An Efficient, Secure Ranked Multi-Keyword Search Protocol for Cloud Computing

Dr. Manavalasundharam[1], Nandhini. T[2], Shivaranjani. M[3], Sowndharya. G[4]

[1]Assistant Professor, Department of Computer Science and Engineering, VCET, Tamilnadu, India.
[2, 3, 4] BE, CSE, VCET, Tamilnadu, India.

Email: manosundar@gmail.com, mandhinithangavel96@gmail.com, shivjani988@gmail.com, sowndhgopal@gmail.com

Received date: 15th April, 2018, Accepted Date: 25th April, 2018.

Abstract - Cloud computing is a new model of IT infrastructure, which can organize resource of computing, storage, and applications, and enable users to a shared pool of configurable computing resources with great efficiency. Attracted by these appealing features, both individuals and enterprises are motivated to contract out their data to the cloud, instead of purchasing software and hardware. So far, abundant works have been proposed under different threat models to achieve various search functionality, such as single keyword search, similarity search, etc. Among them, multi-keyword ranked search achieves more and more attention. In this project, we proposed a secure and ranked multi-keyword search protocol in a multi-owner cloud model over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. A tree-based index structure is constructed, and efficient multi-keyword ranked search is proposed. Due to the use of the tree-based index structure, the proposed scheme can deal with the deletion and insertion of documents flexibly.

Keywords - Cloud Service Provider, PRMSM, Multi-Keyword Search Protocol, Tree-based index structure, SRMSP.

I. INTRODUCTION

Cloud computing is a technology that is changing the way IT hardware and software are designed and purchased. As a new model of computing, cloud computing provides abundant benefits including easy access, decreased costs, quick deployment, and flexible resource management, etc. Enterprises of all sizes can leverage the cloud to increase innovation. Despite the abundant benefits of cloud computing, for privacy concerns, individuals and enterprise users are reluctant to outsource their sensitive data, personal health records, and government confidential files, to the cloud. This is because once sensitive data are outsourced to a remote cloud; the corresponding data owners lose the control of these data. Cloud Service Providers (CSPs) would promise to ensure owners' data security using mechanisms like virtualization and firewalls. However, these mechanisms do not protect owners' data privacy. Encryption on sensitive data before outsourcing can preserve data privacy. However, this method is impractical because it will cause communication overhead. Therefore, developing a secure search service over encrypted cloud data is of paramount importance.

II. SYSTEM MODELS

A. Existing System

Searchable encryption schemes enable the clients to store the encrypted data in the cloud and execute keyword the search over ciphertext domain. Due to different cryptography primitives, searchable encryption schemes can be constructed using public key based cryptography or symmetric key based cryptography. These early works are single keyword boolean search schemes, which are very simple regarding functionality. A general approach to protecting the data confidentiality is to encrypt the data. Searchable encryption schemes enable the client to store the encrypted data in the cloud and execute keyword search. So far, 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. These are significant works as it is highly possible that the data owners need to update their data on the cloud server. A secure multi keyword search method which utilized local sensitive hash functions to cluster the similar documents. The LSH algorithm is suitable for similar search but cannot provide.

Compared with the single-owner scheme, a multi-owner scheme will have many new challenging problems. First, in the single-owner scheme, the data owner has to stay online to generate trapdoors for data users. How-

ever, when a large amount of data owners are involved, they must stay online simultaneously to produce trapdoors would seriously affect the usability of the search system. Second, since no one would be willing to share our secret keys with others, different data owners would prefer to use their secret keys to encrypt their secret data. Consequently, it is very challenging to perform a secure, and efficient search over the data encrypted with different secret keys. Other, when multiple data owners are involved, we should ensure efficient user registration and revocation mechanisms, so that our system enjoys excellent security and scalability.

