Preserving Privacy using Third Party Auditor in Cloud for Data Storage

YogeshShinde
Department of Computer Engineering
Dr. D. Y. Patil SOET
Lohgaon, Pune

Ramesh M. Kagalkar
Department of Computer Engineering
Dr. D. Y. Patil SOET
Lohgaon, Pune

ABSTRACT
Cloud computing is a internet based thing or next generation in information technology. Users store their large amount of data on a cloud server at the remote place without worrying about storage correctness and integrity of data. Security is viewed as one of the top positioned open issues in cloud computing. In most of the before proposed schemes, RSA algorithm was used for storage security. AES being faster in encryption and decryption as compared to RSA. The proposed system makes use of AES algorithm to maintain data integrity at the untrusted server. The client can alternative to a Third Party Auditor (TPA) to check the integrity of outsourced data and be worry free because user does not physical present at all time. The proposed storage security scheme also assures recovery of data files, in case of data loss or corruption. To recover of that block or file to maintain data availability in the cloud server. It supports data dynamics where the user can perform different operations on files such as insert, delete and update as well as batch auditing, where multiple cloud client requests for storage correctness will be handled simultaneously which decrease communication and also computing cost. To expand the user level safety, proposed procedure supplied a click on point based graphical password scheme and One Time Password (OTP) on the time of uploading the file.

General Terms
Cloud Computing, Security and Reliability.

Keywords
Cloud Storage, Data Availability, Data Auditing, Graphical password, Privacy and Security.

1. INTRODUCTION
Cloud computing is one of the most growing field in research and technology in today's world. It uses hardware and software as computing resources to provide service through the internet. It also provides various service models such as Platform as a service (PaaS), Infrastructure as a Service (Iaas), Software as a Service (SaaS), Storage as a Service (STaaS), Security as a Service (SECaaS) and Data as a Service (DaaS). Cloud storage becomes an growing attraction in cloud computing model, which allows client to store their data on cloud and access them anywhere without any risk. The advantages can be listed as on demand self service, worldwide network access, location independent resource pooling and faster resource elasticity. Cloud data storage permits client to collection their files at remote place and reduces local storage maintenance and management. However, their protocol lacks in maintaining privacy of data which is one of the issue for the cloud data storage. Security is viewed as one of the top positioned open issues in cloud computing [1].

The clients oblige that their information stay secure over the cloud and they need to have a strong assurance from the cloud servers that provider store their data correctly without tampering or partially deleting because the internal operation details of service providers may not be known to the cloud users. Thus, an efficient and secure scheme for cloud data storage has to be in a position to ensure the data integrity and confidentiality. Encrypting the data before storing in cloud can handle the confidentiality issue. However, verifying integrity of data is a difficult task without having a local copy of data or retrieving it from the server. Due to this reason the straightforward cryptographic primitives cannot be applied directly for protecting outsourced data. Besides a naive way to check the data integrity of data storage is to download the stored data in order to validate its integrity, which is impractical for excessive I/O cost, high communication overhead across the network and limited computing capability. Therefore, efficient and effective mechanisms are needed to protect the confidentiality and integrity of the users data with minimum computation, communication and storage overhead [2].

The confidentiality and integrity of the outsourced data in clouds are of paramount importance for their functionality. The reasons are the provider whose purpose is mainly to make a profit and maintains a reputation, has intentionally hide data loss an incident which is rarely accessed by the users. The malicious cloud provider might modify data or is able to easily access the entire private files and sell it to the biggest competing company. External attacker who intercepts and captures the communications is able to know the users private data as well as some important business secrets. Internal and external threats are major problem in cloud server [3].

To verify the user data, remote data integrity checking protocol is used. In this mechanism, the client generates some hash. Using challenge response protocol, client send request to cloud server to check integrity of files blocks. Then the server generates responses and sent to third party or verifier. As of late, a few researchers have proposed distinctive varieties ofremote data integrity checking under different cryptography schemes [4]. Cloud storage is support to dynamic file blocks operations. The files are stored on server are not only accessed by client but also perform some block level operations such as update, delete and insert. Hence, it is important to develop a framework which is more secure and efficient to support dynamic audit services. To achieve this, Merkle Hash Tree (MHT) is used. Original file is divided into small part like tree structure. In which leaf node indicates hash of blocks [5], [6].

