Security Maintaining Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing

Rahul S Badhekar,Devika K Shete, Reshma K Shinde, Prof.S.N.Dhage

Abstract— The initiation of cloud computing, it has become progressively popular for data owners to outsource their data to public cloud servers while permitting data users to retrieve this data. For privacy concerns, secure searches over encrypted cloud data have motivated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing. In this paper, we propose schemes to deal with Security Maintaining Ranked Multi-keyword Search in a Multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files,The propose a novel Additive Order and Privacy Preserving Function family. To protect the attackers from snooping secret keys and imagining to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. Also, PRMSM supports efficient data user revocation. Extensive experiments on real-world datasets confirm the efficacy and efficacy of PRMSM.

Index Terms— PRSM,Cloud computing,Ranked keyword search, multiple owners, Privacy preserving, dynamic secret key, Cryptography, Authentication.

I. INTRODUCTION

Cloud computing is the extensive visualized vision of computing as a service, where cloud customers can slightly store their data into the cloud so as to enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources. Its great flexibility and economic savings are motivating both individuals and enterprises to outsource their local complex data management system into the cloud. To protect data privacy and combat unsolicited accesses in the cloud and beyond, sensitive data, e.g., emails, personal health records, photo albums, tax documents, financial transactions, etc., may have to be encrypted by data owners before outsourcing to the commercial public cloud this, however, obsoletes the traditional data utilization service based on plaintext keyword search. The trivial solution of downloading all the data and decrypting locally is clearly impractical, due to the huge amount of bandwidth cost in cloud scale systems. Moreover, aside from eliminating the local storage management,storing data into the cloud serves no.purpose unless they can be easily searched and utilized. Thus, exploring privacy-preserving and effective search service over encrypted cloud data is of paramount importance. Considering the potentially large number of on-demand data users and huge amount of outsourced data documents in the cloud, this problem is particularly challenging as it is extremely difficult to meet also the requirements of performance, system usability and scalability. On the one hand, to meet the effective data retrieval need, the large amount of documents demand the cloud server to perform result relevance ranking, instead of returning undifferentiated results. Such ranked search system enables data users to find the most relevant information quickly, rather than burdensomely sorting through every match in the content collection. Ranked search can also elegantly eliminate unnecessary network traffic by sending back only the most relevant data, which is highly desirable in the "pay-as-youuse" cloud paradigm. For privacy protection, such ranking operation, however, should not leak any keyword related information. On the other hand, to improve the search result.

Accuracy as well as to enhance the user searching experience, it is also necessary for such ranking system to support multiple keywords search, as single keyword search often yields far too coarse results. As a common practice indicated by today's web search engines (e.g., Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most relevant data. And each keyword in the search request is able to help narrow down the search result further. "Coordinate matching", i.e., as many matches as possible, is an efficient similarity measure among such multi-keyword semantics to refine the result relevance, and has been widely used in the plaintext information retrieval (IR) community. However, how to apply it in the encrypted cloud data search system remains a very challenging task because of inherent security and privacy obstacles, including various strict requirements like the data privacy, the index privacy, the keyword privacy, and many others. In the literature, searchable encryption is a helpful technique that treats encrypted data as documents and allows a user to securely search through a single keyword and retrieve documents of interest. However, 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 accommodate such high service-level requirements like system usability, user searching experience, and easy information discovery. Although some recent designs have been proposed to support Boolean keyword search as an attempt to enrich the search flexibility, they are still not adequate to provide users with acceptable result ranking functionality. Our early work has been aware of this problem, and provided a solution to the secure ranked search over encrypted data problem but only for queries consisting of a single keyword. How to design an efficient encrypted data search mechanism that supports multi-keyword semantics without privacy breaches still remains a challenging open problem.

