Concealment of Data Storage in Cloud Using Watermarking

S.Mythili
M.E CSE,
KRCT,
Trichy,Tamil Nadu, India.

E.Srimathi
M.E CSE,
KRCT,
Trichy,Tamil Nadu, India.

S.Priyanka
M.E CSE,
KRCT,
Trichy,Tamil Nadu, India.

Abstract ——**The Cloud computing is a latest technology which provides various services through internet. The Cloud server allows user to store their data on a cloud without worrying about correctness & integrity of data. Cloud data storage has many advantages over local data storage. User can upload their data on cloud and can access those data anytime anywhere without any additional burden. The User doesn't have to worry about storage and maintenance of cloud data. But as data is stored at the remote place how users will get the confirmation about stored data. Hence Cloud data storage should have some mechanism which will specify storage correctness and integrity of data stored on a cloud. The major problem of cloud data storage is security. Many researchers have proposed their work or new algorithms to achieve security or to resolve this security problem. In this paper, we propose a new innovative idea for Concealment of data storage using watermarking with public auditing in cloud computing. It supports data dynamics where the user can perform various operations on data like insert, update and delete as well as batch auditing where multiple user requests for storage correctness will be handled simultaneously which reduce communication and computing cost.**

Keywords——*Public Auditing, Watermarking, TPA, Security*

I. INTRODUCTION

Cloud Computing is using hardware and software as computing resources to provide service through internet. Cloud computing provides various service models as platform as a service (PaaS), software as a service (SaaS), Infrastructure as a service (Iaas), storage as a service (STaaS), security as a service (SECaaS), Data as a service (DaaS) & many more. Out of this Paas, SaaS and IaaS are most popular.

Cloud computing has four models as Public cloud: though which the service is available to all public use. Private cloud: Through which service is available to private enterprise or organization. Community Cloud: It allows us to share infrastructure among various organizations through which we can achieve security, compliance and jurisdiction. This can be managed internally or by a third-party and hosted internally or externally. Hybrid cloud: it is a combination of public and private cloud. Cloud computing has many advantages as: we can easily upload and download the data stored in the cloud without worrying about security. We can access the data from anywhere, any time on demand. Cost is low or pay per usage basis. Hardware and software resources are easily available without location independent. The major disadvantages of cloud computing is security.

A. Security Issues

The security is a major issue in cloud computing. It is a sub domain of computer security, network security or else data security. The cloud computing security refers to a broad set of policies, technology & controls deployed to protect data, application & the associated infrastructure of cloud computing. Some security and privacy issues that need to be considered are as follows

1) *Authentication:* Only authorized user can access data in the cloud
2) *Correctness of data:* This is the way through which user will get the confirmation that the data stored in the cloud is secure
3) *Availability:* The cloud data should be easily available and accessible without any burden. The user should access the cloud data as if he is accessing local data
4) *No storage Overhead and easy maintenance***:** User
5) doesn't have to worry about the storage requirement & maintenance of the data on a cloud
6) *No data Leakage:* The user data stored on a cloud can accessed by only authorize the user or owner. So all the contents are accessible by only authorize the user.

1

7) *No Data Loss:* Provider may hide data loss on a cloud for the user to maintain their reputation.

In cloud computing, cloud data storage contains two entities as cloud user and cloud service provider/ cloud server. Cloud user is a person who stores large amount of data on cloud server which is managed by the cloud service provider. User can upload their data on cloud without worrying about storage and maintenance. A cloud service provider will provide services to cloud user. The major issue in cloud data storage is to obtain correctness and integrity of data stored on the cloud. Cloud Service Provider (CSP) has to provide some form of mechanism through which user will get the confirmation that cloud data is secure or is stored as it is. No data loss or modification is done.

