Madridge Journal of Bioinformatics and Systems Biology

Research Article

Open Access

Privacy Preserving Keyword Search for Multiple Data Users in Cloud

R Naresh[1]*, N Deepa[2] and **P Pandiaraja[2]**

[1]*Associate Professor, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India*
[2]*M. Kumarasamy College of Engineering, Karur, Tamil Nadu, India*

Article Info

***Corresponding author:**
Naresh Ramu
Associate Professor
SRM Institute of Science and Technology
Kattankulathur, Tamilnadu
India
Mobile: 91-80566 62701
E-mail: naresh.r@ktr.srmuniv.ac.in

Received: November 23, 2018
Accepted: November 30, 2018
Published: December 5, 2018

Citation: Naresh R, Deepa N, Pandiaraja P. Privacy Preserving Keyword Search for Multiple Data Users in Cloud. *Madridge J Bioinform Syst Biol.* 2018; 1(1): 1-4.
doi: 10.18689/mjbsb-1000101

Published by Madridge Publishers

Abstract

By the start of cloud computing as well as its benefits the data owners submits their files to cloud service providers which allow the data users to download the searched files. For maintaining the privacy of the files andto have secure search over encrypted files. We have developed the privacy preserving keyword search for multiple data user's incloud. There are many existing papers which supports the same privacy as well as the secure search by using the single owner model. But, utmost cloud servers in real time do not just support one data owner, as a substitute, the cloud server will maintain multiple data owners to enjoy the advantages provided by the cloud. In this research paper, we propose a protocol which will support the short and secure trapdoor to have the secure search over the cloud. By using our protocol even the cloud service providers will not able to find the corresponding search request as well as the data user trapdoor. We have developed the secure trapdoor phase to have the secure search over the cloud. We have developed the cloud matching phase to have the ranked results. We also have developed the decryption phase to decrypt the encrypted files. The performance evaluation and the security evaluation are presented to prove that our protocol is better in terms of computation complexity as well as the communication complexity.

Keywords: Cloud computing; Privacy; Multi data owner; Keywords; Trapdoors; Communication complexity; Computation complexity.

Introduction

Cloud computing is a surprising technology that is varying as a mode that the hardware in addition to software are constructed and acquired [1]. As an innovative prototype of computing, the cloud computing delivers plentiful advantages like easy way of accessing reduced costs, fast deployment in addition to flexible maintenance of resources. All size of companies can outsource the documents to the cloud to raise invention also partnership.

Cloud service providers will give assurance to the data owners by using the files security methods like virtualization as well as firewalls. Though, these methods do not support data owners' privacy, since the cloud itself holds full control over the data owners' files. By encrypting the files before outsourcing can save the data privacy beside cloud. Still, file encryption creates a challenging issue for utilizing the data from the cloud.

Song et al. [2] first define and resolve the difficulty of secure search over encrypted files. They recommend the idea of searchable encryption. Search on an encrypted dataset just as on a plaintext dataset. Searchable encryption is additionally developed by Goh E et al. [3-7].

However, these schemes care mostly on single or Boolean keyword search. Extending these methods for ranked multi-keyword search will suffer from heavy computation complexity and storage complexity. Secure search over encrypted files placed over the cloud is first initiated by Wang et al. [8] and additionally established by Cao N et al [9-14]. Extending the single data owner system to the multi data owner scheme will create more difficulties. In the single data owner system, when the data user requires submitting a keyword search, the data users ask the secret keys from the data owner for generating the trapdoors. If there are numerous data owners then asking the keys from all the data owners will be difficult for the data users.

Initially, it is not necessary for all the data owners to stay in online continuously to deliver the keys whenever the data user submits the query. Moreover if the data owners were in offline means they cannot deliver the keys to the data user when the search request is submitted. Another issue is whenever the data user has to submit the search request towards the encrypted files placed over the cloud by various data owners, definitely the data users has to submit generate the different trapdoors. If the data user submits more trapdoor to the various data owner individually means it would create more communication as well as computation complexity. Another resolution is to share a secret key among whole data owners. When the data user trapdoor and the data owner encrypted keyword matched then the cloud will do the matching process to return the best related search results to data users without leaking any secret information from the files.

The main contributions of this paper are listed as follows:

- ✓ The secure keyword generation for encrypting as well as for decrypting the files.
- ✓ We define a trapdoor for secure keyword search over encrypted cloud data, which is used for preserving the privacy. This trapdoor not only permits the cloud server to perform secure ranked keyword search without knowing the actual data of both search request as well as trapdoors, but also permits data owners to encrypt keywords through their own keys. Finally the authenticated data users can submit the search request without aware about the data owner keys. We implement the experiments on existing world datasets to prove the efficiency of our proposed schemes.
- ✓ The cloud matching phase to match the trapdoor and the data owner index.