In order to resolve the issue of the integrity of outsourced data, many techniques are developed under the system and security model. User store his data at remote location for that security of that data is major concern in cloud computing. It does not present at all the time to check the integrity of that data. For that purpose user put an external third party to check the integrity of files at any time or periodically. Original file is divided into small blocks, before sending data to cloud it will

be encrypted format. One time passwords are often referred to as a secure and stronger form of authentication. It can be used by user to access their private information [7], [8], [9].

Graphical image point passwords are conveyed into use for more memorability and to decrease the tendency of choosing insecure passwords. When using images as security passwords should increase total password security. The image will be displayed on the user login screen and the user will have to click on a some of the regions. When the user clicks on the correct region, then the user will be authenticated. In the graphical image click point password system, the user has to select memorable locations in an image as a security password [10].

2. LITERATURE SURVEY

Recently, much work has been done in the area of cloud security. Majority of them focus on checking the integrity of user files stored in cloud. Cloud computing providing big infrastructure to store and execute client data. There is no need of to make own the infrastructure. The main benefits are to reduce capital expenditure.

The author proposes an approach which consists on Rivest Shamir Adleman (RSA) based hash function for integrity verification of the stored data at remote server. Using this scheme, it is possible for the client to perform multiple challenges using the same metadata. But the limitation of this scheme lies in the computational complexity at the server [4].

The researcher describes a technique in which the data stored remotely across multiple sites can be safeguarded. The main scheme based on algebraic signature. The main disadvantage of this scheme is that the computation complexity at the client side and server side takes place at the cost of the linear combination of file blocks. Also, the security of this scheme remains unclear [11].

The storage services and sharing of resources over networks are become popular but the data stored at untrusted servers has received more attention in cloud. The provable data possession model for remote information checking supports big data sets in widely distributed storage systems. The main disadvantage of this scheme is that an overhead of generating metadata is imposed on the client [12].

A scheme called, "Proofs Of Retrievability(POR)", proposed by Juels and Kalisiki focuses on static archival of large files. To ensure data possession and retrievability, it makes use of spot checking and error correcting codes. This scheme cannot be used for public databases. The disadvantages of this scheme are that, the number of queries clients used is fixed priority. Preprocessing of each file is needed prior to storage at the server. The scheme cannot be used for public databases and not support public auditability i.e. it supports only two parties auditing, which is not efficient because neither the client nor the cloud service provider can give assurance to provide balance auditing [13].

Cloud computing is next generation in the information technology. It provides various services to end user such as allow storing their data on the cloud. The auditing report not only ensures strong cloud storage guarantee but also at the same time achieves fast data error localization i.e. the identification of server misbehavior. For that purpose flexible distributed storage integrity auditing mechanism is used, utilizing the homomorphic token and distributed erasure coded data. It supports dynamic data verification and resilient against byzantine failure and malicious data modification attack [14].

The main system based on proxy provable data possession method. The advantage of this scheme is that the efficient user controlled data management. The disadvantages of that scheme is to public verifiability may cause intruders collision on data [15].

Integrity checking main concern in cloud computing. For that purpose the researcher developed system which makes use of MHT and Advanced Encryption Standards (AES) algorithm to maintain integrity of confidential data. The proposed storage security scheme also assures recovery of data, in case of data loss or corruption, by providing a recovery system [16].

3. SYSTEM OVERVIEW

In the proposed system our main focus is on preserving privacy of outsourced data using TPA for secure cloud storage. Confidential data security in cloud is one of the serious issues with cloud storage facility. The input to the system can be any type of files. The user stores their confidential data at the remote place and deletes copy of that data. For this, auditing of the data is necessary to assure client safety of his data. To overcome this problem of data security, proposed systems introduce an AES based storage integrated. Following figure 1 gives idea about block diagram of proposed system.

