International Journal of Engineering & Technology, *7 (2.24) (2018) 311-315*

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Secure and Efficient Query Processing Using Randomized Encryption and De-duplication on the Cloud

Mohith Lalita Kumar Parvataneni, Eswara Sai Nath Adusumalli, Rajeshbabu.C

Computer Science Engineering, SRM University, Kattankulathur
**Corresponding Author Email: mahinewton007@gmail.com*

Abstract

Cloud computing [1] is the most upcoming area of the IT industry that helps the users to get rid of the hardware resources and complexity in storage and computational power. All the people started using cloud which led to many security concerns relating to the data confidentiality and integrity. This became a challenge to the widespread of the new cloud computing technology. Many measures are taken to improve the cloud security and then there came the concept of cryptography to upgrade the cloud security. Today, the main problem is to sustain data privacy [2] against unreliable cloud service providers and to provide correct query results to the authenticated users. Existing approach provide confidentiality using only one symmetric encryption algorithm which generates only one secret key to both encrypt and decrypt. But this is not so secured as the attacker can easily guess the algorithm and thereby find the key. In this journal, we proposed a randomized encryption technique in which the files are randomly encrypted by three strong algorithms AES, Triple DES and Blowfish [3] which improves the security and we are also implementing new techniques to improve security like key generation via OTP method, etc. In addition to this, we are also going to solve the storage issue of redundant or duplicate files consuming the cloud storage by using a file-level de-duplication technique which can eliminate the duplication of files uploaded into the cloud, thereby saving storage and reduce the need to buy extra storage.

Keywords: *Cloud security, Data Confidentiality, AES, Triple DES, Blowfish, De-duplication, Randomized Encryption.*

1. Introduction

In the present scenario, Cloud Computing is one of the latest emerging areas of the IT Infrastructure with most powerful network and storage facilities which are being used by almost all the users across the world. This provides a lot of storage for free of cost and other resources which may be helpful for the organizational activities of various small scale and large scale Industries and Companies.

Even though cloud is providing free storage and computational resources, there comes a security concern when the whole data is in the hands of cloud service providers. This lead to various security issues that are needed to be resolved to gain trust among the users. When the data is transferred between the users and individuals, there is a security risk that all the shared files are open to the attackers. So, in recent times, there came a new idea called Cryptography to strengthen the Cloud Security by solving the problems of Data Privacy and Integrity. Even till today, data confidentiality is a primary security issue which needs a lot of upgradation to prevent the loss of data and other computational resources. Data Confidentiality means that the file content should not be revealed out to the untrustworthy users. Data integrity means that the is not corrupted and accessed only by the users who are treated as trusted. For preserving this confidentiality and integrity, many cryptography techniques are put forward to save the data from unknown attacks using a concept called data encryption. Data encryption is a concept in which the data is secured using a key. Many different techniques are now been developed using this concept of encryption to improve the Cloud Security. Redundant storage of files is important storage issues of Cloud Computing which consume a lot of disk space. This can be eliminated by the concept of de-duplication that does not allow duplicate files to be stored repeatedly in the cloud.

2. Literature Survey

Cloud Computing [1] is a technology which is used for storage of data and various computations. Using cloud computing the user data need not be stored in his own personal computer or any external storage device rather can be stored in the cloud raising many security concerns or issues.

Though there are many techniques [2] of data security and privacy in cloud computing, data confidentiality and integrity are of major concern till date. Data security and privacy issues are seen in both the software and hardware of the cloud architecture model. Data security is of major concern and the total data will be in the hands of cloud service providers using cloud computing service model.

In the cloud environment, the total organizational data is processed in plain text and is stored under the surveillance of cloud providers arising cloud security issues.

All these security issues have led to a new concept in cloud computing called Cryptography to enhance the Cloud Security using several encryption algorithms both Symmetric and Asymmetric algorithms [3]. Symmetric encryption algorithm uses only a single key namely private key for encryption whereas Asymmetric algorithms use two keys namely public and private. Symmetric encryption is very rapid than Asymmetric encryption. Most frequently used Symmetric encryption techniques include DES and AES. DES is not so secure because of its key size of 56 bits, hence there came to a better enhancement to DES in the form of Triple DES with a key length of 168 bits making it 3 times more robust than DES.

