International Journal of Engineering & Technology, 7 (2.24) (2018) 243-248

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Multi-Keyword Ranked Search in Cloud Storage using Homomorphic Indexing

K. Renugha [1*], P. Shanthi [2], A. Umamakeswari [3]

[1,2,3]Department of CSE, School of Computing, SASTRA Deemed to be University, Thanjavur.
**Corresponding Author Email: renugha.88@gmail.com*

Abstract

In the cloud environment, the main issue is outsourcing of the information to the cloud service provider and outsider. Consider this, the cloud tenant store data in an encrypted form to achieve data security and privacy. The data owner needs the secure information sharing from the cloud and without leak of access pattern to the eavesdroppers. XOR homomorphic encryption searchable algorithm along with ranking is proposed to provide the security over the network. In addition our scheme provides secure Multi-keyword ranked search over encrypted data. Efficient ranked search algorithm returns the relevant document based on the results for the given multiple keywords. The experimental results prove that the system is efficient.

Keywords: *XOR Homomorphic encryption scheme (HOM), cloud storage, data security, keyword search, On-demand key, access pattern, Multi-keyword ranked search.*

1. Introduction

Cloud computing is the key technology in present years, so it provide the more advantages to corporate organisation. Cloud service serves large amount of resource and gives elastic storage to the enterprise or the start-up companies. The people or the IT companies need cloud due to free access of the data anywhere and reduced cost of infrastructure. The people store data on the cloud due to enormous storage capacity and the backup. IT people have the large data growth in online transaction over the business, so they need another data storage which has been tremendously increased.

Business data are very much precious to the any organization; it will be more reliable, highly available and fault-tolerance from cloud service provider and also more safeguard from the attackers. The data owners need to check the data confidentiality and privacy of data from the unauthorised users. The existing searchable encryption Song's scheme [4] provides data security and data confidentiality. But the main disadvantage is it leaks some of the data due to the passive attackers. Exclusive-or (XOR) homomorphism encryption scheme with Multi-keyword ranked search can be implemented to resolve the outsourcing data to the cloud. XOR Homomorphic encryption searchable [1] algorithm is presented to protect the data privacy from unwanted access in cloud environment.

The main mechanism involved in scheme [3] is first; this technique gives protection of data by pre-encrypting the single or multiple keywords and performing randomization over the XOR operation in bitwise and secure data leak on pattern which can be accessed frequently. Secondly, on-demand key evaluation on searching of data from the untrusted environment using homomorphic key and this evaluation can reduce key dependency. Finally, it protects from the passive attacker .Comparing security requirements on proposed schemes with other Searchable encryption schemes (Table.1). The Multi keyword ranked encryption scheme [2] provides secure transaction and the data owners can perform searching on data by using either single or multiple keywords. The cloud storage have large amount of document, retrieving the desired document is very difficult and hence ranking method is used that improves scalability and reduces the searching time.

2. Related Work

The Exclusive-Or homomorphic encryption technique is discussed in this paper and also provide the information on the secure data upload and secure data searching [1]. In past years, the data owners' information can be leaked to the outsider and cloud service provider through passive attacker. The security analysis is additionally proposed for reliability on the secure transaction [2]. This paper suggests the information on searching the multiple keywords and also it explains the concept of the access pattern [3]. The Secure searchable technique is discussed in this paper and also provides the drawback on the leaking of the data pattern due to the storing of keys in the cloud [4].This paper gives information on the pseudo-randomized function instead and a normalized public cryptosystem. In addition to that it provides more information on the semantic security [5].

In the mobile computing, an efficient, secure multi-keyword search is implemented. The proposed scheme provides the information on trapdoor privacy and also gives the solution for the problem of trapdoor unlike ability [6]. The proposed scheme reveals the secure transaction between the cloud user and the cloud server and also provides the information on of data in the cloud environment [7]. This paper provides the information on the hiding pattern on searching the data in the third party environment and also discussed multiple keyword searching [8,13,15]. The complexity on searching over the ranking of data on multiple keyword search can be discussed and also provide information on the searching time optimal search [9,10].

