ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 4, April 2017

A Robust Encryption and Searching System to Share Confidential Data onto the Cloud

[1] Mr. Thejas S, [2] Dr. G. Raghavendra Rao
[1][2] Department of Computer science and Engineering National Institute of Engineering, Mysuru, India

***Abstract*— as cloud computing provides storing and sharing records via a cloud service because of on demand resource access and elasticity. For secure record sharing requires encryption of records before outsourcing. Key management system is one of the important aspects of the system. For sharing large number of documents, each record needs to be encrypted using different keys. Since to decrypt these records requires all keys which is very different task. Hence this approach uses aggregated key system, which aggregates all the keys. Later at the time of decryption of all records it requires one aggregated key. Considering large number of records in the cloud. For the efficient access of these records, it is necessary to search for the records over encrypted file. For this purpose , required to build for the index file. Multi keyword entries searches for the relevance record by matching keywords with the index file.**

***Index Terms*— Cloud, health Records, record decryption, record encryption**

I. INTRODUCTION

As data sharing is one of the features of cloud, there is a need of data to be shared among different user. Cloud storage is storage, where user can upload the files or data via internet. The usage of cloud is been increasing now a days.

Health record (HR) is outsourced to be stored at a third party like cloud. The HR system allows user to create as well as maintain and control their own health data and can add or modify their health information in the particular health data. Data security is the main concern while storing any data on the cloud.

Using an HR system, data owner can share their information with one another. The HR system can be access the information from anywhere through the web. In spite of the fact that there is some protection problems with regards to store the classified information on third party cloud service provider. Therefore, by using cryptography technique reduces the protection problem. In cryptography, plain text is then encrypted into cipher text using some form of key before outsourcing. Then which will in turn be decrypted using the same key.

Since the cryptosystem uses different keys for different record to encrypt. If a large number of records needs to be encrypted and uploaded. Then at the time of decryption, it requires all set of keys. Every time generating different keys and sending over the internet is not secure, and increases the system complexity. Hence the HR system, uses the aggregated key scheme, which is a unique single key. By sharing single aggregated key is sufficient to decrypt the record. This approach reduces the system overhead.

This system is also provides the searching for the relevant record over the encrypted data. For the relevant record searching, the system needs to build the index file. Index file Contains the set of keywords in the binary format. Every time the user searches keywords for the relevant records, keywords match with the index file and displays the top relevant records.

II.LITERATURE SURVEY

This section describes the various existing schemes.

A. Oruta

This approach allows auditing the shared data on the cloud.

This method uses three parties.

Cloud Server: used for storing data and signature.

User : can access or modify shared data

Third party audit (TPA): used to verify the integrity of the shared data.

In this system, which uses two users. Original user is the one who creates the data and stored in the cloud. Group user access or modify the shared data which are created by original user. Third party audit verifies the integrity of the shared data without fetching the whole document. Here the signature identity is kept separate from the TPA. The main disadvantage of this approach is not possible to distinguish the signature identity on each block which leads the identity privacy.[1]

B. Storing Shared data on the cloud via Security-Mediator

This method is used to generate verification metadata on shared data. The objective in this approach is data privacy, public verification, signing efficiency. Here the data owner sends the blinded message to security mediator, the security mediator appends signature and sends back to data owner, data user asks data to the cloud server by sending challenge, cloud server replies data by sending the response. The disadvantage is, it utilizes just a single security mediator which may fail or less dependable and adding more security mediator increases the system overhead.[2]

C. Attribute based encryption scheme

This approach uses the concept of public and personal domain. Public domain is used by the trusted authority (TA) for managing attributes. Personal domain is used by the record owner which acts as a TA. Record owner can encrypt the record using the access structure. By permitting users who have the attributes that fulfils the access structure can decrypt the data. The main drawback is the use of the single TA, which could lead to a single point of failure.

As a means of providing a solution for the above method is an ABE based system which uses multiple authorities. Thus, users belonging to the public domain can ascertain required attributes from the relevant attribute authority (AA) while users in the private domain ascertain the attributes from the record owner. The drawback in the ABE is that it requires to acquire no less than one attribute from each attribute authority for the proper encryption function. [3][4][5]

D.Patient centric access control scheme

In this method, data owner divides data into various numbers of files and then provides different encryption keys for each file. This paper uses the symmetric encryption scheme. Every time data owner requires to transfer keys to data users if wants to access data. This process makes the system inefficient in the cloud. Encryption keys are increases for every category. The disadvantage is transferring keys every time is a difficult task.[6]

III. PROPOSED SYSTEM

A. System model

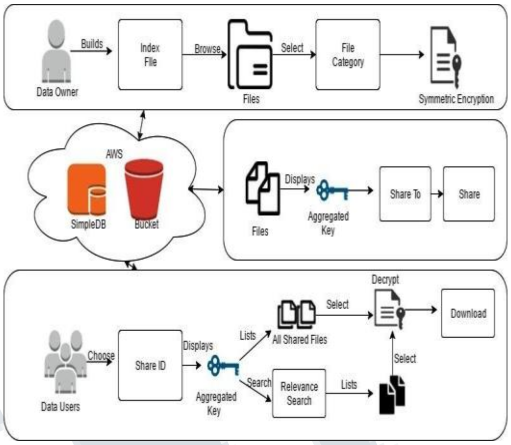

Fig 1. Proposed approach

Initially data user and data owner needs to register to the system. At the time of data owner registration it automatically generates unique master key and bucket name. Master key is used for the encryption process and bucket is used to store all files that are uploaded by the particular registered data owner. Once the registration completes, the one can login to the system for the further process.

