An Analysis of Access Control Mechanism with Authentication of Anonymous user and Deduplication of Data in Decentralized Clouds

Imran D. Tamboli
PG Scholar at MITAOE, Alandi

Ranjana R. Badre
Associate Professor at Computer Science and Engineering Department,
MITAOE, Alandi

ABSTRACT
The internet framework and thousands of users are dealing with cloud for sensitive information over cloud computing nowadays providing exciting features due to the services .In terms of security and the access control mechanisms cloud server management is challenging task due to sensitivity of data deploy in clouds.

For the Key Distribution and also data administration when a course of fine grained access control on data is demanded by the users and the scaling factor must be well enough the cloud server suffers with the processing overhead. To maintain scalability, data confidentiality as well as fine graininess of access control mechanisms at the same time on the risk of uncertainty is the main issues. As based on quality of data the system provides and generates access policies and then afterward gives severs by maintaining the security and encryption of data the permission of data owner and modifier to unauthorized cloud severs by maintaining the security and encryption of data. By taking combination of decentralized key policy and attribute Based Encryption (KP-ABE) this thing can be overcome. The proposed system will be robust and secure. The technique referred is known as Deduplication of data (removal of repeated copies), also removal of copies of continuously repeating data is necessary, one of the most important data compression technique widely used in cloud storage to recover the space and bandwidth of cloud. A big support is provided by using convergent encryption technique to the protection of the confidentiality of sensitive data by performing authorized duplicate check in hybrid cloud storage architectures.

Keywords
Cloud Storage, Access control, Key Distribution Centre, Data Deduplication, KP-ABE.

1. INTRODUCTION
The Cloud computing is taking charge of Internet services as computing standard in facilities over the Internet users. The hiding policies and the imagined policies for the data over the internet services are provided by the cloud computing environment [1]. Here, for the farm out the calculations and the also the storage to server using the internet user can use the cloud computing. The data which has been stored over the cloud server using the cloud computing environment is highly sensitive and responsive for the malicious attacks. for e.g. applications for social networking sites such as Facebook , twitter and medical records which are kept within the cloud .For that purpose the high level security tasks and the privacy policies are needed for the data and information over cloud computing environment.

The cloud that it should not be interfaced with the outsourced data before initiating his transactions over the internet, user should first verify. Hence, there is need of confidentiality to avoid the identification of the users from cloud or other user [5].The data which is outsourced is the check of the cloud over the internet, and the cloud is only responsible for the service it provides to the user. The validation of the user who saves the data on cloud is also verified by cloud administrator. By using the phenomenon of the Access control the permission is given to those users having the permission to access over the cloud. Huge information of various applications can be kept over the cloud servers for privacy purpose. Applications like social networking where the user saves the personal data, the Access control plays the very important practice to give the personal access to specific user. Hear the authorized users can be given the Access control with help of access control system in cloud.

The Data deduplication technique of the system is also focuses on the security and space reduction required in the cloud servers. The deduplication data matching technique is used to remove the significant repeated data, over the stored information in cloud server [6], due to this technique the storage space and the bandwidth of cloud can be reduced and space complexity is achieved. The protection of the confidentiality of responsive data and deduplication is done by the convergent encryption technique. Different from traditional deduplication systems, the differential privileges of users are forehead considered in duplicate check algorithm besides the data itself. The better data storage consumption and the data transfer to reduce the number of bytes that must me transmit are achieved by using the defined technique. Rather keeping multiple data or file copies with the same information, deduplication takes out unnecessary data and information by keeping only one physical copy of source file and referring other unnecessary data to that copy of original file by providing the link to the original file.

Deduplication executed at the places at either block level or [3] file management level. In block level duplication checks it and determines the instances of repeated blocks and removes it by providing the pointer at the block for another user. In file level it checks for the same file, if present keeps only one copy by removing the other one by providing the reference to previous one.

2. EXISTING SYSTEM
All schemes use ABE. Usually alive work based control in cloud on access is centralized in nature. There is no need of authentication as key used is symmetric. Privacy preserving is provided for valid control in cloud on access. Well as, a single key distribution center (KDC) where secret keys and attributes are distributed to all users is used by authors for centralized approach.

