International Journal of Recent Trends in Engineering & Research
ISSN : 2455-1457 Impact Factor : 3.344

SECURE OUTSOURCING OF DATA OVER CLOUD WITH RANKED SEARCH

R. Leelarani
Assistant Professor,Department of Computer Applications,
Alpha Arts and Science College, Porur, Chennai – 116

Abstract: The advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in cloud, it is crucial for the search service to allow multi-keyword query and provide result similarity ranking to meet the effective data retrieval need. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely differentiate the search results. In this paper, I define and solve the challenging problem of Privacy-Preserving Multi-Keyword Ranked Search over Encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. It uses the multikeyword semantics known as "coordinate matching", i.e., as many matches as possible, to capture the similarity between search query and data documents. I propose a basic MRSE scheme using ABS(Attribute Based Encryption System) for encryption, and then significantly improve it to meet different privacy requirements using two algorithms OABS1 (Optimized Attribute Based Encryption System1), OABS2 (Optimized Attribute Based Encryption System2).

I. INTRODUCTION

Cloud Computing is a technology that uses the internet and central remote servers to maintain data and applications. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more efficient computing by centralizing data storage, processing and bandwidth. The cloud computing model is comprised of a front end and a back end. These two elements are connected through a network, in most cases the Internet. The front end is the vehicle by which the user interacts with the system; the back end is the cloud itself. The front end is composed of a client computer, or the computer network of an enterprise, and the applications used to access the cloud. The back end provides the applications, computers, servers, and data storage that creates the cloud of services.

1.1 Objective:

The system is expected to give the following security and performance guarantees as follows.

- Multi-keyword Ranked Search: To design search schemes which allow multi-keyword query and provide result similarity ranking for effective data retrieval, instead of returning undifferentiated results.
- Privacy-Preserving: To prevent the cloud server from learning additional information from the dataset and the index, and to meet the basic privacy requirements.
- Efficiency: Ranked search should ensure privacy and also low communication and computation overhead .

II. LITERATURE SURVEY

2.1. A privacy-protecting file system on public cloud storage," Cloud and Service Computing.
Author: Zhonghua Sheng; Zhiqiang Ma; Lin Gu; Ang Li.
Year: 2011

With the development of cloud-based systems and applications, a number of major technical firms have started to provide public cloud storage services, and store user data in datacenters strategically positioned across the Internet. However, when users store private data in shared datacenters, they lose control over how the data are stored and accessed. Multiple classes of personnel may access the physical storage media and potentially read the data. While strong cryptographic methods can protect user files from unauthorized accesses, they incur computational overhead, and make it difficult for the infrastructure provider to optimize the storage space with effective compression and deduplication. To provide strong protection on user data, we design a new file system called BIFS (Bit-Interleaving File System). Focusing on the privacy protection of the on-disk state, BIFS re-orders data in user files at the bit level, and stores bit slices at distributed locations in the storage system. While providing strong privacy protection, BIFS still retains part of the regularity in user data, and thus enables the infrastructure provider to perform a certain level of space optimization (e.g., compression). We implement BIFS on the Amazon Simple Storage Service (S3), and examine its performance characteristics. The comparison with several existing network or Internet-based file systems shows that BIFS provides robust file system functions with satisfactory throughput on S3.

2.2. Secure Rank-Ordered Search of Multi-keyword Trapdoor over Encrypted Cloud Data.
Author: Ibrahim, A.; Hai Jin; Yassin, A.A.; Deqing Zou,
Year: 2012

