DOI: http://dx.doi.org/10.26483/ijarcs.v8i7.4234

ISSN No. 0976-5697

Volume 8, No. 7, July – August 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

DATA INTEGRITY CHECKING FOR SECURING CLOUD STORAGE USING MODIFIED BLOWFISH ALGORITHM

Dr. G.Sekar
Asst. Prof. PG & Research Dept. of Computer Science
Dr. Ambedkar Govt. Arts College, Vyasarpadi
Chennai, India

A.Sathya
Research Scholar, PG & Research Dept. of Comp. Science
Dr. Ambedkar Govt. Arts College, Vyasarpadi
Chennai, India

Abstract: Cloud computing, an emerging technology is being widely used for outsourcing the data into the cloud instead of storing it in the local physical storage. It can be accessed by either individual user or group of users. Cloud service providers must concern about the privacy, security and data integrity of the outsourced data. Service availability failure and the data loss is possible is due to the malicious intruders in the cloud environment. As the clients no longer have physical possession of data, the integrity and security of data become the major concern in the cloud computing. One of the important security concerns is to verify the integrity of data stored on cloud. To maintain the integrity of data, the user needs the help of a Third Party Auditor (TPA). The TPA checks the integrity of data bases on demand and releases audit reports that helps the user to evaluate the risk of the services they are using. TPA has the capability to check the integrity of data which is not easy for the end users. This paper elaborates new public and private auditability schemes which are used by TPA to audit the remotely stored data in the cloud and hence provide comparative analysis of existing integrity check techniques with the new technique for cloud data storage.

Keywords: Cloud computing, Data Integrity, Third Party Auditor, Public and Private Auditability, Cloud storage model.

INTRODUCTION

In traditional computing, the user can access the data only on the current system. If the user has to access the data in another system, he needs to save it in an External Storage like pen drive, CD Drive. This formidable task is avoided in the case of Cloud Computing, because the only requirement is a system with Internet and Cloud User Interface. Data is not restricted to only one system or network. Cloud services help to reduce the space and cost (both physical space and storage space) needed for data storage [1]. New era of cloud computing moves the data and computing from desktop to large datacenters.

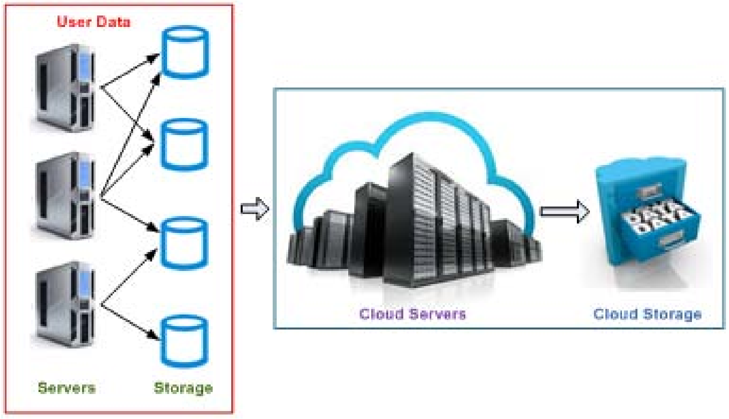

Figure 1. Data movement to the cloud

Cloud computing is combination of parallel, grid and distributed computing. A disruptive technology, like cloud computing, can impact "how" to audit the integrity of outsourced data in cloud storage. Figure 1 shows the data moving from traditional storage to cloud storage. Cloud Computing is also characterized by

- 3 Cloud Service Models
- 4 Cloud Deployment Models
- 5 Essential Characteristics