PRMSM, privacy preserving ranked multi-keyword search protocol in a multi-owner in the cloud model. To enable cloud servers to perform the secure search without knowing the actual value of both keywords and trapdoors, we construct a novel secure search protocol. As a result, different data owners can use different keys to encrypt their files and keywords. Authenticated data users can issue a query without knowing secret keys of those data owners. To rank the search results and preserve the privacy of scores between keywords and files, we propose an Additive Order and Privacy Preserving Function (AOPPF), which helps the cloud server return the most relevant search results to data users. To prevent the attackers from eavesdropping secret keys, we propose a dynamic secret key generation protocol. As a result, attackers who perform illegal searches would be easily detected. Furthermore, when we want to revoke a data user, PRMSM perform efficient data user revocation. Extensive experiments on real-world datasets confirm the efficacy and efficiency of our proposed schemes.

B. PRMSM - Dynamic Multi-Keyword Ranked Search Model/Algorithm

To enable secure, efficient, accurate and dynamic multi-keyword ranked search over outsourced encrypted cloud data under the above models, our system's design goals are,

Dynamic - The proposed scheme is designed to provide not only multi-keyword query and accurate result ranking, but also a dynamic update on document collections.

Search efficiency - Achieve sub-linear search efficiency by exploring a special tree-based index and an efficient search algorithm.

Privacy-preserving – This is designed to prevent the cloud server from learning additional information about the index tree, and the query. The specific privacy requirements are summarized as follows,

1) Index and query confidentiality. Including keywords in the index and query, TF values of keywords stored in the index, of query keywords, must be protected from cloud server;

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

3) Keyword privacy. The cloud server did not identify the keyword in query, index or document collection by analysing the term frequency. This scheme is not designed to protect access pattern, i.e., the sequence of returned documents.

C. Disadvantages of Existing System

- However, all existing schemes are limited to the single-owner model. Compared with the single-owner scheme, developing a full-fledged multi-owner scheme will have many new challenging problems.
- Huge cost regarding data usability. For example, the existing techniques for keyword-based information retrieval, which are widely used on the plaintext data, cannot be applied to the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical.
- Existing System methods not practical due to their high computational overhead for both the cloud sever and user.

D. Proposed System

For privacy concerns, secure searches over encrypted cloud data have motivated several research works under the single owner model. However, most cloud servers do not just serve one owner; but they support multiple owners to share the benefits brought by cloud computing. In enhancement work, we focus schemes to deal with secure data sharing; The group signature scheme enables users to anonymously use the cloud resources allows data owners to securely share their data files with others including new users. The authentication server computes the revocation parameters and gives the resulting public available by migrating them to the cloud. Such a design can reduce the computation overhead of users to encrypt files. Specially, the computation overhead of users for encryption operations and the ciphertext size is constant and independent of the revocation users. To enable cloud servers to perform a secure search without knowing the actual data of both keywords and trapdoors, we will systematically construct a novel secure search protocol.

E. A Secure and Ranked Multi-keyword Search Protocol (SRMSP) in a multi-owner cloud model/algorithm

In addition of PRMSM model, we propose SRMSP model, which is a variant of the short group signature scheme, to achieve anonymous dynamic group access control, as it supports efficient membership data access with revocation member list verification without a key update or key sharing from owners.

F. Proposed system offers unique features as follows.

- Any user in the group can store and share data files with others by the administration server.
- The encryption complexity and cipher texts are independent of the number of revoked users.
- User revocation (trapdoors) can be achieved without updating the private keys of the remaining users.
- A new user can directly decrypt the files stored in the administration server before participation.
- A dynamic multi-owner data sharing scheme - It implies that any user in the group can securely share data with others by the trusted administration server.
- The proposed scheme is can support dynamic groups efficiently. Specifically, newly granted users can directly decrypt data files uploaded before their participation.
- User revocation can be easily achieved through a novel revocation list without updating the secret keys of the remaining users.
- We provide secure and privacy-preserving access control for the users, which guarantees any member in a group to anonymously utilize the cloud server resource.
- Moreover, the real identities of data owners can be revealed by the administration server when disputes occur.
- Define a multi-owner model for privacy preserving keyword search over encrypted cloud data.

