International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163
Volume 1 Issue 5 (June 2014) http://ijirae.com

Retrieving Secure Data from Cloud Using OTP

M. Karthika*
Department of Information Technology
SNS College of Technology

J. Vasuki
Department of Information Technology
SNS College of Technology

S. Sugashini
Department of Information Technology
SNS College of Technology

Abstract— Define and solve the problem of effective and secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency), thus making one step closer towards practical deployment of privecy- preserving data hosting services in Cloud Computing. To improve the security for the data retrieval from cloud environment, the One Time Password is used. The One Time Passwod is sent to the user mail to view the original data. The Model exhibits the Querying Process over the cloud computing infrastructure using Secured and Encrypted Data access and Ranking over the results would benefit the usre for the getting better results.

Keywords— Ranked search, Privacy- Preserving, Cloud Computing, One Time Password, Querying Process

I. INTRODUCTION

Cloud computing, a critical pattern for advanced data service, has became a necessary feasibility for data users to outsource data. Controversies on privacy, however, have been incessantly presented as outsourcing of sensitive information including emails, health history and personal photos is explosively expanding. Reports of data loss and privacy breaches in cloud computing systems appear from time to time.

Furthermore in cloud computing, data owners may share their outsourced data with a number of users, who might want to only retrieve the data files they are interested in. One of the most popular ways to do so is through keyword-based retrieval. Keyword-based retrieval is a typical data service and widely applied in plaintext scenarios in which users retrieve relevant files in a file set based on keywords. However, it turns out to be a difficult task in cipher text scenario due to limited operations on encrypted data. Besides, in order to improve feasibility and save on the expense in the cloud paradigm, it is preferred to get the retrieval result with the most relevant files that match user's interest instead of all the files which indicates that the files should be ranked in the order of relevance by user's interest and only the files with the highest relevance's are sent back to users.

A series of searchable symmetric encryption schemes have been proposed to enable search on cipher text. Traditional SSE schemes enable users to securely retrieve the cipher text but these schemes support only Boolean keyword search i.e., whether a keyword exists in a file or not, without considering the difference of relevance with the queried keyword of these files in the result. To improve security without sacrificing efficiency schemes presented in show that they support top-k single keyword retrieval under various scenarios. Authors are made attempts to solve the problem of top-k multi-keyword over encrypted cloud data. These schemes, however suffer from two problems - Boolean representation and how to strike a balance between security and efficiency. In the former, files are ranked only by the number of retrieved keywords which impairs search accuracy. In the latter, security is implicitly compromised to tradeoff for efficiency which is particularly undesirable in security-oriented applications.

Preventing the cloud from involving in ranking and entrusting all the work to the user is a natural way to avoid information leakage. However, the limited computational power on the user side and the high computational overhead precludes information security. The issue of secure multi-keyword top-k retrieval over encrypted cloud data thus is: how to make the cloud do more work during the process of retrieval without information leakage.

Introduce the concepts of similarity relevance and scheme robustness to formulate the privacy issue in searchable encryption schemes and then solve the insecurity problem by proposing a two-round searchable encryption (TRSE) scheme. Novel technologies in the cryptography community and information retrieval community are employed including homomorphic encryption and vector space model. In the proposed scheme, the majority of computing work is done on the cloud while the user takes part in ranking which guarantees top-k multi-keyword retrieval over encrypted cloud data with high security and practical efficiency.

II. PROBLEM FORMULATION

The main threat on data privacy roots in the cloud itself. To ensure privacy, users usually encrypt the data before outsourcing it onto cloud which brings great challenges to effective data utilization. However, even if the encrypted data utilization is possible, users still needs to communicate with the cloud and allow the cloud operate on the encrypted data which potentially causes leakage of sensitive information.

A. Existing System

To search a file in the internet, make a query to the i nternet server. Internet will retrieve the most number of visited file which is called as number of Hits. Till now any Search engine will retrieve the links to the user based on the frequent number of Clicks or Hits made by the user. So ranking proves is achieved using this methodology only. Even some times irrelevant data would be ranked for the user which is of no use. Furthermore in cloud computing, data owners may share their outsourced data with a number of users, who might want to only retrieve the data files they are interested

© 2014, IJIRAE- All Rights Reserved

in. One of the most popular ways to do so is through keyword-based retrieval. Keyword-based retrieval is a typical data service and widely applied in plaintext scenarios in which users retrieve relevant files in a file set based on keywords. However, it turns out to be a difficult task in cipher text scenario due to limited operations on encrypted data.

