Public Auditing for Secure Cloud Storage based on DHT

Raziqa Masood[1*], Nitin Pandey[1], Q.P. Rana[2]

[1]Amity Institute of Information Technology,
Sector-125, Noida-201313 (UP), India

[2]Jamia Hamdard University,
Hamdard Nagar, New Delhi – 110062, India

[*]Corresponding author: raziqa.masood@s.amity.edu

Abstract
Today cloud computing has been the most popular service enjoyed by people due to the easy maintenance provided by it. Cloud computing is cost-efficient and people pay according to the services they use. Many organizations are using cloud storage and the reason behind it is that the outsourcing services are provided by the cloud computing. Most of people do not trust the legality of the services provided by cloud (CSPs i.e. cloud service providers) because they are afraid of the security breach of their data. The public auditing of the data by their owners is a technique that can maintain the trust of people on cloud services. This research paper is about cloud storage services based on the distributed hash table (DHT).This is required for dynamic auditing of information as this is new two-dimensional data and Third-party Auditor (TPA) is responsible for recording the information to do dynamic auditing and the dimensional data is located at TPA.
The computational costs gets reduced when the authorized information is migrated to the two-dimensional data and the Cloud service provider shifts it to the TPA DHT has many structural advantages and the services can be updated efficiently. The comparison with the present system is also made and is assured that it is the security system for the cloud storage. To secure the data information by blinding it, random masking is provided as a proof for securing process. The authentication is done via hashing technique and integrity and performance checks are made with this authentication process.

Keywords: Cloud computing; distributed hash table; public auditing; privacy and dynamic auditing.

1- Introduction

In the world of cloud computing, storing data on the cloud and using its services like making services on the cloud and many other things. It provides high data outsourcing services to users. [1-5] Cloud computing is cost-efficient and provides high performance.
The professional cloud services are growing and are in demand of many users. Many organizations have their eyes on using professional cloud services and the development of cloud services is growing and is expected to grow at a larger rate in the future too. [6]
The cloud services as mentioned have high securities but still, researchers say that it experiences security issues and those security issues and this technology has big security challenges around.

* Corresponding Author

The users before using the cloud storage services are in a dilemma to know whether the provider of the cloud services meets the legal expectations of users related to data security.
The auditing is the distributed auditing in which peer to peer architecture is followed and the outsourced data is organized with the distributed hash table mechanism. This system has proven security and the auditing of the information is carried out in a detailed and secure manner.
These issues related to security concern arise because there are no traditional ways or methods through which the owners of the data who are using cloud computing's storage services can check or investigate the integrity of their data. [7] The data owners does not believe blindly on the cloud storage services because the data loss and the lost integrity of data can be hid by the Cloud service Providers (CSP). In the cloud environment, it becomes very impractical for the data owners to perform the verification check.
Placing an auditor is a better solution for reduction of cloud users and the third party an auditor is capable and expertise in performing cloud operations and should be placed among the providers of the services and the users who own data.
Each data that has been outsourced is audited by the third-party auditor. The auditing is classified into two different types i.e. private and public auditing.
The integrity and of data is enhanced by the usage of auditor and the cloud service providers can make some arbitrations using auditing services.
As above mentioned the classification of auditing, the integrity of the data being at remote can be checked by using private auditing and in this classification, direct verification is established between the data owners and cloud service providers and the cost for all these processes is very low.
There is a trust issue among the owners of the CSPS and the auditing can no more provide convincing results for this. Frequent carrying out the audits may result in the overhead increase and that overhead may or may not be affordable.
DHT provides efficient handling of a large amount of data for outsourcing. Nowadays there are many schemes into existence and most of those schemes are not enough or mot sufficient for securing the privacy of data.
Others can misuse the data at the auditor so it should be extracted and secured from misuse.
The data gets vulnerable and easily accessible so the owners should take steps towards this unauthorized leakage of the data.

