ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

*Visit: www.ijirset.com*

**Vol. 7, Special Issue 4, April 2018**

# A Secure Cloud Media Center Application with Secure Deduplication and Anticollusion Attack

Selvi.K[1], Deepika.J[2] , Rohini.S [3]

UG Student, Department of CSE, T.J.S Engineering College, Chennai, India [1,2]

Assistant Professor, Department of CSE, T.J.S Engineering College, Chennai, India.[3]

**ABSTRACT:** In cloud computing, users can share data among group members. It might cause to the collusion attack for an unsecured cloud. Data deduplication is one of the techniques which used to solve the repetition of data. Our proposed system prevents the replication of files and multimedia. To prevent the unauthorized use of data accessing and create duplicate data on cloud to encrypt the data before stored on cloud server. CloudMe is proposed for cloud storage. All files of data owners are encrypted using AES algorithm and stored in real cloud.

## I. INTRODUCTION

Cloud Computing is an innovative technology that is revolutionizing the way we do computing. The key concept of cloud computing is that you don't buy the hardware, or even the software, you need anymore, rather you rent some computational power, storage, databases, and any other resource you need by a provider according to a pay-as-you-go model, making your investment smaller and oriented to operations rather than to assets acquisition.
**Data owner:** This is a client who owns data, and wishes to upload it into the cloud storage to save costs. A data owner encrypts the data and outsources it to the cloud storage with its index information, that is, a tag. If a data owner uploads data that do not already exist in the cloud storage, he is called an initial uploader; if the data already exist, called a subsequent uploader since this implies that other owners may have uploaded the same data previously, he is called a subsequent uploader. Hereafter, we refer to a set of data owners who share the same data in the cloud storage as an ownership group.

**Cloud service provider:** This is an entity that provides cloud storage services. It consists of a cloud server and cloud storage. The cloud server deduplicates the outsourced data from users if necessary and stores the deduplicated data in the cloud storage. The cloud server maintains ownership lists for stored data, which are composed of a tag for the stored data and the identities of its owners. The cloud server controls access to the stored data based on the ownership lists and manages (e.g., issues, revokes, and updates) group keys for each ownership group as a group key authority. The cloud server is assumed to be honest-but-curious.

## II. PROBLEM DEFINITION / EXISTING SYSTEM

In existence private key distribution is based on the secure communication channel, In this case, which user have private key can share data unfortunately revoked user also can share data. Revoked user means who have changed their membership. Therefore, secure communication channel is a strong assumption but difficult to use.
In existing system, Cloud storage is not efficiently utilized. Even in google drive duplication of files is possible. Replica of data is possible. In Existing technique disk failure rate may be very high so data may be lossed. We may not have any copy of the data, it leads to data fail in Cloud computing which leads to losses the original data. In existing

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

*Visit: www.ijirset.com*

**Vol. 7, Special Issue 4, April 2018**

the file will not be stores with minimum replication. Cloud storage consumption is very high with high cost. Also our literature survey states deduplication attempt towards multimedia files is not been initiated.

## III. PROPOSED SYSTEM

- The users can securely obtain their private keys from group manager.
- User send request to group manager for access the wanted group, at that time our system provide individual secure key to user without activation.
- Then group manager see the requests and activate the keys after confirm them.
- After user's private key gets activation, then only user can access the group.
- In our proposed system the group manager performs the below tasks when an new user joins the group or a user has left the particular group,
  - *Update the whole user name list.*
  - *Generate a secure key and encrypt the key without activation and send to the updated user list.*
  - *Update the rights in the cloud server.*
- We proposed public cloud named CloudMe for data storage.
- Group manager makes sure that the revoked users cannot access the file if they conspire with untrusted cloud.
- Media files like image and videos were eliminated in the recent surveys, we proposed to develop a unique architecture to prevent collusion attack and avoid replica of text, image and videos.

## IV. MODULES AND ITS DESCRIPTION

- Privacy Preserving
- KNN
- Deduplication
- SVC
- Key distribution & Access control
- Collusion attack
- Encryption (AES)
- Cloud Storage.

**PRIVACY PRESERVING:**

In this paper we address the issue of **privacy preserving** data mining. Specifically, we consider a scenario in which two parties owning confidential databases wish to run a data mining algorithm on the union of their databases, without revealing any unnecessary information. The **privacy preserving data mining** techniques are classified based on distortion, association rule, hide association rule, taxonomy, clustering, associative classification, outsourced **data mining**, distributed, and k-anonymity, where their notable advantages and disadvantages are emphasized.

**KNN**:

In pattern recognition, the **k-nearest neighbor's algorithm**(KNN) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space.KNN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The KNN algorithm is among the simplest of all machine learning algorithms.

**DEDUPLICATION:**

Deduplication is a process to improve data quality by removing redundant or repetitive information from data in storage to improve storage utilization, simplify **ETL,** and optimize data transfers.Organizations often do not have visibility into the sources or causes of redundant data. Thus they have no way of knowing how much redundant data is

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

*Visit: www.ijirset.com*

**Vol. 7, Special Issue 4, April 2018**

costing them. For example, a retailer can waste a lot of money sending multiple copies of the same catalog or campaign to one prospective customer. By deduplicating the data ahead of time the company can prevent waste.

