IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 4, Ver. VII (Jul.-Aug. 2017), PP 01-05
www.iosrjournals.org

Data Storage on Cloud Using Hybrid Encryption with One Time Password

*Munavvara Tahaseen[1], Maniza Hijab[2], Sofia Mariam Hussain[3]
[1](Department of Information Technology, Muffakham Jah College of Engineering and Technology, India)
[2](Department of Computer ScienceEngineering, Muffakham Jah College of Engineering and Technology, India)
[3](Department of Information Technology, Muffakham Jah College of Engineering and Technology, India)
Corresponding Author: *Munavvara Tahaseen*

Abstract : *The arrival of cloud computing has been so extensive that it has become the most sought-after idea for data owners to outsource their data to the cloud thereby allowing users to retrieve it. Although traditional methods are proof enough that the models created based on symmetric key encryption perform well, a significant drawback is vulnerability of the security of the key which when once leaked, discloses confidential data of the data owner. Another drawback is the process of data user authentication, which is very tedious where a long conversation is carried out between the administration server and the user to verify the identity of the user. In this paper, we initiate schemes that deal with improvising the traditional methods as well as improvising the process of data user authentication. To minimize the complexity of the traditional system and enhance the security we implement hybrid encryption, which is an overtly suitable replacement to the traditional symmetric key encryption where, the former utilizes a combination of symmetric and asymmetric encryption techniques. As a result, the first drawback has curbed. This is because, asymmetric encryption has two keys; a public key and a private key respectively and one cannot easily guess the private key from the public one. Therefore, even if the public key used for encryption leaks, the information is secure In order to provide a modern approach for authentication of users, we introduce OTP (One Time Password) mechanism, which is a very secure, and an easy way to verify the identity of the users.*

Keywords: *Symmetric Encryption, Asymmetric Encryption, Hybrid Encryption, Cloud Computing, One Time Password, Security*

Date of Submission: 14-08-2017 Date of acceptance: 26-08-2017

I. Introduction

Using cloud computing users can store data on remote servers without the need to have infrastructure. But this raises security issues as the data is hosted on third party servers which may be un-trusted and the data may be sensitive such as e-mails, personal health records, government documents etc. To overcome this problem, data owners encrypt their data before outsourcing onto the cloud. Conventional searching is based on Boolean search which is not applicable on cloud as the data is encrypted. Users can search for interested files using keyword search. As data is encrypted, traditional keyword search can't be applied to cloud. Many searchable encryption techniques have been proposed that allow users to securely search over encrypted data through keywords, but they support only Boolean search not considering file relevance. Ranked keyword searches, as proposed by [3][4][5][8][9] [10] [11] [12] [13] [14] improve searching efficiency by returning files based on their rank as per user relevance. They combine both cryptography and information retrieval techniques.

In this paper we propose HEOTP, a hybrid encryption with One Time Password in cloud. Section II provides the background of the different searching techniques proposed so far. Section III defines the basic scheme; Section IV reviews the related works. In Section V we introduce HEOTP. Section VI concludes the paper.

II. Background

2.1 Cryptography

Outsourcing data on the cloud is a very sensitive issue as the data may be confidential which should not be leaked. Data leakage may be a serious issue. This raises the need for information security. As a result cryptography has to be used, which has been in use largely in the domain of military and government. Confidentiality of data can be provided by applying various encryption procedures like public key encryption or private key encryption. Symmetric encryption is a traditional and one of the best-known techniques. A secret key i.e. a string of random numbers, a word or just any number is applied to the data in order to change the

content in a particular manner. This could be as simple as shifting each letter in the alphabet by a number of places. As long as the sender and the receiver are both aware of the secret key, they can encrypt and decrypt the data that utilizes this key. The key referred here is known as the Symmetric Key. Only one key is used in Symmetric Encryption i.e. the key used for encryption and decryption by the sender and the receiver respectively, remains same. One significant advantage that symmetric encryption has given us is speed. It consists of algorithms which perform encryption and decryption techniques at a good speed. However, one disadvantage of private key encryption is that it involves a logistics problem of conveying the Symmetric Key. The main drawback is that it lacks security under some condition, especially during the process of transmission of data rather than just keeping it safe. The Symmetric key is to be kept safe. If someone else comes to know about the key, then data is not secure and the confidentiality is lost. Also, it is not preferable to use Symmetric Key where public network is used for sharing the key. The chance for modification and malicious insertion is high in private key encryption techniques.

