A Secure Cloud Media Center Application With Secure Deduplication And Anticollusion Attack

V.Aravindan[1] , S.Dinesh[2], G.Rajasekaran[3]
[1, 2 , 3]Dept of Computer Science
[1, 2, 3] JEPPIAAR SRR Engineering College, Chennai

Abstract- *In cloud computing, users can share data among group members with the characters of less maintenance and little management cost. Sharing data must have security guarantees, if they are out sourced. Sharing data while providing privacy preserving is still a challenging problem, when change of the membership. It might cause to the collusion attack for an unsecured cloud. For existing technique, security of key distribution is based on the secure communication channel, however, to have such channel is a strong assumption and is difficult for practice. We propose a secure data sharing scheme for dynamic users. Key distribution done without any secure communication channels and the user can get the individual key from group manager. Data deduplication is one of the techniques which used to solve the repetition of data. Our proposed system prevents the replication of files and media file like images, videos. The deduplication techniques are generally used in the cloud server for reducing the space of the server. To prevent the unauthorized use of data accessing and create duplicate data on cloud the encryption technique to encrypt the data before stored on cloud server. Cloud me is proposed for cloud storage. All files of data owners are encrypted using AES algorithm and stored in real cloud. Thus we present a secure system architecture design as our initial effort towards this direction, which bridges together the advancements of video coding techniques and secure deduplication. Our design enables the cloud with the crucial deduplicationfunctionality to completely eliminate the extra storage and bandwidth cost.*

Keywords- Cloud Computing, Key distribution, File encryption, deduplication, AES Algorithm, SVC Algorithm.

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.Cloud computing can be defined as a model for enabling ubiquitous, convenient and on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort from the user side and minimal service provider interaction. Cloud computing is considered the evolution of a variety of technologies that have come together to change an organizations' approach for building their IT infrastructure. Actually, there is nothing new in any of the technologies that are used in the cloud computing where most of these technologies have been known for ages. It is all about making them all accessible to the masses under the name of cloud computing. Cloud is not simply the latest term for the Internet, though the Internet is a necessary foundation for the cloud, the cloud is something more than the Internet. The cloud is where you go to use technology when you need it, for as long as you need it. You do not install anything on your desktop, and you do not pay for the technology when you are not using it. The cloud can be both softwareand infrastructure. It can be an application you access through the Web or a server like Gmail and it can be also an IT infrastructure that can be used as per user's request.

II. 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 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.Our scheme have fine-grained access control, any user in the group can use the source in the cloud and revoked users cannotaccess the cloud again after they are revoked.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.In thiswork, we show a securesystem design along this direction, which aims to bring together the advancements of video coding techniques and secure deduplication of text file, image and videos getting uploaded in the cloud environment to optimize the storage space.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.Finally the files are stored in public cloud named CloudMe.

- ***Architecture Diagram***

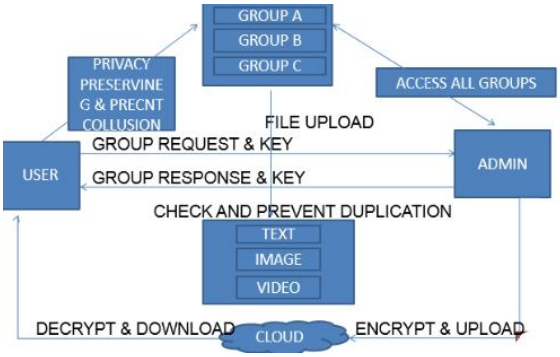

Fig.[1] Architecture diagram

`

IV. MODULES

1. Privacy Preserving
2. Key distribution & Access control
3. Deduplication
4. Encryption (AES)

1)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.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.

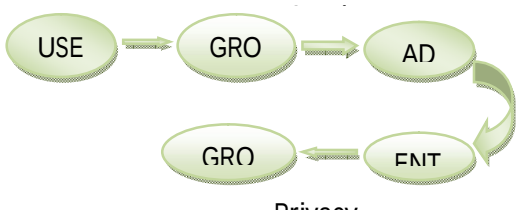

Fig.[2] privacy preserving technique

2)Key distribution and 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.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.

