CONFIDENTIALITY CONSERVING MULTI-KEYWORD RANKED SEARCH ABOVE ENCRYPTED CLOUD DATA

Pothuri Keerthi[1], Dr.P.Varaprasada Rao[2]
Department of Computer Science and Engineering.
Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India.

Abstract

Public cloud is now a rapid growing trend for storing user's data. Most of the users now a day's are storing their personal and professional data on the public cloud. As they outsource their confidential data on cloud, they lose physical possession of their private data. Public cloud is one way to store user's data but it is not trusted storage service. In this paper, two of the privacy preserving issues about accessing the cloud data has been identified i.e., acuteness of keywords sent in queries and the data fetched as a result of those queries. Both of these should be hidden from intruders. To keep the documents private, they should get encrypted before outsourcing to the cloud. Privacy of the data has been achieved using symmetric key cryptography algorithm, we have used Two fish. To fetch the documents of user interest, user should fire a query consisting of multiple keywords which will in turn return the top k ranked documents. As we don't want to disclose neither keyword from query nor query pattern, we have developed fully privacy preserving system by encrypting search pattern as well as secret key. Indexing has been developed to build an index of keywords from documents. Index will be used to retrieve documents in response to search query by using the principle of keyword matching. This paper has analyzed and implemented Lucene indexing algorithm. Ranking of the results has been developed so as to improve the search result correctness as well as to embellish the user searching experience.

Index Terms: Cloud computing, ranked keyword search, several owners, confidentiality preserving, dynamic hidden key

I. INTRODUCTION

Cloud storage system is a set of storage servers, and provides long-term storage services over the Internet. Storing data in a third party's cloud system causes grave to connect to over data secret. Normal hidden schemes defend data secret but have some limitation to functionality of the storage system because a few operations are supported over hidden information.

Building a grave storage system that is compatible with several functions is endurance when system is distributed. Service providers of cloud would pledge to owners data security using phenomenon like virtualization and firewalls. These phenomenon do not protect owner's data confidentiality from the CSP(Cloud service providers) itself, since the CSP control whole of cloud hardware, software, and owners data. Hiding the sensitive data before sent outside can store data confidentiality against CSP. Data hidden makes the conventional data utilization service based on plaintext keyword search a very challenging problem. A solution to this problem is to download all the hidden data and create the original data using the hidden key, but this is not practical because it creates extra overhead. In this paper, we suggest when search multiple owner, multiple keywords provides the confidentiality and show the result in ranking form to make easy for cloud servers to perform safe search excluding knowing the real value of

both keywords and trapdoors, we properly build a novel safe search rule.

So that distinct keys can be used by various data owners to hide their files and keywords. Genuine data users can get a query excluding knowing confidential keys of these various data owners. To rank the search results and preserve the confidentiality of relevance scores between keywords and files, we suggest a family which preserves confidentiality,that helps the cloud server return the most relevant search results to data users without revealing any sensitive information. To protect from disclosing the result a novel dynamic secret key generation protocol and a new data user authentication rule is proposed.

The main contributions of this paper are listed as follows:

• We define search data on clued that data is hidden format and also providing the confidentiality when searching the multiple keywords.

• We suggest a data user authentication rule, which stops attackers to disclose hidden key and only genuine data user can do search.

• We suggest an approach that performs multiple keyword search and rank them properly.

• We suggest an Additive Order and Confidentiality Preserving Function family (AOPPF) which allows the cloud server produce the file that rank properly.

• We supervise experiments on real-world Datasets to verify the effectiveness and capability of our suggested schemes.

II. RELATED WORK

We have again the issue of easy to search symmetric encryption, which gives permission to a client to store its data on an external server in such a way that it can search without disclosing the data. We generate more affords to add new security and new work. Motivated by subtle problems in all previous security definition for SSE, we propose new definitions and point out that the existing notions have significant practical disadvantages contrary to the natural use of easy to find encryption.