Security in cloud computing can be addressed in many ways as authentication, integrity, confidentiality. Data integrity or data correctness is another security issue that needs to be considered. The proposed scheme [15] specifies that the data storage correctness can be achieved by using SMDS (Secure Model for cloud Data Storage). It specifies that the data storage correctness can be achieved in 2 ways as 1) without trusted third party 2) with trusted third party based on who does the verification.
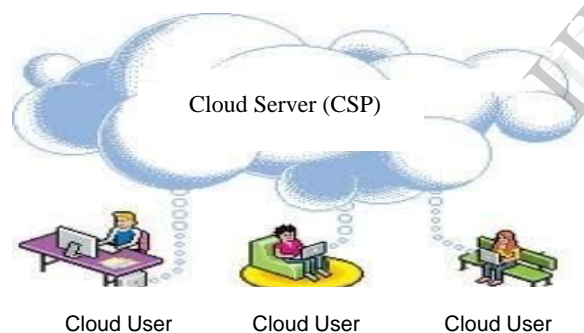

Fig 1: Cloud Architecture

It provides data confidentiality in two stages as 1) Data at rest 2) Data in transmission.

1) *Data at rest:* Symmetric key encryption technique (i.e. AES, TDES, and DES) are recommended which are secure but more time consuming.

2) *Data in transmission:* Secure Socket Layer (SSL) protocol is used for integrity verification. It uses a two different hash function such as Secure Hash Algorithm (SHA1) for digital signature and Message Digest (MD5) is a cryptographic hash function which is used to check the data integrity.

Balkrishna and Hoka address problem of access control using cryptographic techniques which degrades performance and increase the computation cost of managing all keys at Cloud Serever and at the user[6][12]. They proposed Diffie Hellman key exchange scheme for sharing symmetric key securely. Researchers of [15] specify way to achieve storage correctness without Trusted Third Party (TTP). Following are major goals of proposed schemes as

- CS neither should learn any information from user's data nor should misuse the same.
- The User selects the encryption option for their data
- Secure key management
- Flexible access right managements
- It aims to achieve light weight integrity verification process for checking the unauthorized change in the original data without requesting a local copy of the data.

It uses public key encryption to encrypt the data to data storage correctness. It achieves the following goals as data confidentiality, security, light weight verification, key management, access right and no data duplication.

The proposed scheme is compared with different cloud service providers like cloudseal, cloud zone, Venus & EPPS. It uses symmetric encryption which provides confidentiality, integrity, verification with low cost. It also provides authentication for data owner and access control through which only authorized user can access the data.

The correctness of data can be violated due to a broad range of both internal and external threats and CSP may hide data loss or damage from users to maintain a reputation. Major security issues associated with cloud user and CSP are as follows

1) *Cloud Service Provider (CSP):* Organization or enterprises provide various services to cloud users. Confidentiality and integrity of cloud data should be maintained by CSP. The Provider should ensure that user's data and application are secured on a cloud. CSP may not leak the information or else cannot modify or access user's content. The attacker can log into network communication [11].

2) *Cloud Server (CS):* The cloud server where data being stored and accessed by cloud data owner or users. Data should not be accessed by unauthorized users, no data modification or no loss of data.

3) *Cloud User:* Attackers can access basic information like username and password [11]. Key management is major issue in encryption techniques. Data dynamic issues need to be considered by CSP.

2

Cloud Computing Threads [11] are as follows:

- Spoofing Identity Theft
- Data Tempering Threat
- Repudiation Attack
- Information Disclosure on up/download Intra-Cloud
- Denial of Service Attack
- Log In

To achieve security, we can handover our data to a third outsource party who will specify the correctness and integrity of the cloud data. Hence, new concept arrives as Third party auditor (TPA) who will audit the user data stored on the cloud, based on the user's request. In this case, the Cloud service provider doesn't have to worry about the correctness and integrity of the data. In this technique, TPA will audit the cloud data to check the integrity or correctness in two ways as: 1) Download all files and data from the cloud for auditing. This may include I/O and network transmission cost. 2) Apply auditing process only for accessing the data but again in this case, data loss or data damage cannot be defined for unaccessed data. Public audit ability allows user to check integrity of outsource data under different system & security models. We cannot achieve privacy as TPA can see the actual content stored on a cloud during the auditing phase. TPA itself may leak the information stored in the cloud which violate data security. To avoid this, Encryption technique is used where data is encrypted before storing it on the cloud.