To reduce the storage space amount and save bandwidth one of the important data for deduplication of solidity which is also beneficial in the cloud. For securely performing of duplicate verifying with disparity privileges the classified cloud is been involved to allow data users at proxy in deduplication of data system.

Disadvantages of Existing System

(A) This system uses the unequal key approach which does not support for authentication.

(B) It is difficult to maintain because the large number of users are keep close by the cloud world.

(C) While giving the privacy to the data is not matched with data deduplication is said to be traditional encoded data.

(D) The similar data copies of different users will lead to different translated texts, making deduplication impossible.

(E) One basic test of cloud storage services is the management which always increasing volume of data.

3. PROPOSED SYSTEM

The problem of realized deduplication of data is made by this system which makes the first try to formally address [1], to give better safety to the data. The validity of the series without interpretation the user's identification before storing the data is suggested to the system that checks. In this design, it also include characteristic of access control in which only responsible users are able to decode the stored information. It also avoids replay attacks and supports formation adaptation, and valuation of data stored in the cloud and also addresses user reversal. It proposed a fully distributed ABE where users could have one or more attributes from each right and need not require a important server. To get over this problem, the decoding task to interchange server, so that the user can calculate with smallest resources.

For more than one correspondence the KP-ABE is a public key cryptography original. In KP-ABE, [2] information is related with attributes for each of which a public key part is described. The set of attributes to the message by scrambling it with the evaluating public key parts the cipher authority associates. Every client is bound for an access structure which is normally represented as an access tree over information attributes, i.e., within hubs of the access tree are control doors and leaf hubs are attached with attributes. Client secret key is generated to return the access structure so the client has the capacity to decode a coded-text if and just if the information attributes accomplish his access structure. The confluent encryption technique has been proposed to encrypt the data before deploying. To have a good data security, the problem of authorized data deduplication is introduced in the proposed system.

As well the data itself different rights of users are considered in replica check. This Schema represents some new deduplication sustaining recognized duplicate check in combined cloud. The scheme is secure in essential of the definitions specified in the proposed security model.[5] It apply a model of this proposed authorized replica check .In this system ,it proposed realized duplicate check scheme incurs minimum overhead compared to normal operations.

Advantages of Proposed System

(A) This system uses the unequal key approach which does not support for authentication.

(B) It is difficult to maintain because the large number of users are keep close by the cloud world.

(C) While giving the privacy to the data is not matched with data deduplication is said to be traditional encoded data. The similar data copies of different users will lead to different translated texts, making deduplication impossible.

(D) One basic test of cloud storage services is the management which always increasing volume of data.

(E) Store and modify the data on the cloud.

(F) During authentication the user identity is protected from the cloud.

The data stored in the cloud performs multiple read and write.

3.1 Access Control Module:

To search the file using the file id and file name this module is used to help client. If the file name is incorrect means it do not get the file, otherwise server ask the public key and get the encryption file.

1) Distributed Key Policy Attribute Based Encryption (KDC SETUP):

For one-to-many correspondences KP-ABE is a public key cryptography primitive. In KP-ABE, information of which a public key part is characterized for each attributes is associated. Encryptor comparing public key parts with scrambling the message to attributes to the set of associates. Every client is l normally characterized which information access tree attributes that is access tree inside of hubs limit doors and leaf are hubs attributes are connected with the access of structure which is assigned.

If the attributes information is able to fulfill his access structure then the client has the ability to decrypt an encrypted-text as the secret key of client is characterized to reflect the structure to access. There are four algorithms where the scheme is been proposed which is defined as below:

Setup:

This algorithm as takes input secure parameters and attribute universe of cardinality N. It is defines as a bilinear group of prime number. Its returns public key and master key which is kept secret by the authority party.

Encryption:

It takes a message, public key and set of attributes. The output is a cipher text.

