Advances in Computing 2017, 7(2): 58-61
DOI: 10.5923/j.ac.20170702.10

Multi Keyword Search over Encrypted Cloud

M. Ambika Ganguli*, Mariyammal Safana M., Prajna S. Maddodi, Seemashree, Sridevi Saralaya

Department of Computer Science Engineering, St Joseph Engineering College, Mangaluru, India

Abstract Cloud computing has been considered as a new model of enterprise IT infrastructure, which can organize huge resource of computing, storage and applications. It enables the users to access the network to a shared pool of configurable computing resources with great efficiency and minimal economic overhead. In the era of big data, huge amount of data produced world-wide is stored on the cloud. Despite the various advantages of cloud services, outsourcing sensitive information to remote servers brings privacy concerns. The cloud service providers who store user's data may access sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing. In this study, we address the problem of privacy preserving using multiple keywords search over encrypted cloud data. Our approach provides multi-keyword search based on coordinate matching.

Keywords AES Algorithm, Coordinate matching, Cloud data security

1. Introduction

Cloud computing enables cloud customers to remotely store their data into the cloud so as to enjoy the on-demand high quality applications and services. Cloud storage services allow the user to access files from any computer, as long as it is connected to the Internet. The user gets few Gigabytes of storage without paying anything at all. The biggest advantage of using cloud storage services is that the user does not have to carry important files everywhere. They can simply login to the cloud storage account and access files from anywhere. Files uploaded to the cloud can be easily shared with any number of users. The privacy concerns in cloud computing motivate the study on secure keyword search [1].

The system model of existing studies consider one data owner, which implies that in their solutions, the data owner and data users can easily communicate and exchange secret information. When numerous data owners are involved in the system, secret information exchanging will cause considerable communication overhead. We explore the problem of secure multi-keyword search for multiple data owners and multiple data users in cloud computing environment. In this paper, we address secure multi keyword search over encrypted cloud data [42].

Without encryption, the files can be accessed by unauthorized users. To avoid unauthorized access, we hide the information using encryption. Encryption is the process of changing information in such a way as to make it unreadable by anyone except those possessing the key which allows them to change the information back to its original, readable form. Single-Keyword search is when a user searches and lists for exactly one term and not its variations. Multi-keyword search is when a user searches and lists for multiple variations of same keyword. Among various multi-keyword semantics, we choose the efficient principle of coordinate matching, i.e. to find as many matches as possible, to capture the similarity between search query and data documents.

2. Related Work

Wang *et al.* studied the problem of secure ranked keyword search over encrypted cloud data [3]. The authors work on the problem of searchable index before outsourcing the encrypted document. The authors propose a single keyword searchable encryption scheme based on ranking. The early work of Sangwan *et al.* solves secure ranked keyword search which utilizes keyword frequency to rank results instead of returning undifferentiated results [4]. However, it only supports single keyword search. Song *et al.* were the first to discover the method for keyword search over encrypted and outsourced data [5]. The authors begin with the idea to store a plaintext documents on data storage such as mail servers and file servers in encrypted form to reduce security and privacy risks. Koutrika *et al.* presented a data cloud in which cloud search is performed on the basis of query summarization approach [6]. The authors have performed a query refinement model based on the summarization. Based on this summarization the query is presented to the web architecture and relatively the search is performed for reliable and effective cloud service. A multimedia search for the cloud architecture is suggested by Daniel E. [7]. In this

* Corresponding author:
ambikaganguli27@gmail.com (M. Ambika Ganguli)
Published online at http://journal.sapub.org/ac

work different multimedia services are suggested such as client PC, phone, TV etc. Knowledge based search is performed to retrieve the multimedia analysis and perform search respective to client request for the particular multimedia service.

3. Methodology

The architecture of the proposed system is provided in Fig. 1. The cloud server allows authenticated users to upload or download data to and from the cloud respectively. The system allows only registered users to interact with the cloud server. The users need to register before they can login to the site. Once they have been authenticated by the system, it allows them to upload/download the documents and perform multi-keyword search. In order to authenticate or authorize the users for this action there needs to be a centralized database that holds their accounts.

