ISSN 2347 - 3983
International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 113- 116 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp23.pdf

Security in Multi keyword Search Over Encrypted Cloud Data

CH.Ellaji[1], D.Shobha Rani[2]
[1]139E1D5806, M.Tech CSE Student,S.V.Engineering College for Women,Tirupati,India
ellaji0808@gmail.com
[2]Assistant Professor, Dept. of CSE, S.V.Engineering College for Women, Tirupati, India,
shobharani.d@svcolleges.edu.in

ABSTRACT

With the advent of cloud computing, the outsourcing of data is more flexible and economic saving, so that there are huge number of data users and documents in the cloud , it is crucial for the search service to allow users provide multi-keyword query and provide result similarity ranking to meet the effective data retrieval need. The searchable symmetric encryption focuses on single keyword search or Boolean keyword search, and rarely differentiates the search results. This leads to a few issues like Single-keyword search without ranking, Boolean- keyword search without ranking, Single-keyword search with ranking.

Key words: cloud computing,multi keyword, data privacy, keyword buffer controller,ranking

1. INTRODUCTION

Cloud computing is a new computing prototype that is built on virtualization, parallel and distributed computing, utility computing and service oriented architecture. For the past many years, cloud computing has emerged as one of the most important and used technology in the IT industry and academic world. The benefits of cloud computing include reduction in costs and in capital expenditures, increased operational efficiencies, scalability, flexibility, immediate time to market and many more. The various cloud computing models are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). It is a subscription-based service, where in one can obtain networked storage space and computer resources. According to the different needs of the user, one can subscribe different types of cloud. Clouds are classified

1.1. Public Cloud

A public cloud can be used by any user with an internet connection and access to the cloud space.

1.2. Private Cloud

A private cloud is established for a specific group or organization and access to this cloud is limited just to at group or organization alone.

1.3. Community Cloud

A community cloud is shared between two or more organizations which have common cloud requirements.

1.4. Hybrid Cloud

A hybrid cloud is fundamentally a combination of at least two clouds, where the clouds included are a blend of public, private, or community

2. SYSTEM DESCRIPTION

We consider a cloud computing system hosting data service, as illustrated in Figure 1, in which three different entities are involved: Cloud server, Data owner and Data user. The cloud server hosts third-party data storage and retrieve services. Since data may contain sensitive information, the cloud servers cannot be fully entrusted in protecting data. For this reason, outsourced files must be encrypted. Any kind of information leakage that would affect data privacy are regarded as unacceptable.

The data owner has a collection of n files C = {f1,f2,…fn} to outsource onto the cloud server in encrypted form and expects the cloud server to provide keyword retrieval service to data owner himself or other authorized users. To achieve this, the data owner needs to build a searchable index I from a collection of l keywords W = {w1,w2,…wl} extracted out of C, and then outsources both the encrypted index I' and encrypted files onto the cloud server. The data user is authorized to process multi-keyword retrieval over the outsourced data. The computing power on user side is limited, which means that operations on user side should be simpli fied. The data user encrypts the query and sends it to the cloud server that returns the relevant files to the data user. Afterwards, the data user can decrypt and make use of the files.

ISSN 2347 - 3983
International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 113- 116 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp23.pdf

Fig1: Architecture for retrieval encrypted data from cloud

3. EXISTING SYSTEM :

The Traditional single key word searchale encryption schemes usually build an encrypte searchble index such that its content hidden to the server unless it is given appropriate trapdoors can be generated via secret keys.The Searchable symmetric encryption (SSE) retrieves encrypted data over cloud.

- The searchable encryption focuses on single keyword search or Boolean keyword search, and rarely differentiates the search results.
- SSE implements server side ranking based on order-preserving encryption(OPE), OPE leaks data privacy.

Disadvantage:

- Single-keyword search without ranking
- Boolean- keyword search without ranking
- Privacy is compromise.

4. PROPOSED WORK

Objective

- To ensure the cloud data user with appropriate encrypted cloud data in a secure and efficient manner.
- Sensitive information protected by data encryption at data owner side.
- Verification of the received files.
- The retrieval process can be support user for multi keyword searching.
- Multi-keyword ranked search retrieves data accurately when compared to single keyword search.

Techniques used

MD5 (Message Digest) hash function

- Among various multi-keyword semantics, the efficient principle of “coordinate matching” is used.