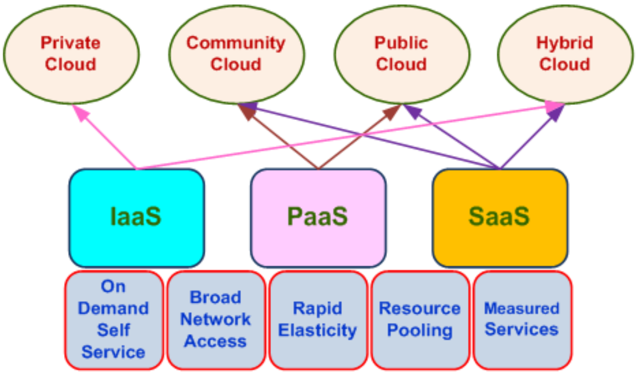

Figure 2. Visual Model of Cloud Computing

Recent years has seen the rapid growth of Big Data in Cloud Computing to effectively process (interrogate) the large data sets or outsourced files, the outsourced data integrity problems in distributed cloud environment begin to attract the attention of people. The previous research of Auditing the integrity of data stored in cloud computing has focused on single cloud, however the research of auditing the integrity of large archival files in distributed cloud storage systems has been overlooked. Figure 2 shows the cloud service models, deployment models and characteristics of cloud computing.

To deal with these auditing challenges, many classic auditing schemes of cloud storage have been proposed. Big data today often deals with very large unstructured data sets and is dependent on rapid analytics, with answers provided in seconds [2]. When comes to deal with the big data, the data integrity is very important aspect. The TPA [3] has experience in checking integrity of the data. The main goal of this research is to provide an efficient RDA (Remote Data Auditing) schemes for third party Auditor verifies the data integrity in distributed cloud server environment. Cloud storage is made up of many distributed resources, but still acts as one. The elasticity of the cloud makes it ideal for big data analytics [4].

Data auditing scheme supports an external auditor to audit the user's outsourced data in the cloud without learning knowledge on the data content. A scheme for auditing outsourced data should be both lightweight and robust. Lightweight means that it does not unduly burden the Storage Service Provider (SSP), this includes both overhead (i.e., computation cost and time) at the SSP and communication between the SSP and the auditor. This goal can be done by checking, in which the auditor randomly pick the block files and checks their integrity and thus reduce the computation time and cost.

A. Data Integrity Checking

Data integrity is an issue in data and computation related context [5]. Cloud users concern about the data should be stored correctly on the cloud server without any changes and if any violation occurs i.e. if the data is get lost or modified, it can be detected. It refers to maintaining consistency of the data all over the cycle. It must remain in the same state. But the integrity of data is at risk in cloud server. Demand efficient storage correctness guarantee without requiring local data copies. Traditional methods for integrity cannot be directly adopted. Retrieving massive data for checking is unpractical (large bandwidth) Allow meaningful tradeoffs between security and overhead causes the communication and computation costs should be low and Integrity checking cost should not outweigh its benefits.

To solve the problem of data integrity checking, many schemes are proposed under various security models. In all these works, great efforts are made to design solutions that meet to reduce communication costs and computation costs in auditing [6]. It contains protocols for data retention specifying the length of data that can be retained. These protocols are consistently and routinely applied to all data entering the system and any enforcement of relaxation will cause error in the data. Data integrity rules causes the error rates to be lower, it results the reliability of outsourced data.

B. Third Party Auditor (TPA)

Traditional cryptographic methods for integrity and accessibility of data are based on the hash functions and on signature schemes cannot work on the outsourced data missing a local copy of data. In addition, it is not a realistic solution for data validation by downloading them due to the exclusive transaction, especially for large-size files. Purpose of the audit is to correct data in a cloud environment which can be terrible and classy or expensive for the normal users. Therefore, it is critical to recognize public audit ability for Cloud Storage Service, so that data owners may cure to a third party auditor (TPA) has ability and that a user does not have to audit the cloud data [7]. This audit service is extensively significant for digital forensics and data assurance in cloud. Impact of downloading the files whose integrity have to check requires high transmission cost to maintain and to decrease the storage risk is very important to take support of a Third party auditor (TPA) who checks the data integrity for the user and helps the user in decrease his risk. Taking into consideration the job of the verifier (TPA) in the model, all the schemes presented before fall into private and public auditability [8]. Methods with private auditability can achieve higher efficiency; public auditability allows anyone, not just the owner, to challenge the cloud server for correctness of data storage while keeping no private information. Provable Data Possession (PDP) and Proof of Retrievability (PoR) Schemes support data dynamics which may lead to security loopholes [9]. Then clients are able to hand over the valuation of the service performance to an independent third party auditor (TPA), without commitment of their computation resources.