The traditional searchable encryption scheme which uses normally uses single keyword search for creating the search index. These indexes can be hidden from the server [11]. Ranking method especially Horvath-Mateme ranked method or the pivot raking method can be introduced for retrieval of the documents. This

technique is providing detailed information on the automatic extraction of concept [12]. The issue on data dependability is explained and also gives information on the confidentiality of data in mail box [14]. An efficient, dynamic and the secure searching scheme can be suggested in this paper. This scheme proposed not only for the multiple keywords search based on ranking but also deals with the dynamic insertion and deletion of files.

The main issue are the cloud provider; i.e user information can be accessed without any authorization. The proposed scheme is Greedy-Depth First searching scheme is used to maximize the efficiency over the linear search. In addition to that, K-NN algorithm is also added to protect the threat model in the cloud environment [16]. Outsourcing mechanism is discussed for the avoiding the leak of the personalized data in the untrusted environment [17]. The proposal scheme also introduced to prevent from the adaptive chosen-keyword attack [18]. The phrase searching can be discussed in this paper. The main usage of phrase searching is used to improve the searched results in an encrypted data and also explained about the multiple keyword searches for increase efficiency in searching.

This paper provides the solution for the fuzzy keyword search problem and the multi keyword search issues. In addition to that, and also provide information about the Euclidean distance which is used to improve the result based on ranking [19]. Security vulnerability on the traditional key encryption scheme with searching keyword is discussed in [20]. For the backup, video sharing , sharing of files to the group members [21] for that reason people need the cloud environment and in this random permutation such as the Staloo`s algorithm is discussed. It performs the uniform distributed random permutation and unbiased shuffle over the uniform distribution. Hence, this secure permutation is more efficient for cipher pad creation. [22].

Table 1: Comparison of security requirements of searchable encryption schemes.

S. No	Parameters	Song`s scheme (Searchable Encryption Scheme)	XOR Homomorphic Encryption Scheme	XOR Homomorphic Encryption Scheme using Multi keyword Ranked Search
1	Keyword Search	Single Keyword Search	Single Keyword Search	Multi Keyword Search
2	Key Storage	On-cloud	On-demand	On-demand
3	Passive Attack	Possible	Not Possible	Not Possible
4	Session key generation	Not used	Yes	Yes
5	Access pattern	Leaked	Not-Leaked	Not-Leaked
6	Hash based indexing	Static	Dynamic	Dynamic
7	Keyword Guessing	Possible	Not possible	Not possible
8	Searching time	Slow	Fast	Very fast
9	Ranked search	Not used	Not used	Used

3. Preliminaries

3.1. Threat model - Passive Attacks and the Adversary

The Passive attackers or the Eavesdropper are the one of the adversary. The main purpose of the passive attackers is the listen the network communication. To guess the information shared from end to end and then, they silently change the information or erase the data. This reveals the wrong information to the receiver. For that, the XOR homomorphic encryption is used to prevent the passive attackers in the network.

3.2. System goals

The main goal is to preserve the access pattern from the cloud storage, so proposed an efficient and dynamic scheme is the XOR Homomorphic Encryption Scheme using Multi keyword Ranked Search for secure transactions.

1. **Access pattern:** The access pattern of the storing data can`t be leaked to the any outsider due to each transaction their should be session key is generated and on- demand key storage.
2. **Secure Index:** The data searching index or user query index are not leaked because the key can encrypted. The same key is used as index of the file. The automatic updating of index if any manipulation carried out in the cloud.
3. **Searching time:** The searching time can be reduced in the user side due to ranking algorithm.

4. Proposed work: XOR Homomorphic Multi-keyword Ranked Search

4.1 System Architecture

The entities involved in system architecture are data owner, cloud storage server and data user. Data owner can store the data in the cloud in an encrypted form. For secure transaction, data can be encrypt before and then upload to the untrusted server for preventing leakage of data access pattern. After uploading to cloud, data owner provide the access control to the data user that is sharing their secret key. Cloud storage server store or retrieve information based on the data owner or data user request (Fig.1).

For further data searching in cloud, initially load the index which used as keyword and then either retrieve information or add new information or delete the information. The keyword may be single or multiple. After the manipulation of data owner, the automatically index can be updated in the cloud. Data user can retrieve the data in the cloud based sharing of key from the data owner. While searching the data in the cloud storage is difficult, for that ranked searchable algorithm is implemented. This algorithm helps to retrieve the document as the user needs based on the ranking.

