ISSN:0975-9646

Suman M et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3239 - 3241

An Approach for Efficient and Secure Retrieval of Encrypted Cloud Data Based On Top-K Multikeywords

Suman M[1], B. Chempavathy[2]

[1]*M.Tech II year (CNE), Dept of ISE*
The Oxford College of Engineering
Bangalore, Karnataka, India

[2]*Asst.Professorof ISEDept*
The Oxford College of Engineering
Bangalore, Karnataka , India

***Abstract*- In cloud computing, due to large number of data users and documents, it is crucial for the search service to allow 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.**
We define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE) by proposing an efficient and more secure algorithm. We also introduce the concept of keyword buffer controller that allows for quick search of documents and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality.

Keyword- Cloud, data privacy, ranking, similarity relevance,keyword buffer controller.

I. INTRODUCTION

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 simplified. 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.

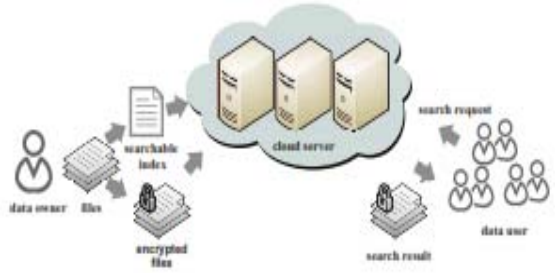


Fig. 1. Scenario of retrieval of encrypted cloud data

II. EXISTING SYSTEM

- 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 compromised.

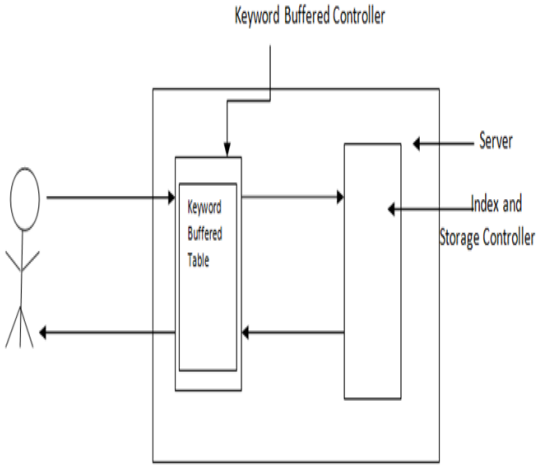
III. PROPOSED WORK

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.
- Employ Two round searchable encryption scheme(TRSE) that supports top-k multikeyword retrieval and address security issue.
- Multi-keyword ranked search retrieves data accurately when compared to single keyword search.

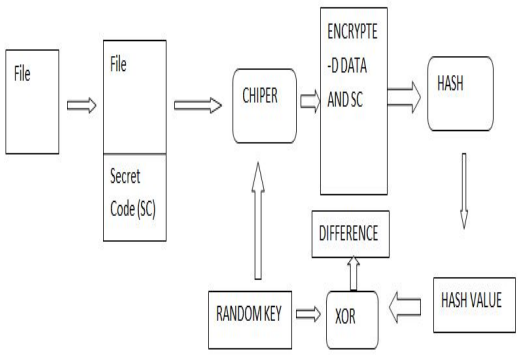
Techniques used

- A Two Round Searchable Encryption(TRSE) scheme is implemented that supports top-k multikeyword retrieval.
- AES(Advanced Encryption Standard) encryption and MD5 (Message Digest) hash function.
- Among various multi-keyword semantics, the efficient principle of "coordinate matching" is used.

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.

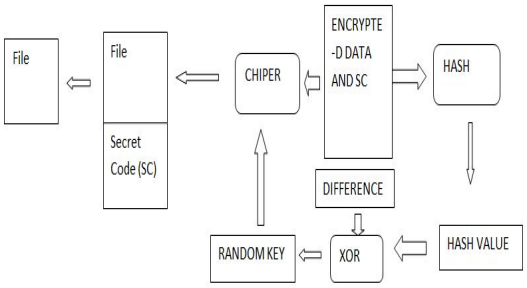


Encoding Operation



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

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.

Overall Operation Steps

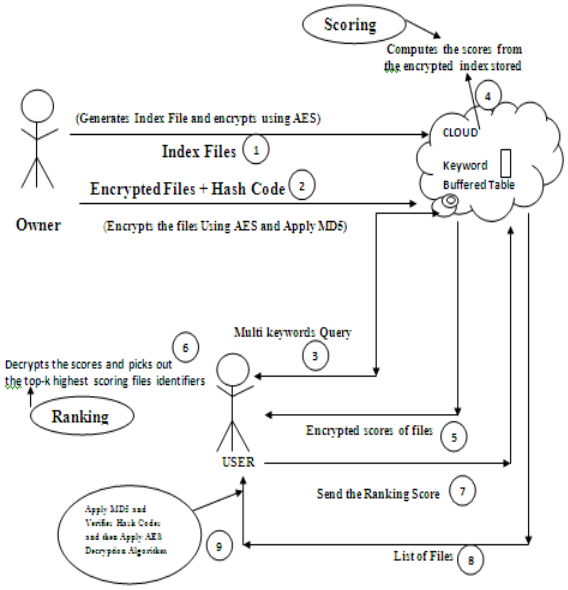


Fig: 2 Architectural Diagram

In the proposed scheme, the data owner encrypts the searchable index with AES (Advanced Encryption Standard) encryption.

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. If verification succeed than it applies the AES decryption algorithm and gets the files.

IV. RESULTS

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

V. CONCLUSION

We have proposed a secured way of accessing files from cloud. We proposed 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 top-k retrieval 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 inCloudComputing,"IEEE Trans. Parallel and Distributed Systems, vol. 22,no. 5, pp. 847-859, May 2011.

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

AUTHORS:

Suman M is pursuing M.Tech in the Computer Network Engineering at The Oxford College of Engineering, Bangalore. She received Bachelor of Technology degree from NIE Institute of Technology, Mysore, in the stream of Computer Science And Engineering. Her research interests are Security in Wireless Sensor Networks and Cloud Computing.

B Chempavathy has done her B.E in computer science from Jaya College of Engineering Affiliated to Madras University and ME from Jaya College of Engineering Affiliated to Anna University. She has worked at Sriram engineering college for one year and at Wipro Technologies for one year. She is currently working as the Assistant Professor in ISE Department of The Oxford College Of Engineering since 3 years. She has guided many M.Tech students in Computer Network Engineering. She has over all 1 year Industry & 4 years of teaching experience.