International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 03
March 2017

A Novel Search Scheme on Dynamic Multi-keyword rank over Encrypted Cloud Data

SANAULLAH[1], MD SHAHZAD ALI[2], BISWADIP BORA[3], FOUZIA SULTANA[4]

[1]B-Tech, Lords Institute Of Engineering And Technology, Mail Id: cmsanaullah@gmail.com
[2]B-Tech, Lords Institute Of Engineering And Technology, Mail Id: shahzadali028@gmail.com
[3]B-Tech, Lords Institute Of Engineering And Technology, Mail Id: biswadipcsebora@gmail.com
[4]Assistant Professor, Lords Institute of Engineering and Technology,
Mail Id: fouziasultana@lords.ac.in

Abstract:-

A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data Due to the incrementing popularity of cloud computing, more and more data owners are incentivized to outsource their data to cloud servers for great accomodation and reduced cost in data management. However, sensitive data should be encrypted afore outsourcing for privacy requisites, which obsoletes data utilization like keyword-predicated document retrieval. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously fortifies dynamic update operations like expunction and insertion of documents. Concretely, the vector space model and the widely-used TF IDF model are cumulated in the index construction and query generation. We construct a special tree-predicated 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 ascertain precise pertinence score calculation between encrypted index and query vectors. In order to resist statistical attacks, phantom terms are integrated to the index vector for visually impairing search results. Due to the utilization of our special tree-predicated index structure, the proposed scheme can achieve sub-linear search time and deal with the effacement and insertion of documents flexibly. Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.

KEY WORDS: - Secure multi-keyword ranked search, Searchable Encryption, Multi keyword, Dynamic Update; Cloud Computing

1. INTRODUCTION

Cloud computing has been considered as an incipient model of enterprise IT infrastructure, which can organize sizably voluminous resource of computing, storage and applications, and enable users to relish ubiquitous, convenient and on-demand network access to a shared pool of configurable computing resources with great efficiency and minimal economic overhead. Magnetized by these appealing features, both individuals and enterprises are incentivized to outsource their data to the cloud, in lieu of purchasing software and hardware to manage the data themselves. Despite the sundry advantages of cloud accommodations, outsourcing sensitive information (such as e-mails, personal health records, company finance data, regime documents, etc.) to remote servers brings privacy concerns. The cloud accommodation providers (CSPs) that keep the data for users may access users' sensitive information without sanction. A general approach to forfend the data confidentiality is to encrypt the data afore outsourcing. However, this will cause a sizably voluminous cost in terms of data usability. For example, the subsisting techniques on keyword-predicated information retrieval, which are widely utilised on the plaintext data, cannot be directly applied to the encrypted data. Downloading all the data from the cloud and decrypt locally is conspicuously impractical. In order to address the above quandary, researchers have designed some general purport solutions with plenarily-homomorphic encryption or oblivious RAMs. However, these methods are not practical due to their high computational overhead for both the cloud server and utilised. On the contrary, more practical special purport solutions, such as searchable encryption (SE) schemes have made categorical contributions in terms of efficiency, functionality and security. Searchable encryption schemes enable the client to store the encrypted data in the cloud and execute keyword search over ciphertext domain. So far, abundant works have been proposed under different threat models to achieve sundry search functionality, such as single keyword search, homogeneous attribute 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 fortify inserting and expunging operations on document accumulation. These are consequential 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.

2. RELEGATED WORK

Existing System

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. All these multi keyword search schemes retrieve search results based on the existence of keywords, which cannot provide acceptable result ranking functionality. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval.

Proposed System

A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data We construct a special tree-based index structure and propose a "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search. 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. Abundant works have been proposed under different threat models to achieve various search functionality, recently, some dynamic schemes have been proposed to support inserting and deleting operations on document collection. This paper proposes a secure tree-based search scheme over the encrypted cloud data, which supports multi keyword ranked search and dynamic operation on the document collection.

3. IMPLEMENTATION

The System and Threat Models:

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.

Known Cipher Text Model

In this model, the cloud server only knows the encrypted document collection C, the searchable index tree I, and the search trapdoor TD submitted by the authorized user. That is to say, the cloud server can conduct cipher text-only attack (COA) in this model.

Known Background Model:

Compared with known cipher text model, the cloud server in this stronger model is equipped with more knowledge, such as the term frequency (TF) statistics of the document collection. This statistical information records how many documents are there for each term frequency of a specific keyword in the whole document collection, as shown in Fig. 2, which could be used as the keyword identity.

Synonym expansion

Are words with the same or similar meanings? In order to improve the accuracy of search results, the A Secure and Dynamic Multi-keyword Ranked extracted from out sourced text documents need to be extended by common synonyms, as cloud customers' searching input might be the synonyms of the predefined A Secure and Dynamic

Multi-keyword Ranked, not the exact or fuzzy matching A Secure and Dynamic Multi-keyword Ranked due to the possible synonym substitution and/or her lack of exact knowledge about the data. A common synonym thesaurus is built on the foundation of the New American Roget's College Thesaurus (NARCT) . Then the keyword set is extended by using the constructed synonym thesaurus.

Rank function

In information retrieval, a ranking function is usually used to evaluate relevant scores of matching files to a request. Among lots of ranking functions, the "TF×IDF" rule is most widely used, where TF (term frequency) denotes the occurrence of the term appearing in the document, and IDF (inverse document frequency) is often obtained by dividing the total number of documents by the number of files containing the term. That means, TF represents the importance of the term in the document and IDF indicates the importance or degree of distinction in the whole document collection.

Cryptography

The art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret key can decipher (or decrypt) the message into plain text. Encrypted messages can sometimes be broken by cryptanalysis, also called code breaking, although modern cryptography techniques are virtually unbreakable.

4. EXPERIMENTAL RESULTS

Fig:-1 Login Screen

Fig:-2 Registration Screen

Fig:-3 Key Screen

Fig:-4 Files Data

5. CONCLUSION

In this paper, a secure, efficient and dynamic search scheme is proposed, which supports not only the accurate multi-keyword ranked search but also the dynamic deletion and insertion of documents. We construct a special keyword balanced binary tree as the index, and propose a "Greedy Depth-first Search" algorithm to obtain better efficiency than linear search. In addition, the parallel search process can be carried out to further reduce the time cost. The security of the scheme is protected against two threat models by using the secure kNN algorithm. Experimental results demonstrate the efficiency of our proposed scheme. 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. Firstly, 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. Secondly, 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.

6. REFERENCE

[1] S. Kamara, and K. Lauter, "Cryptographic cloud storage,"

[2] Grzonkowski, and P. M. Corcoran, "Sharing cloud services: user authentication for social enhancement of home networking,"

[3] P.A. Cabarcos, F.A. Mendoza, R.S. Guerrero, A.M. Lopez, and D. Diaz-Sanchez, "SuSSo: seamless and ubiquitous single sign-on for cloud service continuity across devices,"

[4] D. Diaz-Sanchez, F. Almenarez, A. Marin, D. Proserpio, and P.A. Cabarcos, "Media cloud: an open cloud computing middleware for content management,"

[5] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing,"

[6] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data,"

[7] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data,"

[8] Q. Chai, and G. Gong, "Verifiable symmetric searchable encryption for semi-honest-but-curious cloud servers,"