The rest of this paper is organized as follows. Section 2 explains the related works. Section 3 describes the notations and our proposed protocol. Section 4 presents security analysis. Section 5 explains the performance analysis. Section 6 describes the conclusions and the future works.

Related Work

In this related work section, we describe three types of work: searchable encryption, secure keyword search in cloud, and cloud matching the encrypted files. Initially the Searchable Encryption is the first effort of searchable encryption was done by Song et al. [2]

D. Song et al. [2] have proposed a method to encrypt every word in a file individually and permit the server to find whether that particular single keyword is present in the file without aware of that exact word. This method suffers from high computational complexity. Goh et al. propose to construct a keyword index for each file using Bloom filter to have a quick search [3]. Curtmola et al. propose to build indices for each keyword using hash tables as amethod for searchable encryption [4]. P. Golle and L. Ballard [6,7] additionally improve the search functionalities of searchable encryption by using conjunctive keyword search. The searchable encryption suits for the single keyword search or boolean keyword search. Extending these methods for multi-keyword search will suffer from heavy computational as well as storage cost.

C. Wang et al [8] have proposed a scheme that downloads top-k related files by using single keyword search. N. Cao and W. sun [9,10] extended the secure keyword search to multi-keyword search. Their styles will convert the words to vectors list and will implement based on the matrix multiplications to hide the real keyword from the cloud server. Xu et al [11] have proposed multi-keyword ranked query over then crypted files which enable a dynamic keyword index to avoid the ranking order being inaccurate by some high frequency words. Li et al. and Chuah et al. [12,13] have proposed fuzzy keyword search over cloud files directing at acceptance of mutually slight typos and arrangement inconsistencies for the data user's search request.

Wang et al [14] likewise proposed privacy guaranteed similarity search method over encrypted cloud data. To increase the search efficiency, they additionally construct a multi way trie-tree for storage of similarity keywords. Altogether the similar keywords in the trie-tree can be efficiently placed in the depth-first search. The system representations of earlier works only reflect one data owner, where the data owner and data users can definitely communicate and inter change secret information. When many data owners are involved in the scheme, secret message communication will create huge communication overhead. This paper varies from earlier works for the prominence of multi data owners in this scheme. This paper excellently relaxes the necessities for data owners and data users, thus our proposed scheme could be appropriate for a large number of cloud users (Figure 1). Furthermore, performance result shows that the proposed scheme is highly efficient and reduced complexity. The initial works of Agrawal et al. [15] have proposed an Order Preserving symmetric Encryption system where the mathematical order of plain texts are saved. Boldyreva et al [16] as a supplementary presented a flexible order preserving encryption. These order preserving encryption systems cannot be openly used in our proposed scheme since we permit data owners to practice dissimilar order preserving mappings to safeguard the privacy of their relevance scores.

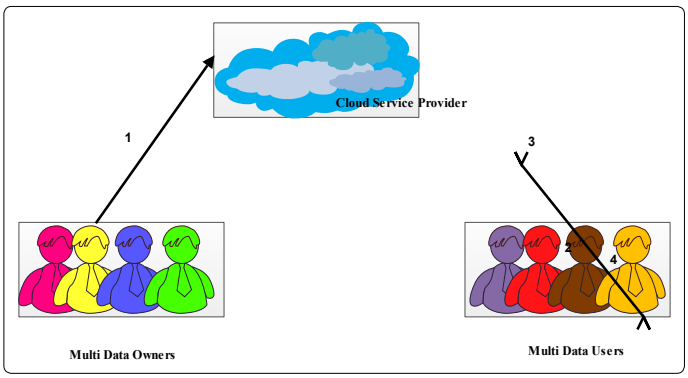

Figure 1. System architecture of privacy preserving keyword search for multiple data users in cloud.

To allow efficient search operation on the encrypted files which is placed over the cloud, data owners initially construct a secure searchable index on the keyword setmined from the files. Then data owners encrypt their files *F* and acquire the equivalent encrypted files *C*. Lastly together the encrypted files as well as indexes are outsourced to the cloud. When an authorized data user needs to search the keywords over the encrypted files stored on cloud servers, the data user initially computes the trapdoor *T* and submits it to the cloud. In our system, every data owner will build their own index. The data user will generate the trapdoor. Authenticated data user can generate his trapdoor based on his required keywords. If the trapdoor and the keyword match then the data user can download the ranked results.

