International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 4, Special Issue 3 (March '2018)
Proceeding of - National Conference on Recent Trends in Information Technology and Communication,
Conference held at Velammal Engineering College,Chennai,India.

CONTENT BASED DUPLICATION AVOIDANCE IN CLOUD USING CP-ABE

Devi K,
Assistant Professor,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Renuka Devi V,
Student ,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Ramyha V,
Student,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Pandeeswari.R,
Student ,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Abstract: Data de-duplication[9] often called intelligent compression or single-instance storage is a process that eliminates redundant copies of data and reduces storage overhead. Data de-duplication techniques ensure that only one unique instance of data is retained on cloud. In this paper, we mainly focus on avoiding duplication in cloud using CP-ABE(cipher text policy -attribute based encryption)and ensure security. CP-ABE[4][8] ensures the high quality randomness in cryptographic key generation with an additional zero-knowledge proof (POK) on the shared cryptographic key and cipher text irrespective of using ABE(Attribute Based Encryption)[1][10]. CP-ABE stores the keys and cipher texts as a pair in the public cloud. Files are listed in the public cloud in order of relevancy to achieve secure de-duplication. File duplication can be identified based on the file content and size. The user can view file information along with key in hidden format. TTP provides key access permission to the requested user, which provides solution for balancing confidentiality and efficiency in performing de-duplication.

Keywords: de-duplication , CP-ABE , zero-knowledge proof , de-duplication.

I.INTRODUCTION

Cloud computing, a convenient way of accessing services, resources and applications over the Internet, shifts the focus of industries and organizations away from the deployment and day-to-day running of their IT facilities by providing an on-demand, self-service, and pay-as-you-go business model. It is, therefore, unsurprising that
cloud computing has continued to increase in popularity in recent times. While cloud computing provides various benefits to users, there are underlying security and privacy risks. One of the challenging paradigm of cloud computing is duplication. Duplication in cloud , a mechanism where data's in cloud been replicated which increases hardware and backup costs. Duplication may cause data traffic and dimnishes storage efficiency. De-duplication [9], a method,effectively eliminates duplication in cloud which also requires a encryption technique to save storage space and network bandwidth by eliminating redundant copies of the encrypted data stored in the cloud. We, have attempted to implement CP-ABE (cipher text-attribute based encryption)[4][8] which can achieve secure de-duplicationwith a random key for each unique file irrespective of Attribute Based Encryption(ABE)[1].

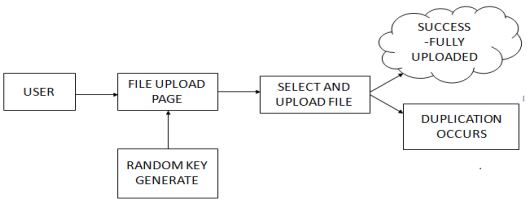
II.OVERVIEW OF OUR WORK

In this paper, we present a cipher text-policy attribute-based encryption (CP-ABE) system which employs supports secure de-duplication. Our main contributions can be summarized as,Firstly, our system acquires the user credentials for data confidentiality and security. The registered user of our system can login to upload or get access through the files using the confidential data's.Secondly, after a victorious registration and login the user may upload a file ,mean-while the random key has been generated for the file being uploaded by the user.Thirdly, we put forth a methodology to modify a plain text into an encrypted cipher text using a symmetric-key algorithm(AES- advanced encryption standard)[5][6][7].At last our system will perceive duplication by examining the contents and size of the file being uploaded irrespective of its name.

III.FEATURES AND MODULE DESCRIPTION

A. AUTHENTICATION AND AUTHORIZATION

The first step before file uploading for a data owner is to register their credentials in-order to achieve high confidentiality of data. Once the registration is complete the data possessor can login, meanwhile the unique key for that file has been generated, the data holder can pick up a file to upload.