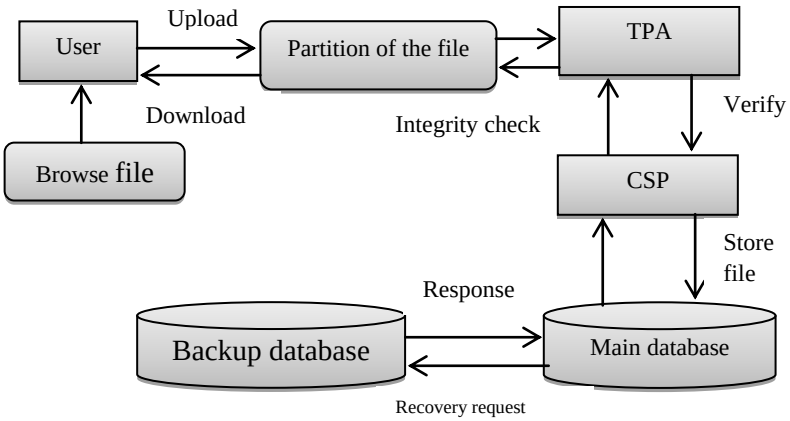

Figure 1: Block diagram of proposed system.

Following section contain brief description of all blocks are present in diagram,

1. Input File: It can be any type of files extensions given as input to the system.

2. Partitioning Files: The input files are divided into small part and these blocks encrypted using AES algorithm before sending to verifier. It helps to store the data effectively in quick manner enhancing easy access to data also when there is need. The original data is complex and there is difficulty in storing it in cloud, so partitioning function is used to make the storage easy in cloud. Original file is also reconstructed when there is need to access the same.

3. Auditor: The integrity verification process consists of where client initiates by forward a request to auditor for auditing the desired file or data. The third party generates a challenge, sends it to the cloud provider and in response, the server generates a proof for the corresponding challenge. The verifier fetches stored hash of that requested file and performs comparison in newly generated hash and stored hash. If both the hash are equal then auditor transfer acknowledgement to

the user. If both are equal then acknowledgement will be that stored data is as it is. If data get corrupted by cloud then acknowledgement will be that your data is corrupted.

4. Backup and Recovery System: The proposed storage security scheme also assures recovery of data, in case of data loss or corruption, by providing a recovery system. Thus the proposed scheme aims at keeping the user data integrated and support data restore. The recovery system adds to the plus points as it contributes to the availability of data which is a very important parameter.

3.1 System Architecture

In this section describe about system model of public auditing using TPA scheme which provides a complete outsourcing solution of confidential data and also integrity checking periodically.

3.1.1 Objective

The objective of this research is to implement an application that helps store the data in remote places with a secure way by allowing the auditor to check the integrity of the file; on download, there should not be any leakage of data and one click point image password authentication scheme developed to improve user level security.

3.1.2 Problem Statement

To develop the system which improve the user level security using graphical click point image passwords and protect the integrity of outsourced data using third party auditor for secure cloud storage. Third party auditor should audit the outsourced data from the cloud, not ask for a copy and also should not create new vulnerability to user data privacy. It checks that data is modified or not if modified that information send to the user.

3.1.3 Proposed Architecture

An architectural description involves modelling and representation of architectures using appropriate mechanisms like architecture description languages and architecture frameworks. The architecture design of proposed system shown in the figure 2. Notice the flow of data between the client, cloud service provider and third party auditor.

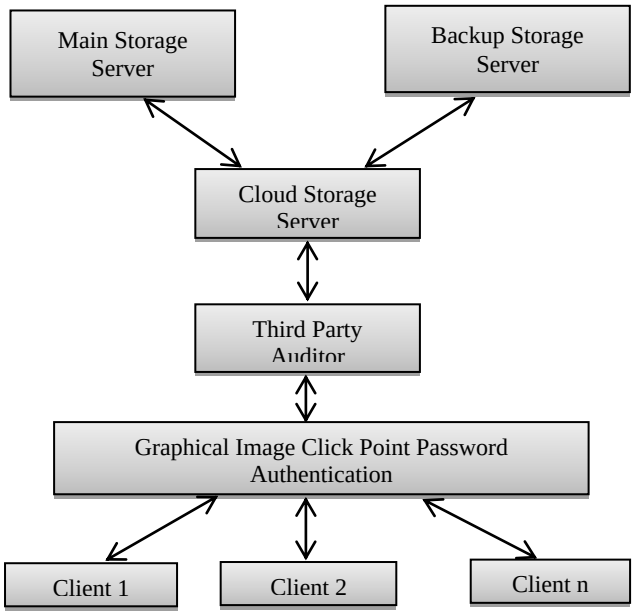

Figure 2: Architectural diagram of proposed system.

The cloud storage model considering here is consists of five main components as illustrated in figure.