Decryption:

It takes as a input cipher text, user secret key and public key. The first computes are a key for each leaf node. Then it aggregates to the results using polynomial interpolation technique and returns the message.

Assured File Deletion:

The file policy may be denied under the request by the customer, when terminating the time of the agreement or totally move the files starting with one cloud onto the next cloud nature's domain. The point when any of the above criteria to exists the policy will be repudiated and the key director will totally evacuate the public key of the associated file. So no user can recreate or regenerate the control key of a repeated file in future. File is certainly erased. To recover the

file, the user can ask for the key supervisor to produce the public key. The user can be must verified the file. The key policy attribute is based encryption (ABE) standard is utilized by access the file and verified attribute connected with the file. The file access control the file downloaded from the cloud will be in the arrangement of read just or write predicted. Every client has connected the approaches for each one file. So the right person can access the correct file. For making file access the key policy attribute based on coded data.

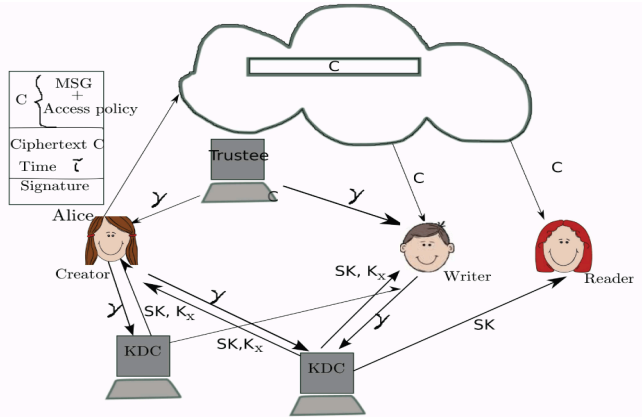

Fig 3.1

2) Secure Deduplication System:
Authorized deduplication is to support; file f of tag is to be determined by file f and the privilege. System calls traditional notation of tag as file token instead whereas to show difference. To generate a file token privilege p is been bounded with a secret key which is KP, which will support authorized access. The token of F is denoted which is access by user which is only allowed with privilege P, where F()=TagGen(F, kp).

It can be also said as the token F() can only be computed where the users with privilege P. The result, when a file will be uploaded as the user with a duplicated token f() then a duplication check will be given from user which will become successful if and only if duplicator also has the file F with privilege P. Therefore the function of token generation is been easily implemented as H(F,kp),where H(_) can be denoted as cryptography hash function.

Safety of Duplicate Check Token:

The system needs to protect, that is, duplicate-check token which is already diminished There are two types of Conflicts, that is, external conflict and internal conflict. As shown below, the external conflict can be viewed as an internal conflict without any privilege. If a user has privilege p, it requires that the conflict cannot fabricate and output a valid duplicate token with any other privilege level p′ on any file F, where p does not match p′. Furthermore, it also requires that if the conflict does not make a request of token with its own privilege from private cloud server, it cannot fabricate and output a valid duplicate token with p on any F that has been queried.

Send Key Algorithm:
Once the key request was received, the sender to be sends the key or he can decline it. With this key and request id which was created at that time of user sending key request the receiver can decrypt the message. Once the key request was receive, the sender can send the key or he can declined. During send key creation, the receiver requests the public key and request id. Using this parameters the receiver can decrypt the message.

4. MATHEMATICAL BACKGROUND

Identities are mapped by Map e: G X G=GT This map satisfies following properties:

1. $e(aP, bQ) = e(P, Q)^{ab}$ *for all P,Q∈G and a,b* $\in \mathbb{z}_q$, $\mathbb{z}_q = \{0, 1, 2 \ldots q\text{-}1\}$

2. *Nondegenerate :* $e(g,g) \neq 1$
$\forall\ a,b \in \mathbb{z}^p$, $e(g^a, g^b) = e(g, g)^{ab}$

To generate secret keys for user j SHA-1 hash function is used and represented as:

$SK[j] = \{\alpha_i, y_i, i \in L_j\}$.