Triple DES is well built but time-consuming. So, there is a necessity for a better encryption algorithm to make the encryption fast and efficient, then there comes AES [4] with a better length of the key varying from 128, 192 and 256 bits. AES operation is based on rounds and no of rounds are based on key length. Nowadays Blowfish [5] encryption algorithm is used for more security with the key length varying from 32 to 448 bits. There are no cryptanalytic attacks raised against AES and Blowfish till date. The most common encryption algorithms used today are AES, Triple DES, and Blowfish.

The other main problem in the cloud is the redundant storage of data in the cloud. There are plenty of different methodologies developed to detect the duplicate files in the cloud. This is done using the concept of de-duplication [6]. There are many de-duplication techniques in which the common ones are File-level and block level. In File-level de-duplication, hash values are compared to eliminate the duplicate files using SHA(secure hashing algorithm).

3. Cloud Security Threats

A. Improper credential management

When the information of all the users of the cloud is not properly stored [13], the attackers can acts as the legitimate users, enter into the system and access, modify or delete files which may cause potential harm to the organization.

B. Account Hijacking

It is a technique through which the malicious attackers enter into the system finding the system vulnerabilities[14] and keep an eye on the activities of the organization, manipulate the data and other illegal activities. He also can steal the credentials which affect the data confidentiality and integrity.

C. Insider attacks

These type of attacks come into picture when the data is solely in the hands of the cloud service providers. Any bad system admin can cause harm to our sensitive information stored in the cloud. This is one of the greatest security risks which cannot be predicted and data breach may also happen due to this.

4. Proposed System

We propose an encryption scheme where the files are randomly encrypted using three different encryption algorithms. This randomized encryption [17] scheme improves the security with an add-on of OTP based private key generation technique and we also include a de-duplication technique which does not support the redundancy in files uploaded into the cloud.

Algorithms Used

Triple DES:

DES encryption algorithm [7]is a weak algorithm with the key length of only 56 bits but is popular. So it is modified as Triple DES[3] which is three times stronger and faster than DES with the key length of 3*56=168 bits making it strong and intense compared to DES. It uses three keys Key1,key2,key3, first to encrypt and then the encrypted data is decrypted using the second key key2 and then the decrypted data is again encrypted using the third key key3.Hence it is referred as an encrypt-decrypt-encrypt process. A user first decrypt using key3, then encrypt with key2, and then decrypt using key1.

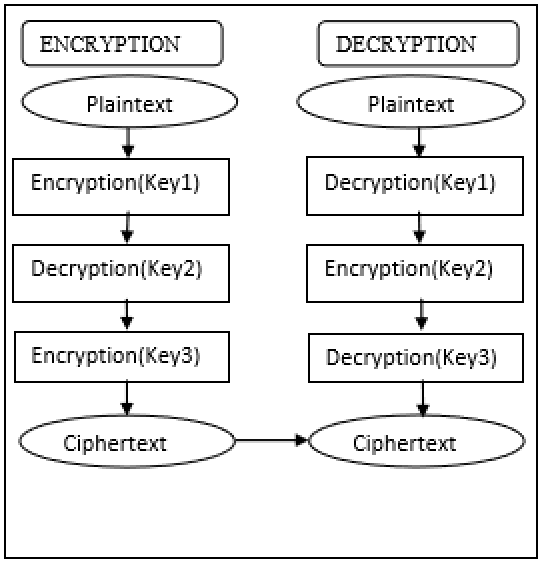

Fig. 1 [3]: Triple DES Encryption

Encryption in TDES:
Ciphertext=encryption(key3)(decryption(key2)(encryption(key1)(plaintext)))
Decryption in TDES:
Plaintext=Decryption(key1)(encryption(key2)(decryption(key3)(ciphertext()))

AES:

AES is a symmetric encryption algorithm [4] which was put forward by NIST. It is stronger and rapid than DES and Triple DES. It is so robust that nocryptanalytic attack against it was recorded. It is the most frequently used technique adopted by almost all cloud users around the world. It is safer compared to DES and Triple DES. In AES, a 128-bit or 16-byte plaintext block size is used for encryption. The length of the key can be 128/192/256 bits. The input is given as one block of 128 bits. This is interpreted as a matrix of 4 × 4-byte squares. This block is copied to the status table, which is changed at each step of encryption or decryption. After the last step, the status is copied to an output matrix. The data encryption consists of N turns, where the number of turns depends on the key length.