Through this, we achieved privacy up to certain extent but which increases complex key management on user side. This technique cannot be long lasting as unauthorized user can easily access original content by using the decryption key which is easily available. Hence to achieve privacy preserving public auditing using TPA for cloud data storage security, researchers have proposed various techniques.

II.   EXISTING SYSTEM

The cloud data storage service contains 3 different entities as cloud user, Third party auditor & cloud server / cloud service provider. Cloud user is a person who stores large amount of data or files on a cloud server. Cloud server is a place where we are storing cloud data and that data will be managed by the cloud service provider. Third party auditors will do the auditing on users request for storage correctness and integrity of data.

The existing system specifies that user can access the data on a cloud as if the local one without worrying about the integrity of the data. Hence, TPA is used to check the integrity of data. It supports privacy preserving public auditing. It checks the integrity of the data, storage correctness.
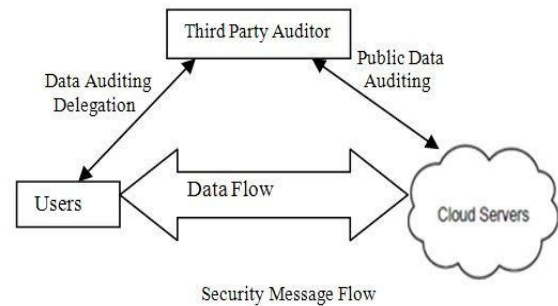

Fig 2: Architecture of Cloud Data storage service

In cloud, data is stored in a centralized form and managing this data and providing security is a difficult task. TPA can read the contents of data owner hence can modify. The reliability is increased as data is handled by TPA but data integrity is not achieved. It uses encryption technique to encrypt the contents of the file.

TPA checks the integrity of the data stored on a cloud but if the TPA itself leaks the user's data. Hence the new concept comes as auditing with zero knowledge privacy where TPA will audit the users' data without seeing the contents. It uses public key based homomorphic linear authentication (HLA) [1], [15] which allows TPA to perform auditing without requesting for user data. It reduces communication & computation overhead. In this, HLA with random masking protocol is used which does not allow TPA to learn data content.

A. Goals

- It allows TPA to audit users' data without knowing data content
- It supports batch auditing where multiple user requests for data auditing will be handled simultaneously.
- It provides security and increases performance through this system.

B. Design Goals

1) *Public audit ability:* Allows third party auditor to check data correctness without accessing local data.
2) *Storage Correctness:* The data stored on a cloud is as it. No data modification is done.
3) *Privacy preserving:* TPA can't read the users' data during the auditing phase.
4) *Batch Auditing:* Multiple users auditing request is handled simultaneously.
5) *Light Weight:* Less communication and computation overhead during the auditing phase.

3

C. Batch Auditing

It also supports batch auditing through which efficiency is improved. It allows TPA to perform multiple auditing task simultaneously and it reduces communication and computation cost. Through this scheme, we can identify invalid response. It uses bilinear signature (BLS proposed by Boneh, Lynn and Shacham) to achieve batch auditing. System performance will be faster.

D. Data Dynamics

It also supports data dynamics where user can frequently update the data stored on a cloud. It supports block level operation of insertion, deletion and modification. Author of [6] proposed scheme which support simultaneous public audability and data dynamics. It uses Merkle Hash Tree (MHT) which works only on encrypted data. It [11] uses MHT for block tag authentication.

III.   LITERATURE SURVEY

A. MAC Based Solution

It is used to authenticate the data. In this, user upload data blocks and MAC to CS provide its secret key SK to TPA. The TPA will randomly retrieve data blocks & Mac uses secret key to check correctness of stored data on the cloud. Problems with this system are listed below as

- It introduces additional online burden to users due to limited use (i.e. Bounded usage) and stateful verification.
- Communication & computation complexity
- TPA requires knowledge of data blocks for verification
- Limitation on data files to be audited as secret keys are fixed
- After usages of all possible secret keys, the user has to download all the data to recomputed MAC & republish it on CS.
- TPA should maintain & update states for TPA which is very difficult
- It supports only for static data not for dynamic data.