1. Client: The client is an entity that has large amount of files which are to be outsourced i.e. to be stored on cloud and accessing the data any time.
2. Cloud Service Provider (CSP): Cloud server maintained by cloud service provider, has consequential storage space and estimation resources to maintain the clients remote data.
3. Third Party Auditor (TPA): The third party auditor or verifier, who has proficiency, capabilities to audit data any time and verifies the integrity of outsourced data in cloud on behalf of users. Based on the audit result, the auditor could release an audit report to user.
4. Main Storage Server: Storage in main server is also called as primary storage. It is area in a computer device in which files are stored and access by processor very efficient manner. The purpose of main storage is maintained the original files or private data of the client are stored.
5. Backup Storage Server: A backup or the process of backing up refers to the copying and archiving of computer data so it may be used to restore the original after a data loss event.

The following are the steps used in the framework

1. User login to the system using text password, image, click point authentication password.
2. In proposed methodology client store confidential data on cloud. At the time of uploading files to cloud it get encrypted by AES algorithms.
3. Original file split (break up) into smaller parts. It helps to store the data effectively in a quick manner, enhancing easy access to data also when there is a need.
4. These blocks are in encrypted format transfer to TPA. It generates a hash of the particular block using MD5 algorithm. That generated hash stored along with TPA.
5. Client checks integrity of files then it will send request to the TPA and auditor forward that request to the cloud.
6. Cloud sever will generate the newly hash of requested file and send that generated hash to the TPA.
7. If the newly hash are equal to stored hash then TPA transfer acknowledgement to the user else file was corrupted or modify by unauthorized person.
8. The auditing reportsends to the user as well as a cloud provider for the improve services.

3.1.3 Security Model

The preserving privacy public auditing scheme consists two phases such as,

1. Setup Phase: Setup phase contain following Algorithms

    1. Key generation module: This algorithm used for generation of public and secret key. The

proposed system use RSA algorithm for generating both key.which takes a large security parameter p and q prime number as input and produces apublic and private key pair (pk,pr) based on RSA algorithm .

2. Signature generation module: This is used for generation of each block hash, metadata and digital signature.

2. Verification Phase: Audit phase contain following Algorithms,

1. Generation of proof module: when user send request for checking integrity of file or blocks.

2. Verify proof module: This is run by TPA for checking integrity of files which is stored on remote place.

3.1.4 Algorithmic Steps

Algorithm for Data Integrity Verification - To protect the data loss, Data integrity verification mechanism is used. It also achieves theeffective storage and retrieval processes. The public auditability mechanism manages erroridentification by comparing hash of files. This ensures data security from unauthorized access.It also increases the performance of uploading and downloading files. The knowledge basedauthentication mechanism is also provided for user level authentication. Dynamic data operation,like insertion, deletion and updating is also done before partitioning the data.The proposed system consists of following algorithmic steps,

Steps

1. Third party auditor generates a random set of challenge.

2. Cloud service provider computes new hash code based on the filename or blocks input.

3. Provider computes the originally stored value.

4. Auditor receives new hash and compares with stored hash.

5. After verification, the auditor can determine whether the integrity is breached.

6. Send acknowledgement to client and auditing report.

3.1.5 Mathematical Model

A mathematical model is a description of a system using numerical concepts and language. LetS be the whole system.

S= {U, F, Sc, Bc, TPA, P, H}

Where,

1. U is set of users.

$U = \{U_1, U_2 \ldots Un\}$

2. F is a set of files.

$F = \{F_1, F_2 \ldots Fn\}$

3. Sc is data storage server, where the files are stored.

4. Bc is back up storage server.

5. TPA is auditor who audits the data.

6. P is a set of file partition. $P = \{P_1, P2\ldots., Pn\}$

7. H is the hash of file block.

$H = \{H_1, H_2, \ldots Hn\}$

Table 1: Activities table for proposed system.

Activity	Relation	Description
Activity 1	U →F	Every user has number of data files.
Activity 2	B→TPA	Block of file sends to third party auditor.
Activity 3	TPA → H	TPA generates hash of each block and store along with.
Activity 4	TPA → CS	TPA sends file to CS for storing purpose.
Activity 5	CS→ TPA	CS sends newly hash to TPA and compares both hashes for integrity checking.
Activity 6	TPA→U	Send acknowledgement of user.