Operation of AES

The schematic of AES structure is given in the following illustration –

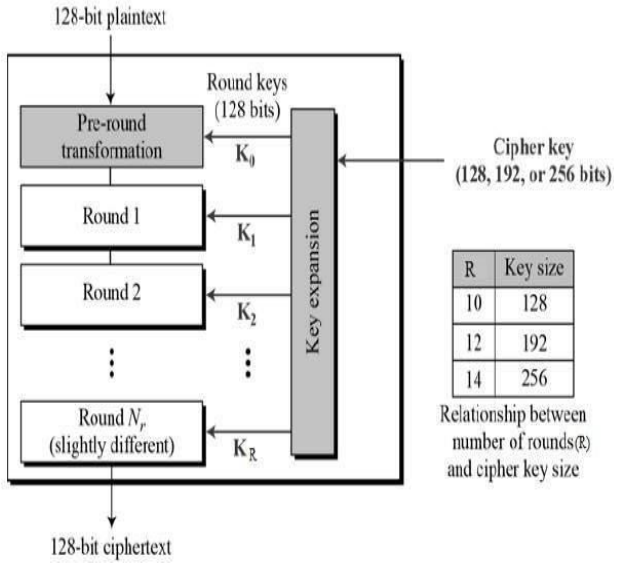

Fig. 2 [17]: AES Encryption

There are four transformation functions
Substitute bytes*:*In this, an S-box is used for performing a block-wise byte-by-byte substitution.
ShiftRows*:* A normal mathematical operation to arrange the rows.
MixColumns: A substitution that uses an arithmetic function and transforms4 bytes of every column.
AddRoundKey*:* A simple XOR is done to the current block of the enlarged key bit-wise.

Algorithm : AES Encryption

```
CipherText(Input[16], Output[16], word[0….43]){
BlockToAState (Input,X)
X ← AddRoundKey (X,word[0….3])
for (round = 1 to 10)
{
X ← SubBytes (X)   X ← ShiftRows (X)
If ( round ≠ 10)X ← MixColumns (X)
X ← AddRoundKey (X, word[4×round, 4×round+3])
}
StateToABlock (X, Output)  }
```

Each round comprises of four sub-processes. The first round procedure is shown in the below figure −

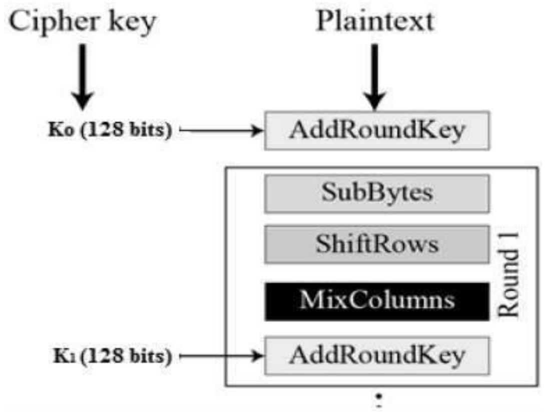

Fig. 3 [22]

Blowfish:

Blowfish is the latest symmetric encryption algorithm[20] which is upcoming and being accepted nowadays. It is more secured algorithm compared to DES, 3 DES, AES etc, It is fast and strong encryption algorithm because it has not been cracked to date. In Blowfish, the data input is of 64 bits with its key length varying from 32- 448 bits. There are total 16 rounds of encryption performed in Blowfish.