B. HLA Based Solution

It supports efficient public auditing without retrieving data block. It is aggregated and required constant bandwidth. It is possible to compute an aggregate HLA which authenticates a linear combination of the individual data blocks.

C. Privacy Preserving Public Auditing Proposed by CongWang

Public auditing allows TPA along with user to check the integrity of the outsourced data stored on a cloud & Privacy Preserving allows TPA to do auditing without requesting for local copy of the data. Through this scheme [1], TPA can audit the data and cloud data privacy is maintained. It contains 4 algorithms as

1) *Keygen:* It is a key generation algorithm used by the user to setup the scheme.
2) *Singen:* It is used by the user to generate verification metadata which may include digital signature.
3) *GenProof:* It is used by CS to generate a proof of data storage correctness.
4) *Verifyproof:* Used by TPA to audit the proofs

It is divided into two parts as setup phase and audit phase.

1) *Setup Phase:* Public and secret parameters are initialized by using keygen and data files f are preprocesses by using singen to generate verification metadata at CS & delete its local copy. In preprocessing user can alter data files F.
2) *Audit Phase:* TPA issues an audit message to CS. The CS will derive a response message by executing Genproof. TPA verifies the response using F and its verification metadata.

TPA is stateless i.e. no need to maintain or update the state information of audit phase. Public key based homomorphic linear authentication with random masking technique is used to achieve privacy preserving public auditing. TPA checks the integrity of the outsourced data stored on a cloud without accessing actual contents. Existing research work of proof of retrievability (PoR) [12] or Proofs of Data Possession (PDP) technique doesn't consider data privacy problem. PDP scheme first proposed by Ateniese et al. used to detect large amount corruption in outsourced data. It uses RSA based Homomorphic authentication for auditing the cloud data and randomly sampling a few blocks of files. A Second technique proposed by Juels as Proofs of retrievability (PoR) allows user to retrieve files without any data loss or corruptions. It uses spot checking & error correcting codes are used to ensure both "Possession" and "Retrievability". To achieve Zero knowledge privacy, researcher [15] proposed Aggregatable Signature Based Broadcast (ASBB).

It provides completeness, privacy and soundness. It uses 3 algorithms as Keygen, Gentag and Audit.

D. Using Virtual Machine

Abhishek Mohta proposed Virtual machines which uses RSA algorithm, for client data/file encryption and decryptions [5]. It also uses SHA 512 algorithm which makes message digest and check the

4

data integrity. The Digital signature is used as an identity measure for client or data owner. It solves the problem of integrity, unauthorized access, privacy and consistency.
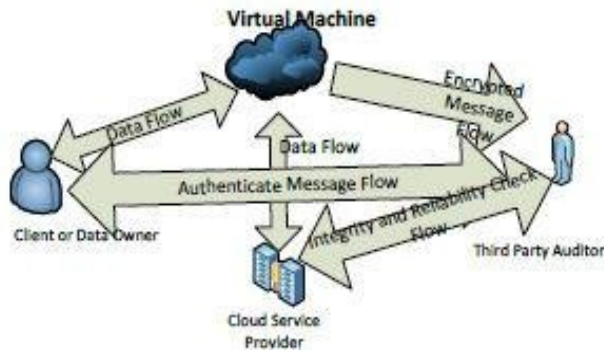

Fig 3: Architecture of Cloud server with CU and TPA

E. Non Linear Authentication

D. Shrinivas suggested Homomorphic non linear authenticator with random masking techniques to achieve cloud security [14]. K. Gonvinda proposed digital signature method to protect the privacy and integrity of data [4]. It uses RSA algorithm for encryption and decryption which follows the process of digital signatures for message authentication.