At this time, introduces user registration, authentication of users, uploading data with encryption, download data with decryption, generation of independent file code for each user, etc.,. The easy to use environment with sophisticated techniques is developed. This will provide maximum efficiency by using secure as well as efficient algorithms such as MD5

It will support multi keyword searching which helps the user to find any related materials that they need. The ranking of the file based on the usage of the file is listed and then the list is show as per the maximum used files. By doing the ranking process at the client side the data leakage can be prevented.

We are opting MD5 encryption algorithm for authentication which is bit more complex when compared to the traditional algorithms in storing the data but it is low level of hacking because of 32 bit encryption.

In the proposed system we have introduced a Keyword Buffered Controller in the server. On receiving a request from a user, the Keyword Buffered Controller stores the keywords in the Keyword Buffered Table and on retrieval of the relative files, it stores the file names in the Keyword Buffered Table along with the keywords. When the next request from a user, Keyword Buffered Controller first check in the Keyword Buffered Table, if it finds the matching encrypted keywords in the table, it fetch out the related file names and sends the content to the user. By doing so the computational task of the cloud is reduced.

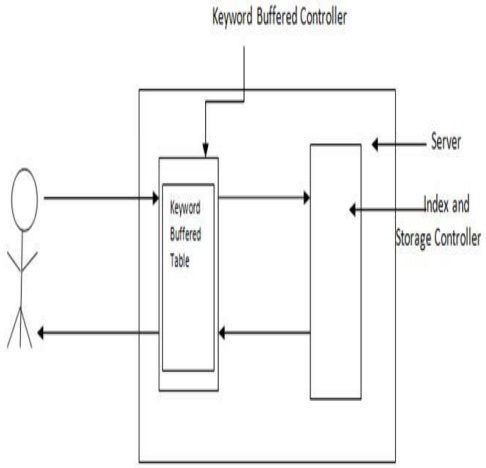

Fig: Keyword buffered controller

4.1 Encoding operation:

A Secret Code (SC) is appended to the file, and the file and SC are encrypted with a random key. A hash value of the encrypted data is computed. The hash value and random key are then combined via bitwise exclusive-or to form a difference, which is appended to the encrypted data

ISSN 2347 - 3983
International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 113- 116 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp23.pdf

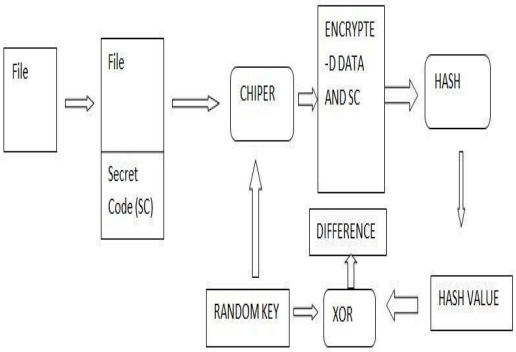

Fig: Process of Encoding operation

4.2 Restoring Data:

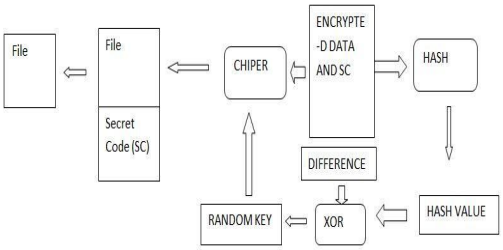

Fig: Architecture for Restoring Data

The first step is to compute the hash *h*, of the encrypted data. Since the last block contains *K*(X-OR)*h*and we know the hash value *h*, we may exclusive-or the last block with the hash to find (*K*(X-OR)*h*(X-OR)*h*). Since *h*(X-OR)*h*equals zero, the result is the random key *K*. The random key is then used to decrypt the encrypted data, and the secret code is checked to detect corruption

4.3 ENCRYPTION ALGORITHM

The algorithm is used to encrypt the file using the key, which will produce the cipher text. The cipher text is actually transmitted to the cloud server. An encryption algorithm also gives the decryption algorithm, which will give the information to receiver how to decrypt the plaintext. The key generation algorithm is used to produce the key that the owner and the receiver need to share. We have analyzed the performance of various encryption algorithm which is suitable for cloud computing.

Procedure of MD5:

The authentication between the owner and the user is done with MD5 message-digest algorithm. The MD5 message-digest algorithm is a cryptographic hash function that makes a 128-bit (16-byte) hash value.MD5 processes a variable-length message into a fixed-length output of 128 bits. The input message is broken up into chunks of 512-bit blocks (sixteen 32-bit words); the message is padded so that its length is divisible by 512. The padding works as follows: First a single bit, 1, is appended to the end of the message. This is followed by as many zeros as are required to bring the length of the message up to 64 bits less than a multiple of 512. The remaining bits are filled up with 64 bits representing the length of the original message, modulo 264.