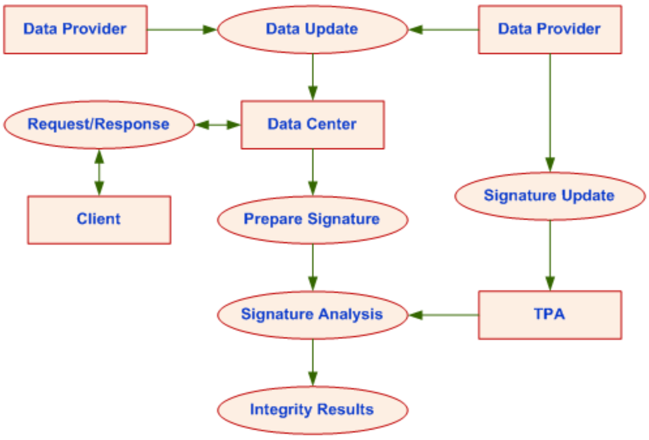

Figure 3. Working of Third Party Auditor

Common problem raised during the process of auditing for the shared data in the cloud is how to preserve identity privacy from the TPA. Public auditability allows an external party other than the user to verify the correctness of remotely stored data. Only trusted third party has the authority to check and to maintain the integrity of the data. Figure 3 shows the working of TPA.

- Third Party should audit the data from the cloud, not ask for a copy
- Third Party should not create new vulnerability to user data privacy
- TPA audit the cloud data storage without demanding the local copy of data, and introduce no additional online burden to the cloud user
- TPA can perform multiple auditing tasks simultaneously
- Achieves batch auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA

C. TPA Audit Service Architecture

There are three different network entities in cloud data audit service architecture which is users, cloud service provider and third party auditor. Figure 4 shows the basic entities of data auditing.

- **Data Owner** has data files to be stored in the cloud and relies on the cloud for data maintenance can be an individual customer or an organization.

- **Cloud Storage Service Provider (CSP)** provides data storage service and has sufficient storage space to maintain users' data.

- **TPA (Third Party Auditor)** is a trusted person who manage or monitor outsourced data under request of the data owner.TPA must be able to make regular check on the integrity and availability of these delegated data at appropriate intervals. TPA must be able to take the evidences for the disputes about the inconsistency of data in terms of authentic records for all data operations

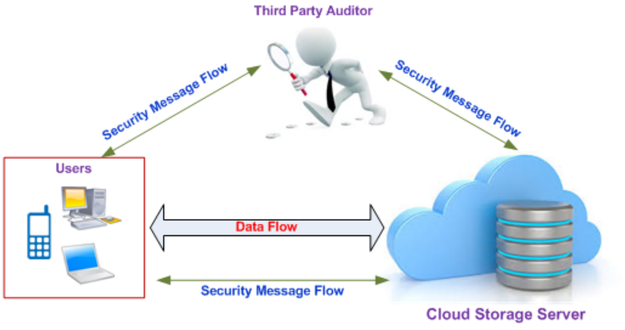

Figure 4. Basic Entities of TPA Audit Service Architecture

Auditing Protocols for TPA considered to be a following security issues:
Audit-without-downloading allows to TPA (or other clients with the help of TPA) to access and check the integrity of cloud data on demand without downloading a copy of entire data.
Verification-correctness helps to make sure that there exists no unethical CSP that can pass the audit from TPA without storing users' data.
Privacy-preserving allows to make sure that there exists no way for TPA to form user data from others during auditing process.
High-performance allows TPA to perform data auditing with minimum overheads in cloud storage, to reduce the communication and computation cost. The parameters for high performance are as follows:

