Multi-Level Data Integrity Service

Sukhpreet Kaur
M.Tech Student
Lovely Professional University
Phagwara, India

ABSTRACT
Cloud computing is an auspicious computing model that empowers on-demand network access to a shared pool of computing resources. One of the services offered by cloud is moving data into the cloud. Cloud Providers offer storage space to consumers for storing their data. Data moves between cloud providers and consumers can be compromised by any malicious user. This new paradigm of data storage service introduces new security challenges. Therefore, an independent service for data integrity auditing is required that allows consumers to verify true integrity and authenticity without compromising the data privacy. In this paper, we have presented a grid based multilevel data integrity service for cloud users.

Keywords
Cloud Computing, Data Integrity, Multi-Level

1. INTRODUCTION
Cloud computing is a very vast concept in today's world. It provides storages, computing, and software as a service to the users on demand basis. "Cloud computing is referred as a combination of interdependent and virtualized machines that are forcibly allocated and represented as single or multiple integrate computing assets" [4]. The problem in cloud computing is security of data integrity. As the data moves in the cloud so ensuring the security at the data level is salient so that organizations who uses clouds must sure that data is guarded wherever it goes [1][5]. Security is considered as a huge issue for the cloud. The survey found that 58 percent of the general population and 86 percent of senior business leaders are excited about the potential of cloud computing, more than 90 percent of these same people are concerned about the security, and privacy of their own data in the cloud. There is a probability where a malicious user can invade the cloud by impersonating a genuine user, there by infecting the entire cloud thus troubling many customers who are sharing the contaminated cloud [2][7]. Data integrity is one of the main security problem which is faced by cloud computing. Any user can access the data resides in a cloud from any location. Cloud does not differentiate between a sensitive data from a common data thus enabling anyone to access those sensitive data's. So there is a lack of data integrity in cloud computing [8].Data integrity is defined as the accuracy and consistency of stored data. The services provided by the cloud should ensure data integrity. As cloud vendor provides storage services on demand basis, but it's lacking of offering assurance of data integrity may avoid its wide adoption by both enterprise and individual cloud users [3][6]. As a result of the importance of Data integrity in Cloud computing this paper focuses on maintaining the authenticity and Data integrity aspect of Cloud computing environment. It proposes a Grid based multilevel Data Integrity Service instead of providing a single level service. The purpose of the proposed new service is to address the security and privacy risks challenges in Cloud computing. We examined two security factors in our proposed service, namely Data Authenticity and Data integrity.

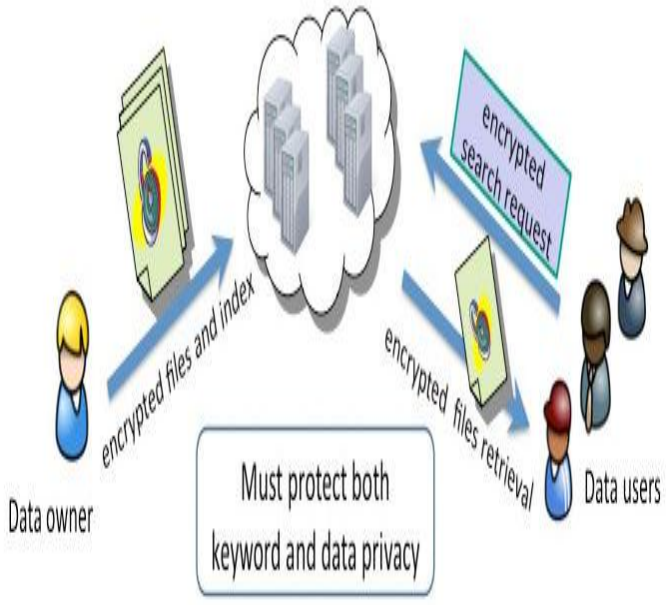

Fig 1: Data Integrity Process

The rest of this paper is organized as follows: Section 2 discusses background and related work. Section 3 describes methodology. Section 4 describes implementation. Section 5 presents results. Section 6 concludes paper and proposes future research directions. Section 7 describes the References.