II LITRATURE SURVEY

1 Privacy Preserving Ranked Multi Keyword Search for Multiple Data Owners in Cloud Computing.AUTHORS: Wei Zhang,
Explore the problem of secure multi keyword search in multi keyword search. PRMSM model in this system searches a keywords without knowing actual data of trapdoors as well as keywords. This system preserves the keywords and files systematically. In this system sum of the relevance scores is used to search result in metric. Authors defined the problem of secure search over encrypted data. Additive Order and Privacy Preserving Function family (AOPPF) is proposed to preserve the privacy of relevant scores of different functions. This system works on Ranked Multi keyword Search over Multi owner, Data owner scalability, Data user revocation and Security Goals of system
2. A view of cloud computing," Communication of the ACM, AUTHORS: M. Armbrust,
The simple figure to evaluate the comparison between cloud computing and conventional computing. It also identifies functional and non functional opportunities of cloud storage.
3. Privacy preserving public auditing for secure cloud Storage AUTHORS: C. Wang, S. S. Chow,
Data security in cloud this paper proposed a Privacy preserving public auditing system. This system handles multiple audit session different users for their outsourced data files. The privacy preserving public auditing sc heme required to design auditing protocol to prevent data from flowing away. Therefore it is not completely solve the problem of privacy preserving in key management. Therefore unauthorized data leaked problem cannot be solved by this system. TPA audit outsourced data when it is required. Authors were utilizes homomorphic linear authenticator and random masking to provide assurance that TPA cannot learn about knowledge of data.
4 Practical techniques for searches on encrypted data AUTHORSD.Song, D.Wagner, and A.Perrig
describes cryptographic schemes for the problem of searching on encrypted data. It also provides proofs of security for the resulting crypto systems. This scheme is provably secure for remote searching on encrypted data using an untrusted server. This system searches data remotely from untrusted server. This system provides the proofs of security that required for crypto systems. This system worked efficiently for query isolation as they are simple and fast. Only O(n) stream cipher required for encryption and search algorithm.

III. PROPOSED SYSTEM

We define a multi-owner model for privacy preserving keyword search over encrypted cloud data. We propose an efficient data user authentication protocol, which not only prevents attackers from eavesdropping secret keys and pretending to be illegal data users performing searches, but also enables data user authentication and revocation. We systematically construct a novel secure search protocol, which not only enables the cloud server to perform secure ranked keyword search without knowing the actual data of both keywords and trapdoors, but also allows data owners to encrypt keywords with self-chosen keys and allows authenticated data users to query without knowing these keys.

We propose an Additive Order and Privacy Preserving Function family (AOPPF) which allows data owners to protect the privacy of relevance scores using different functions according to their preference, while still permitting the cloud server to rank the data files accurately. We conduct extensive experiments on real-world datasets to confirm the efficacy and efficiency of our proposed schemes. system of search as show in below Fig.1

System Implementation consist of various parts described as follows: We are implementing our project by using Java Technology and MySQL database.
Various components of our system are:
1. Data Owner.
2. Data user.
3. Application server.
4. Cloud server.

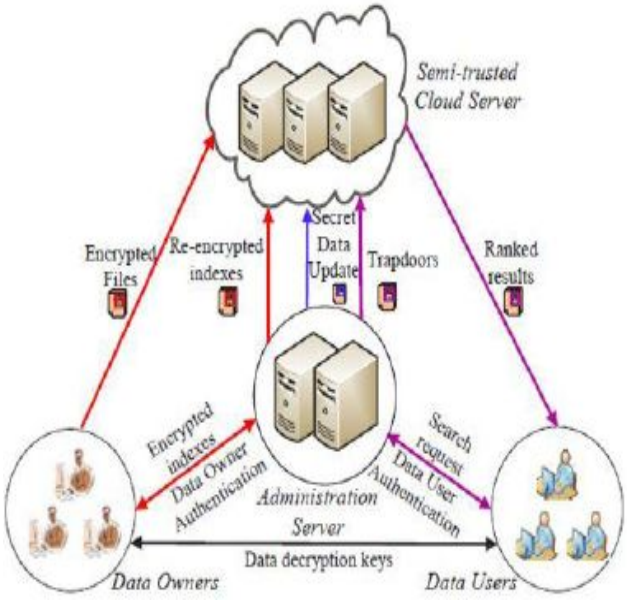

Fig.1.Proposed system architecture.