Encryption:
Blowfish encryption algorithm is split into 2 parts.
1. Key Expansion
2. Data encryption

Key Expansion:
In this process, the length of the key which is 48 bits is changed into 4168 bytes.

Data encryption:
The data encryption procedure implements some simple steps using simple mathematical function 16 times.
The encryption process of this is elucidated using the figure below:

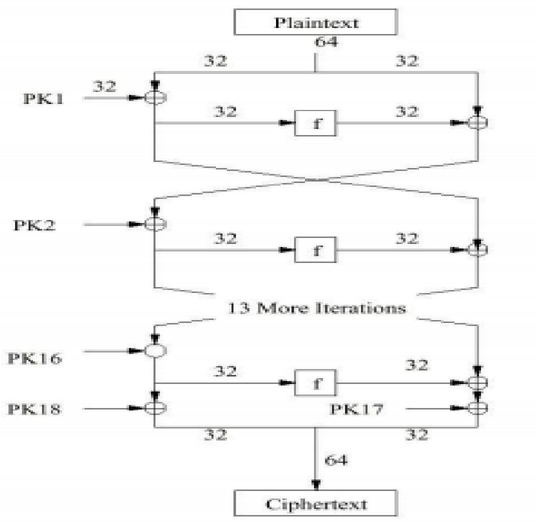

Fig. 4 [7]: Blowfish Encryption

Algorithm:Blowfish Encryption

```
            Y is divided into two halves of 32-bits: Y1, Y2
                for ( j = 1to 16 )
                      y1 = Y1 XOR Pi
                      y2 = F(Y1) XOR y2
                      Swap Y1 and y2
                      Swap Y1 and y2 (Undo the last swap.)
                      y2 = y2 XOR P17
               y1 = y1 XOR P18
              Recombine y1 and y2
```

Private Key Generation Via OTP Method:

This is the newly introduced technique to enhance the security of the cloud using time schedulers inserted into the web application using simple SQL queries. These queries are used for timely generation of the secret key(owner key). To clearly elucidate this technique, when the user requests the secret key from the owner, the owner must respond with the secret key to the authorized users. Even though the intruder tries to get the possession of the secret key to break the encryption, as in the proposed technique key will be available only for a particular amount of time. So this timely generation of the secret key using OTP improves the security by restricting the key to a particular amount of time as like the bank transactions are being secured by using OTP.

De-Duplication Technique:

Using this technique, duplicate or redundant files are not allowed to be uploaded to the cloud. In this technique, we use a simple hashing algorithm which allocates the index values to the files and on a comparison, the duplicate files are detected i.e files with same index value are not allowed into the cloud.

When the owner uploads a file the de-duplication generates a hash value as an index, if the index is not found in the already created index table, the file is non-redundant else duplication is detected and the file is not uploaded.

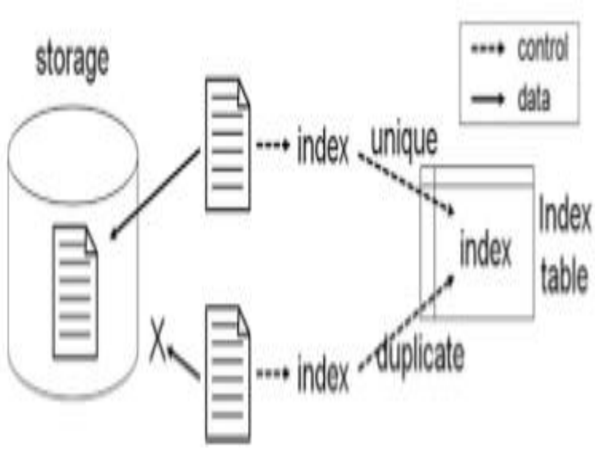

Fig. 5 [6]: File-level de-duplication

Algorithm used: SHA-512(**Secure Hash Algorithm**)

Description:

This is the most [9] used hash function nowadays. The hash value of this algorithm is of 512 bits and the data input is of 1024 bits. The no of computational steps is 80.