2. BACKGROUND AND RELATED WORK
Cloud storing facilities are not protected by natural surroundings. There are various risks involved in it: threat of data expose to unauthorized users on the cloud or by the cloud supplier themself (data privacy); to data altering by the unauthorized user on the cloud or by the cloud worker themself (data reliability) and to disowning of data by unauthorized users on the cloud or by the cloud worker themself (data availability). This section provides a summary of related methods and techniques applied in cloud computing environment [4].

2.1 Cloud- RAID Technique
Cloud-RAID used to enhance the confidentiality, integrity and obtainability of data kept in repository of cloud. In this, the data is distributed among several cloud service providers to achieve the repetition of data among several cloud service providers to reduce the failure or loss of data. AES, SHA algorithms are used for this purpose. Encryption is also performed for maintain the integrity of data and keys used for

encryption must be kept confidential [9]. The RAID Techniques are used to overcome various limitations in cloud storage such as:

- **Security:** As provider might be honorable, but mischievous insider's grounds security. This problem can be tackle by encoding and encrypting the innovative data and by dispensing the fragments evidently through multiple providers. In this way, no one of the storage retailers is in a complete ownership of the client's data.
- **Data lock-in:** The fragments of data are stored among several providers taking into concern user expectations concerning the price and availability of the hosted content. Only a fraction of the total quantity of data is stored on every cloud provider. So, transferring one provider for another cost only a fraction of what it would be else.
- **Service Availability:** Administration of computing assets by a single company infers the threat of a single point of failure. This issue is solved by storing the data on multiple clouds [10]. Whereby no single entire copy of the data resides in one position and only a subset of providers requires being obtainable in order to recreate the data.

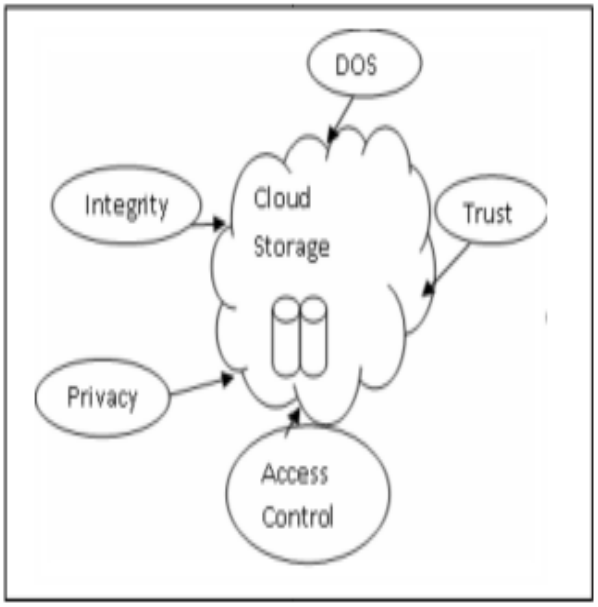

Fig 2: Cloud Computing Issues

2.2 Integrity Check Algorithm

It is a fine grained data integrity check scheme. This scheme is based on hash value of the assets. The hash value of initial data is stored and after the data is transmitted, on the receiver side again the hash value of data is calculated. After that the previously calculated hash value matches with the hash value calculated on receiver side. If both the hash value matches it means the data is not compromised by third party. Fine grained means that huge amount of data is split into small parts and then the integrity check method is applied on it. This strategy can shorten check value for data integrity to lessen storage and increase check proficiency of multi-error data objects [11].