Disadvantages: They only give the assurance to security for users that fulfill all their searches at once. We notice this limitation by introducing stronger definition that guarantee security even when users perform more realistic searches. Analysis give guidance to the choice of the size of cipher text space, and at the end suggest a unique and efficient transformation that can be applied to any OPE scheme. Our deep study shows that the transformation yields a scheme with more result safety, in that the scheme oppose the one-wayness and window one-wayness attacks. We opened the new way on how to get this notion, but more efficient variant is certainly required. Second, how to construct SCF-PEKS scheme secure against keyword guessing attacks without requiring bilinear pairing operations would be very interesting.

III. SYSTEM ARCHITECTURE

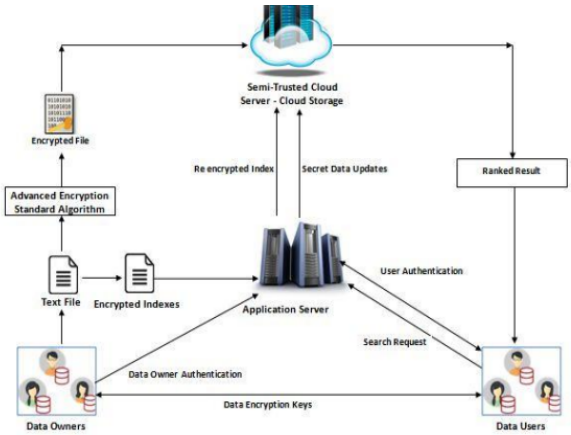

Fig1: System Architecture

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: Data Owner, Application Server, Data user, Cloud Server. Data Owners will upload the data into cloud servers and now application server encrypts the file so that they are more secured. At the same time it assigns key indexing to the file and maintain all details. It will also do the authentication of data owners and users. Application server update the keys and maintain the rank for files. On request of user, server will send from the cloud server to the end user.

First end user needs to authenticate from system server and then he will send the request to server, based on his request server will get high ranked results.

IV. MODULES OR IMPLEMENTATION

a. Data Owner:

Data owner have the set of files, they create the index file and send that file to the application server. Finally Data owner encrypts file and send encrypted file to the cloud server as well as sends the encryption key to thedata user.

b. Application server:

Application server re-encrypt the index file of authenticated user and send that re-encrypted file to the cloud server.

c. Data user:

Data user send keywords to search the words in the application server, application server send that request to the cloud server if the data user is authenticated by creating the trapdoor.

d. Cloud server:

The cloud server searches the encrypted index of each data owner and returns the related set of encrypted files after receiving the trapdoors.

V. APPLICATIONS

- Military Applications
- Health Applications
- Environmental Applications
- Home Applications
- Commercial Applications
- Area Monitoring
- Healthcare Monitoring
- Web Applications

VI. DESIGN GOALS

To enable privacy preserving ranked multi-keyword search in multi-owner and multi-user cloud environment, our system design should simultaneously satisfy security and performance goals.

a. Ranked Multi-keyword Search over Multi-owner:

The proposed scheme should allow multi-keyword search over encrypted files which would be encrypted with different keys for different data owners. It allows the cloud server to rank the search results among different data owners and return the top-k results.

b. Data owner scalability:

The proposed scheme should allow new data owners to enter this system without affecting other data owners or data users, i.e., the scheme should support data owner scalability in a plug-and-play model.

c. Data User revocation:

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.

d. Security Goals:

The proposed scheme should achieve the following security goals: 1) Keyword Semantic Security (Definition 1). We will prove that PRMSM achieves semantic security against the chosen keyword attack. 2) Keyword secrecy.