F. Using EAP

S. Marium proposed use of Extensible authentication protocol (EAP) through three ways hand shake with RSA. They proposed identity based signature for hierarchical architecture. They provide an authentication protocol for cloud computing (APCC) [11]. APCC is more lightweight and efficient as compared to SSL authentication protocol. In this,
Challenge –handshake authentication protocol (CHAP) is used for authentication. When make request for any data or any service on the cloud. The Service provider authenticator (SPA) sends the first request for client identity. The steps are as follows

1) When Client request for any service to cloud service provider, SPA send a CHAP request / challenge to the client.
2) The Client sends CHAP response/ challenges which is calculated by using a hash function to SPA
3) SPA checks the challenge value with its own calculated value. If they are matched then SPA sends CHAP success message to the client.

Implementation of this EAP-CHAP in cloud computing provides authentication of the client. It provides security against spoofing identity theft, data tempering threat and DoS attack. The data is being transferred between client and cloud providers. To provide security, asymmetric key encryption (RSA) algorithm is used.

B. Dhiyanesh proposed Mac based and signature based schemes for realizing data audit ability and during auditing phase data owner provides a secret key to cloud server and ask for a MAC key for verification [11].

C. Wang proposed an effective and flexible distributed schemes as Homomorphic token with distributed verification of erasure coded data proposed scheme achieves an integration of storage correctness insurance and data error localization i.e. identification of misbehaving server [13].
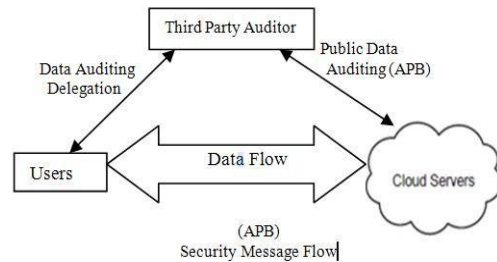

Fig 4: Automatic Protocol Blocker

G. Using Automatic Protocol Blocker

Balkrishna proposed efficient reed Solomon technique for error correction which check data storage correctness[13].

Kiran Kumarproposed automatic protocol blocker to avoid unauthorized access [14]. When an unauthorized user access user data, a small application runs which monitors user inputs, It matches the user input, if it is matched then it allow user to access the data otherwise it will block protocol automatically. It contains five algorithms as keygen, SinGen, GenProof, VerifyProof, Protocol Verifier. Protocol Verifier is used by CS. It contains three phases as Setup, Audit and PBlock.

H. Random Masking Technique

Jachak K. B. proposed privacy preserving Third party auditing without data encryption. It uses a linear combination of sampled block in the server's response is masked with randomly generated by a pseudo random function (PRF) [15].
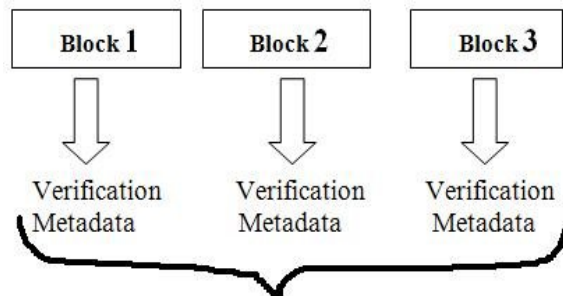

Fig 5: Homomorphic Authenticator

5

Researchers of [14] use the concept of virtual machines, The RSA algorithm is used to encode and decode the data and SHA 512 algorithm is used for message digest which check the integrity of information

Dr. P.K. Deshmukh uses the new password at each instance which will be transferred to the mail server for each request to obtain data security and data integrity of cloud computing [14]. This protocol is secure against an untrusted server as well as third party auditor. Client as well as trusted third party verifier should be able to detect the changes done by the third party auditor. The client data should be kept private against third party verifier. It supports public verifiability without help of a third party auditor. This protocol does not leak any information to the third party verifier to obtain data security. This proposed protocol is secure against the untrusted server and private against third party verifier and support data dynamics. In this system, the password is generated and that will be transferred to email address of the client. Every time a key is used to perform various operations such as insert, update delete on cloud data. It uses time based UUID algorithm for key generation based on pseudo random numbers. If an intruder tries to access the users' data on a cloud, that IP address will be caught and transferred to the user so that user will be aware of.

