ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

Advanced MPSE Scheme for Searching Shared and Encrypted Data

Gokula Nath G[1], Syamamol T[2]

Department of Computer Science, Mangalam College of Engineering, M G University, Ettumanoor, Kottayam, India[1,2]

Abstract*:* Encryption is a well established technology for protecting sensitive data. Multi Party Searchable Encryption is a scheme in which multiple users store and shared their data with each other. The scheme consists of two entities: A server and a set of users. Achieving multi-keyword searching is challenging in the scheme and also it is challenge to have a secure searchable encryption due to the key sharing between set of users. The proposed scheme, advanced Multi Party Searchable Encryption allows a key server to solve the key sharing problem. The new scheme allows multi-keyword searching with homomorphic encryption and also enables searching keyword in the form of checksum. Moreover, the evaluations show the speed of proposed scheme compared with the old MPSE scheme with respect to searching and encryption/decryption.

Keywords: Multi Party Searchable Encryption (MPSE), Homomorphic Encryption, checksum, MD5, DES, RSA.

I. INTRODUCTION

Searchable Encryption enables users to perform keyword based searches on an encrypted database just as in normal database transactions [2]. The scheme limited to the single user setting where the data owner who generate the database allows a single user to perform searches on it. To support multi-user searches [1], share the secret key for database searching among all users. The scheme allows only one user to upload the data, though multiple users are able to search.

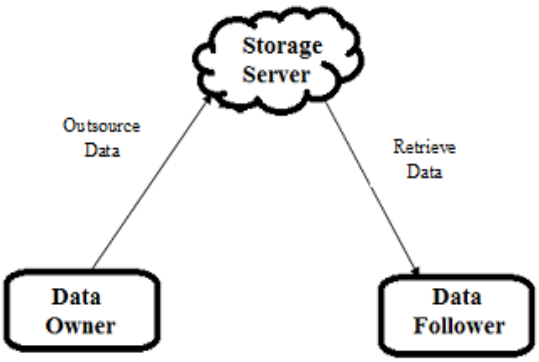

Fig.1. Basic Searchable Encryption Scheme

MPSE scheme [1], an approach to the single user searching problem was introduced, which allows every user to build an encrypted index for each of her documents and store it at the server. Then delegate the server to search on their behalf by issuing a trapdoor. The index contains list of encrypted keywords, as well as some authorization information which selectively authorizes other users to search over this index. In the proposed scheme, an individual user can act as a data owner and/or a data follower. Data owner is the user who uploads the file and allows others to search. Data follower is the user who will be authorized by others to search their data.

MPSE scheme doesn't support multi-keyword searching and the user who want to search, send the keyword in the form of index or ID. For that the key should be shared between the data owner and the data follower in a secure way. File updation is another one important problem for MPSE scheme. This paper presents a new approach to provide multi-keyword searching and it allows file updation. The keyword searching can also be done with checksum conversion.

Most of the existing approaches discussed with searchable encryption schemes [9] and follow up by many others [7] [6] [4]. In the work of Bao et al. a new party namely, a user manager [3] is introduced into the system to manage multiple users' search capabilities. In this, the user manager needs to be fully trusted since it is capable of submitting search queries and decrypting encrypted data. Most of these work discussed about order preserving encryption [8] [5], where the cipher text preserve the order of the plaintext so that every entity can perform an equality comparison.

II. RELATED WORK

Searchable encryption provides a promising direction in solving the privacy problem when outsourcing data to the server. Such schemes allow users to store their data in encrypted form at an untrusted server [9], and then allow the server to search on their behalf by issuing a trapdoor [1]. When a user Alice acts a follower and searches another user's data e.g. Bob's, then Bob can figure out the keyword in Alice's query if he colludes with the server. As such, Alice may not want a random user Eve to authorize her to search his data. It is crucial to have a secure and efficient procedure for Alice to allow another user Bob to authorize her to search his data.

Almost all existing schemes [10] only consider the scenario where a single user acts as both the data owner and the querier. However, most databases in practice do not just serve one user; instead, they support search and write operations by multiple users. Extending a single-user scheme to a full-fledged multi-user scheme by sharing secret keys among all users is a naive approach with several serious shortcomings. First, there is no feasible means to determine the originator of a query in a

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

provable manner, since all queries are generated from the same key. This becomes unacceptable when accountability of queries is desired by the database application. Secondly, user revocation can be prohibitively expensive. In a multi-user application, user revocation is a routine procedure. For a key-sharing based scheme, revocation often implies a new round of key distribution involving all non-revoked users. Obviously, this is not scalable for large and dynamic systems where user revocation may occur frequently.