2.3 Homomorphic Encryption Method

This method relies on elliptic curve cryptography. A data ownership method is applied to maintain various actions performed on data. The third party auditor is responsible for confirming and reconstructing the data on account of client. Elliptic curve cryptography is a public key cryptography. The security structure centered on elliptical curve cryptography contains four algorithms [13]:

- **key generation algorithm** to create public and private key
- **Signature generation algorithm** where signature is created for each block of data whose integrity is to be checked.
- **proof generation algorithm**
- **Verify proof algorithm**

This security system produces the signature based on elliptic curve cryptography algorithms which is then used to maintain the confidentiality and integrity of data. It is also responsible for maintaining the reliability of data in a periodic manner. The Elliptic curve cryptosystem security scheme (ECCSS) is recommended to guarantee the reliability of information on the distant server. This strategy applies the conceptions of the provable data possession (PDP) to create the data procedures vibrant. The customer can review the information on the cloud themself instead of taking it each time. Homomorphic encryption is implemented on the signs to enhance to the level of safety. This permits the customer to do its tasks on the cipher text themselves instead of decrypting it. Client encrypts the info and delivers the public key to the TPA, who conveys out the procedures on its behalf [14].

2.4 Bilinear Aggregate Signature Method

To confirm the exactness of data, a Third Party Auditor (TPA), in support of the cloud consumer is used to prove the reliability of the data kept in the cloud [16]. The method of bilinear aggregate signature is used to accomplish batch auditing. Batch auditing lessens the overhead in computation in cloud.

2.4.1 Batch Auditing:

TPA holds numerous reviewing responsibilities upon numerous consumers' desires. It is very time overwhelming.

A allocation of audit is secure, if

- The secrecy of the records is sheltered beside the TPA and the CSP.
- The data holder can validate whether TPA has definitely accomplished the review task identified by the data holder.
- It uses a batch signature structure called BLS signature algorithm to complete benign batching

2.4.2 The BLS algorithm Comprises of Three Segments:

- Generation Phase: In this, source elects the private key x and public key y.
- Signing Phase: Signing segment makes the hash function and produces a sign for message m.
- Verification Phase: In this point, the receiver first calculates hash function, matches it with the signature of m, if it matches, then the communication is reliable [15].

2.5 Dynamic Password Authentication and Certificate-based Authorization

In cloud, data protection comprises results with cryptography, Public Key Infrastructure effective with Dynamic Password to guarantee the authentication, integrity and confidentiality of elaborate files and transportations. The explanation grants a horizontal level of facility accessible to all associated objects that appreciates a safety mesh, within which critical confidence is preserved [12].

2.5.1 Certificate-based Authorization

Certificates that are being delivered by a PKI capability can be used for imposing access controller in the Web surroundings. These credentials are dispensed by a certification consultant that turns as a trust midpoint in the inclusive Web surroundings. A reliable document assists as a consistent object that inaugurates an individual's individuality, authorizations and accountabilities. Belief can be observed as a series from the end consumer, to the application vendor, reliance the substructure supplier. This official document is used in grouping with the service supplier's official document to create a protected SSL connection among them, thus converting substituted data and ensuring their safety by the cloud substructure. Trusted Third Party facilities inside the cloud, leads to the creation of the required Faith level and offers ultimate resolutions to reserve the confidentiality, integrity and authenticity of data and transportations. The trustworthy third party can be depending on for [17]:

• Generating Security Domains

• Short and Great level confidentiality.

• Cryptographic Departure of Data.

• Certificate-Based Approval.

• Server and Client Authentication

2.6 Batch Audit Scheme with Dynamic Data Support

The difficulty of confirming the reliability and safety of data storing in Cloud Computing and offers an operative and elastic Batch Audit scheme with dynamic data support to diminish the overheads in computation.

2.6.1 Dynamic Data Process

There are various dynamic data operations that are to be done at cloud by user:

- Update Operation: For new data update, each data block should be updated spontaneously from previously existing value to new updated value.
- Append Operation: It is very effectual method in our suggested scheme. The procedures such as add, variation, and eliminations by client are handled by space amount order for operative ID of data reliability in cloud catalog. So if there is any such alteration/tampering by violence then customer can give guarantee to the data reliability [18].
- Deletion Operation: This deletion operation based on customer's effort on his data kept in cloud server using his login operation.