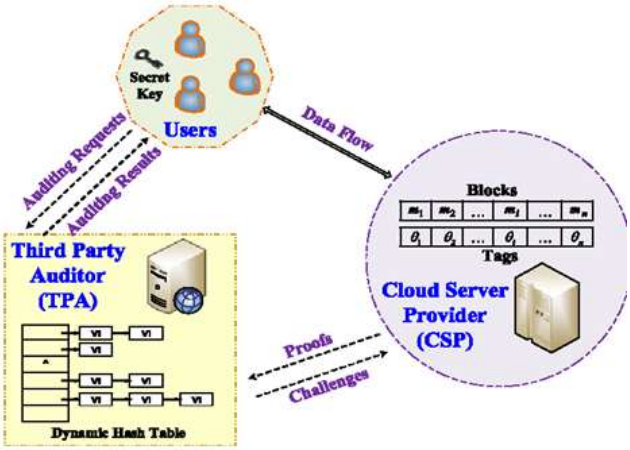

Fig 1: DHT and TPA image

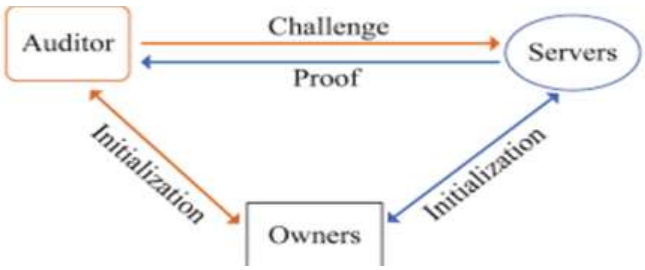

Fig 2 Flow of the System

The other classification of the auditing is public auditing and below the paper has been concluded with the public auditing techniques using distributed hash tables. There are some problems in public auditing and those problems are mentioned below:

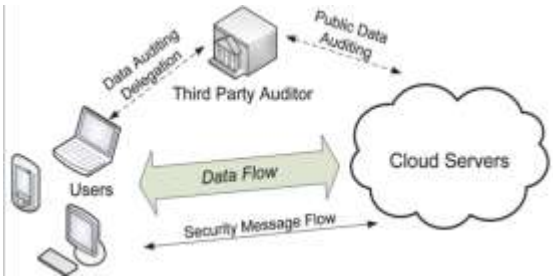

Fig 3: Public auditing

2- Preserving the Privacy

What a data owner expects is the privacy of its data and that privacy is the major concern for which the cloud services are used and still some issues related to privacy are faced in it also.
Data privacy protection (DPP) is the major concern and while using the third-party auditor in public auditing, the preservance of the privacy is the center of the problem. [8]
This problem can be mitigated by paying attention toward the data exploitation and encryption. During the verification of the data leakage cannot be prevented.
Thus for the data auditing it is required to keep a safe mechanism besides it rather than encryption method to maintain and preserve the data privacy.

3- Batch Auditing

The batch auditing is the process by which the public auditing can be made efficient and expandable. The role of a third-party auditor is that it should support the auditing of the batch in a very cost-effective manner.

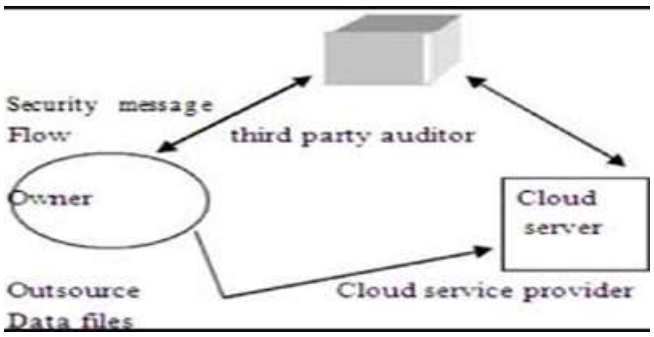

Fig 4: Showing batch auditing over cloud storage

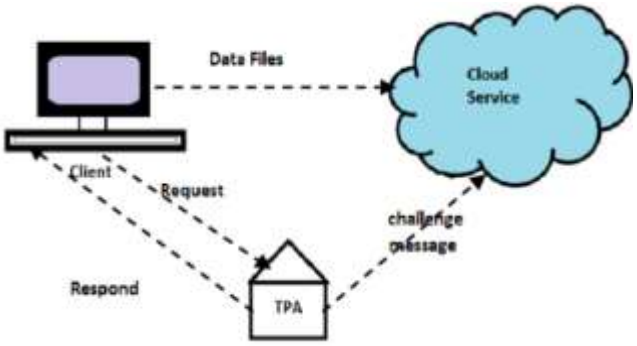

Fig 5: Storage service security on behalf of user upon request.