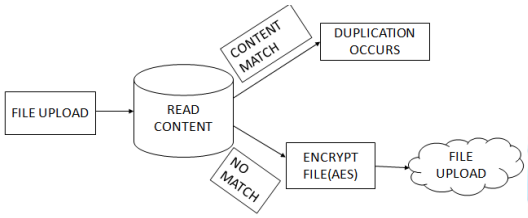


B. ENCRYPTION AND DUPLICATE DETECTION

After the file has been uploaded, our CP-ABE system uses an AES(advanced encryption standard) [5][6][7]which has been briefed on section 4,to encrypt the uploaded data with a symmetric key. After encryption the file has been stored in the cloud which will be monitored by admin . The second user when tries to upload a file our system will examine the file by verifying its CONTENT and SIZE irrespective of names since the files with same names may have different contents which doesn't really a duplication.

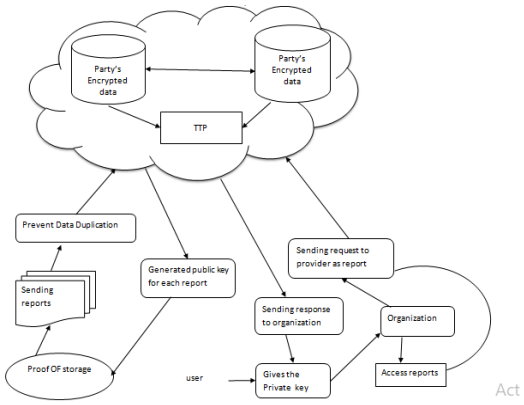
C. FILE ACCESS WITH A KEY REQUEST

Once the data holder is prompted with a duplication message, the user had a choice of requesting a key access to download a file and accessing it. Admin will decide whether to accept or decline the request. Once the request has been accepted the key and the download option for a file will be noticeable by the user and can access the file. The another unique idea of our work comes here, once the request for a particular has been accepted the random key for the file will get recast.



IV. SYSTEM ARCHITECTURE

The overall schematic system diagram represents the data flow from the user login and eliminating duplicate file and till retrieving the required files. The main purpose of CP-ABE is to protect the sharing of user's data by trusted third party without their knowledge which can be achieved by sharing of files with user's private attributes.



V. ALGORITHMIC TECHNIQUES

A) Random key generation:

1. **Authentication**

Authenticate(credentials):
If credentials match:
proceed to file upload
File Upload(file)
Else if a new user:
proceed to register page
Authenticate(credentials)
Else
proceed to error page

2. **File uploading:**

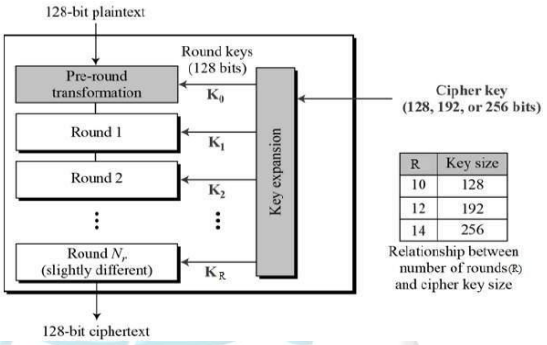
File Upload(file):
While(File Upload(file)):
If Validate(file!=Existing-files):
key=Random Key Generator()
Encrypt(file, key)
successfully upload a file.
Else:
error duplicate entry.
Access_File(file)

B) Encryption Technique

AES:

AES is a Symmetric key symmetric block cipher methodology with 128-bit data and 128/192/256-bit keys to based on 'substitution–permutation network'. AES treats plain texts as a bytes rather than treating them as a bits.

AES SCHEMATIC STRUCTURE



Encryption

```
Cipher(byte in[4*Nb], byte out[4*Nb], word w[Nb*(Nr+1)])
begin
byte state[4,Nb]  state = in
AddRoundKey(state, w[0, Nb-1])
for round = 1 step 1 to Nr–1
SubBytes(state)
ShiftRows(state)
MixColumns(state)
AddRoundKey(state, w[round*Nb, (round+1)*Nb-1])
end for
SubBytes(state)
ShiftRows(state)
AddRoundKey(state, w[Nr*Nb, (Nr+1)*Nb-1])
out = state
end
```