Notations

S.NO	Symbols	Description
1		Master secret key
2		Public key of the data owner
3		Data user identity number
4		Trapdoor generated by the Data user
5		Hash function
6		Search keyword of the data user
7		Prime number *n*
8		Random number selected by the data owner
9		Data owner encrypted keyword
10		Cloud matching the trapdoor and the keyword

Keyword generation (Figure 2)

When the data user register towards the data owner to access the files, after successful registration and authentication the data user will receive the master secret key. Where is the master secret key to decrypt the downloaded files. The term refers to the public key of the data owner. The term refers to identity number of the data user.

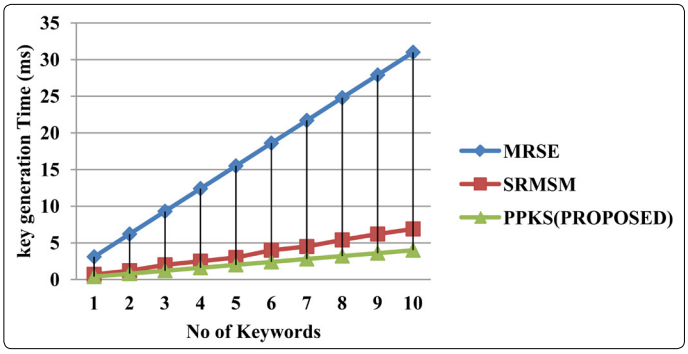

Figure 2: Key generation.

Trapdoor Generation (Figure 3)

The data user will register himself towards the data owner. The search keyword of the data user is . The data user uses the hash function to have the secure search over the encrypted files. Where *mod n* refers to the prime number selected by the data user.

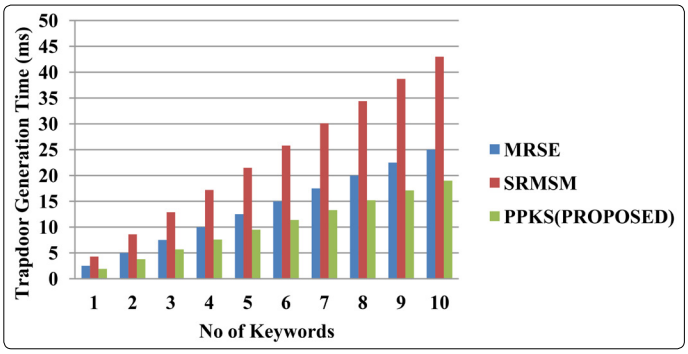

Figure 3: Trapdoor generation.

Data Owners keyword encryption (Figure 4)

The data owner will encrypt the files and place the files over the cloud. For encrypting the files the data owner will choose the random number s and the private key of the data owner where *k* refers to the private key and *i* refer to the n number of the data owners. Finally the data owner will submit the encrypted files over the cloud.

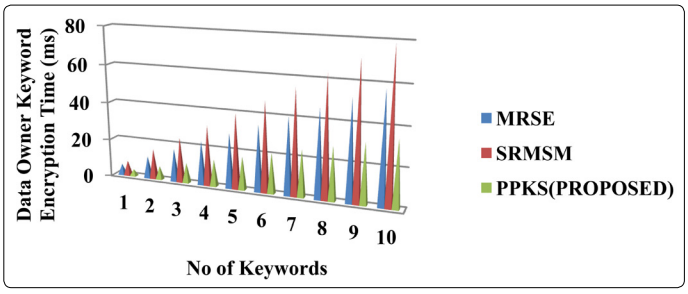

Figure 4: Keyword encryption.

Cloud matching the trapdoor and keyword (Figure 5)

When the cloud service provider receives the trapdoor submitted by the data user and the keyword submitted by the data owner, once if the trapdoor and the keyword matches then the cloud will provide the ranked results to the data user.

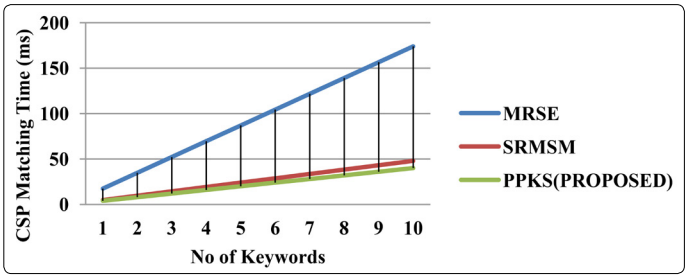

Figure 5: CSP Matching.

Security Evaluation

A. Man in the middle attack

The man in the middle attack is not possible since the keyword is hashed and the search request is submitted towards the cloud service provider. Without knowing the

hashing algorithm the attacker cannot decrypt the file. Thus man in the middle attack is impossible.

B. Impersonation attack

The data user will register towards the data owner. While registering the data user will submit the identity of the data user. The data owner using its public key will generate the master secret key. So if the attacker impersonates like the legal data user then without having the master secret key he cannot decrypt the downloaded files.