Asymmetric encryption makes use of two keys; a public key and a private one. While the public key is used for encoding, the private key is used for decoding. The public key is based on the private key but it is difficult to trace back to the private key. It requires a lot of effort to work out the private key from the public one. This is because one would have to factor a very large prime number which is beyond the computational abilities in a reasonable time frame. So even if the public key is leaked, it is difficult to guess the private key from the public one information is secure. The basic idea of Asymmetric Encryption is that if a person A wants to send a message to a person B that only the person B is allowed to read, person A has to use the public key given by person B to encrypt the respective message which only the person B's private key can decrypt.

Hybrid Encryption allows us to combine the benefits of symmetric and asymmetric encryption. Hybrid Encryption is a suitable and a better replacement. [20] A hybrid cryptographic technique combines the benefits of asymmetric and symmetric encryption. The benefits include security and speed. The process of hybrid encryption begins with the sender obtaining the recipient's public key. Let's name sender as A and recipient as B. Once A obtains B's public key, it generates a symmetric key for the data encapsulation scheme. A then uses the symmetric key and encrypts the message using the data encapsulation. Also, using B's public key, A encrypts the symmetric key under the key encapsulation scheme. Finally, A sends both encryptions to B. Once B receives the cipher texts, it uses its private key to decrypt the symmetric key and uses the symmetric key to decrypt the message both of which are in the encapsulation segment.

2.2 Information retrieval

Information retrieval is used to find relevant documents not simple finding matches to patterns. A retrieval strategy is an algorithm that takes a query and finds its similarity with the documents. For computing similarity coefficient stop-words (i.e. words which are insignificant like verbs, articles etc.) are eliminated before index creation. The remaining terms are keywords and for each keyword a weight is assigned. Various techniques have been proposed for computing term weight. We use the following formula for computing the similarity score or the relevance score of term t in document fd;

(1) Relevance score (t, fd) = (1/ |fd|). (1+ln fd, t)

fd,t denotes the Term Frequency of term t in file fd; |fd| is the length of file fd, obtained by counting the number of indexed terms.

An inverted index is prepared for the document collection. Inverted index consists of keywords and for each keyword there is a posting list which consists of document identifiers and relevance score of the keyword in the document. Whenever an end user searches for a document he can request for top n documents and these documents are displayed ranked on their relevance score.

III. Basic Scheme

Cloud computing enables users to store data on the cloud server. For security issues, the data to be outsourced is encrypted. Different searching techniques have been proposed to securely search over cloud data. The basic idea of all the searches is same. The data owner creates an index for his file collection before encrypting it. After encrypting the file collection, the data owner also encrypts the index file and then outsources the encrypted file collection along with the encrypted index onto the cloud server. The data users search for a file on the cloud server using keywords. The keyword is given to a trapdoor, which encrypts the keyword, which is then searched on the cloud server. The files with the keyword are returned as per their relevance score.

IV. Related Work

In [1] the index generation in SSE-1 is very complex and cumbersome as it uses an array, look-up table and linked list to create the index, SSE-2 uses padding in index to have equal length posting list for each keyword which makes its accuracy less compared to SSE-1. In [2] RSSE uses the idea of SSE and introduced the use of information retrieval and cryptography together in cloud. It uses relevance score to rank the

documents. SSE-1 and SSE-2 in [1] are limited to single user search and multi-user SSE is limited to single keyword search. [4] Introduced multi keyword search whose implementation process is very complex. This search approach is limited to data stored by single data owner. [5], [6] introduced searching on data stored by multiple data owners. [6] Uses an administrative server which provides additional layer of security to the index file and also checks the identity of the user. The use of administrative server adds additional communication overhead and hence the search process is slow.

