree I, and finally returns the corresponding collection of top-k ranked encrypted documents. Besides, upon receiving the update information from the data owner, the server needs to update the index I and document collection C according to the received information.

Our contributions are summarized as follows: 1) We design a searchable encryption scheme that supports both the accurate multi-keyword ranked search and flexible dynamic operation on document collection. 2) Duetothespecialstructureofourtree-basedindex, the search complexity of the proposed scheme is fundamentally kepttologarithmic. Andinpractice, the proposed scheme can achieve higher search efficiency by executing our "Greedy Depth-first Search" algorithm. Moreover, parallel search can be flexibly performed to further reduce the time cost of search process.

V. IMPLEMENTATION

In this part of the testing each of the conditions were tested to both true and false aspects. And all the resulting paths were tested. So that each path that may be generate on particular condition is traced to uncover any possible errors. This type of testing selects the path of the program according to the location of definition and use of variables. This kind of testing was used only when some local variable were declared. The definition-use chain method was used in this type of testing. These were particularly useful in nested statements. In this type of testing all the loops are tested to all the limits possible. The following exercise was adopted for all loops:

- All the loops were skipped at least once.
- For nested loops test the inner most loop first and then work outwards.

For concatenated loops the values of dependent loops were set with the help of connected loop.

Modules:

1. Login Module

In this we are signing in to the application , If the credentials are correct then it will open homepage otherwise it will alert to enter correct details.

2. Keyword Encryption Module

The input text we are providing is encrypted by using several algorithmic techniques.

3. Data Integration

We are combining technical and business processes used to combine **data** from disparate sources into meaningful and valuable information. A complete **data integration** solution delivers trusted **data** from various sources.

4. Encryption

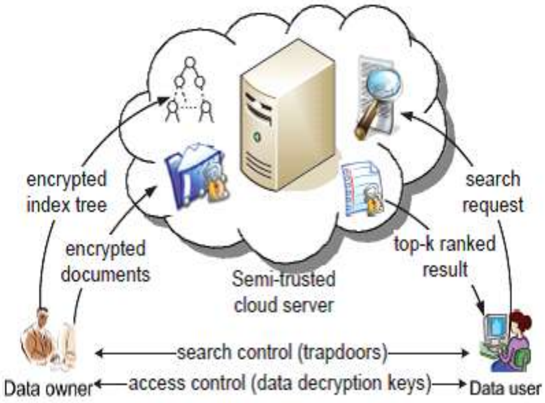
Here we are actually doing the encryption using cryptography.

5. Index Encryption

Single values level **encryption** of the **index** reveals sensitive information, such as frequencies of the **index** values.

Whole **Index** level **encryption** ensures that information about the indexed data cannot be leaked, since the **index** is **encrypted** as one unit.

VI. SYSTEM ARCHITECTURE



All encryption and decryption of data as well as keyword search and displaying multiple results with high weightage is done by system only.

VII. CONCLUSION

There are still many challenge problems in symmetric SE schemes. In the proposed scheme, the data owner is responsible for generating updating information and sending them to the cloud server. Thus, the data owner needs to store the unencrypted index tree and the information that are necessary to recalculate the IDF values. Such an active data owner may not be very suitable for the cloud computing model. It could be a meaningful but difficult future work to design a dynamic searchable encryption scheme whose updating operation can be completed by cloud server only, meanwhile reserving the ability to support multi-keyword ranked search. In addition, as the most of works about searchable encryption, our scheme mainly considers the challenge from the cloud server. Actually, there are many secure challenges in a multi-user scheme. First, all the users usually keep the same secure key for trapdoor generation in a symmetric SE scheme. In this case, the revocation of the user is big challenge. If it is needed to revoke a user in this scheme, we need to rebuild the index and distribute the new secure keys to all the authorized users. Second, symmetric SE schemes usually assume that all the data users are trustworthy. It is not practical and a dishonest data user will lead to many secure problems. For example, a dishonest data user may search the documents and distribute the decrypted documents to the unauthorized ones. Even more, a dishonest data user may distribute his/her secure keys to the unauthorized ones. In the future works, we will try to improve the SE scheme to handle these challenge problems.

REFERENCES

[1]. K. Ren, C. Wang, and Q. Wang, "Security challenges for the public cloud," IEEE Internet Comput., vol. 16, no. 1, pp. 69–73, Jan-Feb. 2012.

[2]. S. Kamara and K. Lauter, "Cryptographic cloud storage," in Proc. Financ. Cryptography Data Secur., 2010, pp. 136–149.

[3]. C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 2009.

[4]. O. Goldreich and R. Ostrovsky, "Software protection and simulation on oblivious rams," J. ACM, vol. 43, no. 3, pp. 431–473, 1996.

[5]. D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. Adv. Cryptol.-Eurocrypt, 2004, pp. 506–522.

[6]. D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, "Public key encryption that allows pir queries," in Proc. Adv. Cryptol., 2007, pp. 50–67.

[7]. D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Secur. Privacy, 2000, pp. 44–55.

[8]. E.-J. Goh, "Secure indexes," IACR Cryptol. ePrint Archive, vol. 2003, p. 216, 2003.

[9]. Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. 3rd Int. Conf. Appl. Cryptography Netw. Secur., 2005, pp. 442–455.

[10]. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," in Proc. 13th ACM Conf. Comput. Commun. Secur., 2006, pp. 79–88.