3. RESULTS AND DISCUSSION

Data security in the cloud is one of the serious issues with cloud storage facility. Client store theirdata in the cloud, delete the local copy of that data and rely completely on the cloud server fordata safety and maintenance. For this, auditing of the data is necessary to assure client safetyof his data. To overcome this problem of data security, we introduce an AES based storage integrated.

3.1 File Upload to Cloud Service Provider

This graph represents the performance of uploading file to our system. X-axis represents the file size in kb and Y-axis represent the time require uploading in the system.

Table 2: Different size of files uploads on cloud.

File size in KB	Time required for upload file(ms)
5 kb	337
10 kb	430
15 kb	440
20 kb	470

Table 2 shows the time taken by the proposed system for uploading file on a cloud. The file considered for uploads are different sizes such as 5 kb,10 kb,15 kb and 20 kb. If the file size is smaller in kB than the minimum time required for uploading particular file. For example, if user has file size is 5 kb than it required 337 ms time for upload that file on cloud. Similarly, user has file sizes are 10 KB, 15 kb and 20 kb file than it required 430 ms, 440ms and 470 ms to upload on server.

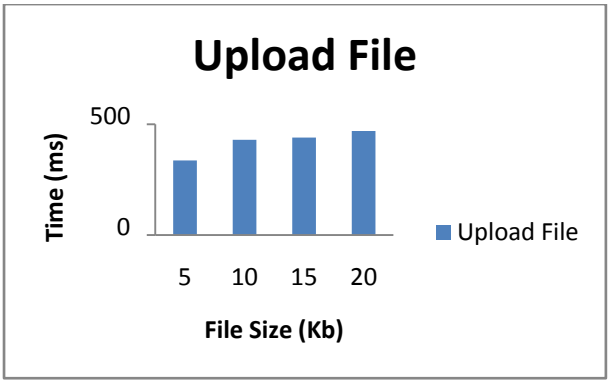

Figure 3: Graph of file uploads on cloud service provider.

Figure 3 shows graph of file upload on cloud. The user data is encrypted using AES andthen stored on the cloud server. User uploads any size of file with minimum time it become moresecure by using one time password at time of file upload.

3.2 Batch Auditing

Batch auditing for multi-client data are used to handle multiple verification sessions. TPAcan perform the many auditing jobs in a batch manner for better efficiency. To achieve batch auditing where multiple delegated auditing tasks from different users can be performedsimultaneously by the TPA.

Table 3: Number of task for batch auditing of system.

Number of task	Time(ms) required for batch
5	125
10	140
15	151
20	172

Table 3 shows time required for batch auditing of multiple tasks. In proposed system, thirdparty auditor performs auditing on number of task at same time. Instead of individual auditing,batch auditing is better because it requires minimum time span to audit data.

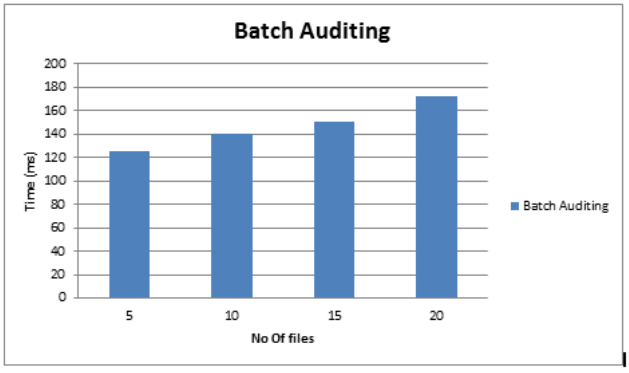

Figure 4: Graph of file uploads on cloud service provider.

Figure 4 shows third party auditor performs audits for multiple users simultaneously andefficiently. Consider nclients having n files on the similar cloud. They have the identical auditorand audit the data from the cloud, not ask for any copy of that data.

4. CONCLUSIONS

It is clear that although the use of cloud computing has rapidly increased. Security is still considered the major issue in the cloud computing environment. Customers do not want to lose their private information as a result of malicious insiders in the cloud. We have seen how the security service which is provided by third party helps in securing data, it provides the facility of data verification and allows data to be shared between designated groups of people. The system proposes a preservingprivacy for data storage security in cloud computing. Graphical image, click point password and one time password achieve the high level of security in authenticating the user over the internet. This is further expand the public auditing using auditor procedure intoa multi user setting, where the auditor can execute many auditing jobs in a batch manner for more proficient.