This algorithm takes the input with a maximum length of 2^128 bits and produces a hash value of 512 bits as the output. The processing of this algorithm takes place in following steps:

- Add the padding bits
- Append length
- Initialize Hash buffer
- Process message in 1024-bit blocks
- After all the bits are processed, generate 512-bit hash value

5. Overall Implementation

The proposed system is an extension of the existing cloud security architecture which uses cryptography. In this system, when a data owner uploads the files into the cloud, the files get randomly encrypted using three strong encryption algorithms AES, Triple DES, and Blowfish which are resistant to crypt-analytic attacks. This randomized encryption does not allow the intruder to reuse the key to break other encrypted files. The key generation is the most important part of encryption, to improve the efficiency of key generation we implement timely access of private key via OTP method. While the owner uploads the file, a duplication detect is implemented to find the redundant files in the cloud using the technique of hashing(by comparing the hash values). This saves the cloud space by blocking the redundant files.

Now, the data which is encrypted is uploaded to the cloud. When someone needs to download a file, he/she must request the cloud key from the cloud admin which is generally the location key and private key from the owner which is the encrypted key. If the user is authorized, he gets both the private key and the cloud key and thereby he is able to download the needed file from the cloud.

Thus, our project mainly focuses on enhancing the cloud security and storage efficiency.

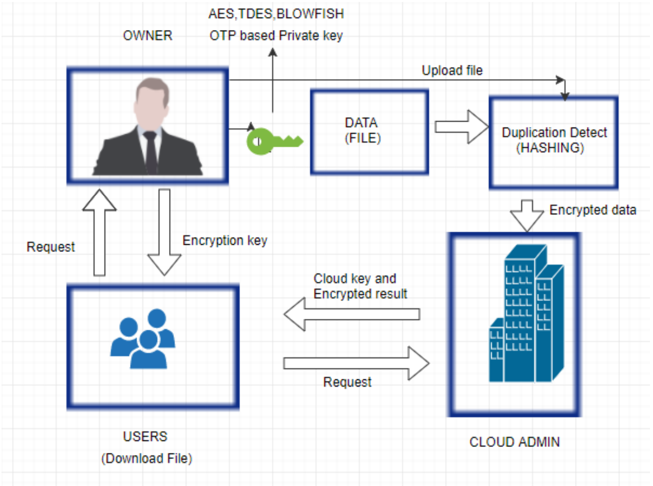

Fig. 6: Proposed Architecture for Randomized encryption and de-duplication

All the above enhancement of the project is done mainly in five modules. They are

- User Interface Design
- Private key generation
- Duplication detect
- Encrypt and store
- Download

User Interface Design:

All the login and registration pages for the users, Owners, and Admin are created.

Private key generation:

In this private key is created depending on the time schedules using simple SQL queries via OTP method.

Duplication Detect:

In this duplicate or redundant files are detected during the owner file upload time.

Encrypt and Store:
The data is encrypted with the private key generated by OTP method and is stored in the cloud.

Download:
The user can download the file by requesting the private key from the owner and cloud key from the cloud admin.

6. Acknowledgement

It is a great honor to express our profound and sincere gratitude to our Computer Science Engineering department, SRM Institute of science and technology, Kattankulathur, for encouraging and giving us support and encouragement by providing good environment to carry out this research work and to gain valuable experiences.

7. Conclusion and Future Enhancement

Cloud Computing has many advantages in terms of storage and computational resources but, cloud security is always a major concern in the Cloud environment. In our system, we tried to enhance the cloud security by employing three strong encryption algorithms Triple DES, AES and Blowfish to protect the files from malicious attackers using a randomized encryption technique. No cryptanalytic attacks were found on AES and Blowfish till date making the encryption stronger and efficient. We also solved the problem of redundant file storage via file level de-duplication.
We can include MD5 algorithm and digital signatures to improve the security by providing data authentication using the concept of hashing and block level de-duplication makes the file storage more effective .