The public key is generated as:

$PK = \left(g, g^{\beta}, e(g,g)^{\alpha}\right)$

5. SYSTEM IMPLEMENTATION

The proposed system consists of the following modules:

1) System Initialization

2) Request for Upload & Download

2.1 User Registration

2.2KDC setup

2.3Attribute generation

2.4Sign

2.5Verify

3) Deduplication

The system can be implemented with the above mentioned modules. The very first module takes responsibility of System initialization. Within system initialization, the system will start and the necessary forms and the basic tasks will be shown to user. Then afterward the user can make a registration or log in to current scenario. Then the user request for the file uploads by providing the set of attributes and policy. The access control will be granted by the key distribution centre and the user gets authenticated. After authentication the user can upload the contents on the cloud which are further encrypted when stored on cloud. Same procedure will be followed for downloading the file from the cloud. In the last module if user wants to upload another file, the deduplication will be checked by the system on file level as well as on block level. If the contents or file name matched then the system will show the deduplication for the file and it will do not upload the file up to contents are unmatched.

The proposed system can have a great combination of access control as well as deduplication over the cloud due to KP-ABE approach and deduplication check with the chunking algorithm. The user can depend upon the system for access control as well as for deduplication check in secure manner.

6. RESULT ANALYSIS

6.1 Experimental Analysis

The experimental analysis in proposed system is on 4 GB RAM and used AES encryption system to encrypt the data before uploading the cloud. As here anonymous authentication system used and it will take time from normal

single step of authentication system but definitely will achieve greater security. The existing scheme (Decentralized Access Control with Anonymous Authentication of Data Stored in Clouds) specifies the time to encrypt the file and it goes exponentially as shown in table 6.1

Here AES used for less memory consumption and less computation time as compared to other scheme, the system compare existing scheme and proposed scheme on the basis of encryption. Our proposed system using AES required less time than existing system as Advanced Encryption Standard not only assures security but also improves the performance shown in figure 6.2

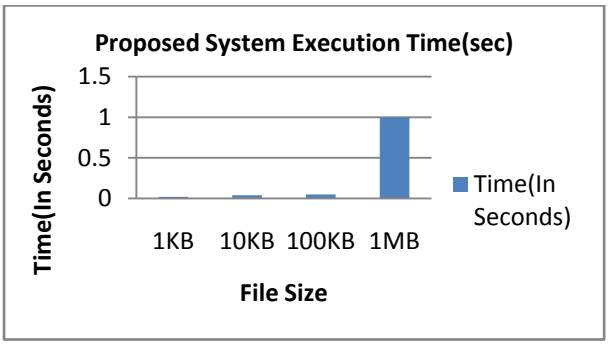

Fig 6.1

Table 6.1 shows time required for existing system and time required for proposed system for file execution, as the proposed system uses AES algorithm for encryption and it is one of the fastest algorithm. The performance is analyses using under various file sizes, time taken to encrypt the file of existing scheme is comparatively more than proposed scheme system.

Table 6.1

Sr. No.	File Size	Existing system Execution Time	Proposed System Execution Time(Sec)
2	1KB	0.8	0.02
3	10KB	1.5	0.04
4	100KB	3	0.5
5	1MB	6	1

6.2 Time Performance

File uploading time is not a constant one. For same size file the time taking for uploading is randomly different. Using the time taken to upload the file one can identify the encryption standard. To confuse the hacker the random time delay is achieved as shown in figure 6.3

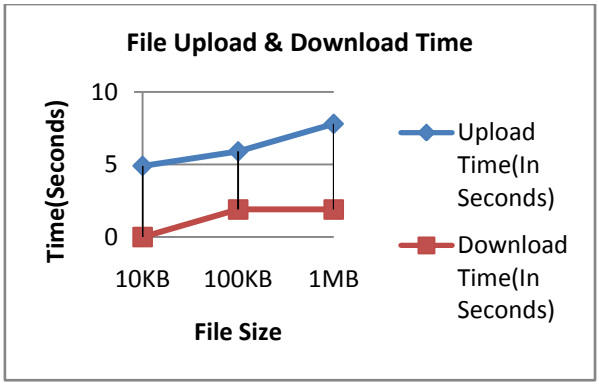

Fig 6.2

File downloading time is also not a constant one. For same size file the time taking for downloading is randomly different. Using the time taken to download the file one can identify the encryption standard. To confuse the hacker the random time delay is achieved as shown in figure6.4