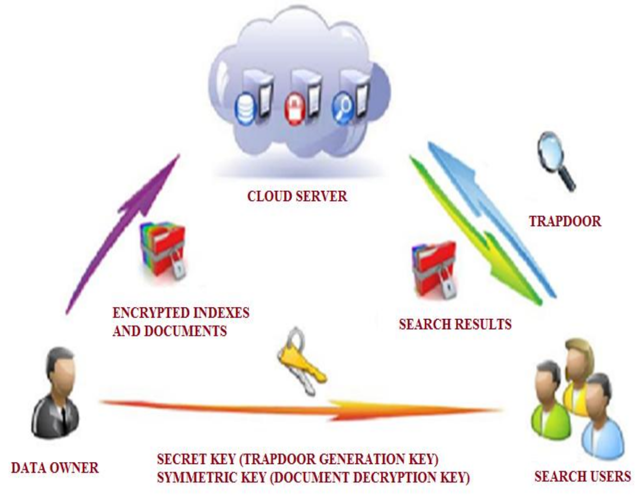
CLOUD SERVER
TRAPDOOR
ENCRYPTED INDEXES AND DOCUMENTS
SEARCH RESULTS
DATA OWNER
SECRET KEY (TRAPDOOR GENERATION KEY)
SYMMETRIC KEY (DOCUMENT DECRYPTION KEY)
SEARCH USERS

Figure 1. Architecture of the system

The proposed work provides data integrity using AES algorithm. Here, the data owner encrypts and uploads the files to the cloud server. The data user can search the files using multiple keywords. By using multi keyword search algorithm, the best result is displayed to the user (Fig. 2). User can request for the file of his interest. The data owner can either accept/reject the request. If the data owner approves the request, a secret key is a randomly generated and is sent to the user's email. The data user/requestor can decrypt and download the file using secret key sent by the data owner. The sequence of interactions is displayed in Fig. 3.

3.1. Encryption

In the proposed system, we have used well known AES algorithm to encrypt/decrypt the data file uploaded by the owner. The Advanced Encryption Standard (AES) is an encryption algorithm for securing sensitive data. AES is a symmetric block cipher. It uses the same key for both encryption and decryption. AES standard states that the algorithm can only accept a block size of 128 bits and a choice of three keys -128, 192, 256 bits. In our application, we have used 128 bits with 10 rounds.

3.2. Uploading and Encryption of the File

The process of uploading a file by the owner consists of generating the key, encryption of a file followed by uploading. The algorithm is provided below:

```
Step 1: Data owner uploads files.
Step 2: If file name is new
            Generate file key for download
            Encrypt file using AES algorithm and Upload the file to the cloud
                        Else if filename already exists
            Overwrite/Stop
```