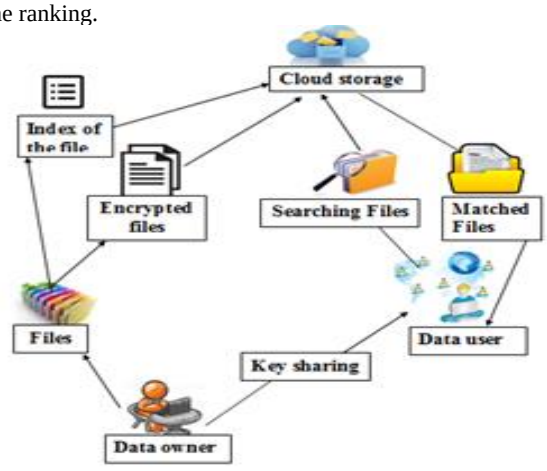

Fig. 1: System Architecture

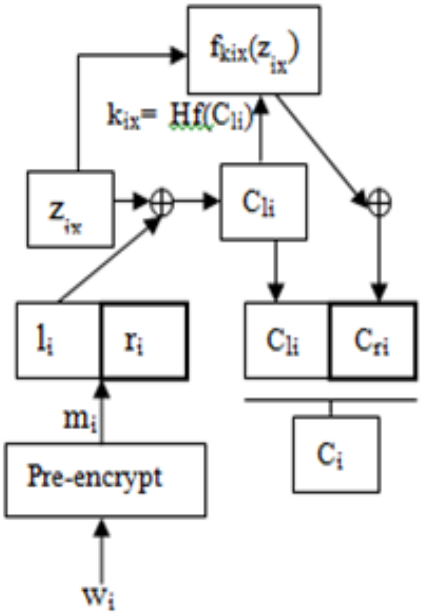

Fig. 2: On-demand key Calculation

4.2. Constructing the secure dictionary for secure uploading and cipher pad generation

The Upload the data in the untrusted server, first focus on the data privacy and security for that previously there should be two types of protection will be carried out. First, the keyword which also used as the index, can be pre-encrypted and randomized using the deterministic symmetric algorithm. Second, the file also encrypted using AES.

Fig.3: Secure files uploaded on the cloud

These two procedures are done and then finally the data can be securely uploaded to the cloud (Fig 3**).** The following steps involved to create the cipher pad for secure uploading of data to the cloud environment (Fig.2):

1. The keyword w_i which has been pre-encrypted using the detrministic symmetric algorithm $m_i = l_i \oplus r_i$ (m_i is the pre-encrypted keyword and l_i or r_i is left/right side bit of m_i) .
2. Generate the random number z_{ix} and xor with the l_i, to create current session permutation key k_{ix}= Hf($l_i \oplus z_{ix}$) . The Homomophic function Hf(.) has been used for enabling k_{ix} to calculated during on evaluation instead storing of key in the cloud.