Table 6.2

File Size	Upload Time(In Seconds)	Download Time(In Seconds)
10KB	4.9	0
100KB	5.9	1.9
1MB	7.8	1.9

The performance of the proposed scheme was analyzed under various file sizes. At first the time performance of the scheme is evolved for different file sizes. Then the cryptographic operation time is evolved. The only achievement of the scheme is, it supports random time duration for any size of files to download/Upload as shown in table 6.2

6.3 Analysis with respect to access policy

The user can decide which user can access his data and which user can only read the data or can modify. There are two types of access policy, the only read policy can permit the user to only read the files contents and can only download the files. User can't make any changes through it. Relation between access policies of various users of proposed system is as shown in figure 6.5

Table 6.3 show the comparison of granting the permission for read or writes access of proposed system, these policies are depends upon the user to grant the permission to another user. Some of the existing system support single read and single write but the proposed scheme support multiple read and write.

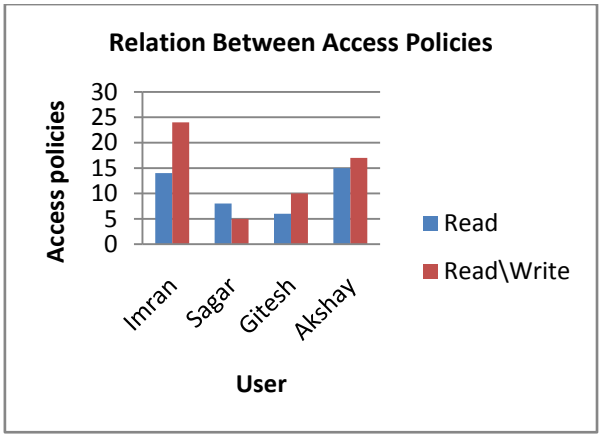

Fig 6.3

Table 6.3

User Name	Read	Read\Write
Imran	14	24
Sagar	8	5
Gitesh	6	10
Akshay	15	17

6.4 Time Performance with respect to access policy

Read/ write access permits the user to read the content of files and if the user feels to make changes it then the file can be modify and user may write it back. These policies granting the facility it depends upon the user itself. It depends wholly on the user to grant the access permission to another user. For analysis of the time performance of access policies, the system consider various size of files, here we measure the time required for read operation and the time required for write on cloud after modify the same file.

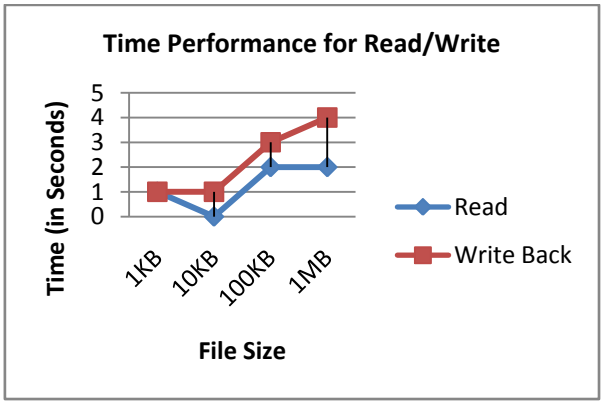

Fig 6.4

The graph in figure 6.6 shows the time required for read operation and time required for write to another user after modifying the same file.

Table 6.4

Sr.No.	File Size	Time Required for Read Operation	Time Required for Write Operatipon
1	1KB	1	1
2	10KB	0	1
3	100KB	2	3
4	1MB	2	4

Here the system consider various file size for read and write access, table 6.4 shows the comparison between the time taken to read the file and time taken by same file size after modify and write it back to the another user. It is observed that the system support random time duration for any size of files. From this it is concluded that the update process has similar characteristics of regular uploading with linear growth or decrease in time depending on increasing or decreasing the size of file.