1. *Scalability:* Cloud is a dynamic environment any number of users can access it. It has huge data storage on cloud server. TPA functions should not be affected by number of cloud users, servers, number of files stored on the data center or the total size of the entire storage. TPA must offer scalable architecture for users to access it.
2. *Dynamic data operation support:* Main advantage of cloud computing and other online storage system is its dynamic data support like insertion, deletion and insertion. TPA should take care of the data stored on cloud may be used or edited by multiple users simultaneously [10].
3. *Batch Auditing in* Third Party Auditor used to improve efficiency. TPA should perform more than one auditing tasks simultaneously and it also reduces communication and computation cost.
4. *Countability:* If the TPA provides a proof of misbehavior for any data, a scheme is countable.

The existing RDA schemes are as below:

- Privacy- Preserving Public Auditing Scheme [11]
- Provably secure Data Possession (PDP) - RSA based HVT [12]
- Proof of Retrievability [13]
- POR based on BLS homomorphic authentication technique
- Dynamic POR based on ORAM technique
- MuR- PDP
- Authenticated Data structure(MHT) with Bilinear aggregate signature [14]
- Multi Replica(MR) with Authenticated Data structure (MHT) [15]
- DCT (Data and Conquer Table) data structure with Algebraic signature [16]

RELATED WORK

The proposed scheme consists of three entities data owner, cloud server storage and TPA. Data owner is responsible for splitting the file into blocks of fixed size, encrypting those using modified blowfish algorithm which was stored in cloud server. When the client or data owner request for data auditing to the TPA, it immediately request for the encrypted data from the cloud server randomly. After receiving the data, it decrypted the data and compare with original data. If they matches with each other means the data is not been get lost or tampered by any attacker. If it does not match then the data integrity [17] has been affected. The result for the data integrity check is provided to the data owner.

D. Modified Blowfish Algorithm

Modified Blowfish is a 64-bit symmetric block cipher with variable length key of 64 bits. The algorithm operates with two parts: a key expansion and a data encryption. The data encryption occurs via a 14-round Feistel network. It is faster than other encryption algorithms when implemented on 32-bit microprocessors with large data sets [18]. The nature of encryption algorithms are that, once any significant amount of security analysis is done, it is very undesirable to change the algorithm for performance reasons, thereby invalidating the results of the analysis. Thus, it is imperative to consider both security and performance together during the design phase.

Table I. Comparison of algorithms based on the parameters

Algorithm	*Block Size*	*Rounds*	*Key*
AES	128,192,256 bits	10,12,14	128 bits
Blowfish	32-448 bits	16 rounds	64 bits
Modified Blowfish	64 bits	14 rounds	64 bits

Advantages of Modified blowfish algorithm:

- Blowfish is license-free and is available free for all uses.
- Blowfish is one of the fastest block ciphers developed to date.
- Blowfish suffers from weak keys problem, still no attack is known to be success.

Many cryptography algorithms have been proposed for securing outsourcing data in cloud storage. In this paper, private key cryptography or Block cipher were used to check the data integrity [19]. This symmetric cipher splits messages into blocks of 64 bits and encrypts them individually. Blowfish can be used only for securing payments efficiently to password management, where it used to protect passwords. It is definitely one of the more flexible encryption methods available. It is known for both its tremendous speed and overall effectiveness as many claim that it has never been defeated. Meanwhile, users have taken full advantage of its free availability in the public domain. Table I shows the comparison of algorithms based on the primary parameters

Algorithm: Modified Blowfish Encryption

```
Step 1: Split large file x into fixed size of blocks
Step 2: Get the block of file randomly and encrypt it
Step 3: Divide block of file content into two 32-bit
        halves: xL, xR
        For i = 1to 14:
            xL = XL XOR Pi
            xR = F(XL) XOR xR
            Swap XL and xR
            Swap XL and xR (Undo the last
            swap)
            xR = xR XOR
            xL = xL XOR P16
            Recombine xL and xR
```

Figure 5 shows the flow of the proposed system. The function F of Feistel Network is shown in Figure 6.