I. Analysis of protocol proposed by C. Wang which contains security flaws:

Researchers of [10] analyses the Protocol proposed by Wang et al and find security flaws in their protocol. A Public auditing protocol is a collection of 4 polynomial time algorithm as (Keygen, TagBlock, Genproof, and CheckProof)

Keygen: User executes Keygen for key generation.

TagBlock: User executes TagBlock to produce verification metadata.

Genproof: Cloud server executes Genproof for proof of possession.

CheckProof: TPA will validate a proof of possession by executing CheckProof.

The Problem with this system is that cloud server might be malicious which might not keep data or might delete the data owned by cloud users and might even hide the data possessions.

1) Data modification tag forging attacks
2) Data lost auditing pass attack
3) Data interception and modification attack
4) Data Eavesdropping and Forgery

This protocol is vulnerable to existential forgeries known as message attack from a malicious cloud server and an outside attacker. The analysis shows that they are not providing any security for cloud data storage.

IV.   PROPOSED SCHEME

The data on the cloud has a minimum concern about sensitive information such as social security number, medical records, bank transaction and shipping manifests for hazardous material. We provide additional security such as watermark technique. These techniques enable single sign-on in the cloud and access control for sensitive data in both public and private clouds.
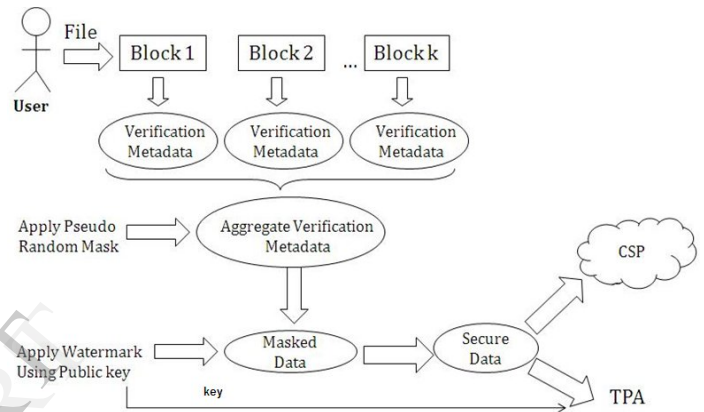

Fig 6: Architecture of proposed system

In the Proposed system we used water marking process, to store the data or images in the cloud server by assigning the public key, and this key and watermarking images are sent to third party and third party have complete authority to check the key and sent it to the server, and there Third Party Auditor must have a public key whenever the data to be retrieved. In the watermarking process, the security level is very high so the data or images cannot be identified by the attackers in the cloud. We also use Compression technique for watermark image to reduce communication overhead.
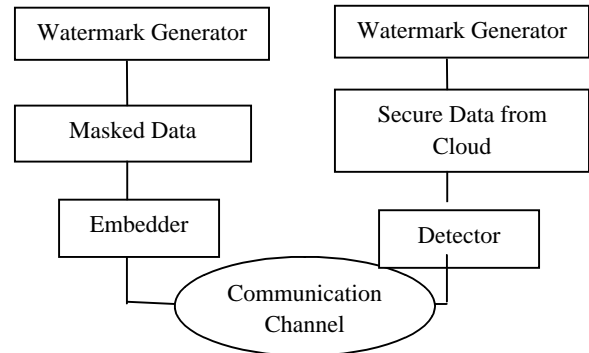

Fig 7: Watermarking Technique

6

The main elements in watermarking process: an embedded, a communication channel and a detector and is shown in Figure. Watermark information is embedded into original image itself, and it is performed in the encryption process for making security on original information. Embedded is similar to encryption process which is used to change content into another format with the help of the secret key. Detector process is also similar to decryption process which is used to perform reverse process of encryption. The watermark information is embedded within the original image before the watermarked image is transmitted over the communication channel, so that the watermark image can be detected at the receiving end.