2.7 RSA based Assumption Data Integrity Check

This strategy is for data reliability checked based on popular RSA safety notion. The convenience of using this structure is that consumer doesn't need to maintain the copy of data on consumer side. It reduces the responsibility of client of maintaining the data. This data integrity check scheme is used to insure and secure the integrity of data. This strategy is used by both data holder and third party verifier to assure the reliability of data on cloud. The cryptographic method RSA is used to insure the reliability of records in cloud. This data integrity check scheme is a fusion of digital signature and verification and validation based in identity. RSA theory means that RSA problem difficult to resolve if the modulus n is adequately big and initiated casually and m is any casual integer. By using this scheme, consumer only needs to maintain the secret key on its machine rather than maintaining the carbon copy of data.

As Data reliability check is an essential safety approach in the cloud computing. This strategy presents a safe and logical strategy, which empowers not only the information holder but also a third-party verifier to validate information reliability. This strategy moreover diminishes the responsibility of keeping the data at client side.

Definition: The RSA problem (RSAP) is the following: given a positive integer n that is a product of two different odd primes p and q (n=pq), a positive integer e such that and an

Integer c, find out an integer m.

RSA Assumption: The RSA hypothesis is that the RSA Problem is hard to resolve when the modulus n is adequately huge and arbitrary originated, and the plaintext m is a casual integer.

This strategy uses variable sized blocks. It divides the data into several chunks of somewhat scope. The cause is that in the creation the direction of G (Generator) is unspecified to the server. Due to affect for safety in the framework of public authentication, the genuine information file should not be disclosed to PKG throughout authenticating manner. This strategy authenticates the chunk tags alternate of innovative data chunks in the verification process. The utmost essential feature of this method is that the actual storing releases in the customer .The consumer only require to hold the secret key excluding the unusual data in his/her PC [19].

3. METHODOLOGY

Proposed "Grid based multilevel Data Integrity Service"

The proposed scheme provides a solution for Data Integrity and Authenticity Problem in cloud environment by introducing Grid based Multilevel Data Integrity service. It is a 2-tier service which provides two levels of security to the user.

3.1 Registration

In this phase, an account is created for each user through which they can upload and download data in cloud environment. Two level securities are provided through Password and Unique Grid for each user. Following steps are followed during Registration process:

Step 1:- In this, User create an account by providing Username, Email Id, Security Question and Answer of its own.

Step 2:- In this, user will create a password of its own with mandatory condition that password includes one Uppercase, One Lowercase, One Special character and One Numeric.

Step 3:- In this, 3*3 unique Grid is generated for each user. Random values are assigned to each column in a Grid using a random function. This 3*3 grid is not directly saved at the server end. Only the hash values with 84 combinations of grid are saved at server end for every user.

