Enabling Secure and Efficient Multi-Keyword Ranked Search Scheme

[1]Amrapali S. Ahire, [2]Shweta D. Dhokte, [3]Prajakta J. Karpe, [4]Sunil B. Palde,
[5]Prof. U. R. Patole

[1,2,3,4] *BE Students: Dept. of Computer Engineering, SVIT College, Chincholi, Nashik, Maharashtra, India*
[5] *Professor: Dept. of Computer Engineering, SVIT College, Chincholi, Nashik, Maharashtra, India*

---***---

Abstract - *With increasing popularity of cloud computing, the data owners outsource their sensitive data to cloud servers for flexibility and reduced cost in data management. To protect data privacy the sensitive data should be encrypted by the data owner before outsourcing which obsoletes data utilization like keyword-based document retrieval. It is essential to develop an efficient and reliable cipher text search techniques, so that data owners can easily access and update cloud data. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which supports both multi-keyword ranked search and dynamic update. We have used the vector space model and the widely-used TF_IDF models are combined in the index construction and query generation. To improve search efficiency, we design tree based index structure which supports insertion and deletion update well without privacy leakage. To encrypt the indexes and query vectors the secure KNN algorithm is used. To calculates relevance score between encrypted index and query vectors our system is efficient. Our scheme achieves optimal search efficiency. Our scheme also reduces communication overhead. Hence the analysis shows security and efficiency of our scheme.*

Key Words: **Cloud computing, Multi-keyword search, data integrity, encrypted cloud data, searchable encryption, dynamic update.**

1. INTRODUCTION

Cloud computing is the use of computing resources that are delivered as a service over a network. The name comes from the common use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with user's data, software and computation. Cloud computing enables access to shared pool configuration system resources and higher-level services. These services typically provide access to advanced software applications and high-end networks of server computers. 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. The cloud service providers (CSPs) that keep the data for users they can access users' sensitive information without authorization. To protect the data confidentiality it is important to encrypt the data before outsourcing. However, this will cause a huge cost in terms of data usability.

1.1 Keyword Search Techniques

Searchable encryption (SE) schemes usually build an encrypted searchable index based on the keyword within the document set, by which its content is hidden to the cloud server. 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.

1.2 Single Keyword Search

A single keyword searchable encryption schemes usually builds an encrypted searchable index such that, it's content is hidden to the server, unless it is given appropriate trapdoors generated via secret key(s). However, it only supports single keyword search. Where anyone with public key can write to the data stored on server, but only authorized users with private key can search. Single keyword search schemes uses encrypted searchable index. These indexes content will be hidden to the server. It is not comfortable enough to express complex information needs is the major drawback of single keyword search.

1.3 Multi-Keyword Search

Multi-keyword Boolean search allows the users to enter multiple query keywords to request suitable documents.

These schemes retrieve search results based on the existence of keywords and cannot provide acceptable result ranking functionality. Multi-keyword search on encrypted cloud data have been investigated in. It provides security and efficient search by using two thread models, cipher-text model and background model. A secure k-Nearest Neighbor (k-NN) algorithm was implemented in MRSE scheme. Efficient privacy-preserving search over encrypted cloud data that utilizes min hash functions to improve the precision rate. The advantages of this scheme are multi keyword search in a single query. The effective ranking capability based on term frequency and inverse document frequency.

2. PROPOSED SYSTEM

We construct a tree-based index structure that proposes a "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search result. The proposed scheme achieves sub-linear search time. System also supports the flexible deletion and insertion of documents. Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.

- 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.

- Abundant works have been proposed under different threat models to achieve various search functionality.

- To resist different attacks in different threat models, we construct two secure search schemes: the basic dynamic multi-keyword ranked search scheme in the known cipher text model, and the enhanced dynamic multi-keyword ranked search scheme is known as background model.

3. SYSTEM ARCHITECTURE

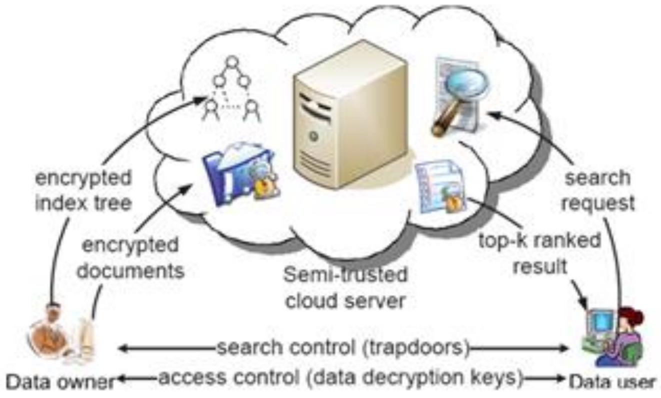

Fig 1. System Architecture of Multi keyword Ranked Search Scheme

3.1 DESCRIPTION OF A SYSTEM

To enable secure, efficient, accurate and dynamic multi data our system has the following

- ***Multi keyword ranked search***: To design a search scheme over encrypted cloud data, which is not only capable of effective multi-keyword search, but also, with the use of vector space model, supports search result similarity ranking.