1. Data Owner: Data owner have the set of files, they create the index file ad send that file to the application server. Finally Data owner encrypt that file and send encrypted file to the cloud server as well as send the encryption key to the data user.
2. Application server: Application server re-encrypt the index file of authenticated user and send that re-encrypted file to the cloud server.
3. Data user: Data user send keywords to search to words the application server, application server send that request to the cloud server if the data user is the authenticated user by creating the trapdoor.
4. Cloud server: Upon receiving the trapdoor, the cloud server searches the encrypted index of each data owner and returns the corresponding set of encrypted files.

IV. ALGORITHM

KeyWord Extraction Algorithm

- Filter word:

- Read textual content and remove unwanted words.

- Tokenize:
  -Chunks are Generated.
  -Generate different output stream.

- Perform Stemming:
  -reducing words to its root.
  ex. Search , searching or searched are consider as a search .
- Calculate Similarity:
  - set of keywords that truly represent original text.
  - treat as unique words.

VI. ADVANTAGES OF PROPOSED SYSTEM

- The proposed scheme allows multi-keyword search over encrypted files which would be encrypted with different keys for different data owners.
- The proposed scheme allows new data owners to enter this system without affecting other data owners or data users, i.e. the scheme supports data owner scalability in a plug-and-play model.
- The proposed scheme ensures that only authenticated data users can perform correct searches. Moreover, once a data user is revoked, he can no longer perform correct searches over the encrypted cloud data.
- To enable cloud servers to perform secure search without knowing the actual value of both keywords and trapdoors, we systematically construct a novel secure search protocol. As a result, different data owners use different keys to encrypt their files and keywords. Authenticated data users can issue a query without knowing secret keys of these different data owners.
- To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a new additive order and privacy preserving function family, which helps the cloud server return the most relevant search results to data users without revealing any sensitive information.
- To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol.

CONCLUSION

In this paper, for the main time we define and solve the problem of multi-keyword ranked search over encrypted cloud data, and launch a variety of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching", Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset show our proposed schemes introduce low overhead on both computation and communication. As our future work, we will explore supporting other multi keyword semantics (e.g., weighted query) over encrypted data, integrity check of rank order in search result and privacy guarantees in the stronger threat model.

REFERENCES

1. Wei Zhang, Student Member, IEEE, Yaping Lin, Member, IEEE, Sheng Xiao, Member, IEEE,Jie Wu, Fellow, IEEE, and Siwang Zhou," Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing".
2. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Communication of the ACM, vol. 53, no. 4, pp. 50– 58, 2010.
3. C. Wang, S. S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacypreserving public auditing for secure cloud storage," Computers, IEEE Transactions on, vol. 62, no. 2, pp. 362–375, 2013. [4] D.Song, D.Wagner, and A.Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE International Symposium on Security and Privacy (S&P'00), Nagoya, Japan, Jan. 2000, pp. 44–55.
4. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. ACM CCS'06, VA, USA, Oct. 2006, pp. 79–88.
5. P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. Applied Cryptography and Network Security (ACNS'04), Yellow Mountain, China, Jun. 2004, pp. 31–45.
6. C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE Distributed Computing Systems (ICDCS'10), Genoa, Italy, Jun. 2010, pp. 253–262.
7. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacypreserving multi-keyword ranked search over encrypted cloud data," in Proc. IEEE INFOCOM'11, Shanghai, China, Apr. 2011, pp. 829–837
8. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in Proc. IEEE INFOCOM'10, San Diego, CA, Mar. 2010, pp. 1–5.
9. P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack," Computers, IEEE Transactions on, vol. 62, no. 11, pp. 2266– 2277, 2013