B. Limitations of Existing System

- Existing multi-keyword search support only Boolean queries, i.e., a file either matches or does not match a query.
- Data storage without cryptographic on cloud will encourage the data theft by the malicious users.

III. PROPOSED SCHEME

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 may regard as unacceptable. The data owner has a collection of n files 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 from a collection of keywords and then outsources both the encrypted index 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 authorized data user at first generates a query. For privacy consideration, which keywords the data user has searched must be concealed. Thus the data user encrypts the query and sends it to the cloud server that returns the relevant files to the data user (e.g. Fig. 1). Afterwards, the data user can decrypt and make use of the files.

A. Advantages

- Ranking based search for users are more convenient
- Proposed cloud storage systems that provide confidentiality, integrity and verifiability of client data against un-trusted cloud provider.

B. System Architecture

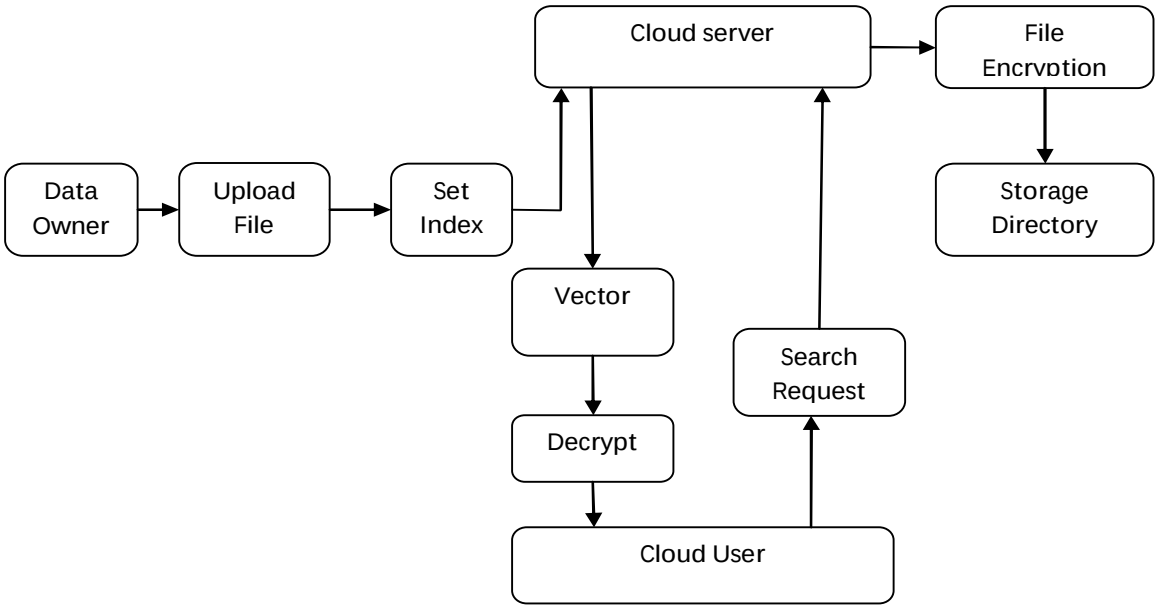

Fig.1 System Architecture

C. Module Description

1) Registration: Admin wants registration and login through user name and password. Data owner and user registration will be carried on (e.g. Fig. 2 & Fig. 4).

2) Encryption: The data is uploaded into the cloud server. The cloud server hosts third-party data storage and retrieve services (e.g. Fig. 1). Encryption is made on the file for protection using AES algorithm. AES has a fixed block size of 128 bits, and a key size of 128, 192, or 256 bits. AES operates on a 4×4 column-major order matrix of bytes

The key size used for an AES cipher specifies the number of repetitions of transformation rounds that convert the input, called the plaintext, into the final output, called the cipher text.