Order-preserving symmetric encryption (OPE) [11] [12] is a deterministic encryption scheme whose encryption function preserves numerical ordering of the plaintexts. OPE has a long history in the form of one-part codes, which are lists of plaintexts and the corresponding cipher texts, both arranged in alphabetical or numerical order so only a single copy is required for efficient encryption and decryption. OPE not only allows efficient range queries, but allows indexing and query processing to be done exactly and as efficiently as for unencrypted data, since a query just consists of the encryptions of a and b and the server can locate the desired cipher texts in logarithmic-time via standard tree-based data structures.

Current security mechanisms are not suitable for organisations that outsource their data management to untrusted servers. Encrypting and decrypting sensitive data at the client side is the normal approach in this situation but has high communication and computation overheads if only a subset of the data is required. New cryptographic schemes have been proposed that support encrypted queries over encrypted data. But they all depend on a single set of secret keys, which implies single user access or sharing keys among multiple users, with key revocation requiring costly data re-encryption. So introduced an encryption scheme [16] where each authorised user in the system has his own keys to encrypt and decrypt data. The scheme supports keyword search which enables the server to return only the encrypted data that satisfies an encrypted query without decrypting it.

Most existing symmetric searchable encryption schemes aim at allowing a user to outsource her encrypted data to a cloud server and delegate the latter to search on her behalf. These schemes do not qualify as a secure and scalable solution for the multiparty setting, where users outsource their encrypted data to a cloud server and selectively authorize each other to search. Due to the possibility that the cloud server may collude with some malicious users, it is a challenge to have a secure and scalable multiparty searchable encryption (MPSE) [1] scheme. The scheme allows every user to build an encrypted index for each of her documents and store it at a cloud server. The index contains a list of encrypted keywords, as well as some authorization information which selectively authorizes other users to search over this index.

III. ADVANCED MULTI PARTY SEARCHABLE ENCRYPTION

In this paper propose a new approach known as Advanced Multi Party Searchable Encryption (Advanced MPSE), in which searching can be done based on ranking of all documents. Ranking based on TFIDF value of all documents. The document with high TFIDF value has the priority to send back to the specific user. The proposed scheme mainly consists of two servers:

Storage server: A server in which all the encrypted documents are stored.

Key Server: A server in which all the public keys and the authorization information are stored.

Advanced MPSE scheme consists of following operations:
1. TFIDF Calculation
2. Checksum Calculation
3. Encryption of documents
4. Searching of keyword
5. Decryption

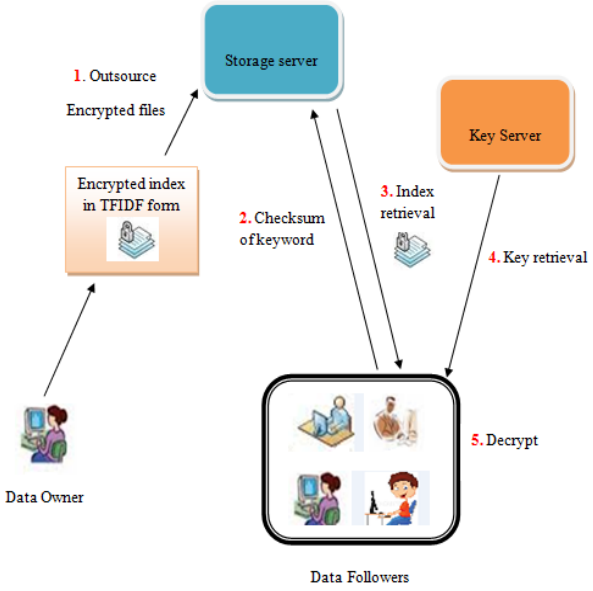

Fig.2. System Architecture

Key sharing is an important problem for MPSE scheme. I.e. the data follower wants a public key for the conversion of his keyword which is to be searched into index form. So this key retrieval should be done via a secure channel and also whenever the user wants the key, the owner should be ready to give the key. So the absence of the data owner affects this process. To overcome this problem, introduce a key server which has the responsibility for this key retrieval. When the follower wants the key, he should contact with the key server. If the user has the permission to access the document, then only the key server send the key for decrypting it.

Every document stored in index form; index in the sense which contains Ids of all the encrypted keywords within that document. Every data owner has the rights to search over his index because he has the corresponding key to decrypt it. There wouldn't be a key sharing between the data owner and follower so that the follower doesn't want to see the owners.

Algorithm:

In Advanced MPSE algorithm, there are mainly four phases: TFIDF calculation, Encryption, Keyword

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

Searching and Decryption. Here encryption of TFIDF value is done based on RSA algorithm and the file encryption is done based on Data Encryption Standard (DES) algorithm [13].

The checksum conversion of every word within each document can be done using MD5 algorithm [14].