V.   CONCLUSION

In this paper, we proposed watermarking technique to conceal the data stored in the cloud Cloud computing security is a major issue that needs to be considered**.** Using TPA, We can verify the correctness and integrity of data stored on a cloud. It uses public key based homomorphic linear authentication (HLA) protocol with random masking to achieve privacy preserving data security. We achieved zero knowledge privacy through random masking technique. It supports batch auditing where TPA will handle multiple users request at the same time which reduces communication and computation overhead. It uses bilinear signature to achieve batch auditing. It also supports data dynamics. It uses Merkle Hash Tree (MHT) for it.

REFERENCES

[1] C wang, Sherman S. M. Chow, Q. Wang, K Ren and W. Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage",IEEE Trasaction on Computers I, vol. 62, no. 2, pp.362-375 , February 2013.

[2] Dr. P. K. Deshmukh, Mrs. V. R. Desale, Prof. R. A. Deshmukh, "Investigation of TPA (Third Party Auditor Role) foe Cloud Data Security", International Journal of Scientific and Engineering Research, vo. 4,no. 2,ISSn 2229-5518, Feb 2013.

[3] KunalSuthar, Parmalik Kumar, Hitesh Gupta, "SMDS: secure Model for Cloud Data Storage", International Journal of Computer applications, vol56, No.3, October 2012

[4] K Govinda, V. Gurunathprasad and H. sathishkumar, "Third Party Auditing for Secure Data Storage in Cloud Through Digital Signature Using RSA", International Journal of Advanced science and Technical Research, vol 4,no. 2, ISSN: 2249-9954,4 August 2012

[5] AbhishekMohta, Lalit Kumar Awasti, "Cloud Data Security while using Third Party Auditor", International Journal of Scientific & Engineering Research, Volume 3, Issue 6, ISSN 2229-8 June 2012.

[6] Balkrishnan. S, Saranya. G, Shobana. S and Karthikeyan. S, "Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud", International Journal of computer science and Technology, vol. 2, no. 2, ISSN 2229-4333 (Print) | ISSN: 0976-8491(Online), June 2012

[7] Lingaraj Dhabale, Priti Pavale, "Providing Secured Data Storage by Privacy and Third Party Auditing In Cloud", International Conference on Computing and Control Engineering, ISBN 978-1-2248-9, 12 & 13 April, 2012

[8] Jachak K. B., Korde S. K., Ghorpade P. P. and Gagare G. J. ,"Homomorphic Authentication with Random Masking Technique Ensuring Privacy & Security in Cloud Computing", Bioinfo Security Informatics, vol. 2, no. 2, pp. 49-52, ISSN. 2249-9423, 12 April 2012

[9] K. Kiran Kumar, K. Padmaja, P. Radha Krishna, "Automatic Protocol Blocker for Privacy-Preserving Public Auditing in Cloud Computing", International Journal of Computer science and Technology, vol. 3 pp, ISSN. 0976-8491(Online), pp. 936-940, ISSN: 2229-4333 (Print), March 2012

[10] Gayatri. R, "Privacy Preserving Third Party Auditing for Dynamic Data", International Journal of Communication and engineering, vol. 1, no. 1, issue: 03, March 2012

[11] S. Marium, Q. Nazir, A. Ahmed, S. Ahthasham and Aamir M. Mirza, "Implementation of EAP with RSA for Enhancing The Security of Cloud Computig", International Journal of Basic and Applied Science, vol 1, no. 3, pp. 177- 183, 2012

[12] D. Shrinivas, "Privacy-Preserving Public Auditing in Cloud Storage security", International Journal of computer science nad Information Technologies, vol 2, no. 6, pp. 2691-2693, ISSN: 0975-9646, 2011

[13] C. Hota, S. Sanka, M. Rajarajan, S. Nair, "Capabilitybased Cryptographic Data Access Control in Cloud Computing", in International Journal of Advanced Networking and Applications, Volume 01, Issue 01, 2011.

[14] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public auditing for storage security in cloud computing," in Proc.of IEEE INFOCOM'10, March 2010.

[15] Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, http://www.cloudsecurityalliance.org