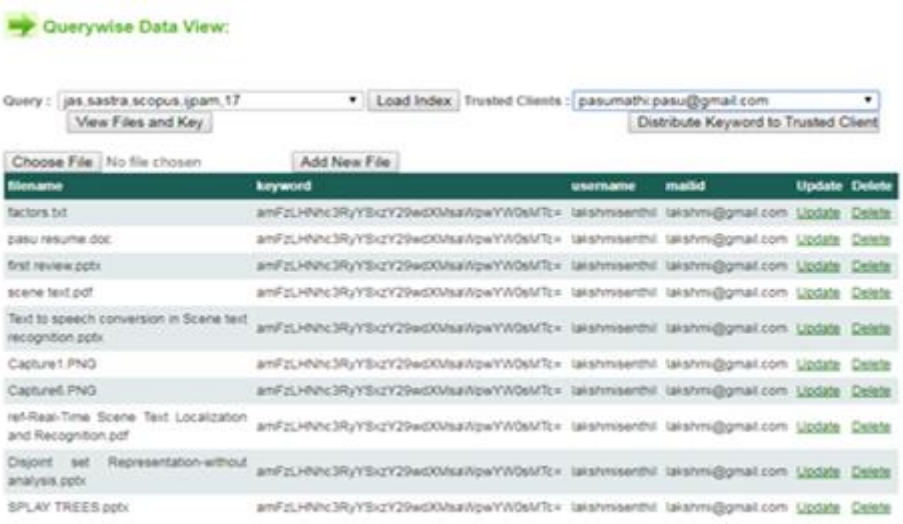

Fig. 4: Loading and updating index sharing to the data user

3. The permutating the z_{ix} with k_{ix}, to form pseudo-random permutating key $f_{kix}(z_{ix})$.

4. Masked ciphertext (C_i) has been formed on ith keyword for any session, $l_i \oplus z_{im} | r_i \oplus f_{kix}(z_{ix})$ and this has been sent to the server for searching of data.

4.3 Secure searching of data

The data can be searched securely on the cloud environment. The data encrypted before upload preventing leakage of data access pattern. After uploading to cloud, data owner provide secret key to the datauser to search the data on the cloud storage. First, load the index which used as keyword and then either retrieve information or add new information or delete the information. Automatically index can be updated in the cloud (Fig.4). Data user can retrieve the data in the cloud based sharing of key from the dataowner.

4.4. Secure data downloading and Ranking

The data owners or data user can perform searching on data by using either single or multiple keyword and the cloud server can provide results based on the user request. Ranked searchable encryption scheme is mainly used to retrieve document based on the ranking in the cloud environment. i.e The cloud storage have large amount of document in that retrieving the desired document is very difficult for that the ranking method is used. Scoring is the way to provide weightage for the relevance document while searching.

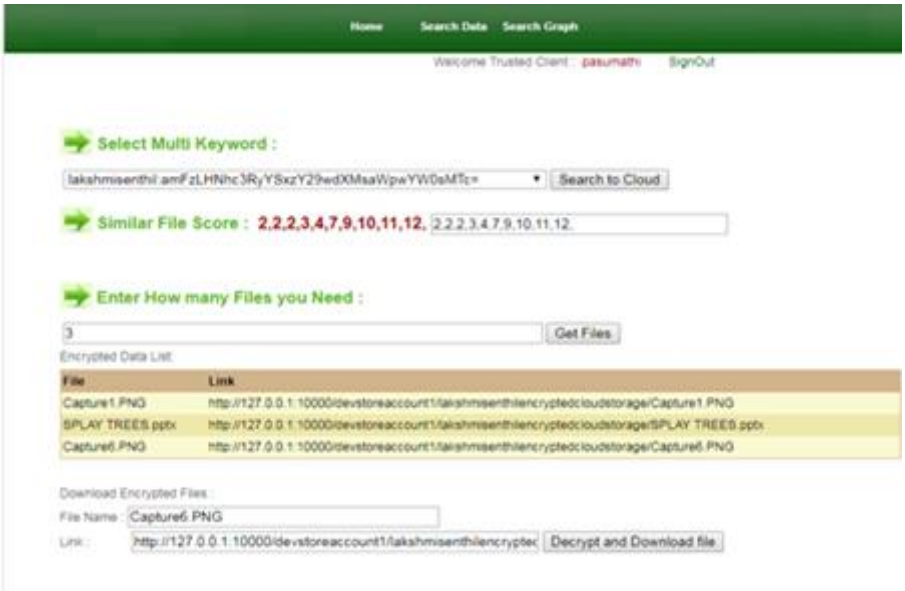

Fig. 5: Top Three files and download the files

The scores can be evaluated on index of data. The resulting set can be formed by matched query and the matched file which can refer as score. This score value is very much important for ranking the document. Scoring evaluation can be performed by the Term Frequency-Inverse Document Frequency. The scoring and ranking of the top-k files can be performed by the retrieval phase. The calculation of the score done by cloud server side and the ranking of top-k documents can be done through data user and are shown in (Fig.5**).**

First, select the documents which can be encrypted by AES algorithm. Scores for each document can be based on query which evaluated and then it return an encrypted score to data user. They get top-k scored files and request data owner for decrypt the keys. Finally, the data owner can provide the decrypted keys to data user by mail-id. The ranked searched method are used to search files based on the user`s query and download files which can be based on the ranking and improve scalability and reduce the searching time.