Advances in cloud computing and Internet technologies have pushed more and more data owners to outsource their data to remote cloud servers to enjoy with huge data management services in an efficient cost. However, despite its technical advances, cloud computing introduces many new security challenges that need to be addressed well. This is because, data owners, under such new setting, loss the control over their sensitive data. To keep the confidentiality of their sensitive data, data owners usually outsource the encrypted format of their data to the untrusted cloud servers. Several approaches have been provided to enable searching the encrypted data. However, the majority of these approaches are limited to handle either a single keyword search or a Boolean search but not a multikeyword ranked search, a more efficient model to retrieve the top documents corresponding to the provided keywords. In this paper, we propose a secure multi-keyword ranked search scheme over the encrypted cloud data. Such scheme allows an authorized user to retrieve the most relevant documents in a descending order, while preserving the privacy of his search request and the contents of documents he retrieved. To do so, data owner builds his searchable index, and associates with each term document with a relevance score, which facilitates document ranking. The proposed scheme uses two distinct cloud servers, one for storing the secure index, while the other is used to store the encrypted document collection. Such new setting prevents leaking the search result, i.e. the document identifiers, to the adversary cloud servers. We have conducted several empirical analyses on a real dataset to demonstrate the performance of our proposed scheme.

2.3. Public Key Encryption with Ranked Multi-keyword Search.
Author: Chengyu Hu; Pengtao Liu
Year: 2013

It is necessary for the cloud server to allow multi-keyword query and provide result ranking to effectively retrieve data from the large number of encrypted data stored in the cloud storage. A related work that support privacy preserving ranked multi-keyword search over encrypted cloud data (MRSE) is not in public-key model. In this paper, the definition of an extension of PEKS called

Public-Key Encryption with Ranked Multi-Keyword Search(PERMKS) is proposed, which means that the receiver could query any subset of keywords and the number of queried keywords appearing in the data can be provided which evaluate the similarity ranking of the data to the search query. Then, a construction of public-key encryption with ranked multikeyword search scheme with sub-linear cipher texts based on anonymous hierarchical identity-based encryption (AHIBE) is put forward and its security is analyzed.

2.4. Fuzzy Keyword Search over Encrypted Data in Cloud Computing.
Author: Jin Li; Qian Wang; Cong Wang; Ning Cao; Kui Ren; Wenjing Lou.
Year: 2010

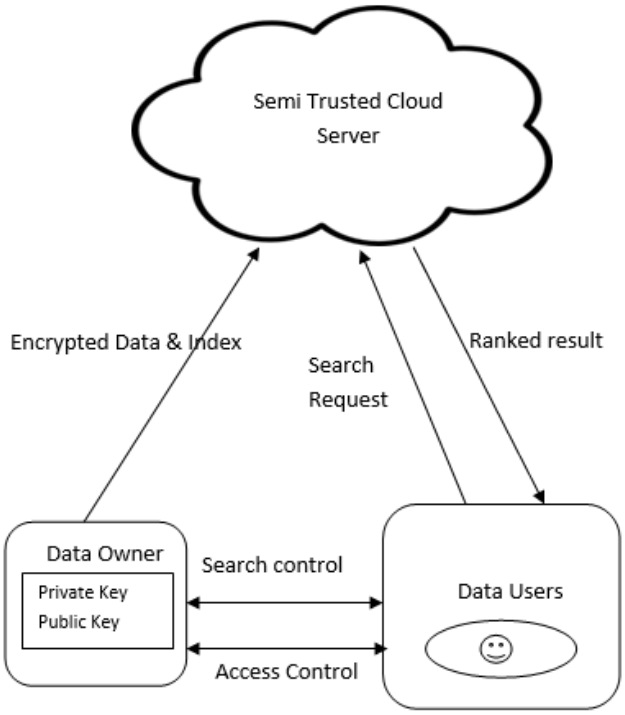
As Cloud Computing becomes prevalent, more and more sensitive information are being centralized into the cloud. For the protection of data privacy, sensitive data usually have to be encrypted before outsourcing, which makes effective data utilization a very challenging task. Although traditional searchable encryption schemes allow a user to securely search over encrypted data through keywords and selectively retrieve files of interest, these techniques support only exact keyword search. That is, there is no tolerance of minor typos and format inconsistencies which, on the other hand, are typical user searching behavior and happen very frequently. This significant drawback makes existing techniques unsuitable in Cloud Computing as it greatly affects system usability, rendering user searching experiences very frustrating and system efficacy very low. In this paper, for the first time we formalize and solve the problem of effective fuzzy keyword search over encrypted cloud data while maintaining keyword privacy. Fuzzy keyword search greatly enhances system usability by returning the matching files when users' searching inputs exactly match the predefined keywords or the closest possible matching files based on keyword similarity semantics, when exact match fails. In our solution, we exploit edit distance to quantify keywords similarity and develop an advanced technique on constructing fuzzy keyword sets, which greatly reduces the storage and representation overheads. Through rigorous security analysis, we show that our proposed solution is secure and privacy-preserving, while correctly realizing the goal of fuzzy keyword search.