1. TFIDF Calculation

Calculate Term Frequency (TF) and Inverse Document Frequency (IDF) for every uploaded document based on each keyword in that document. Then rank all documents based on this TFIDF value.

Document having high TFIDF value has highest rank. This ranking can be done for every searching. Then encrypt these values by using RSA algorithm [15]. TFIDF calculation can be done using following formula:

$$\text{TF} = \frac{Total\ num.of\ documents}{Doc.containing\ particular\ word}$$

$$\text{IDF} = \frac{Num.of\ times\ the\ keyword\ appears}{Total\ num.of\ words\ in\ a\ doc.}$$

2. Encryption

The uploaded document can encrypt and stored at the storage server. Otherwise it is visible by all other users. This encryption can be done using the symmetric key algorithm DES [13]. The encryption key can be generated from RSA algorithm.

3. Searching

Searching can be done by using checksum of keyword or keywords. This checksum conversion based on Message Digest algorithm, MD5 [14]. Every word in a document converts to checksum after uploading it to the server. Then this checksum-word list shared with users. The users who want to search send checksum of corresponding keyword. Then the document with highest rank i.e. with highest TFIDF value can send back to the user.

One of the important features of proposed scheme is, multi-keyword searching. It is possible with homomorphic algorithm [1]. It is the type of encryption, which allows taking two pieces of cipher text and performing an operation on them which results in the cipher text of the concatenation of two respective pieces of plain text.

4. Decryption

Decryption is also done using the inverse form of DES algorithm. Use cipher text and the public key as the inputs to DES but use the keys are in reverse order. That is, use K16 on the first iteration, K15 on the second until K1 which is used on the 16th and last iteration.

Decryption is done only when the private key is match with the public key. These keys distributed by the key server only if the corresponding user has the permission to access it.

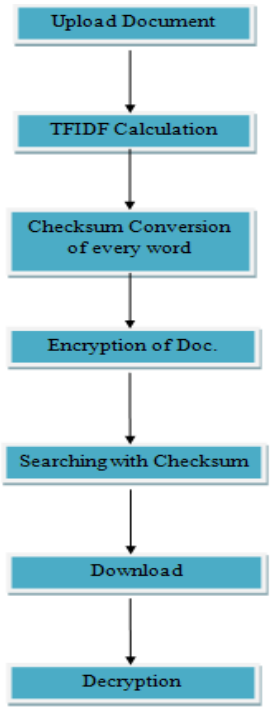

Fig.3. Flow chart

IV. EXPERIMENTATION AND RESULTS

The performance of the proposed MPSE scheme compared with the old scheme with respect to keyword searching and the encryption/decryption.

Experimental setup Simulations are performed up to N = 100 pdf files. Then encryption and decryption are performed and was chosen in the range [0 to 4 sec]. And also searching of 100,000 and 1,000,000 keywords are performed for both old and new MPSE scheme.

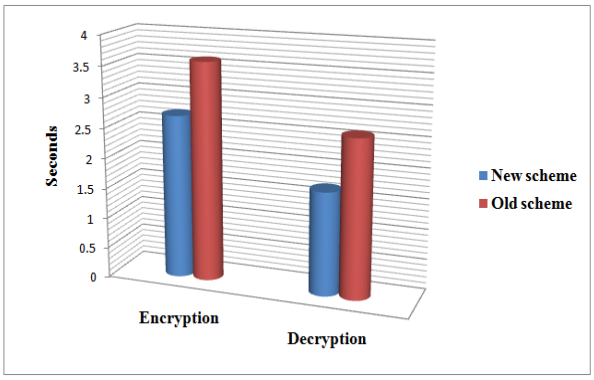

Fig: (a)

First experiment measured the performance of bulk encryption/decryption. The data set used was 100 pdf files and the total size was 99.2 MB. Each document was associated with 10 keywords, i.e. 1000 keywords to encrypt in total. The files were encrypted with DES. The results are shown in Fig: (a). Third experiment was to measure the search times with various sized databases. Here used databases containing 100,000 keywords and 1,000,000 keywords. The results are shown in fig: (b).The new scheme requires less time for keyword searching than the old MPSE scheme.

DOI 10.17148/IJARCCE.2015.4739

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

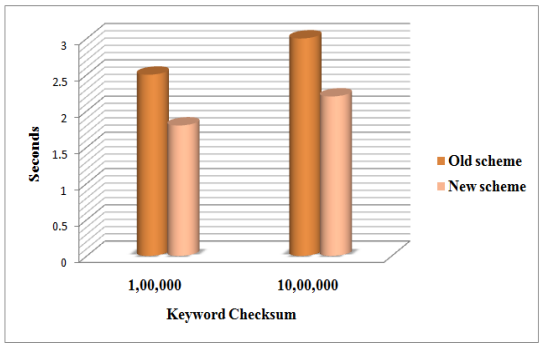

Fig: (b)
Fig.4. Performance Evaluations

From the graph above, it should be clear that Advanced MPSE method is much faster than old MPSE scheme in terms of keyword searching and encryption/decryption. In traditional MPSE, for searching a word, matching can be done for all document stored at the server. But for advanced scheme, the document with highest rank should send back to the specific user. So new method should take less time for the keyword searching and also it is much secured than the old scheme in terms of key sharing.