Performance Evaluation

The performance analysis section is implemented based on the Python programming language on a PC with 3.2 GHZ Pentium Dual Core CPU and memory: 2GB.

Conclusion

In this paper, we resolve the problem of computation complexity of secure multikeyword search for multiple data owners also multiple data users in the cloud environment. When compared to the earlier works, our scheme enables authenticated data users to achieve secure and efficient search over multiple data owners' files. To enable the cloud to execute secure search among multiple owners' files encrypted with different secret keys. We have developed a secure trapdoor for having secure search over the cloud. We have developed the data owner constructing the encrypted keyword for the files collection. Moreover, we show that our approach is computationally efficient even for large data sets it also suit for large keyword sets.

References

1. Armbrust M, Fox A, Griffith R, et al. A view of cloud computing. *Communication of the ACM*. 2010; 53(4): 50-58. doi: 10.1145/1721654.1721672
2. Song DX, Wagner D, Perrig A. Practical techniques for searches on encrypted data. *IEEE International Symposium on Security and Privacy (S&P)*. 2000; 44-55.
3. Goh E. Secure indexes. 2003.
4. Curtmola R, Garay J, Kamara S, Ostrovsky R. Searchable symmetric encryption: improved definitions and efficient constructions. *ACM, VA*. 2006; 79-88. doi: 10.1145/1180405.1180417
5. Boneh D, Di Crescenzo G, Ostrovsky R, Persiano G. Public key encryption with keyword search. In: LNCS, 3027. Springer-Verlag. 506-522.
6. Golle P, Staddon J, Waters B. Secure conjunctive keyword search over encrypted data. *Applied Cryptography and Network Security (ACNS)*. 2004; 31-45.
7. Ballard L, Kamara S, Monrose F. Achieving efficient conjunctive keyword searches over encrypted data. *Information and Communications Security (ICICS)*. 2005; 414-426. doi: 10.1007/11602897_35
8. Wang C, Cao N, Lou W. Secure ranked keyword search over encrypted cloud data. IEEE Distributed Computing Systems (ICDCS). 2010; 253-262. doi: 10.1109/ICDCS.2010.34
9. Cao N, Wang C, Li M, Ren K, Lou W. Privacy-preserving multi-keyword ranked search over encrypted cloud data. *IEEE INFOCOM*. 2011; 829-837. doi: 10.1109/TPDS.2013.45
10. Sun W, Wang B, Cao N, et al. Privacy preserving multi-keyword text search in the cloud supporting similarity based ranking. *IEEE ASIACCS*. 2013; 71-81.
11. Xu Z, Kang W, Xu C. Efficient multi-keyword ranked query on encrypted data in the cloud. *IEEE Parallel and Distributed Systems (ICPADS)*. 2012; 244-251. doi: 10.1109/ICPADS.2012.42
12. Li J, Wang Q, Wang C, Cao N, Ren K, Lou W. Fuzzy keyword search over encrypted data in cloud computing. *IEEE INFOCOM*. 2010; 1-5.
13. Chuah M, Hu W. Privacy-aware bed tree based solution for fuzzy multi-keyword search over encrypted data. *IEEE 31th International Conference on Distributed Computing Systems (ICDCS)*. 2011; 383-392. doi: 10.1109/ICDCSW.2011.11
14. Wang C, Ren K, Yu S, Raje Urs KM. Achieving usable and privacy-assured similarity search over outsourced cloud data. *IEEE INFOCOM*. 2012; 451-459. doi: 10.1109/INFCOM.2012.6195784
15. Agrawal R, Kiernan J, Srikant R, Xu Y. Order preserving encryption for numeric data. *ACM SIGMOD*. 2004; 563-574.
16. Boldyreva YL, Chenette N, AO Neill. Order-preserving encryption revisited: Improved security analysis and alternative solutions. *Advances in Cryptology (CRYPTO)*. 2011; 578-595.
17. Yi Y, Li R, Chen F, Liu AX, Lin Y. A digital watermarking approach to secure and precise range query processing in sensor networks. *IEEE INFOCOM*. 2013; 1950-1958.
18. Jung TX, Li Y, Wan Z, Wan M. Privacy preserving cloud data access with multi-authorities. *IEEE INFOCOM*. 2013; 2625-2633. doi: 10.1109/INFCOM.2013.6567070
19. Li M, Yu S, Cao N, Lou W. Authorized private keyword search over encrypted personal health records in cloud computing. *IEEE 31th International Conference on Distributed Computing Systems (ICDCS)*. 2011; 451-459. doi: 10.1109/ICDCS.2011.55
20. Singhal A. Modern information retrieval: A brief overview. *IEEE Data Engineering Bulletin*. 2001; 24: 35-43.