III. METHODOLOGIES

The system uses Encryption and Decryption mechanisms. In Encryption the key generation initially was done by using the algorithm ABS(Attribute Based Encryption System) algorithm. It used various parameters of the user like session id, login time and various parameters of files like file size, the type of file, last updated time etc. Then the same key was used for Decryption also. It is Symmetry key system. Now in this system two more algorithms are used. They are OABS1(Optimized Attribute Based Encryption System1) and OABS2(Optimized Attribute Based Encryption System2). They are used for improving the performance by minimizing the Encryption time.

Instead using all the attributes in key generation, OABS1 and OABS2 uses only few attributes. Those attributes are selected by applying logical AND operation, so the randomness also maintained. It will not generate the same key more than once. Thus it is difficult to the intruders to find out the key.

3.1 System Architecture diagram:



3.2 The existing work lacks the following:

- Single-keyword search without ranking
- Single-keyword search with ranking
- Key generation will be a time consuming process.
- Keys can be easily predicted by the intruders.

3.3 Proposed Work:

We define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. Among various multi-keyword semantics, we choose the efficient principle of "coordinate matching". The performance will be optimized by using two algorithms Optimized Attribute Based Encryption System for key generation, so that the time taken for key generation is reduced. At the same time randomness maintained.

Advantages:

I. Multi-keyword ranked search over encrypted cloud data (MRSE).
II. "Coordinate matching" by inner product similarity.
III. Performance will be increased by reducing the time taken for key generation.
IV. It will not generate duplicate keys.

IV. MODULE DESCRIPTION:

4.1 Data user module:

In this module the following processes are carried out by the data user.

Registration Process:

In this process the user register himself by giving the first name, last name, user id, pass word, email id and phone number. He also specifies the subscriber's type. Then he submits this. The given user id is verified if it is already exist, the system will give a warning message. Otherwise an account will be created.

Login Process:

In this process the user gives user id and pass word, the system will verify the subscriber type, so an user session is created.

User Session:

In user session He can access the File Store for viewing the downloaded files, downloading and searching the files in the cloud. He also uses Cryptosystem for Decryption of the cipher text in the cloud into the plain text. He also able to manipulate his profile details, like changing password, viewing his profile and logout.

4.2 Data owner module :

In this module the following processes are carried out by the data owner.

Registration Process:

In this process the data owner register himself by giving the first name, last name, user id, pass word, email id and phone number. He also specifies the subscriber's type. Then he submits this. The given admin id is verified if it is already exist, the system will give a warning message. Otherwise an account will be created.

Login Process:

In this process the admin gives user id and pass word, the system will verify the subscriber type, so an admin session is created.

Admin Session:

In admin session, He uses Cryptosystem for Encryption of the plain text into the cipher text then it will be uploaded into the cloud. He can access the File Management for viewing the downloaded and uploaded file, He can perform operations like upload and download. He can use the Cryptosystem for encrypting and decrypting the files. He can also delete the unwanted files. He also able to manipulate his profile details, like changing password, viewing his profile and logout.

4.3 Cryptosystem module :

This module will perform the major operation of the system Encryption and Decryption.