**SVC:**

**Scalable Video Coding** (**SVC**) is the name for the Annex G extension of the H.264/MPEG-4 AVC video compression standard. SVC standardizes the encoding of a high-quality video bit stream that also contains one or more subset bit streams. A subset video bit stream is derived by dropping packets from the larger video to reduce the bandwidth required for the subset bit stream. The subset bitstream can represent a lower spatial resolution (smaller screen), lower temporal resolution (lower frame rate), or lower quality video signal. H.264/MPEG-4 AVC was developed jointly by ITU-T and ISO/IEC JTC 1. These two groups created the Joint Video Team (JVT) to develop the H.264/MPEG-4 AVC standard.

**KEY DISTRIBUTION & ACCESS CONTROL:**

- Group manager takes charge of system parameters generation, user registration, and user revocation. In the practical applications, the group manager usually is the leader of the group. Therefore, we assume that the group manager is fully trusted by the other parties.
- Group members (users) are a set of registered users that will store their own data into the cloud and share them with others. In the scheme, the group membership is dynamically changed, due to the new user registration and user revocation.
- Once the user is revoked, The group manager creates the new encryption key for the specific group and transmits in an encrypted format. Second the group manger updates the whole data list in the cloud server. Third the group manages updates the user list and activates the key for access.

**COLLUSION ATTACK:**

- The user leaving a particular group are termed as revoked users. The revoked users can not be able to get the original data files once they are revoked even if they conspire with the untrusted cloud. Thus our proposed system detects the revoked users and protects the data confidentiality and privacy.

**ADVANCED ENCRYPTION STANDARD** (**AES**):

The **Advanced Encryption Standard** (**AES**), also known as **Rijndael** (its original name), is a specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001.

AES is a subset of the Rijndael cipherdeveloped by two Belgian cryptographers, Joan Daemen and Vincent Rijmen, who submitted a proposal to NIST during the AES selection process. Rijndael is a family of ciphers with different key and block sizes.

**AES :PSEUDOCODE**

```
Cipher(byte in[16], byte out[16], key_arrayround_key[Nr+1])
begin
byte state[16];
state = in;
AddRoundKey(state, round_key[0]);
for i = 1 to Nr-1 stepsize 1 do
SubBytes(state);
ShiftRows(state);
MixColumns(state);
AddRoundKey(state, round_key[i]);
end for
SubBytes(state);
```

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

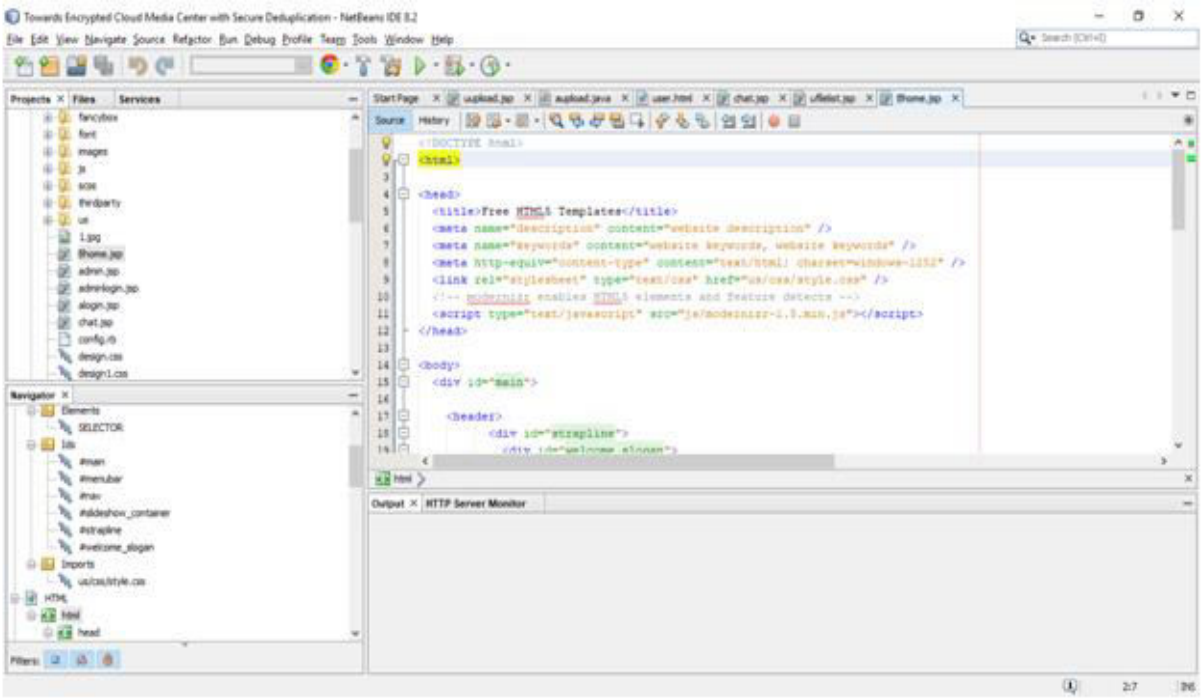
International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

*Visit: www.ijirset.com*

**Vol. 7, Special Issue 4, April 2018**

ShiftRows(state);
AddRoundKey(state, round_key[Nr]);
End



**CLOUD STORAGE:**

The group user can upload the files in real cloud server named cloudMe. Duplication of files are checked and the files is been uploaded in the cloud server. To get a file, the user needs to send a request to the cloud server. The cloud server will also check the user's identity before issuing the corresponding file to the user. During file access the user key has to match by the group manager and the requested file can be downloaded by the group users.

## REFERENCES

1. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Commun. ACM, vol. 53, no. 4, pp. 50–58, Apr. 2010.
2. S. Kamara and
3. K. Lauter, "Cryptographic cloud storage," in Proc. Int. Conf. Financial Cryptography Data Security, Jan. 2010, pp. 136–149.