V. HEOTP

The model consists of four entities namely; data owners, data users, administration server and cloud server. The data owners have their own collection of files. They first build a searchable secure index on keywords extracted from the file collection and encrypt the symmetric keys used to decrypt each file and submit the index and the encrypted symmetric keys to the administration server. They now encrypt their files and outsource the encrypted files to the cloud server. Once the administration server obtains the index, as in [6] it further re-encrypts index and submits it to the cloud server. After successful user authentication, the data user is allowed to send its public key to the administration server along with the desired search request. After receiving the user's public key and the search request, the administration server further re-encrypts the search request as a trapdoor and submits it to the cloud server along with a unique ID assigned to the user. The cloud server, having received the ID and the trapdoor successfully, looks for the encrypted index of every data owner and keeps the corresponding set of files ready for the user with the unique ID. Meanwhile, the administration server encrypts the symmetric key that is required to decrypt the files using the end user's public key and sends it to the end user.

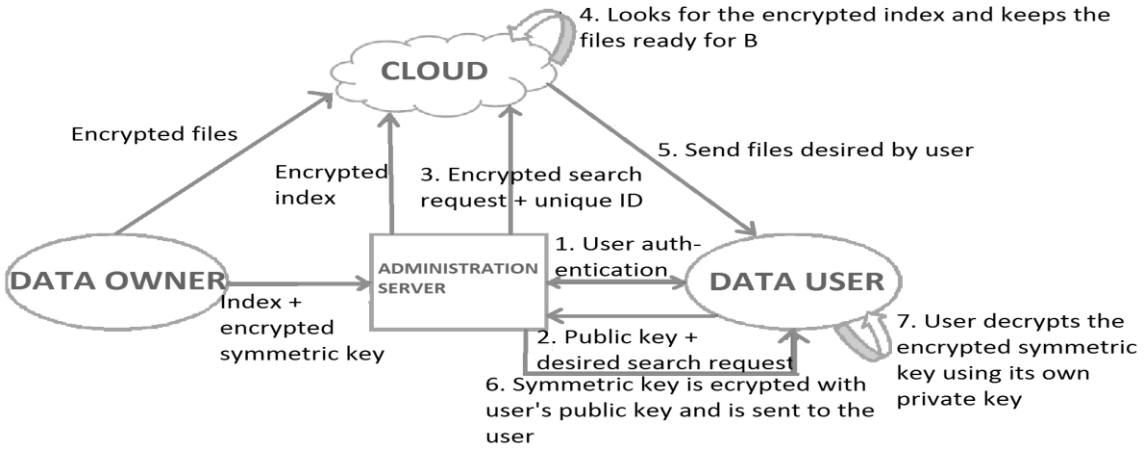

Fig.1 Architecture of HEOTP

5.1 User authentication

To prevent unauthorized users to access the data stored on cloud, and to ensure that the communication between end user and administrative server is not eavesdropped, authentication is provided between end user and the administration server. In this section we present user authentication assuming the administration server as A, and the user as B.

1. B sends a request to A.
2. A sends an OTP to B along with a timestamp.
3. Once B receives the OTP, it has to decrypt it and send it back to A within the stipulated time frame, failing to do so would turn the OTP invalid.
4. When B is successful in returning the correct OTP to A within the stipulated time, it is authenticated.
5. If B was unable to send the OTP to A in the required time limit, due to say Man in The Middle attacks or any other faulty actions, the OTP would turn invalid and B would fail to get authenticated. In this case, it would have to request A again after a few minutes for a new OTP. After say about 3 unsuccessful trials, B would be blocked to send any more requests to A.

This process prevents malicious users from accessing the data stored on the cloud. This process is simple as compared to data authentication in [6] which requires the construction of historical data.