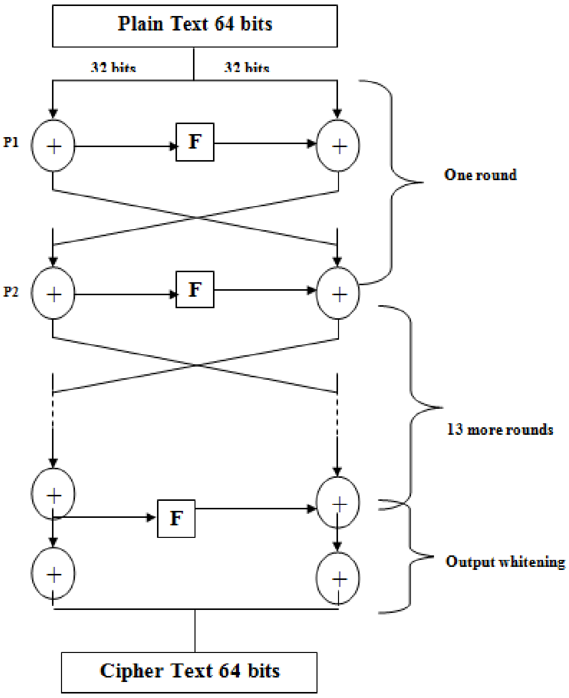

Figure 5. Overall flow of the proposed system

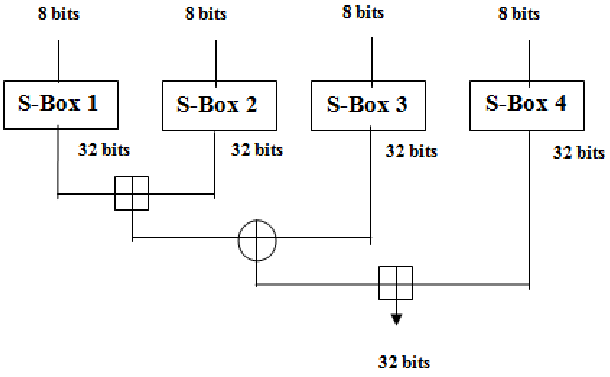

Figure 6. Feistel Network (Functions of F)

E. Performance Analysis

Performance analysis of proposed method is more efficient compare to existing AES scheme based on block size, key size and number of rounds.

Computation time:

Proposed scheme's required time to encrypt and decrypt the outsourced data considered to be as computational time. Modified Blowfish algorithm effectiveness is proved by its low computational time. Figure 7 shows the encryption/decryption time taken by AES and modified Blowfish algorithm. Figure 8 shows the variation of CPU time for both proposed and existing AES. The CPU time for AES is 1.38 ms and for proposed Modified Blowfish algorithm is 0.86ms. The optimal steps deduction in proposed scheme provides 80% in CPU time. Table II and Table III show the performance analysis based on the block size and CPU time.

i) Block size: Data and key consumes the block size in cloud architecture. Proposed scheme provides the low space consumption for block. The block size for AES is 128 bits and new scheme cipher block is 64 bits. The optimal steps deduction in proposed scheme provides 50% in block size.

ii) Key size: key size refers number of bits consumes the key generation. Proposed scheme provides the low space consumption for key size. The key size for AES is 128 bits and for new scheme is 58 bits

iii) Number of rounds: The number of rounds for AES is 16 and for new scheme is 14.

iv) Cycles / blocks: The New scheme number of cycles per block is less compared to existing scheme.