3)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 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. **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.

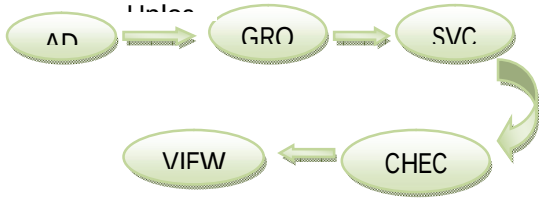

Fig.[3] Deduplication

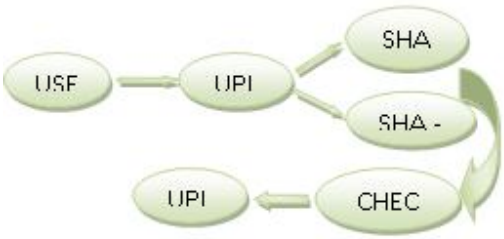

Fig.[4]deduplication using svc

4)Encryption(AES)

The **AdvancedEncryptionStandard** (**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, JoanDaemen 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 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 finitefield.AES consists of several rounds of several processing steps that include substitution, transposition and mixing of the input plaintext and transform it into the final output of ciphertext.

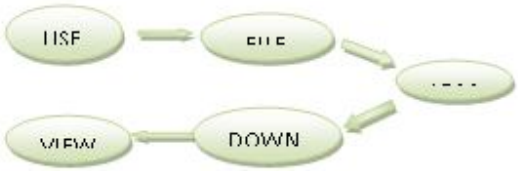

Fig.[5]Encryption and cloud storage

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

V. LIMITATION OF EXISTING SYSTEM

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 very original data. In

existing the file will not be stores with minimum replication. Cloud storage consumption ishigh with highcost.Also our literature survey states deduplication attempt towards multimedia files is not been initiated.
increasing the probability than an attacker eventually obtains the keys.

VI. APPLICATIONS

1) It can be used in various social networks sites such as face book , whatsapp , YouTube etc…
2) It can be used in industries and educational institutes .

VII. CONCLUSION

Our proposed scheme provides a possible way to fight against immoral interference with the right of privacy. We hope more schemes can be created to protect cloud user privacy. Dynamic ownership management is an important and challenging issue in secure deduplication over encrypted data in cloud storage. Security analysis demonstrates that our schemes are secure in terms of insider and outsider attacks specified in the proposed security model.Also our proposed system effectively overcomes the collusion attack. Also prevention of storing a same file like text file, image and videos will optimize the user storage space and saves user storage space and cost

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] K. Lauter, “Cryptographic cloud storage,” in Proc. Int. Conf. Financial Cryptography Data Security, Jan. 2010, pp. 136–149.
[3] R. Swaminathan, Q. Wang M. Kallahalla, E. Riedel, , and K. Fu, “Plutus: Scalable secure file sharing on untrusted storage,” in Proc. USENIX Conf. File Storage Technol., 2003, pp. 29–42.
[4] E. Goh, H. Shacham, N. Modadugu, and D. Boneh, “Sirius: Securing remote untrusted storage,” in Proc. Netw. Distrib. Syst. Security Symp., 2003, pp. 131–145.
[5] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, “Improved proxy re-encryption schemes with applications to secure distributed storage,” in Proc. Netw. Distrib. Syst. Security Symp., 2005, pp. 29–43.
[6] M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryption and secure deduplication,” in Proc. 32nd Annu. Int.Conf. Theory Appl. Cryptographic Techn., 2013, pp. 296–312.
[7] M. Bellare, C. Namprempre, and G. Neven, “Security proofs for identity-based identification and signature schemes,” J. Cryptol., vol. 22, no. 1, pp. 1–61, 2009.
[8] M. Bellare and A. Palacio, “Gq and schnorr identification schemes: Proofs of security against impersonation under active and concurrent attacks,” in Proc. 22nd Annu. Int. Cryptol. Conf. Adv. Cryptol., 2002, pp. 162–177.
[9] S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider, “Twin clouds: An architecture for secure cloud computing,” in Proc. Workshop Cryptography Security Clouds, 2011, pp. 32–44.