6.5 Analysis of KDC with respect to the distribution of keys

In single key distribution environment the keys of all user is maintained by the single KDC. The single KDC is responsible for all encryption and decryption .Hence the load on the single key distribution centre is too much to handle such as a large amount of key as the user of cloud are increase day by day .

Hence the burden on one KDC is divided into more than one KDC .As the one KDC is a single point failure may break down the transaction. Here we consider some existing scheme for comparison between the numbers of KDC. Table 7.5 shows the comparisons of our scheme with the other scheme, we compare according to key management of our scheme with existing scheme.

Table 6.5

Sr.No.	Scheme	Approach	No.of KDC	Key Management
1	Secure and Efficient access to outsource the data	Centralize	Single	Poor
2	Realizing finegrained and flexible access control to outsourced data with attribute based cryptosystemsAttribute based data sthgaring	Centralize	Single	Poor
3	Attribute based data sharing with attribute revocation	Centralize	Single	Poor
4	**Proposed Scheme**	**Decentralize**	**Two**	**Efficent**

6.6 Time Performance of KDC

In proposed scheme Attribute Based Encryption (ABE) system is proposed which is decentralized in nature. Key distribution is done in a decentralized manner which

distributes the access policy and attributes of a user. The proposed scheme prevents from replay attacks and it also supports user revocation.

Algorithm describes the sharing of secured key to the user for sharing the document on cloud, MD5 algorithm is used in system for generating of 32- bit key, MD5 algorithm takes input as the unique attributes of the user for generating of 32-Bit key. Here for measure the performance of KDC system measure time taken by the KDC for generating the keys for various users, as shown in figure 6.7.

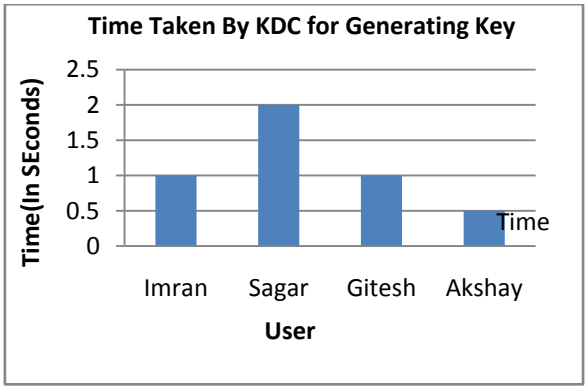

Fig 6.5

Here various users are considered for measuring the performance of KDC. Md5 is fastest hashing an algorithm and there is no way to easily decrypt the hash value represent by MD5, it generates 32- bit hash value for the KDC for generating the keys for various user.

The cryptographic operation time is evolved. The only achievement of the scheme is, it supports random time duration as shown in table 6.6

6.7 Comparison of Proposed Scheme with Existing Scheme

The system scheme is compared with other access control schemes and show that our scheme supports many comparison of our Scheme with existing access control schemes features that the other schemes did not support. 1-W-M-R means that only one user can write while many users can read. M-W-M-R means that many users can write and read.

Table 6.6

Reference No.	Centralized/ Decentralized	Type of access control	User revocation	Deduplication
1	Centralized	Symmetric Key	no	No
2	Centralized	ABE	no	No
3	Centralized	ABE	no	No
4	Decentralized	ABE	yes	No
5	Centralized	ABE	no	No
6	Decentralized	ABE	yes	No
Proposed System	**Decentralized**	**ABE**	**yes**	**Yes**

It shows that most schemes do not support many writes which is supported by this scheme. Proposed system scheme is robust and decentralized; most of the others are centralized. The system scheme also supports privacy preserving authentication, which is not supported by others. Most of the schemes do not support user revocation, which our scheme does. Comparison of proposed scheme with the existing Scheme on the basis of various parameters is as shown in table 6.7

6.8 Analysis with respect to Deduplication

The current system is compared with the previous systems in terms of the deduplication of files. The files which are already uploaded by user are compared with the newly updated files. If contains and constrains are same then the file has been given with the reference number and the space has been reduced by avoiding number of replicas of same information. Here, the comparison graph shows the total number of files uploaded by a particular user and if the same user tried to upload the same content file again then the file will not be uploaded and it will show the total number of DE duplicated files.