As shown in the Fig 1. Data owner is one who uploads the data into the cloud i.e., patient. The system is automatically builds an index file and that is stored in a cloud owner bucket. Index file contains the binary data of the file type category. Data owner, by choosing the category of the file and the required record to upload it displays the symmetric encryption key and performs encryption before uploading the record to cloud storage i.e., Amazon S3 buckets. Hence all the records that are stored in cloud storage are secure. For every file that are uploaded on cloud generates keys.

Once the uploading records in cloud storage finishes, the data owner needs to share the records with data user. For this, data owner needs to choose the available member ID and the required records that are already uploaded on cloud storage to share. Then it displays the one aggregated key. Finally all the records that are shared is available the data user.

Data user is one who uses data, i.e., Doctor, Nurse, Insurance broker. Once the data user logged in to the system, it automatically lists all Patient ID's that are shared records. By choosing the patient ID, it displays the aggregated key and the list of all shared records. Data user can decrypt and download those by choosing records. And also data user can search for the particular encrypted records in the available list by choosing the record category. It displays the top relevance ranked records. Then decrypt and download those records.

IV. ALGORITHMS

A. Rijndael algorithm

Rijndael algorithm is the advanced encryption standard algorithm and an iterated block cipher. It uses the same key for both encryption and decryption. Rijndael algorithm supports the block sizes of 128 bits and key sizes of 128, 192, 256 bits. It is the best among the combination of security, performance, efficiency, ease of implementation and flexibility.

Initially to encrypt the data master key is used, which is given by the user. These master keys are usually small in size and easy to hack. Hence, to become complex key, salt data are added to the master key. In the proposed approach, complex key generates an initialization vector. The first block of the data is encrypted by using the initialization vector along with the master key. The remaining blocks are encrypted with the help of previous encrypted block along with master key.

***Steps*:**

Step 1: Byte Sub Transformation
Each byte of the block is replaced by its substitute in an S-box. Each byte is treated *independently* Single S-box is used for the entire state.

Step 2: Shift Row Transformation
Each row of the state is shifted cyclically a certain number of steps. The number a row is shifted can't be the same.

Step 3: Mix Column Transformation
State columns are treated as polynomials over GF(28). Each column is multiplied by modulo x4 + 1 by a fixed polynomial c(x) = `03` x3 + `01` x2 + `01`x + `02`

Step 4: Round Key Addition XOR round key with state.

B. Cosine similarity search algorithm

The distance between any two vector d1 and d2 captured by cosine of the angle X between them is called cosine similarity. The cosine similarity of two vectors is a discretionary numerical measure of how comparative two vectors are on a size of [0 to 1]. Here value 1 being that the vectors is indistinguishable.

The cosine similarity of two vectors d1 and d2 is defined as,

$$Cos(d1, d2) = \frac{dot(d1, d2)}{\| d1 \| \| d2 \|} \text{………(1)}$$

Where,

$$dot(d1, d2) = (d1[0] * d2[0]) + (d1[1] * d2[1]) + \ldots$$

$$\| d1 \| = \sqrt{(d1[0]^2) + d1[1]^2 + \ldots)}$$

This algorithm is used for the purpose of searching the multiple keywords in the system to filter out the relevant records. Queried keywords matches with the index file and top value are displayed in the system. Hence this achieved by listing top relevant records.

V. CONCLUSION

In the proposed system, secure record sharing onto the cloud by using symmetric key cryptosystem. Health records are encrypted before outsourcing. Hence, data security is achieved. Using this system, distribution of single aggregated key is sufficient for sharing a large number of records. It can also possible to search for the relevant records by matching with the index file by resulting the top relevant records.

IFERP connecting engineers... developing research

REFERENCES

[1] B. Wang, B. Li, and H. Li, "Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud," in IEEE Cloud, June 2012, pp. 295-302

[2] Dynamic Audit Services for Outsourced Storages in Clouds Yan Zhu, Member, IEEE, Gail-Joon Ahn, Senior Member, IEEE, Hongxin Hu, Member, IEEE, Stephen S. Yau, Fellow, IEEE, Ho G. An, and Chang-Jun Hu.

[3] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, “Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute Based Encryption,” IEEE Transactions on Parallel and Distributed Systems, vol. 24, no. 1, pp. 131–143, Jan. 2013.

[4] L. Ibraimi, M. Asim, and M. Petkovic, “Secure Management of Personal Health Records by Applying Attribute-Based Encryption,” in Proceedings of the 2009 6th International Workshop on Wearable Micro and Nano Technologies for Personalized Health, Jun. 2009, pp. 71–74.

[5] J. Bethencourt, A. Sahai, and B. Waters, “Ciphertext-Policy Attribute Based Encryption,” in Proceedings of the 2007 IEEE Symposium on Security and Privacy, May 2007, pp. 321–334.

[6] Kuo-Hsuan Huang, En-Chi Chang, and Shao-Jui Wang, “A Patient-Centric Access Control Scheme for Personal Health Records in the Cloud, ” Fourth International Conference on Networking and Distributed Computing, IEEE 2014.

IFERP connecting engineers... developing research