VII. PERFORMANCE EVALUATION

we measure the efficiency of PRMSM, and compare it with its previous version, Secure Ranked Multi-keyword Search for Multiple data owners in cloud computing (SRMSM), and the state-of-the-art, privacy-preserving Multi-keyword Ranked Search over Encrypted cloud data (MRSE), side by side. Since MRSE is only suitable for the single owner model, our PRMSM and SRMSM not only work well in multi-owner settings, but also works well in MRSE on many aspects.

a. *Evaluation Settings:*

We conduct performance experiments on a real data set, the Internet Request For Comments dataset (RFC). We use Hermetic Word Frequency Counter to extract keywords from each RFC file. After the keyword extraction, we compute keyword statistics such as the keyword frequency in each file, the length of each file, the number of files containing a specific keyword, etc. We further calculate the relevance score of a keyword to a file based on these statistics.

b. Evaluation Results:

i. Index Construction

ii. Trapdoor Generation

iii. Re-encryption by the administrations server

iv. search

VIII. CONCLUSION AND FUTURE WORK

In this paper,secure multi-keyword search for multiple data owners and multiple data users in the cloud computing environment is performed. Different from prior works, our schemes enable authenticated data users to achieve secure, convenient, and efficient searches over multiple data owner's 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 owner's data encrypted with different secret keys, we systematically construct a novel secure search protocol. To rank the search results and preserve the confidentiality of relevance scores between keywords and files, we propose a novel Additive Order and Confidentiality Preserving Function family. Moreover, we show that our approach is computationally efficient, even for large data and keyword sets.

REFERENCES

1.R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definition and efficient constructions," in Proc. ACM CCS'06, VA, USA, pp. 79–88, Oct. 2006.

2. R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in Proc. ACM SIGMOD'04, Paris, France, pp. 563–574, Jun. 2004.

3. D. B. et al., "Public key encryption with keyword search secure against keyword guessing attacks without random oracle," EUROCRYPT, vol. 43, pp. 506–522, 2004.

4 M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange,J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, "Huntableencodeion revisited: Consistency properties, relation to anonymous ibe, and extensions," J. Cryptol., vol. 21, no. 3, pp. 350–391, 2008.

5 J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword hunt over encodeed data in cloud computing," in Proc. of IEEE INFOCOM'10 Mini-Conference, San Diego, CA, USA, March 2010.

6 D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. S. III, "Public key encodeion that allows pir queries," in Proc. of CRYPTO, 2007.

7 P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword hunt over encodeed data," in Proc. of ACNS, 2004, pp. 31–45.

8 L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword huntes over encodeed data," in Proc. of ICICS, 2005.

9 R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserv-ing encryption for numeric data," in Proc. ACM SIGMOD'04,Paris, France, Jun. 2004, pp. 563–574.

10 A. Boldyreva, N. Chenette, Y. Lee, and A. O, "Order-preserving encryption revisited: Improved security analysisand alternative solutions," in Proc. Advances in Cryptology(CRYPTO'11), California USA, Aug. 2011, pp. 578–595.

11 Y. Yi, R. Li, F. Chen, A. X. Liu, and Y. Lin, "A digitalwatermarking approach to secure and precise range queryprocessing in sensor networks," in Proc. IEEE INFOCOM'13,Turin, Italy, Apr. 2013, pp. 1950–1958.

12 R. A. Popa, F. H. Li, and N. Zeldovich, "An ideal-security pro-tocol for order-preserving encoding," in Security and Privacy(SP), 2013 IEEE Symposium on. IEEE, 2013, pp. 463–477.

13 F. Kerschbaum and A. Schroepfer, "Optimal average-complexity ideal-security order-preserving encryption," inProceedings of the 2014 ACM SIGSAC Conference on Computerand Communications Security. ACM, 2014, pp. 275–286.

14 P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme underkeyword guessing attack," Computers, IEEE Transactions on,vol. 62, no. 11, pp. 2266–2277, 2013.

15 B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preservingmulti-keyword fuzzy search over encrypted data in the cloud," in IEEE INFOCOM, Toronto, Canada, May 2014, pp.2112–2120.