III. MODULES IN THE PROPOSED METHOD

A. Module Description

- Authentication server module
- Data owner/user module
- Cloud storage server module

B. Authentication Server Module

Authentication Server takes charge of system parameters generation, user registration, user revocation. In the given example, the authentication server is acted by the administrator of the system. Therefore, we assume that the authentication server is fully trusted by the other parties. Group users are registered under authentication server for to create, delete or access data's in the cloud. Authentication server distributes public key for every registered owners/user which is used for connecting cloud. User's details with specific signature periodically revoke to the cloud system for request user authorization.

C. User revocation

User revocation is performed by the authentication server via a public available revocation list, based on which group user's can encrypt their data files and ensure the confidentiality against the revoked users. The authentication server compute the revocation parameters and make the result public available by migrating them. Such a design can significantly reduce the computation overhead of users

D. Traceability

Anonymity guarantees that group users can access the cloud without revealing the real identity. It represents an effective protection for user identity, it access potential inside attack risk to the system. The authentication server should have the ability to reveal the real identities of data owners.

E. Data Owner/User Module

Owners/users are a set of registered users they will store their data into the cloud server and share them with others in the group. The group membership is dynamically changed, due to new user's participation in the system. Data owners have a collection of files F. To enable efficient search operations on these files which will be encrypted, data owners first build a secure searchable index on the keyword set extracted from F, and then they submit to the administration server.

F. File Access

Upon receiving I, the administration server re-encrypts the owner's file and outsources the re-encrypted index to the cloud server. Any group user can store and search data files with others in the group by the cloud.

G. Cloud Storage Server Module

The cloud server provides data storage and search services to data owners and data users. After verifying the user connection under signature, user can able to access the particular owner's data with respect to owner's private key is reference to the user identity (IDdata). So the cloud verifies whether the request user is in the revoke list and it provide permission to access the data else throw unauthorized user request.

H. Search Request

Once a data user wants to search k keywords over these encrypted files storing in the cloud server, computes the trapdoors for the users and submits them to the administration server. Once the data user is authenticated by the administration server then follows to re-encrypt the trapdoors.

I. Search Result

Upon receiving the trapdoor T , the cloud server searches the encrypted index I of owner and returns the corresponding encrypted files. To improve the file retrieval accuracy data user would tell the cloud server and cloud server would return the top-k relevant files to the data user.

Apart from that we are using elliptic curve cryptography algorithm for key generation, distribution, encryption and decryption.

- Elliptic curve cryptography (ECC) is an approach to public-key cryptography.
- ECC algorithm is used for key generation, distribution, encryption and decryption between group manager, group members and cloud server authentication and communication.
- Elliptic curve with 160-bitgroup order, which provides a competitive security level with 1,024-bit RSA (Elliptic curve is better than RSA in terms of key computational cost).

IV. CONCLUSION

In this paper, we explore the problem of secure multi-keyword search for multiple data owners and multiple data users in the cloud computing environment. Different from prior works, these enable data users to achieve secure, convenient, and efficient searches over multiple data owners' data. To efficiently authenticate data users and detect attackers who steal the secret key and perform illegal searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. To enable the cloud server to perform secure search among multiple owners' we systematically construct a novel secure search protocol. To rank the search results, we propose a novel additive order and privacy preserving function family. Moreover, we show that our approach is computationally efficient, even for large data and keyword sets. As our future work, on one hand, we will consider the problem of secure fuzzy keyword search in a multi-owner paradigm. On the other hand, we plan to implement our scheme on the commercial clouds.

REFERENCES

[1] M. Airburst, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rifkin, I. Stoics, and M. Zachariah, "A view of cloud computing," Communes. ACM, Vol. 53, No. 4, pp. 50-58, 2010.

[2] C. Wang, S. S. Chow, Q. Wang, K. Reno, and W. Lou, "Privacy-preserving public auditing for secure cloud storage," IEEE Trans. Compote., vol. 62, no. 2, pp. 362-375, 2013.

[3] D. Song, D. Wagner, and A. Perry, "Practical techniques for searches on encrypted data," in Proc. IEEE Int. Sump. Security Privacy, pp. 44-55, 2000.

[4] E. Groh. (2003). Secure indexes [Online]. Available: http://eprint.iacr.org/