5. Experiment Analysis

The Xor homomorphic multi-keyword ranked search can be implemented for secure searching of data on the cloud environment. The real-time performance results can be depicted on graph (Fig.6). The graph shows the performance evaluation on default grep searching, HOM searching at server and client-side and Opti-HOM searching at server and client-side.

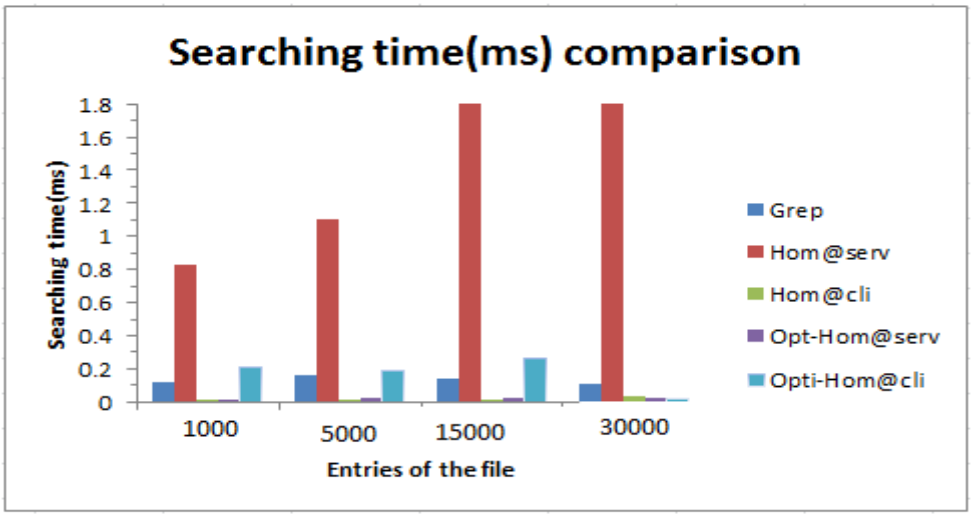

Fig. 6: Performance on comparison of searching time

The proposed system has the modules such as the 1.secure building index data protections on data owner, 2.the secure data loading and uploading, 3.secure data searching at cloud server and 4.secure downloading and the ranking at data user side. In our graph, module 1 and 3 can be depicted. The module 2 is expensive so it can be preceded by onetime-offline process. The each and every encypted key and the corresponding, docId which can be store as value and also its index are used as keys (x-axis referred as index of the document).

The queries are passed through the cloud storage, and it can be attached on the extra index of the keyword, also hash index of key assisted to cloud storage to filterate the many entries on keyword list. The searching time in millisecond is referred as y-axis. The cormparing the secure searching time with default grep, HOM searching and Opti-HOM searching. Finally, the performance can be evaluated experimentally and its verified.

6. Conclusion

The Xor homomorphic multi-keyword ranked search is proposed to provide the protection over the data on searching performed on the cloud server. The proposed search mechanism introduces a randomization based technique for every session, for that the data access leakage pattern should be preserved and there should any eavesdroppers are possible. Because the data protection method on data owner can be implemented before upload of data on the cloud server. In cloud server, first load the index and then perform any manipulation such as add or delete or update of index. The data owner can share the key to the data user to access the file. The data user can retrieve top-k files based on the ranking. Hash indexing is used to calculate searching performance. The evaluation of searching can be done by on-demand calculation. The ranked searched algorithm used to retrieve document based on the ranking in the cloud environment. This proposed search does not leak data pattern to eavesdropper. The proposed searchable encryption can be verified experimentally and results can be displayed on real time in the cloud server.

References

[1] Ren SQ, Tan BH, Sundaram S, Wang T, Ng Y, Chang V, “Aung KM,Secure searching on cloud storage enhanced by homomorphic indexing”, Future Generation Computer Systems, 2016.

[2] Xiuxiu Jiang , Jia Yu, Jingbo Yan , Rong Hao, “Enabling efficient and verifiable multi-keyword ranked searchover encrypted cloud data”, Information Sciences, PP.403–404.2017.