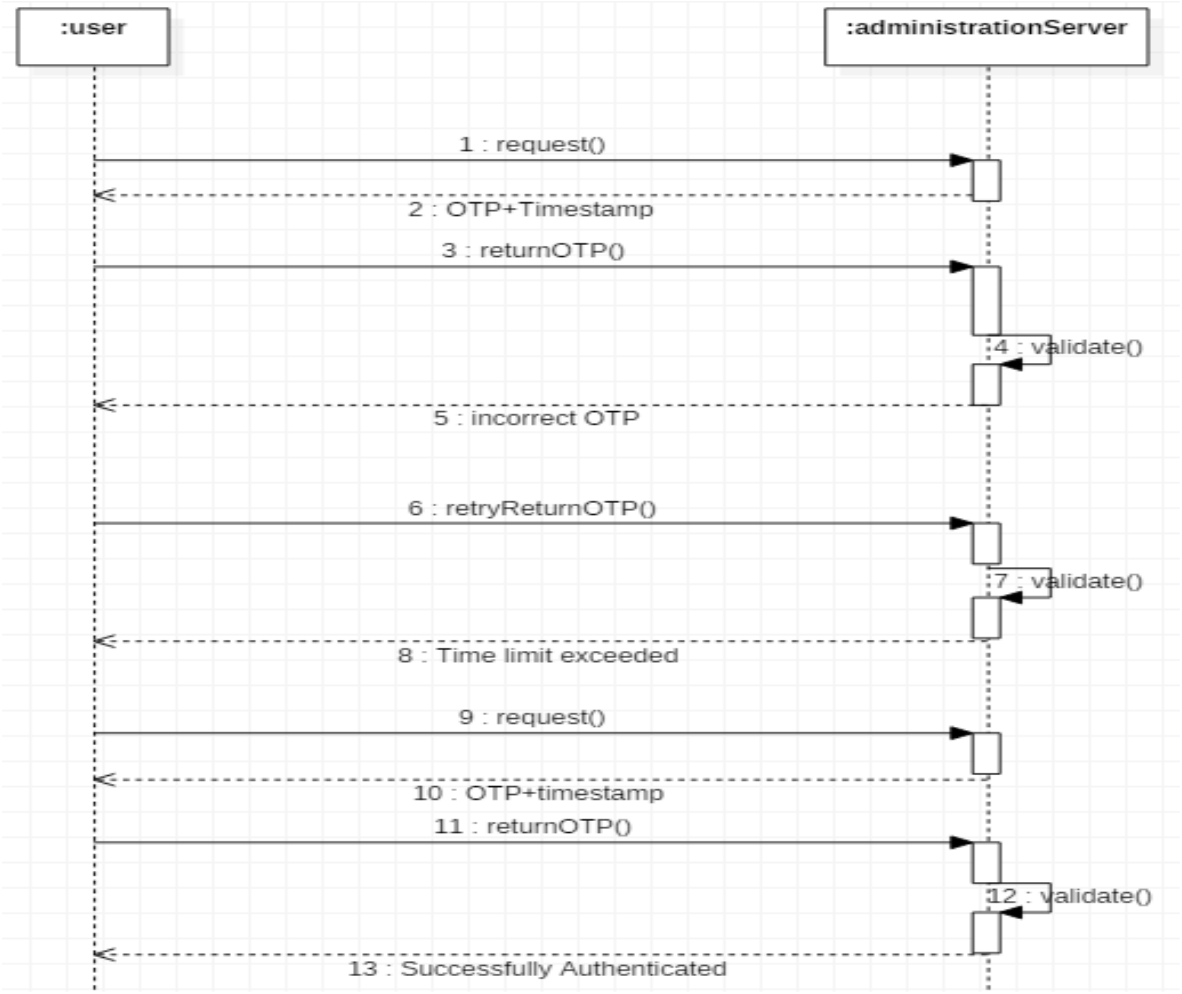

Fig.2 sequence diagram for user authentication

5.2 Keyword encryption

Keywords are extracted from the file collection by the data owners. The keywords are encrypted with symmetric keys of data owners. Each data owner has his own secret keys. The administrative server is responsible for exchange of secret keys with the end users. This reduces the overhead on the end users who are relieved from the burden of contacting different data owners. Administrative server is assumed to be trusted and secure as some certificate authority in public key encryption [6]. We assume the cloud server is not malicious but curios as in [6], [17].