4- Auditing Dynamically:

A dynamic provable data possession has been presented by Erway (DPDP) in which the original provable data possession scheme has been extended but public auditing is not supported by it. [9]
There is often need of updating the data dynamically and this cloud system has not just worked a warehouse for data storage but it supports the changes and dynamics in data.
The both auditing requirements were fulfilled by the scheme based on Merkle Hash Tree (MHT).
This research makes the following novel contributions.
With the help of distributed hash tables, the TPA can perform data auditing in a better and efficient cost-effective manner and the data can be updated efficiently. [10]
The security of this system has been evaluated and the comparisons of the auditing performance have been done.
A scheme supports all the three problems of public auditing mentioned above i.e. the privacy-preserving, batch, and dynamic auditing.

5- Related Work

Juels presented a related work on “Proof of Retrievability (PORS) and in this work its mentioned, the error-correcting code can retrieve the stored data. Third-party auditing is not supported by PoRs and it is a typical private auditing solution.
The first original Public auditing for the provable data possession was presented and in this study, the outsourced data can be checked with RSA as its base keeping integrity as its prime concern.
The public auditing is preferred over the private auditing because it reduces the overhead to users and the third-party auditor is independent.
The other factors that are concerning the data storage on the cloud are clearly mentioned in this paper i.e. batch auditing, privacy-preserving, etc.
A very early research work on auditing on the cloud was proposed in the year 2007 [12] known as the ‘Proof of Retrievability (PoR)”. A coding technique scheme for error-correcting utilized to ensure that user data was retrievable. PoR supports solution for private auditing but there is no support for using TPA in PoR.
In a similar study [11] the authors presented a public auditing system which was called provable data possession (PDP). The authors generated RSA based authentication tags. The data integrity for the auditing was proposed using randomly sampling [13]. A framework for secure PDP schemes was presented in [14]. The authors proposed homomorphic encryption schemes [15].

Many important factors which include privacy protection, dynamic auditing, and batch auditing, are required for efficient and robust cloud auditing system.

Another researcher extended the above PDP model by introducing a rank based skip list for authentication. This idea was proposed to support dynamic auditing [16]. They also proposed a model for dynamic auditing for the TPA by combining data verification algorithm with corresponding.
MHT was also employed for dynamic auditing in [17] using TPA. This model proposed batch verification and privacy-protection in addition to public auditing. But the problem with such approaches is that there is a high computational costs for TPA and the update and verification also involves high communication overheads.
After some years a research study proposed a new public auditing system [18] which introduced index-hash-table (IHT). Instead of CSP, the properties of the data and integrity are maintained in the party of TPA. Due to this modification the auditing efficiency is improved.
But the updation operation needs to transfer at least half of data elements to the IHT because of its structure.
Therefore another research study proposed dynamic hash table (DHT) for TPA based public auditing schemes [19] to remove the issues. Another novel structure was proposed in [20] which was based on location array and an information table which was doubly linked. This new scheme was able to improve the computational and communication overheads.

Another research study [21] proposed designing a secure cloud auditing protocol which was based on secure network coding scheme.

6- Security Analysis

It is based on some theorems:

6-1- Unforgeability of BLS-HVAs

This theorem has been taken from the work of a researcher known as Wang and in his study, it has been proved the unforgeability of HVA and with assumption BLS has been mentioned to be secure.

6-2- Unforgeability of the Proof:

The auditing proof cannot be forged and is regarded as completely infeasible in case CSP does that.

6-3- Immunity from replacing attacks

A specified block with its tag cannot be replaced to pass the verification by Cloud service provider.