References

[1] Ronald Kurtz and Russell Dean Vines "Cloud Security(A Comprehensive Guide to Secure Cloud Computing)" WILEY-INDIA

[2] K.Hashizume, D.G.Rosado, E.B.Fernandez, "An analysis of Security issues for cloud computing", Journal Internet Services and Applications, Vol.4,2013,pp.1-13.

[3] MD Asif Mushtaque, Harsh Dhiman, and Shahnawaz Hussain "Evaluation of DES, TDES, AES, Blowfish and Two fish Algorithm: Based on Space Complexity" International Journal of Engineering Research &Technology(IJERT).

[4] Daemen, J., and Rijmen, V. "Rijndael: The Advanced Encryption Standard." Dr.Dobb's Journal, March 2001.

[5] Gurujeevan Singh, Ashwani Kumar and K.S. Sandha, "A Study of New Trends in Blowfish Algorithm", International Journal of Engineering Research and Applications, Vol.1, Issue 2,pp.321-326.

[6] Kim, D.; Song, S.; Choi, B,-Y."Data Deduplication for Data Optimization for Storage and Network Systems"http://www.springer.com/978-3-319-42278-7

[7] Mitali and Vijay Kumar "A Survey on Various Cryptography Techniques" International Journal of Emerging Trends & Technology in Computer Science(IJETTCS)

[8] Ayesha M.Talha and Ibrahim Kamel "Facilitating Secure and Efficient Spatial Query Processing on the cloud" IEEE

[9] William Stallings "Cryptography and Network Security(Principles and Practice)" Pearson Education sixth Edition.

[10] E.Surya and C.Divya,"A Survey on Symmetric Key Encryption Algorithms", International Journal of Computer Science &Communication Networks, Vol.2(4),475-477.

[11] Singh, S Preet, and Maini, Raman "Comparision of Data Encryption Algorithms", International Journal of Computer science and Communication, vol.2, No.1, January-June 2011, pp.125-127.A.

[12] Monika Agarwal, Pradeep Mishra, " A Comparative Survey on Symmetric Key Encryption Techniques", International Journal of Computer Science and Engineering(IJCSE), Vol.4 No.05 May 2012, PP877-882.

[13] Cloud Security Alliance. Top Threats to Cloud Computing, Cloud Security Alliance, 2010.

[14] https://googleweblight.com/i?u=https://www.incapsula.com/blog/top-10-cloud-security-concerns.html&hl=en-IN

[15] Atul karate "Cryptography and Network Security", Tata McGraw-Hill Companies,2008.

[16] https://blog.demofox.org/2012/09/15/cryptography-101-encryption-symmetric-keys/

[17] Miss Shakeeba S.Khan and Prof. Ms.R.R.Tuteja "Cloud Security Using Multilevel Encryption Algorithms" International Journal of Advanced Research in Computer and Communication Engineering vol. 5, Issue 1, January 2016.

[18] https://www.tutorialspoint.com/cryptography/triple_des.html

[19] Tingyuan Nie, and Teng Zhang,"A Study of DES and Blowfish Encryption Algorithm", IEEE, 2009.

[20] Bruce Schneier. "The Blowfish Encryption Algorithm Retrieved", October 25, 2008.

[21] Hui Cui and Robert H. Deng "Attribute-Based Storage Supporting Secure Deduplication of Encrypted Data in Cloud" IEEE

[22] https://www.tutorialspoint.com/cryptography/advanced_encryption_standard.htm

[23] http://www.tutorialspoint.com/cryptography/cryptography_quick_guide.htm

[24] S.V.Manikanthan and K.Baskaran “Low Cost VLSI Design Implementation of Sorting Network for ACSFD in Wireless Sensor Network”, CiiT International Journal of Programmable Device Circuits and Systems,Print: ISSN 0974 – 973X & Online: ISSN 0974 – 9624, Issue : November 2011, PDCS112011008.

[25] T. Padmapriya and V.Saminadan, “Handoff Decision for Multi-user Multiclass Traffic in MIMO-LTE-A Networks”, 2nd International Conference on Intelligent Computing, Communication & Convergence (ICCC-2016) – Elsevier - Procedia of Computer Science, vol. 92, pp: 410-417, August 2016.