Table II. Performance analysis based on the block size

Algorithm	*File size*	*Encryption/decryption in ms (per bits)*
AES	100KB	1.37293 for 128 bits
Modified Blowfish	100KB	0.85057 for 64 bits

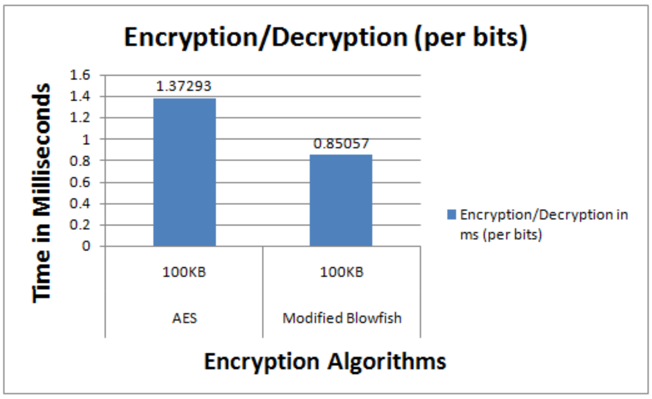

Figure 7. Block Size Analysis

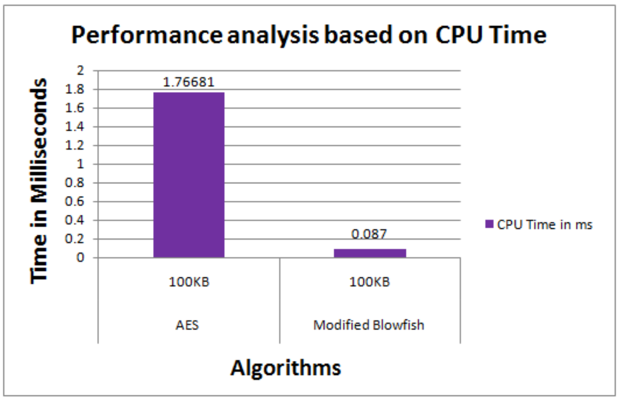

Figure 8. CPU time analysis

Table III. Performance analysis based on the CPU time

Algorithm	*File size*	*CPU Time in ms*
AES	100KB	1.76681
Modified Blowfish	100KB	0.087

The comparison of existing methods of Data auditing schemes is shown in the table IV.

Table IV. Comparison of Existing methods of Data auditing schemes for securing cloud storate

Reference papers	*Methods*	*Security pattern*		*Cryptography Model*	*Supports Batch Auditing*	*Maintaining Integrity of Data*
		Data Auditing	*Dynamic structure*			
Third party public auditing scheme for cloud storage[3]	*AES with RSA digital signature (SHA 2 for generating hash value)*	*YES*	*NO*	*AES algorithm*	*YES*	*NO*
Dynamic remote data auditing for securing big data storage in cloud computing [1]	*RSA with DCT data structure*	*YES*	*YES*	*RSA algorithm*	*YES*	*YES*
New Public Integrity Auditing Scheme for Cloud Data Storage Using Mac and Symmetric key cryptography algorithms [10]	*MAC and AES*	*YES*	*NO*	*AES algorithm*	*YES*	*YES*
Efficient ID based Public auditing for the outsourced data in cloud storage [15]	*ID Based authentication with homomorphic authenticator technique*	*YES*	*NO*	*Private key generator based auditing*	*NO*	*NO*

CONCLUSION

This paper proposed a new scheme called Modified Blowfish algorithm for checking data integrity and compared various methods to check integrity of outsourced data on cloud. It is observed that various methods for integrity checking are already available. So depending upon the file size, Third party auditor divides the file and encrypted all those segmented files. When data owner requests to check the data integrity, TPA will get the random file from segmented files and decrypt it to check the integrity. The performance of new scheme is analyzed by parameters like CPU time taken, time taken for encryption and decryption, key size etc. The results show that the proposed algorithm overrides the existing ones in proving and maintaining the integrity of data.