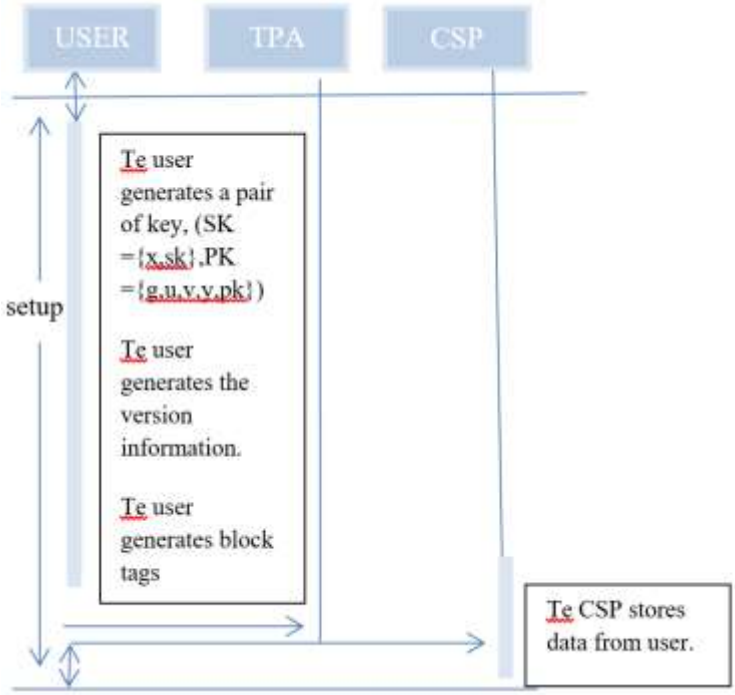

Fig: 6 Workflow of dynamic verification for privacy protection

7- Evaluation of the performance

7-1- Communication cost

Here is a table for communication cost

Table no: 1 table for communication cost

Schemes	Cost(verification) of communication	Cost(Updating) of communication

MHT	cO (Logn)	O(Logn)
DAP	O(c)	O(1)
DHT-PA	O(c)	O(1)

8- Protocol for preserving privacy

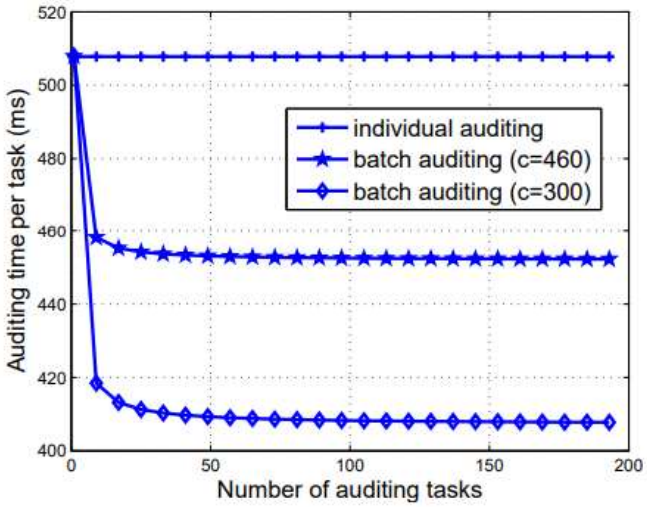

Fig7: Auditing time comparison between batch auditing and individual auditing

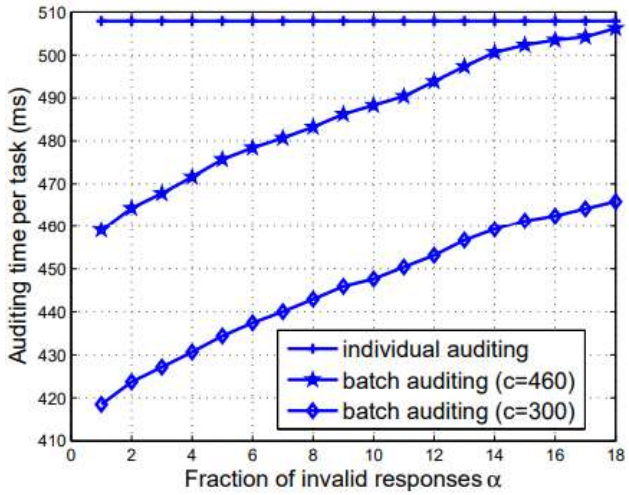

Fig8: Auditing time comparison between batch auditing and individual auditing and the alpha fraction responses are invalid.