[3] K. Renugha, P. Shanthi,A ,“ novel secure searching on cloud storage enhanced by homomorphic indexing”, International Journal of Pure and Applied Mathematics, vol 117 No. 16, PP.533-541, 2017.

[4] Song DX, Wagner D, Perrig A, “Practical techniques for searches on encrypted data”, In Security and Privacy, 2000.

[5] Boneh D, Di Crescenzo G, Ostrovsky R, Persiano G, “Public key encryption with keyword search”, InEurocrypt, Vol. 3027, pp. 506-522, 2004.

[6] Haiping Huang, Jianpeng Du, Hui Wang and Ruchuan Wang, “A Multi-keyword Multi-user Searchable Encryption Scheme Based On Cloud Storage”, IEEE TrustCom-BigDataSE-IS,2016.

[7] Chen R, Mu Y, Yang G, Guo F, Wang X, “Dual-server public-key encryption with keyword search for secure cloud storage”, IEEE transactions on information forensics and security, pp.789-98, 2016.

[8] Mathew P, Babu SS, “Secure Fuzzy Multi-Keyword Ranked Search over Encrypted Cloud Data”, International Journal of Innovative Research in Computer and Communication Engineering, 2015.

[9] Strizhov M, Ray I, “Multi-keyword similarity search over encrypted cloud data”, IFIP International Information Security Conference, Springer, pp. 52-65, 2014.

[10] Craig Macdonald , Iadh Ounis, “The influence of the document ranking in expert search”, Information Processing and Management, pp. 376–390, 2011.

[11] Wang C, Cao N, Ren K, Lou W, “Enabling secure and efficient ranked keyword search over outsourced cloud data”, IEEE Transactions on parallel and distributed systems, pp.1467-79, 2012.

[12] Alain Materne, Gershom Sleightholme, “Methods of ranking search results for searches based on multiple search concepts carried out in multiple databases”, World Patent Information, 2014 .

[13] Amit Praseed, Remya K Sudheesh and Dr. K.Chandrasekaran, “Efficient Privacy Preserving Ranked Search over Encrypted Data”, IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2015.

[14] Rodrigues J, Ferreira B, Domingos H, “TMS: a trusted mail repository service using public storage clouds”, InProceedings of the 8th Workshop on Middleware for Next Generation Internet Computing, 2013.

[15] Goh EJ, “Secure indexes”, IACR Cryptology ePrint Archive, 2003.

[16] Zhihua Xia, Xinhui Wang, Xingming Sun and Qian Wang, “A Secure and Dynamic Multi-Keyword Ranked Search

Scheme over Encrypted Cloud Data", IEEE transactions on parallel and distributed systems, vol. 27, no. 2, 2016.
[17] Bing Wang, Shucheng Yu,Wenjing Lou and Y. Thomas Hou, "Privacy-Preserving Multi-Keyword Fuzzy Search over Encrypted Data in the Cloud", IEEE INFOCOM, 2014.
[18] Hongwei Li, Dongxiao Liu, Yuanshun Dai, Tom H. Luan,And Xuemin (Sherman) Shen, "Enabling Efficient Multi-Keyword Ranked Search Over Encrypted Mobile Cloud Data Through Blind Storage", IEEE transactions on emerging topics in computing, 2015.
[19] Chang YC, Mitzenmacher M, "Privacy preserving keyword searches on remote encrypted data", InACNS, Vol. 5, pp. 442-455, 2005.
[20] Liu C, Zhu L, Wang M, Tan YA, "Search pattern leakage in searchable encryption: Attacks and new construction", Information Sciences, pp. 176-88, 2014.
[21] Durstenfeld R, "Algorithm 235: random permutation", Communications of the ACM, 1964.
[22] T. Padmapriya, V.Saminadan, "Performance Improvement in long term Evolution-advanced network using multiple imput multiple output technique", Journal of Advanced Research in Dynamical and Control Systems, Vol. 9, Sp-6, pp: 990-1010, 2017.
[23] S.V.Manikanthan and V.Rama"Optimal Performance of Key Predistribution Protocol In Wireless Sensor Networks" International Innovative Research Journal of Engineering and Technology ,ISSN NO: 2456-1983,Vol-2,Issue –Special –March 2017.