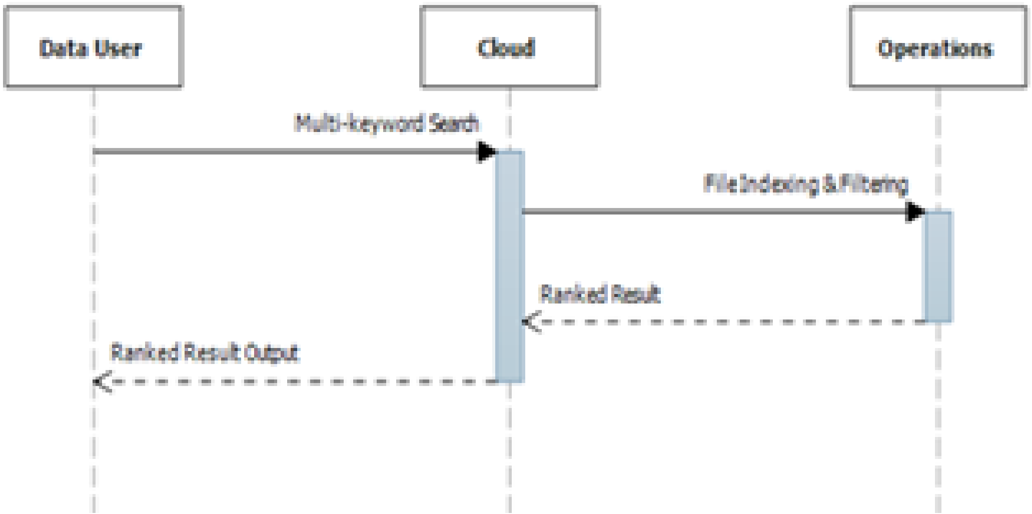

Figure 2. Sequence diagram for multi-keyword search

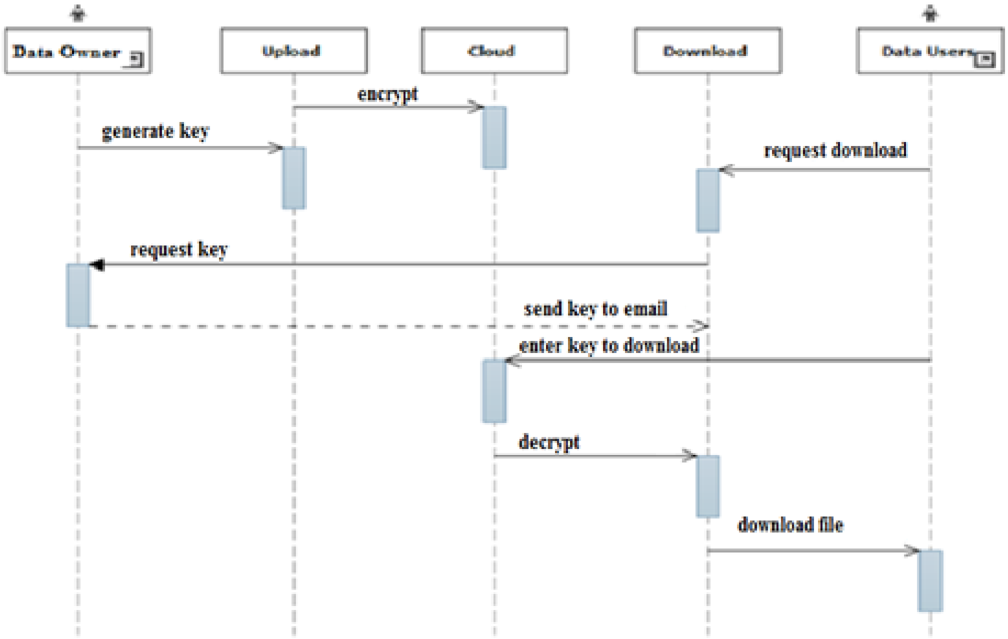

Figure 3. Interaction between Data owner and Data User

3.3. Multi Keyword Search Algorithm

In the proposed system, we have used multi keyword search algorithm for searching multiple keywords in single query. When a data user is interested in reading/viewing any file of interest, he enters related multiple keywords. The keywords are split in array of strings. Every word in the String Array is compared with the first record. Add this file to the Search Output List and displays the output list to the user. The algorithm is given below:

```
Step 1: Input Multi-Keywords to the System
Step 2: Split the words into a String Array
Step 3: For every word in the String Array
            Compare the word with first record (file)
              If Found
                 Add this file to the Search Output List
              Else Continue
Step 4: End For
Step 5: Display the output list to the user
```

3.4. Download/Decrypt File

Once the output list is displayed to the requestor/user, the user has to request the file from the data owner. If the data owner is willing to share the file, he/she sends a key to the requestor. Using this key, the requestor can download and decrypt the file. The algorithm is provided below:

```
Step 1: Data user requests key to download the file
Step 2: If key is valid
               Download and decrypt the file
          Else if the key is invalid
               Error while downloading
Step 3: End If
```

4. Limitations and Future Work

The short comings of our approach are:

- Search is based on the keywords present in the name of the file only. Searching cannot be performed for keywords present in author, subject, etc.
- The retrieval speed of file depends on the speed of the internet.
- The performance time for retrieval of files depends on the key size of the AES algorithm, used for encryption and decryption.

As a part future work, we have to work on ranking criteria which is specified on the basis of implicit and explicit properties such as security, availability etc. We have to also explore supporting other multi keyword semantics over encrypted data.

5. Conclusions

In the proposed system data owners store their data in encrypted form on the cloud. The authenticated data requestors search for the required data using multiple keywords. Based on the search results the requestors request the data owners for permission to download data. The data owners provide the key to download and decrypt the files to authenticated users. The proposed Multi keyword search over encrypted cloud data enables users to achieve secure and efficient searches over multiple data owner’s data.

REFERENCES

[1] Makkar, Vimmi, Sandeep Dalal, MDU DCSA, and MDU Rohtak DCSA. International Journal of Computer Applications & Information Technology Vol. 3, Issue I June-July 2013 (ISSN: 2278-7720) Pp. 1-10.

[2] Örencik, Cengiz, and Erkay Savaş. "Efficient and secure ranked multi-keyword search on encrypted cloud data." In Proceedings of the 2012 Joint EDBT/ICDT Workshops, pp. 186-195. ACM, 2012.

[3] Wang, Cong, Ning Cao, Jin Li, Kui Ren, and Wenjing Lou. "Secure ranked keyword search over encrypted cloud data." In Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference on, pp. 253-262. IEEE, 2010.

[4] Sangwan, Nikita, and Sunita Sangwan. "Ranked Keyword Search in Cloud Computing: An Innovative Approach." (2014). International Journal of Computational Engineering Research,Vol, 03, Issue, 6, no.39-44.

[5] Song, Dawn Xiaoding, David Wagner, and Adrian Perrig. "Practical techniques for searches on encrypted data." In Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on, pp. 44-55. IEEE, 2000.

[6] Koutrika, Georgia, Zahra Mohammadi Zadeh, and Hector Garcia-Molina. "Data clouds: summarizing keyword search results over structured data." In Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology, pp. 391-402. ACM, 2009.

[7] Rose, Daniel E. "Cloud Search and the democratization of information retrieval." In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval, pp. 1022-1023. ACM, 2012.