8-1- DHT structure

The index hash table contains the index numbers, random values, block numbers and this is a single dimensional array. The computational costs and the overhead of the communication gets reduced using this Index-hash table scheme. The cloud service Provider is not used and instead the IHT is used for auditing properties. The operations of updation and deletion in the IHT due to its structure are inefficient.

The block number modification is inevitable during the insertion and deletion in blocks as the block number gets modified. This inefficiency may cause the cost overhead. To overcome these inefficiencies, a new system design was needed and the design was DHT which is same like IHT but a few dissimilarities are there. It reduces the unwanted cost head. The latest information of the user's data gets tracked for auditing and it is a two dimensional data structure.

The organization of the file is like linked lists are organized and index number is on every file as an identification. The block and file operations are carried out by the DHT in which the insertion and deletion along with modification are included.

To secure the data information by blinding it, random masking is provided by TPA as a proof for securing process. Perfect audits are possible so different strategies are needed to build to achieve perfect auditing.

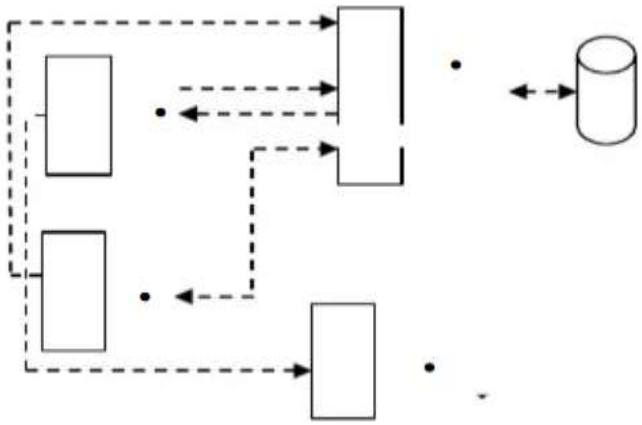

Fig 9: Architecture of DHT

9- Selection of Some algorithms

Some algorithms are necessary to implement and they have advantages.

Encryption and decryption for preserving the security is important so generating some algorithms will help in preserving the cloud storage. The symmetric and asymmetric parties are responsible for the securing and preserving the cloud storage.

For encryption and decryption part of the information, keys are already generated and less computation power is required.

In asymmetric, there are two different keys for encryption and decryption and the keys should not be compromised either. Secret key is the private key that should not be compromised.

9-1- Selection of SHA-1 for security

The fixed length is mapped with the variable length inputs and outputs. The values of the inputs are compared and answers are recorded and the inputs does not show similarity. It means the inputs are never same.

The authentication is done via hashing technique and integrity and performance checks are made with this authentication process. For preserving security the hash algorithm SHA- 1 has been mentioned in this paper.

The hashing tasks will be performed by hash algorithm SHA-1. Its full form is secure hash algorithm. National Security Agency of United Nations has developed this hashing algorithm.

The security system is basically a cryptography system and data hash is counted in this cryptographic function.

There are types of SHA like SHA-0 and SHA -1an the use of these are wide for error correction and weakness detection.160-bit message is the final result of the SHA technology.

Ensure the security of data on cloud

Step1: Choose the file to upload

User will choose the file that he wants to update or upload and then selecting the file it will be updated. The details of the file need to be filled so that it first gets to the TPA and then the encrypted data of user is generated.

Step2: Request sent by TPA

Third party click on the file generates a new file in which details are already mentioned. The request from the user is initiated, it is seen by the TPA and two types of requests are experienced by the TPA from user.