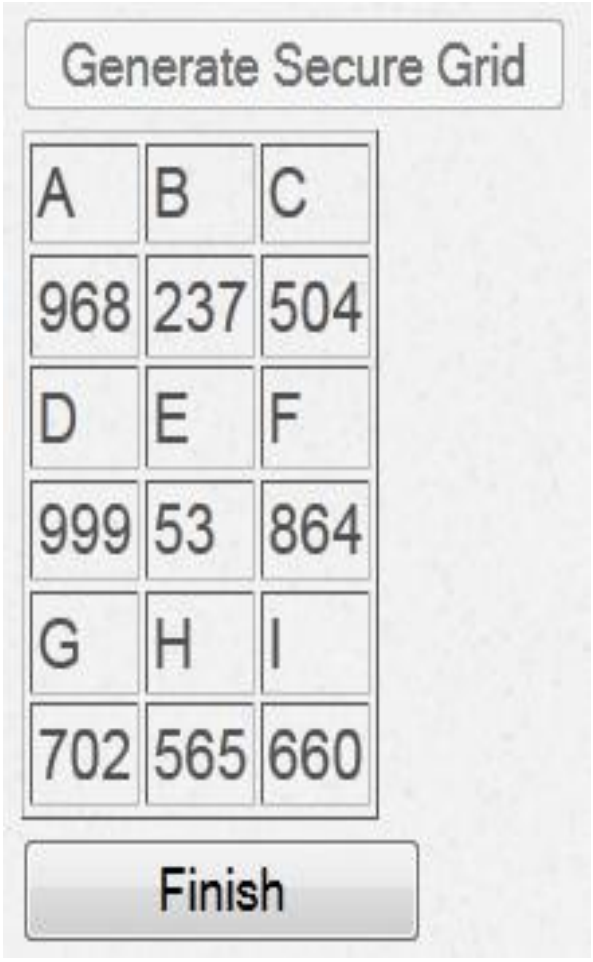

Fig 3: Snapshot of Grid Generation

Step 4:- After a generation of Grid, an email is sent to user which includes registration detail in it. Security answer saved in database as a hash value. Now, user is responsible for privacy of data inside the email (i.e. it's upon user to maintain the secrecy of grid from others).

3.2 Login

After the registration process, only registered users are allowed to access this service. In this phase, 2-tier securities are provided to user. Following steps are followed during Login process.

Step 1:- This is the first level of security. In this step, authenticated users login into their accounts by entering Username and Password. Only three attempts are provided to user for entering password. If user enters each of three time wrong password, then user account suspended temporarily. For getting access again to his/her account user has to click on Forgot Password tab.

Step 2:- We use SHA-3 algorithm for maintaining Data Integrity. During uploading, hash of the file is generated by SHA-3 algorithm and three random hash values from grid will be attached with it. During downloading, first grid values will be asked by the system and if user enters correct values only then file will be downloaded. Only three attempts will be provided to user for downloading a file. After this, hash of the file compared and if file retrieved properly then that means data integrity maintained.

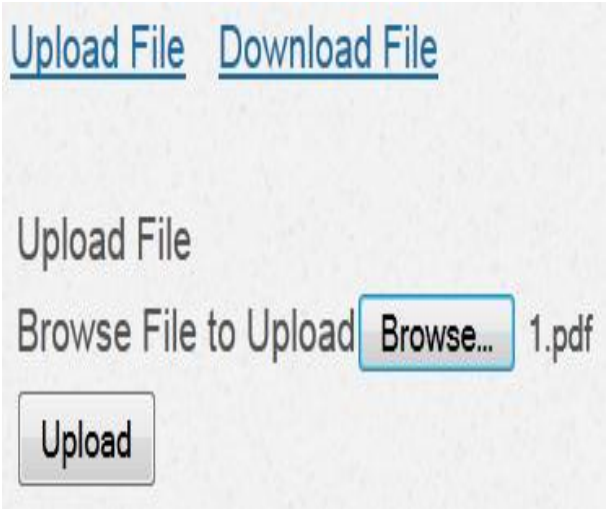

Fig 4: Snapshot of Downloading File Using Grid Hash Values

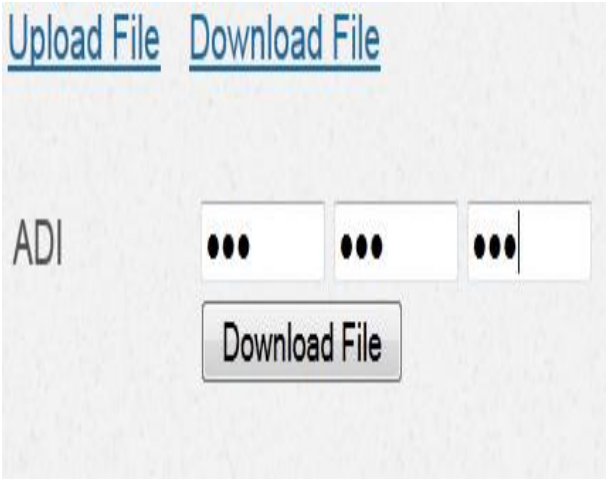

Fig 5: Snapshot of Downloading File Using Grid Hash Values

3.3 Forgot Password

If user fails to provide right password at the time of login or right values of grid then account of the user suspended and grid will be dropped permanently. For accessing account again user has to use the Forgot password phase. In this, two options are provided to user:

User has to enter either Email or Security question & answer. In email, user has to enter the email id and then a link for changing password and generating a new grid is sent on to that mail which is provided by user at the time of registration. If user enters security question and answer correctly then without sending a link on the mail, user will be automatically take to the change password and generate new grid webpage.

4. IMPLEMENTATION

For implementation, we use .Net framework in which front-end is Asp.net and back-end is SQL server 2008 R2 edition. In Asp.net we designed interface for users to register themselves. After registration they can login there accounts and access service options (i.e. uploading and downloading files). There is an option of forgot password which will serve as a back-up for users account. With the help of this users again get access to their accounts which were suspended by service provider due to security reasons. In case of SQL server, 84 combinations of hash of grid values are stored for each user separately in a separate table. Each user has its own table in which hash of grid values are stored.

5. RESULTS

5.1 No of Possible Combinations of a Grid:

$$^{999}C_9$$

$$= 2.8309511397 * 10^{21} \text{ or}$$

$$2830951139700000000000$$

This value shows that this number of different grid combinations (i.e. From A to I) are possible. The grid combination remains same for each user but value inside combinations will be different for each and every user.

5.2 No of Possible different Grids for Each User:

$$^{999}C_9 * {}^{9}C_9$$

$$= 6.882042845 * 10^{25} \text{ or}$$

$$68820428450000000000000000$$

This value represents the number of grids that can be provided to number of users and each grid is unique to a specific user. Each user has to maintain its own grid for accessing this service.

5.3 No of Possible Breaking Point:

$$^{9}C_3 + {}^{999}C_3$$

$$= 1.658341675 * 10^{14} \text{ or}$$

$$165834167500000$$

This value represents the number of tries that a brute force has to apply for breaking grid combinations. But our service provides only three tries to each user for uploading or downloading files. If user enters each time a wrong value then his or her grid will be deleted permanently and that user has to generate a new grid for accessing this services. So, it is very difficult to break a grid by using brute force attack.

Multiple users can upload and download their files. Above values show that these many grids can be created and in case of breaking point it is such a large numbers that it cannot be broken by brute force. During uploading, hash of grid values are attached randomly with files and during downloading user has to enter those three values correctly. If user enters correct value then he/she able to download the file and if user enter wrong values then file will not be download. Only three attempts are given to user for downloading a file. Data integrity is maintained in a manner that no one else can download file without the grid and even at the server end only the hash values are kept and hash values are attached with file while upload so even if someone tries to download the file at server end, to access the file grid authentication needs to be provided and therefore the integrity is maintained.

6. CONCLUSION AND FUTURE WORK

There are three basic characteristics of cloud data security: confidentiality, integrity and availability. As the data resides in cloud and moves on demand basis then there is major threat for integrity of data in cloud. In this paper, we proposed a 2-tier Data integrity service to address the data integrity fears in the cloud storing provision. The service consists of two levels: Username-Password and Unique Grid. This service provides full privacy and security to user data because grid is not directly saved at server end. Values of grid are hashed and saved at server, so, even at the server end no one can know the values of grid. In case, if server gets compromised then integrity of user's data still maintained with this service. The purpose is to provide authenticity and integrity in public cloud.

In future, different combinations for grid can be made like 2*3 matrixes, 4*4 matrixes for enhancing the security of data.