V. CONCLUSION

Advanced Multi Party Searchable Encryption introduced as a new version of Multi Party Searchable Encryption Scheme which supports multi-keyword searching by means of homomorphic algorithm. The proposed scheme also avoids key sharing with the help of a key server. All the uploaded documents can be stored in the encrypted form. Based on TFIDF value every document assign with a rank. Then the document with highest TFIDF value can send back to the corresponding user. The follower doesn't want to see the data owners. Searching keyword given in the form of checksum to avoid the need of a key for index conversion. File updation can also possible with advanced MPSE scheme. The evaluations demonstrate the viability of the proposed mechanisms in compared with MPSE scheme.

ACKNOWLEDGEMENT

The first author would like to thanks all those people, who guided and supported. Without their valuable guidance and support, this task was not possible and also likes to thank colleagues for their discussions and suggestions.

REFERENCES

[1] Qiang Tang, "Nothing is for Free: Security in Searching. Shared and Encrypted Data," IEEE Trans. On Information Forensics And Security, Vol. 9, No. 11, Nov. 2014, pp. 1943-1952.
[2] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric encryption: Improved definitions and efficient constructions," in Proc. 13th ACM Conf. Comput. Commun. Security, 2006, pp. 79–88.
[3] F. Bao, R. H. Deng, X. Ding, and Y. Yang, "Private query on Pract. Encrypted data in multi-user settings," in Proc. 4th Int . Conf. Inf. Security Experience, vol. 4991. 2008, pp. 71–85.
[4] C. Bösch, Q. Tang, P. Hartel, and W. Jonker, "Selective document retrieval from encrypted database," in Proc. 15th Inf. Security Conf. (ISC), vol. 7483. 2012, pp. 224–241.
[5] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. 3rd Int. Conf. Appl. Cryptography Netw. Security, vol. 3531. 2005, pp. 442–455.
[6] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in Proc. IEEE 28th Int. Conf. Data Eng., Apr. 2012, pp. 1156–1167.
[7] M. Raykova et al., "Usable, secure, private search," IEEE Security Privacy, vol. 10, no. 5, pp. 53–60, Sep./Oct. 2012.
[8] E. Shen, E. Shi , and B. Waters, "Predicate privacy in encryption systems," in Proc. 6th Theory Cryptography Conf. Theory Cryptography, vol. 5444. 2009, pp. 457–473.
[9] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for 2000, searches on encrypted data," in Proc. IEEE Symp. Security Privacy, May pp. 44–55.
[10] F. Bao, R. H. Deng, X. Ding, and Y. Yang, "Private query on Pract. Encrypted data in multi-user settings," in Proc. 4th Int. Conf. Inf. Security Experience, vol. 4991. 2008, pp. 71–85.
[11] A. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill, "Order-preserving symmetric encryption," in Advances in Cryptology—EUROCRYPT (Lecture Notes in Computer Science), vol. 5479, A. Joux, Ed. Berlin, Germany: Springer-Verlag, 2009, pp. 224–241.
[12] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2004, pp. 563–574.
[13] W. C. Barker and E. B. Barker, "Sp 800-67 Rev. 1. Recommendation for the Triple Data Encryption Algorithm Block Cipher," technical report, Nat'l Inst. of standards and Technology, 2012
[14] The MD5 Message-Digest Algorithm, RFC1321.
[15] Maha TEBAA, Saïd EL HAJJI, Abdellatif EL GHAZI, "Homomorphic Encryption Applied to the Cloud Computing Security" Proceedings of the World Congress on Engineering 2012 Vol I. WCE 2012, July 4, 2012, London, U.K
[16] C. Dong, G. Russello, and N. Dulay, "Shared and searchable Encrypted data for untrusted servers," in Proc. 22nd Annu. IFIP WG 11.3 Works. Conf. Data Appl. Security XXII, vol. 5094. 2008 pp 127– 143.

BIOGRAPHIES

Gokula Nath G, Department of Computer Science & Engineering, Mangalam college of Engineering, Ettumanoor, Kerala, India.

Syamamol T, Assistant Professor, Department of Computer Science and Engineering, Mangalam college of Engineering, Ettumanoor, Kerala, India.