In future, this application must be simulated with the real cloud and check whether it works exactly in the same way and helps the group access of data for the user become secured. Third party auditor notification regarding data within the specific time. It also included within the cloud with less overhead.

5. ACKNOWLEDGMENTS

I would like to thanks Head of Computer Engineering Department, Prof. ArtiMohanpurkarand my project guide Prof. Ramesh M. Kagalkar.The success of this project has throughout depended upon an exact blend of hard work and unendingco-operation and guidance, extended to me by the superiors at our college.I am indebted to Prof. RoshaniRaut (Ade), ME Co-ordinatorand Dr. UttamKalwane, Principal whose constant encouragementand motivation inspired me to do my best.

6. REFERENCES

[1] Cong Wang ,Sherman S.M. Chow ,Qian Wang ,KuiRen and Wenjing Lou ,"Privacy- Preserving Public Auditing for Secure Cloud Storage",IEEE Transactions on Computers,Vol. 62, No. 2, Feb. 2013.

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy Preserving Public Auditing for Storage Security in Cloud Computing", Proc. IEEE INFOCOM 10, Mar. 2010.

[3] H. Takabi, J.B.D. Joshi, and G. Ahn, "Security and Privacy Challenges in Cloud Computing Environments",Article in IEEE Security and Privacy, Vol. 8, No.6, Nov-Dec. 2010.

[4] Y. Deswarte, J.-J. Quisquater, and A. Saidane ,"Remote integrity checking ", In Proc. Of Conference on Integrity and Internal Control in Information Systems (IICIS),Vol. 3 , Nov. 2003.

[5] Betzy K. Thomas, M. NewlinRajkumar,"A Dynamic Public Auditing Security Scheme To Preserve Privacy In Cloud Storage",International Journal of Software and Hardware Research in Engineering,Vol.2,Issue 1,Jan.2014.

[6] Giuseppe Ateniese, Roberto Di Pietro, Luigi V. Mancini, and Gene Tsudik., "Scalable and Efficient Provable Data Possession*",*Proc.SecureComm 08, Sept. 2008.pp. 24-31.

[7] YogeshShinde ,OmprakashTembhurne ,"A Review of Protect The Integrity of Outsourced Data using Third

Party Auditing for Secure Cloud Storage ", International Journal of Science and Research (IJSR),ISSN(Online):2319-7064,Vol-3, Issue 10 , Oct.2014.

[8] YogeshShinde ,AlkaVishwa,"Privacy Preserving using Data Partitioning Technique for Secure Cloud Storage",International Journal of Computer Applications (0975 8887), Vol. 116 ,No. 16, Apr.2015.

[9] YogeshShinde , AlkaVishwa,"Public Auditing Security Scheme To Preserving Privacy For Secure Cloud Storage", Fourth Post Graduate Conference for Computer Engineering students (cPGCON) , Mar.2015.

[10] FarnazTowhidi, Maslin Masrom ,"A Survey on Recognition-Based Graphical User Authentication Algorithms ",International Journal of Computer Science and Information Security,Vol.6, No.2, Nov. 2009.

[11] T.Schwarz and E.L. Miller, "Store, forget, and check: Using algebraic signatures to check remotely administered storage,In Proceedings of *ICDCS* . IEEE Computer Society, 2006.

[12] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and D. Song, "Provable Data Possession at Untrusted Stores",Proc. 14th ACM Conf. Computer and Communication Security(CCS 07), PP. 598-609, 2007.

[13] A. Juels and J. Burton, S. Kaliski, "PORs: Proofs of Retrievability for Large Files",Proc. ACM Conf. Computer and Comm. Security(CCS 07*)*, pp. 584-597, Oct. 2007.

[14] Cong Wang,QianWang,KuiRen, Ning Cao , and Wenjing Lou "Toward Secure and Dependable Storage Services in Cloud Computing",IEEE Transaction On Service Computing,Vol. 5,No.2,Apr-Jun. 2012.

[15] Huaqun Wang, "Proxy Provable Data Possession in Public Clouds ",IEEE Transactions On Services Computing, Vol. 6,ISSN: 1939-1374 , No. 4, Oct-Dec 2013.

[16] Poonam M. Pardeshi, Prof. Bharat Tidke,"Improving Data Integrity for Data Storage Security in Cloud Computing",International Journal of Computer Science and Information Technologies, Vol. 5, May 2014.