Decryption

```
InvCipher(byte in[4*Nb] , byte out[4*Nb], word w[Nb*(Nr+1)])
begin
byte state[4,Nb]
state = in
AddRoundKey(state, w[Nr*Nb, (Nr+1)*Nb-1])
for round = Nr-1 step -1 downto 1 InvShiftRows(state)
InvSubBytes(state)
AddRoundKey(state, w[round*Nb, (round+1)*Nb-1])
InvMixColumns(state)
end for
InvShiftRows(state)
```

InvSubBytes(state)
AddRoundKey(state, w[0, Nb-1])
out = state
end
Rounds N r = 6 + max{N b , N k }
N b = 32-bit words in the block
N k = 32-bit words in key
AES-128: 10
AES-192: 12
AES-256: 14

C) Cipher text- Policy Attribute Based Encryption (CP-ABE) Scheme

CP-ABE (Cipher text-Policy Attribute-Based Encryption) , user's private key has a connection with some set of attributes which enables the data holders to share their encrypted data using cloud storage with authorized users by hiding the control policies.

SYSTEM PERSPECTIVE

1: procedure DEDUPPROVE(αs, kc, αc, {c1, . . . , cn}, I,Q)
2: c ← 0, t ← ∅, ζ ← 1, l ← 1
3: while ζ ≤ n do
4: δ ← 0
5: while ζ <ıjl do
6: δ ← δ + c_ , ζ ← ζ + 1
7: pop the first element in Q
8: t ← t ∪ {fkc (iklikvi) + αcαsδ}
9: c ← c + c_
10: l ← l + 1, ζ ← ζ + 1
11: return c, t

D) Recasting of Random Key:

Recast(file):
If file request accepted:
key=Random Key Generator()
file=key

VI. ADVANTAGES

- Efficient user revocation for cloud storage system.
- To solve data privacy issues.
- Reduces Data traffic.

VII. CONCLUSION

Implemented with J2EE, MYSQL and an IDE Eclipse. Using CP-ABE scheme the encrypted data can be kept confidential even if the server storage is untrusted whereas ABE fails to provide secured de-duplication than our system.Also using AES to encrypt and decrypt a file may make our system more robust against hacking. The unique method which takes more advantage is avoiding duplication based on the content and size which reduces network traffic and will be more efficient. Even though the file with same name can be uploaded if it contains different content.

VIII. FUTURE ENHANCEMENT

Additionally, we outsource operations with high computation cost to E-CSP and D-CSP to reduce the user's computation burdens.

IX.REFERENCES

[1]. Hui Cui, Robert H. Deng, Yingjiu Li, and GuoweiWu,"Attribute-Based Storage Supporting Secure Deduplication of Encrypted Data in Cloud",23 January 2017.
[2]. K. R. Choo, J. Domingo-Ferrer, and L. Zhang, "Cloud cryptography: Theory, practice and future research directions," Future Generation Comp. Syst., vol. 62, pp. 51–53, 2016.
[3]. K. R. Choo, M. Herman, M. Iorga, and B. Martini, "Cloud forensics: State-of-the-art and future directions," Digital Investigation,vol. 18, pp. 77–78, 2016.
[4]. Guangbo Wang and Jianhua Wang " Research on Ciphertext-Policy Attribute-Based Encryption with Attribute Level User Revocation in Cloud Storage",23 May 2017.
[5]. R.Manjushaand R.Ramachandran"Comparative study of attribute based encryption techniques in cloud computing,13 November 2014 .
[6]. ArwaAlrawai ,Abdulrahman Alhothaily, ChunqiangHu,Xiaoshuang Xing, Xiuzhen "An Attribute-Based Encryption Scheme to Secure Fog Communications",23 May 2017.
[7]. Daniyal M. Alghazzawi ,Syed Hamid Hasan, Mohammed Salim Trigui "Advanced Encryption Standard - Cryptanalysis research"12 June 2014 .
[8]. Kaitai Liang "A Cipher text-Policy Attribute-Based Proxy Re-Encryption with Chosen-Cipher text Security"2009.
[9]. Mark W. Storer Kevin Greenan "Secure Data Deduplication,2008.
[10]. Allison Lewk "Decentralizing Attribute-Based Encryption Search Scheme over Encrypted Cloud Data",2015.