7. REFERENCES

[1] V. Nirmala, R.K. Sivanandhan, Dr. R.Shanmuga Lakshmi "Data Confidentiality and Integrity Verification using User Authenticator scheme in Cloud" Green High Performance Computing, 2013.

[2] A M Talib, Rodziah Atan, Rusli Abdullah, Masrah Azrifah, Azmi Murad "CloudZone: Towards an Integrity Layer of Cloud Data Storage Based on Multi Agent System Architecture" IEEE Conference on Open Systems, 2011.

[3] slideshare, Available: http://www.slideshare.net/ronak2454/issues- in-cloud-computing.

[4] RajKumar Buyya, James Broberg, Andrzej Goscinski "Cloud Computing Principles and Paradigms" (1st ed.). Hoboken, New Jersey, USA: Wiley, 2011.

[5] Lee Kangchan "Security Threats in Cloud Computing Environments" International Journal of Security and Its Applications Vol. 6, No. 4, October, 2012.

[6] B. Priyadharshini, P. Parvathi "Data Integrity in Cloud Storage" IEEE International Conference On Advances In Engineering, Science And Management (ICAESM-2012) March 30, 31, 2012.

[7] Meena S, Esther Daniel, Dr. N. A Vasanthi "Survey on Various Data Integrity Attacks in Cloud Environment And the Solutions" International Conference on Circuits, Power and Computing Technologies, 2013.

[8] Irfan Gul, Atiq ur Rehman, Islam M Hasan "Cloud Computing Security Auditing".

[9] Maxim Schnjakin, Dimitri Korsch, Martin Schoenberg, Christoph Meinel "Implementation of a Secure and Reliable Storage above the Untrusted Clouds" the 8th International Conference on Computer Science & Education (ICCSE 2013) April 26-28, 2013. Colombo, Sri Lanka.

[10] Rizwana Shaikh, M. SasiKumar "Security Issues in Cloud Computing: A Survey" International Journal of Computer Applications, Volume 44 - No 19, April, 2012.

[11] Yun Yang, Lie Wu, Yulin Yan, Cong Xu "Fine-Grained Data Integrity Check for Power Cloud Computing" 5th International Conference on BioMedical Engineering and Informatics, 2012.

[12] Osama Harfoushi, Bader Alfawwaz, Nazeeh A. Ghatasheh, Ruba Obiedat, Mua'ad M. Abu-Faraj, Hossam Faris "Data Security Issues and Challenges in Cloud Computing" Communications and Network, 2014.

[13] Tamal Kanti Chakraborty, Anil Dhami, Prakhar Bansal, Tripti Singh "Enchanced Public Auditability & Secure Data Storage in Cloud Computing" 3rd IEEE International Advance Computing Conference, 2013.

[14] Kevin Hamlen, Murat Kantarcioglu, Latifur Khan, Bhavani Thuraisingham "Security Issues for Cloud Computing" International Journal of Information Security and Privacy, April-June 2010.

[15] Miss. M. Sowparnika1, Prof. R. Dheenadayalu "Improving data integrity on cloud storage services" International Journal of Engineering Science, 2013.

[16] Rohit Bhadauria, Sugata Sanyal "Survey on Security Issues in Cloud Computing and Associated Mitigation Techniques" International Journal of Computer Applications, Volume 47– No.18, June 2012.

[17] Saurabh Sharma, Prof. Ashok Verma, Prof. Satpal Singh, Vimmi Pandey "Data Protection in the Cloud: dynamic Password Authentication and Certificate-Based Authorization" International Conference on Cloud, Big Data and Trust 2013, Nov 13-15, 2013.

[18] Khaba M.V and M. Santhanalakshmi, "Remote Data Integrity Checking in Cloud Computing" International Journal on Recent and Innovation Trends in Computing and Communication, June 2013.

[19] Zhang Jianhong and Chen Hua "Security Storage in the Cloud Computing: A RSA- based Assumption Data Integrity Check without Original Data" International Conference on Educational and Information Technology, 2010.