The graph is derived from real time environment and the user attempted the number of instances to upload the files. And according to this analysis the system shows the deduplication and the reduction in space in cloud as shown in figure 6.6

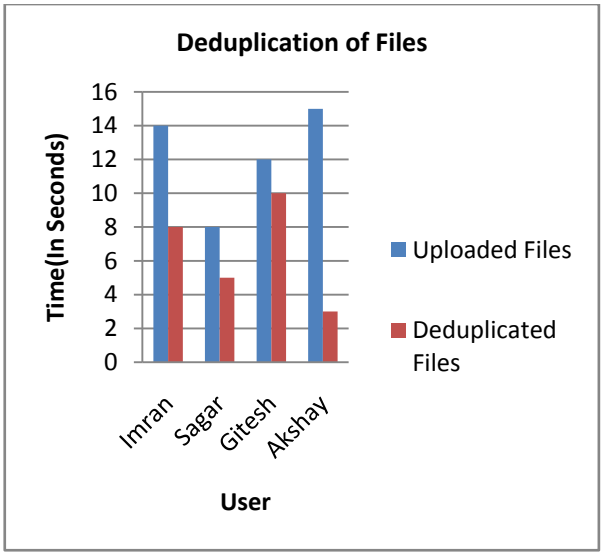

Fig 6.6

7. CONCLUSION

The proposed system gives a hybrid access control technique with unknown authentication, which provides user revocation and prevents replay attacks. Hence, the current system provides the anonymous authentication and the data deduplication technique for saving the cloud space and the data redundancy is maintained for secure cloud usage.

8. REFERENCES

[1] SushmitaRuj, Milos Stojmenovic, and AmiyaNayak, "Decentralized Access Control with Anonymous Authentication of Data Stored in Clouds", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL.

[2] H.K. Maji, M. Prabhakaran, and M. Rosulek, "Attribute-Based Signatures: Achieving Attribute-Privacy and Collusion-Resistance," IACR Cryptology ePrint Archive, 2008.

[3] M. Green, S. Hohenberger, and B. Waters, "Outsourcing the Decryption of ABECiphertexts," Proc. USENIX Security Symp., 2011.

[4] H. Li, Y. Dai, L. Tian, and H. Yang, "Identity-based authentication for cloud computing," in *CloudCom*, ser. Lecture Notes in Computer Science, vol.5931. Springer, pp. 157–166, 2009.

[5] S. Ruj, M. Stojmenovic and A. Nayak, "Privacy Preserving Access Control with Authentication for Securing Data in Clouds", *IEEE/ACM International Symposium on Cluster,Cloud and Grid Computing*, pp. 556–563, 2012.

[6] Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, WenjingLou,"A Hybrid Cloud Approach for Secure Authorized Deduplication",IEEE Transactions on Parallel and Distributed Systems, 2014.

[7]"Libfenc: The Functional Encryption Library,"http://code.\google.com/p/libfenc/, 2013.

[8] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, "Toward Secure and Dependable Storage Services in Cloud Computing," IEEE Trans. Services Computing, vol. 5, no. 2, pp. 220-232, Apr.-June 2012.

[9] http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-cs-01-en.pdf, 2013.

[10]http://securesoftwaredev.com/2012/08/20/xacml-in-the-cloud,2013.

[11] K. Yang, X. Jia, and K. Ren, "DAC-MACS: Effective Data Access Control for Multi-Authority Cloud Storage Systems," IACR Cryptology ePrint Archive, p. 419, 2012.

[12] A.B. Lewko and B. Waters, "Decentralizing Attribute-Based Encryption", Proc. Ann. Int'l Conf. Advances in Cryptology(EUROCRYPT), pp. 568-588, 2011.