The numbers of cycles of repetition are as follows:

- 10 cycles of repetition for 128-bit keys.
- 12 cycles of repetition for 192-bit keys.
- 14 cycles of repetition for 256-bit keys

© 2014, IJIRAE- All Rights Reserved

3) Random Key Generation: While searching the required data on cloud, the encrypted files are displayed. To open the encrypted files the user needs to enter the key. The random key was generated at server side by PRNG and sent to the user mail. The random key value generated by PRNG evaluation algorithm at server side of the system (e.g. Fig. 6). The generated random key that is one time password is sent to the user mail id.

4) Decryption: To avoid the leakage of sensitive information, one time password (OTP) is used. This OTP is used to see data in cloud ,when search a file and tends to see the file the OTP will send to email and get the OTP and apply to see the file (e.g. Fig. 7*)*..The OTP is valid for some specified time not more than that. If the time expires, the password should not be valid.

D. Algorithms:

1) Top – K select Algorithm: Since the complexity of the INSERT algorithm is O (K), the overall complexity of TOPKSELECT algorithm is O(NK). Note that k, which denotes the number of files that are most relevant to the user's interest, is generally very small compared to the total number of files. In case of large value of k, the complexity of the TOPKSELECT algorithm can be easily reduced to O (n log k) by introducing a fixed-size min-heap.

2) AES Algorithm: AES is based on a design principle known as a substitution-permutation network, and is fast in both software and hardware. Unlike its predecessor DES, AES does not use a Feistel network. AES is a variant of Rijndael which has a fixed block size of 128 bits, and a key size of 128, 192, or 256 bits. By contrast, the Rijndael specification per se is specified with block and key sizes that may be any multiple of 32 bits, both with a minimum of 128 and a maximum of 256 bits.

AES operates on a 4×4 column-major order matrix of bytes, termed the state, although some versions of Rijndael have a larger block size and have additional columns in the state. Most AES calculations are done in a special finite field. Each round consists of several processing steps, each containing five similar but different stages, including one that depends on the encryption key itself. A set of reverse rounds are applied to transform cipher text back into the original plaintext using the same encryption key.

E. Database Design:

Design Table 'userinfo' in 'SEO' on '(LOCAL)'

Column Name	Data Type	Length	Allow Nulls
username	varchar	20	✓
password	varchar	20	✓
email	varchar	50	✓
question	varchar	100	✓
answer	varchar	100	✓

Fig.2 Database design of User Information

Design Table 'uploadinfo' in 'SEO' on '(LOCAL)'

Column Name	Data Type	Length	Allow Nulls
keyword	varchar	20	✓
semantickeyword	varchar	20	✓
filepath	varchar	500	✓
owner	varchar	100	✓
rkey	varchar	100	✓
score	numeric	9	✓

Fig.3 Database design of Uploaded Information

F. Simulation Results:

The performance of the system is improved by avoiding the data leakage. The security provided to the data which has been uploaded and retrieval increased to 2-3%. The random key generation provides dynamic password which avoids the hacking and improves the security. The authentication from two different environments provides improved security for data retrieval over the encrypted cloud.