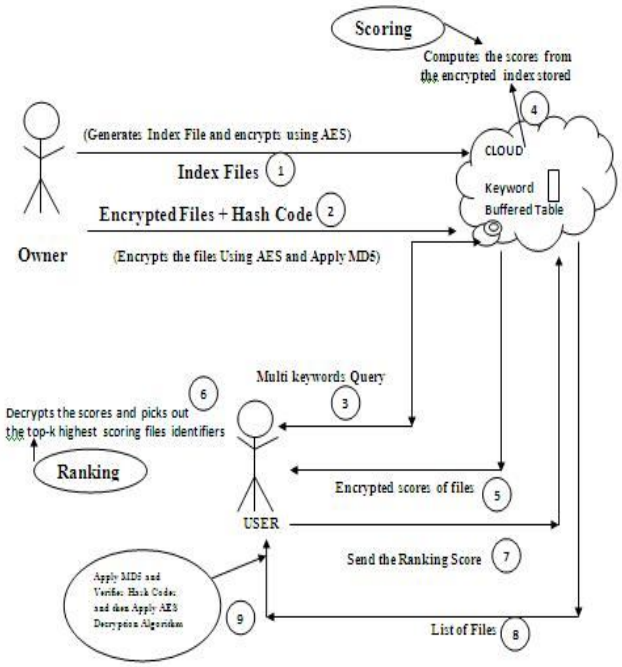

Fig: Architectural Diagram

The Data owner applies MD5 (Message Digest) on the encrypted data and transfers to the cloud.
When the cloud server receives a query consisting of multikeywords from data user, it computes the scores from the encrypted index stored on cloud and then returns the encrypted scores of files to the data user. Next, the data user decrypts the scores and picks out the top-k highest scoring files identifiers to request to the cloud server. The retrieval takes a two-round communication between the cloud server and the data user. We thus name the scheme the Two round searchable encryption scheme, in which ranking is done at the user side while scoring calculation is done at the server side.
When user receives the files and corresponding hash scores, it applies MD5 hash checking on the received files and generates the hash code and verifies the files by comparing the hash code with the received from the cloud.

ISSN 2347 - 3983
International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 113- 116 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp23.pdf

5. RESULTS

- Secured multi keyword retrieval over encrypted cloud data.
- Similarity relevance and scheme robustness.
- A server-side ranking SSE scheme.
- data privacy.
- Cost efficient.
- Time efficient.

6. CONCLUSION

In this paper, proposed a secured way of accessing files from cloud, and a secured and reliable scheme for data owner to provide better services to the users. The owner side encryption scheme and index file generation helps the data user to get secure and protected data with better QOS. To improve the QOS a client side ranking process has been adopted. Searching the query in the index file rather than the file system cloud server can give quick response very quickly. Our proposed scheme fulfills the security requirements of multi keyword search over the encrypted cloud data.

References

[1] Cloud Security Alliance, "Top threats to cloud computing," http://www.cloudsecurity alliance.org, 2010.

[2] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. of ACM CCS, 2006.

[3] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. of ICDCS, 2010.

[4] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy, 2000.

[5] D. Boneh, G. Crescenzo, R. Ostrovsky and G. Persiano, "Public-key encryption with keyword Search," in Proc. of Eurocrypt, 2004.

[6] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A. L. Varna, S. He, M. Wu, and D. W. Oard, "Confidentiality-preserving rank-ordered search," in Proc. of the Workshop on Storage Security and Survivability, 2007.

[7] K. Ren, C. Wang, and Q. Wang, "Security Challenges for thePublicCloud,"IEEE Internet Computing, vol. 16, no. 1, pp. 69-73,2012.

[8] M.Belare, A.Boldyreva, and A.O'Neil, "Deterministic and efficiently searchable encryption," in Proceedings of rypto 2007, volume 4622 of LNCS. Springer- Verlag, 2007.

[9] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "EnablingPublicVerifiability and Data Dynamics for Storage Security in Cloud Computing, "IEEE Trans. Parallel and Distributed Systems, vol. 22,no. 5, pp. 847-859, May 2011.

[10] Sun-Ho Lee and Im-Yeong Lee,"Secure Index Management Scheme on Cloud Storage Environment" ,International Journal of Security and Its Applications Vol. 6, No. 3, July, 2011.