We use the one-many-order preserving encryption as in [17] for preserving the order of relevance score of keywords. The rank of a file can be obtained by adding the relevance score of keywords in the search query which appears in the file.

VI. Conclusion

In this paper we have proposed the use of hybrid encryption with OTP for outsourcing data on cloud. The proposed approach successfully overcomes the drawbacks of the previous papers where the security of the key was at risk, is now ensured with total safety and the tedious user authentication process is now simple and secure. We are in the implementation phase of outsourcing on private cloud using our scheme.

References

[1] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. ACM Conf. Computer and Comm. Security (CCS'06), 2006.

[2] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", IEEE Transactions on Parallel and Distributed Systems, vol.23, No.8, Aug.2012.

[3] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE Distrib.Comput.Syst., Genoa, Italy, Jun. 2010 elevated channel lowtemperature poly-Si TFT," IEEE Electron Device Lett., vol. 20, pp.569–571, Nov. 1999.

[4] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy- preserving multi-keyword ranked search over encrypted cloud data," in Proc.IEEE INFOCOM, Shanghai, China, Apr.2011.

[5] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy- preserving multi-keyword ranked search over encrypted cloud data," IEEE Trans. Parallel Distrib. Syst., vol. 25.

[6] Wei Zhang, Y Lin, Sheng Xiao, Jie Wu Siwang Zhou, "Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing" IEEE Transactions on Computers Vol.65, No.5, May 2016.

[7] A. Boldyreva, N. Chenette, and A. O'Neill, "Order- preserving encryption revisited: Improved security analysis and alternative solutions," in Proc. 31st Annu. Conf. Adv. Cryptol., Aug. 2011.

[8] Cong Wang, Ning Cao, Kui Ren, W Lou "Enabling Secure and Efficient Ranked Keyword Search Over Outsourced CloudData" IEEE Transactions on Parallel and Distributed Systems, vol 23, No.8, August 2012.

[9] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure knn computation on encrypted databases," in Proc. of SIGMOD, 2009.

[10] Q. Liu, C. C. Tan, J. Wu, and G. Wang, "Efficient information retrieval for ranked queries in cost-effective cloud environments," in Proc. IEEE INFOCOM, 2012.

[11] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H.Li,"Verifiable privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 11.

[12] Z. Xu, W. Kang, R. Li, K. Yow, and C. Xu, "Efficient multikeyword ranked query on encrypted data in the cloud,"in Proc. IEEE 19th Int. Conf. Parallel Distrib. Syst.,Singapore,Dec. 2012.

[13] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in Proc. IEEE INFOCOM, San Diego, CA, USA, Mar. 2010.

[14] Ke Li,, Weiming Zhang, Ce Yang and Nenghai Yu."Security Analysis on One-to-Many Order Preserving Encryption-Based Cloud Data Search",IEEE Transactions on Information Forensics and Security,2015.

[15] de.slideshare.net

[16] Wei Zhang, Y Lin, Sheng Xiao, Jie Wu Siwang Zhou, "Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing" IEEE Transactions on Computers, 2015.

[17] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data," Proc. IEEE 30th Int'l Conf. Distributed Computing Systems (ICDCS '10),2010.

[18] www.cs.ucla.edu

[19] Wei Zhang, Sheng Xiao, Yaping Lin, Ting Zhou, SiwangZhou."Secure Ranked Multi-keyword Search for Multiple Data Owners in Cloud Computing", 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2014.

[20] https://en.wikipedia.org/wiki/Hybrid_cryptosystem

IOSR Journal of Computer Engineering (IOSR-JCE) is UGC approved Journal with Sl. No. 5019, Journal no. 49102.

Munavvara Tahaseen. "Data Storage on Cloud Using Hybrid Encryption with One Time Password ." IOSR Journal of Computer Engineering (IOSR-JCE), vol. 19, no. 4, 2017, pp. 01–05.