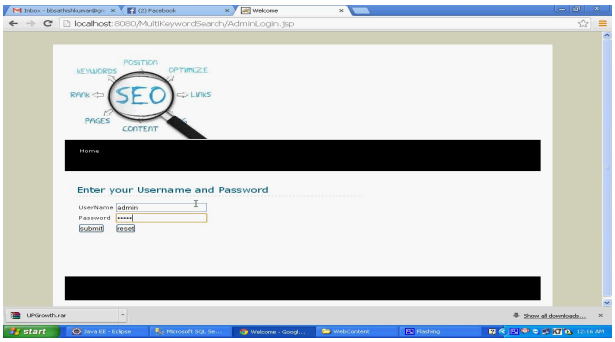

Fig.4 Registration

Fig.5 Uploading Information

© 2014, IJIRAE- All Rights Reserved

Fig.6 Key Generation

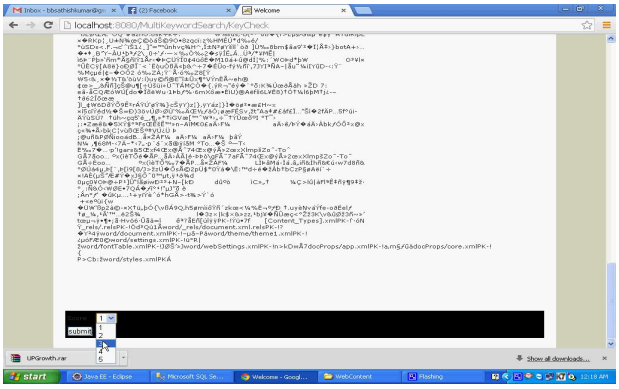

Fig.7 Decryption

IV. CONCLUSIONS

Motivating and solving the problem of secure multi-keyword top-k retrieval over encrypted cloud data. Define similarity relevance and scheme robustness. The performance of the system is improved by avoiding the data leakage. The security provided to the data which has been uploaded and retrieval increased. The random key generation provides dynamic password which avoids the hacking and improves the security. The authentication from two different environments provides improved security for data retrieval over the encrypted cloud.

Storing data on remote cloud storage makes the maintenance affordable by data owners. When multiple data owners are involved, the aspects of membership and data sharing need to be addressed. The proposed scheme provides privacy and complexity while handling the data sharing over cloud. Here the security is enhanced by means of Random Key encryption technique and AES.

ACKNOWLEDGMENT

We are grateful to thank for suggestion and the work which was supported by the department faculties of Information Technology of SNS College of Technology, Coimbatore.

REFERENCES

1. AHN, “Romney hits Obama for security information leakage,” http://gantdaily.com/2012/07/25/romney-hits-obama-for-security-information-leakage/, 2012
2. Amazon.com, “Amazon s3 availability event: July 20, 2008,” http://status.aws.amazon.com/s3-20080720.html, 2008.
3. Armbrust.M, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and M. Zaharia. “A view of cloud computing,” Communication of the ACM 53 (4): 50 58, 2010.
4. Arrington.M, “Gmail disaster: Reports of mass email deletions,” http://www.techcrunch.com/2006/12/28/gmail-disasterreports-of-mass-email-deletions/,December 2006.
5. Cao.N, Wang.C,Li.M Ren.K,and Lou.W,”Privacy-Preserving Multikeyword Ranked Search Over Encrypted Cloud Data,”Proc.
6. Cloud Security Alliance, “Top threats to cloud computing,” http://www.cloudsecurity alliance.org, 2010.
7. Curtmola.R, Garay .J.A, Kamar.S, and Ostrovsky.R,”Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions,”Proc
8. Cormen.T.H,Leiserson .C.E,Rivest.R.L ,and Stein.C ,Introduction To Algorithms,PP.856-887.MIT Press and McGraw-Hill,2001.
9. Dubin.D,”The Most Influential Paper Gerard Salton Never Wrote,”Library Trends,Vol.52,no.4,pp,748-764,2004.
10. Eman M.Mohamed,Hatem.S Abdelkader, Sherif El-Etriby “ Data Security Model for Cloud Computing” ICN 2013 : The Twelfth International Conference on Networks.
11. Gentry.C, “Fully Homomorphic Encryption Using Ideal Lattices,”Proc.
12. Leslie, “NSA has massive database of Americans’ phone calls,” http://usatoday30.usatoday.com/news/washington/2006-05-10/.
13. Randeep Kaur Chhabra, Prof. Ashok Verma “Strong authentication system along with virtual private network: A secure cloud solution for cloud computing” International Journal of Electronics and Computer Science Engineering 1566” Online at www.ijecse.org” ISSN- 2277-1956.\
14. RAWA News, “Massive information leak shakes Washington over Afghan war,” http://www.rawa.org/temp/runews/2010/08/20/massive-information-leak-hakeswashington-over-afghan-war.html, 2010.
15. Regev.O, “New Lattice-Based Cryptographic Constructions,”ACM.J Vol.51,no.6,PP.899-942,2004.
16. Vishal Paranjape, Vimmi Pandey,” An Improved Authentication Technique with OTP in Cloud Computing” International Journal of Scientific Research in Computer Science and Engineering Research Paper Vol-1, Issue- ISSN-2320-7639.
17. Wang.C, Cao.N, li. J, Ren.K, and Lou.W, “Secure Ranked Keyword Search Over Encrypted Cloud Data,”proc.

© 2014, IJIRAE- All Rights Reserved