The checks made by the user are because for integrity of the data on cloud and the other system is of downloading the file.

The downloading request of the file gets transferred to the cloud storage provider on time so that the fill gets transferred to the third part auditor.

Step3: forwarding of request by TPA

TPA forwards the check and downloads to the CSP. TPA can login the system and use the authenticated account. The data sent to the cloud is transferred to TPA.

Step4: CSP views the data file

Cloud works to store the data and it shows the data of those files that are already stored. The cloud storage provider manages the data.

The cloud storage provide stores the data on cloud. Hash of the file is generated when the request is sent to the TPA to check the integrity.

The cloud transfers the file to user when the request for the file is made by the user and TPA is responsible for the transferring of the request to the cloud provider.

10- Conclusion

Today if we start counting then almost everyone has utilized the services. No one has remained behind using the cloud services and these services have provided a lot of benefits to the common users.

A lot of people have been attracted to the services that cloud computing provides. Any technology we use has some main obstacles in its way that somewhere causes issues or hampers its growth and such an issue with the cloud has been clearly mentioned in this paper.

The cloud Service providers cannot be trusted because the data owners cannot trust whether their data is safe from breaches.

To strengthen the data security there is a need to build the trust of users on the cloud service providers and in this paper, I have clearly mentioned the distributed hash table techniques which is two-dimensional and dynamic auditing is possible in this distributed hash table.