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

- ***A. Privacy-preserving:*** The proposed scheme is designed to prevent the document collection, index trees and query vectors.

- ***B. Keyword Privacy:*** The proposed system ensures the privacy of keyword. The cloud server could not identify the specific keyword in query, 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. IMPLEMENTATION

This project uses following modules:

Data Owner: This module helps the owner to register their details and this also includes login details. This module helps the owner to upload his file with encryption using RSA algorithm. This ensures the files to be protected from unauthorized user. Data owner has a collection of documents *F ={f*1; *f*2; *:::; fn}* that he wants to outsource to the cloud server in encrypted form. The data owner the securely distributes the key information of trapdoor generation and document decryption to the authorized data users. Data owners would define the access policy and compute the authorization cipher text for each document.

Data User: 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: This module is used to help the server to encrypt the document using RSA Algorithm and to convert the encrypted document to the Zip file with activation code and then activation code send to the user for download the files. Cloud server stores the encrypted document collection *C* and encrypted searchable tree index. After receiving the trapdoor TD from the data user, the cloud server executes search over the index tree I, returns the corresponding collection of top-k ranked encrypted documents. The server needs to update the index I and document collection C according to the received information.

5. CONCLUSIONS

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. binary tree as the index that proposes a "Greedy Depth-first Search" algorithm to obtain better efficiency than linear search. For efficiency aspect, we propose a tree based index structure. The tree based searchable designed in our scheme can support dynamic update well, only accessing portion of the index tree. The security of the system is ensured with two threat models by using the secure KNN algorithm. Extensive experiments show that the proposal can achieve better efficiency in terms of functionalities and computational overhead compared with the existing ones.

6. ACKNOWLEDGEMENT

We take this opportunity to express our hearty thanks to all those who helped us in the completion of the paper. We express our deep sense of gratitude to our guide Prof. U. R. Patole, Asst. Prof., Computer Engineering Department, Sir Visvesvaraya Institute of Technology, Chincholi for his guidance and continuous motivation. We gratefully acknowledge the help provided by him on many occasions, for improvement of this project report with great interest. We would be failing in our duties, if we do not express our deep sense of gratitude to Prof. K. N. Shedge, Head, Computer Engineering Department for permitting us to avail the facility and constant encouragement. Lastly we would like to thank all the staff members, colleagues, and all our friends for their help and support from time to time.

REFERENCES

[1] D. Liu, H. Li, Y. Yang, and H. Yang, "Achieving multi-authority access control with efficient attribute revocation in smart grid," in Proceedings of ICC, 2014, pp. 634–639.

[2] Q. Zheng, S. Xu, and G. Ateniese, "Vabks: Verifiable attribute-based keyword search over outsourced encrypted data," in Proceedings of INFOCOM. IEEE, 2014

[3] C. Wang, N. Cao, J. Li, K. Ren, and W. J. Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data," in 30th International Conference on Distributed Computing Systems, Genova, Italy, pp. 253-262, 2010.

[4] S. Kamara and K. Lauter, Cryptographic cloud storage, in Financial Cryptography and Data Security. Springer, 2010, pp. 136149.

[5] Wang, Cong, Ning Cao, Kui Ren, and Wenjing Lou. "Enabling secure and efficient ranked keyword search over outsourced cloud data," IEEE Transactions on parallel and distributed systems 23, no. 8 (2012): 1467-1479.

[6] S. Kamara and K. Lauter, "Cryptographic cloud storage," in Financial Cryptography and Data Security. Springer, 2010, pp. 136–149.

[7] K. Ren, C.Wang, Q.Wang et al., "Security challenges for the public cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, 2012.

[8] H. Li, Y. Yang, M. Wen, H. Luo, and R. Lu, "Emrq: An efficient multi-keyword range query scheme in smart grid auction market." KSII Transactions on Internet and Information Systems, vol. 8, no. 11, pp.3937–3954, 2014.

[9] H. Li, D. Liu, Y. Dai, T. H. Luan, and X. S. Shen, "Enabling efficient multi-keyword ranked search over encrypted cloud data through blind storage," IEEE Transactions on Emerging Topics in Computing, 2014,DOI:10.1109/TETC.2014.2371239.

[10] Handa, R. and Challa, R.K., "A cluster based multi-keyword search on outsourced encrypted cloud data ," International Conference on Computing for Sustainable Global Development (INDIACom), pp. 115-120, 2015.

[11] C, Orencik and E. Savas, "An efficient privacy-preserving multi-keyword search over encrypted cloud data with ranking,", Springer Distributed and Parallel Databases, pp. 119–160, 2014.

[12] Pasupuleti Syam Kumar, Subramanian Ramalingam, and Rajkumar Buyya. "An efficient and secure privacy-preserving approach for outsourced data of resource constrained mobile devices in cloud computing." Journal of Network and Computer Applications 64 (2016): 12-22.

[13] Sun, Wenhai, Bing Wang, Ning Cao, Ming Li, Wenjing Lou, Y. Thomas Hou, and Hui Li. "Verifiable privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking." IEEE Transactions on Parallel and Distributed Systems 25, no. 11 (2014): 3025-3035.