REFERENCES

[1] Sookhak, M., Gani, A., Khan, M. K., & Buyya, R. (2017). Dynamic remote data auditing for securing big data storage in cloud computing. Information Sciences, 380, 101-116.

[2] Desai, C. V., & Jethava, G. B. (2014). Survey on Data Integrity Checking Techniques in Cloud Data Storage. International Journal, 4(12).

[3] Swapnali, and Sangita Chaudhari. "Third Party Public Auditing scheme for Cloud Storage." Procedia Computer Science 79 (2016): 69-76.

[4] Liu, C., Ranjan, R., Yang, C., Zhang, X., Wang, L., & Chen, J. (2015). MuR-DPA: Top-down levelled multi-replica merkle hash tree based secure public auditing for dynamic big data storage on cloud. IEEE Transactions on Computers, 64(9), 2609-2622.

[5] Goyal, Renuka, and Navjot Sidhu. "Third Party Auditor: An Integrity Checking Technique for Client Data Security in Cloud

Computing." International Journal of Computer Science & Information Technologies 5 (2014).

[6] Attas, Dalia, and Omar Batrafi. "Efficient integrity checking technique for securing client data in cloud computing." IJECS 8282.6105 (2011): 11.

[7] Morghade, Rahul K., and Sonal Honale. "International Journal Of Engineering Sciences & Research Technology Data Storage Security in Cloud Computing Using Third Party Auditor (TPA)".

[8] Han, Shuai, and Jianchuan Xing. "Ensuring data storage security through a novel third party auditor scheme in cloud computing." Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on. IEEE, 2011.

[9] Hsien, W. F., Yang, C. C., & Hwang, M. S. (2016). A Survey of Public Auditing for Secure Data Storage in Cloud Computing. IJ Network Security, 18(1), 133-142.

[10] Jenifer, Ms A. Emily, and Ms S. Karthigaiveni. "New Public Integrity Auditing Scheme for Cloud Data Storage Using Mac And Symmetric Key Cryptographic Algorithms." International Journal of Applied Engineering Research 11.3 (2016): 1894-1899.

[11] Wang, Cong, et al. "Privacy-preserving public auditing for secure cloud storage." IEEE transactions on computers 62.2 (2013): 362-375.

[12] Ateniese, Giuseppe, et al. "Scalable and efficient provable data possession." Proceedings of the 4th international conference on Security and privacy in communication netowrks. ACM, 2008.

[13] Juels, Ari, and Burton S. Kaliski Jr. "PORs: Proofs of retrievability for large files." Proceedings of the 14th ACM conference on Computer and communications security. Acm, 2007.

[14] Wang, Qian, et al. "Enabling public auditability and data dynamics for storage security in cloud computing." IEEE transactions on parallel and distributed systems 22.5 (2011): 847-859.

[15] Zhang, Jianhong, and Qiaocui Dong. "Efficient ID-based public auditing for the outsourced data in cloud storage." Information Sciences 343 (2016): 1-14.

[16] Liu, Chang, et al. "X" Future Generation Computer Systems 49 (2015): 58-67.

[17] Sookhak, Mehdi, et al. "A review on remote data auditing in single cloud server: Taxonomy and open issues." Journal of Network and Computer Applications 43 (2014): 121-141.

[18] Aurora, Tanjyot, and Parul Arora. "Blowfish Algorithm." International Journal of Computer Science and Communication Engineering IJCSCE Special issue on “Recent Advances in Engineering & Technology” NCRAET 3.4 (2013): 238-243.

[19] Balakrishnan, S., et al. "Introducing effective third party auditing (tpa) for data storage security in cloud." International Journal of computer science and Technology 2.2 (2011): 397-400.