Encryption :

The encryption is done by using any one of the following algorithms.

- ABS Algorithm
- OABS- I Algorithm
- OABS- II Algorithm

ABS Algorithm:

It comprises the following activities. Reading the file in original text format, collect the attributes such as admin session id, file type, number of files, time of file uploading. Then based on these details it will create a Encryption key. Then it passes the key to encryption system, it will encrypt the plain text into the cipher text. During the encryption process it uses the UTF 8 character encoding.

OABS Algorithm :

It is also similar to the previous one algorithm Except few changes. Reading the file in original text format, collect the attributes such as admin session id, file type, number of files, time of file uploading. It will create an Encryption key by selecting only few attributes. It will not use all the attributes for key generation then it will be a time consuming process. By using AND gate it will randomly select few attributes then by using them it will generate an Encryption key. The similar mechanism it will follow to produce the decryption key also. However the uniqueness is maintained. No duplication will occur.. Then it passes the key to encryption system, it will encrypt the plain text into the cipher.

Decryption :

The cipher text is taken as input by the Decryption system, then it will request the key from the data base. Then it will decrypt the cipher text into plain text, the plain text is stored into a document.

4.4 Rank search module:

In this module the user can perform searching of the files in the cloud by entering multiple keywords. The files are compared with the keyword using the concept of "Coordinate matching". While displaying the results preference is given to the keywords in the order they are entered. It will also rank the outputs by applying Binary Sort. Then it will display the results based on the rank.

V. CONCLUSION & FUTURE ENHANCEMENT:

5.1. Conclusion:

This paper defines and solves the problem of multi-keyword ranked search over encrypted cloud data, and establishes a variety of privacy requirements. Among various key generation algorithms, it uses the efficient Attribute Based Encryption System, optimized Attribute Based Encryption system. It also uses the multi-keyword semantics, "coordinate matching," i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity measure. For meeting the challenge of supporting multi-keyword semantic without privacy breaches, it proposes a basic idea of MRSE using secure inner product computation. It reduces the time taken for key generation in the Encryption process of Cryptosystem module. It uses two optimized ABS algorithms that reduces the time taken for key generation at the same time the encryption will be done in a secure and efficient way.

5.2 Future Enhancement :

In the future work, the integrity of the rank order in the search result can be checked assuming the cloud server is untrusted.

REFERENCES

i. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM, pp. 829-837, Apr, 2011.
ii. L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A Break in the Clouds: Towards a Cloud Definition," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50-55, 2009.
iii. N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.
iv. S. Kamara and K. Lauter, "Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptograpy and Data Security, Jan. 2010.
v. A. Singhal, "Modern Information Retrieval: A Brief Overview," IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35-43, Mar. 2001.
vi. I.H. Witten, A. Moffat, and T.C. Bell, Managing Gigabytes: Compressing and Indexing Documents and Images. Morgan Kaufmann Publishing, May 1999.
vii. D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000.
viii. E.-J. Goh, "Secure Indexes," Cryptology e Print Archive, http://eprint.iacr.org/2003/216. 2003.
ix. Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving KeywordSearches on Remote Encrypted Data," Proc. Third Int'l Conf.Applied Cryptography and Network Security, 2005.
x. R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. 13th ACM Conf. Computer and Comm. Security(CCS '06), 2006.
xi. D. Boneh, G.D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public Key Encryption with Keyword Search," Proc. Int'l Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2004.
xii. M. Bellare, A. Boldyreva, and A. ONeill, "Deterministic and Efficiently Searchable Encryption," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.

xiii. M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, "Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous Ibe, and Extensions," J. Cryptology, vol. 21, no. 3, pp. 350-
xiv. 391, 2008.
xv. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy Keyword Search Over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM, Mar. 2010.
xvi. D. Boneh, E. Kushilevitz, R. Ostrovsky, and W.E.S. III, "Public Key Encryption That Allows PIR Queries," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.