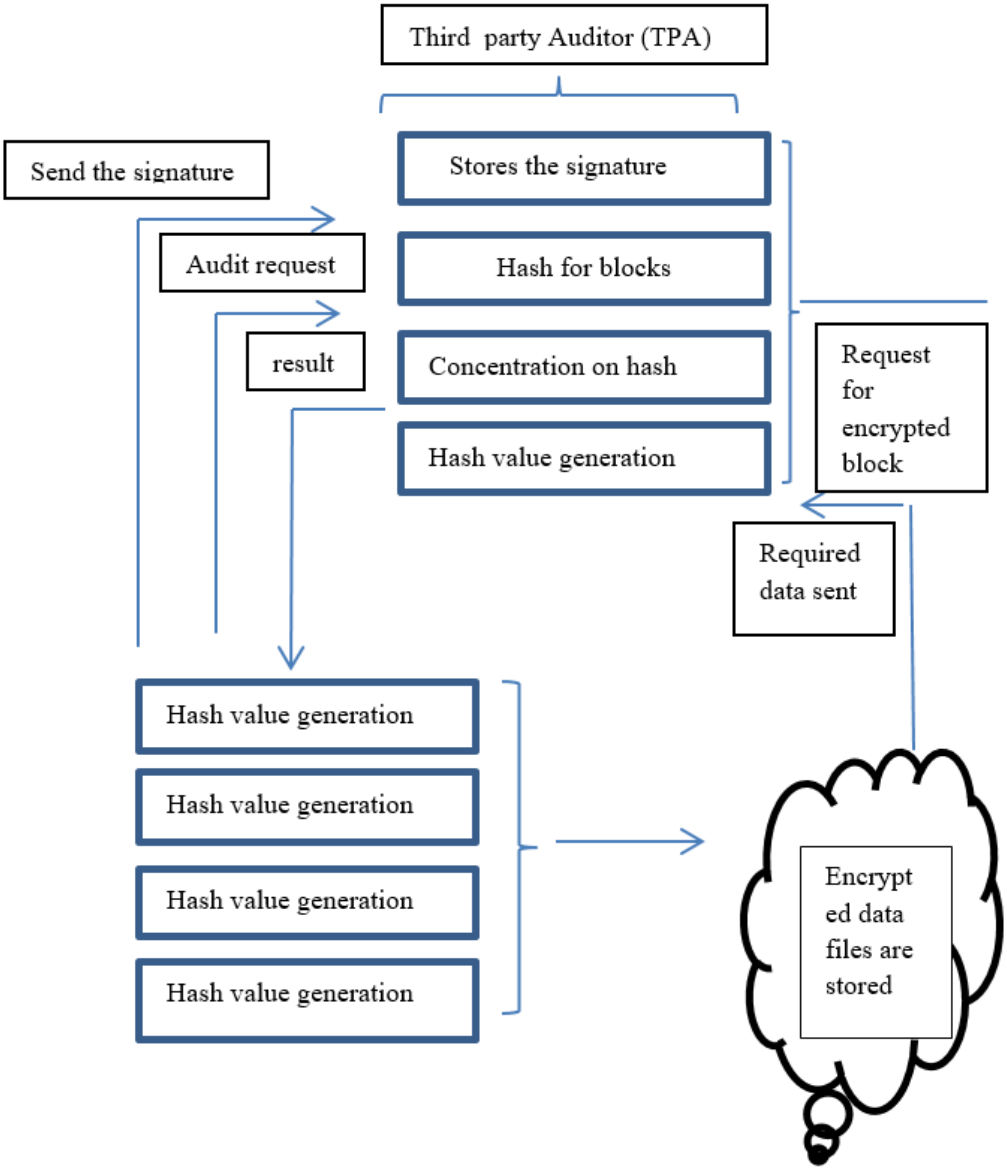

Fig 10: The proposed auditing scheme architecture

* Corresponding Author

Secure data storage is due to data auditing in cloud and user verifies the data by himself. Auditing is used for verification. No one other than the owner or user can verify the auditability and all the capabilities and expertise. The cloud and data audibility does not need to do anything with local copy of data.

Some efficient methods are required for the data security of the cloud storage so that integrity is maintained.

There is not a single method that can help in audits so that perfect data security can be attained. A new and more effective scheme is needed to develop for different types of data stored on the cloud.

References

[1] S. Subashini, V. Kavitha, “A survey on security issues in service delivery models of cloud computing”, *Journal of Network and Computer Applications*, Volume 34, Issue 1, 2011, Pages 1-11, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2010.07.006.

[2] Saurabh Singh, Young-Sik Jeong, Jong Hyuk Park, “A survey on cloud computing security: Issues, threats, and solutions”, *Journal of Network and Computer Applications*, Volume 75, 2016, Pages 200-222, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2016.09.002.

[3] Q. He, Z. Li and X. Zhang, "Study on Cloud Storage System Based on Distributed Storage Systems," 2010 International Conference on Computational and Information Sciences, 2010, pp. 1332-1335, doi: 10.1109/ICCIS.2010.351.

[4] S. Rizvi, A. Razaque and K. Cover, "Cloud Data Integrity Using a Designated Public Verifier," 2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems, 2015, pp. 1361-1366, doi: 10.1109/HPCC-CSS-ICESS.2015.277.

[5] C. Wang, K. Ren, W. Lou and J. Li, "Toward publicly auditable secure cloud data storage services," in IEEE Network, vol. 24, no. 4, pp. 19-24, July-August 2010, doi: 10.1109/MNET.2010.5510914.

[6] A. -E. Bouaouad, A. Cherradi, S. Assoul and N. Souissi, "Architectures and emerging trends in Internet of Things and Cloud computing: a literature review," 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET), 2020, pp. 147-151, doi: 10.1109/IC_ASET49463.2020.9318269.

[7] Simou, S., Kalloniatis, C., Gritzalis, S., and Mouratidis, H. (2016) A survey on cloud forensics challenges and solutions. Security Comm. Networks, 9: 6285– 6314. doi: 10.1002/sec.1688.

[8]Tengfei Tu, Lu Rao, Hua Zhang, Qiaoyan Wen, Jia Xiao, "Privacy-Preserving Outsourced Auditing Scheme for Dynamic Data Storage in Cloud", Security and Communication Networks, vol. 2017, Article ID 4603237, 17 pages, 2017. https://doi.org/10.1155/2017/4603237.

[9] Q. Wang, C. Wang, K. Ren, W. Lou and J. Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," in IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 5, pp. 847-859, May 2011, doi: 10.1109/TPDS.2010.183.

* Corresponding Author

[10] E. Daniel, N. A. Vasanthi, "LDAP: a lightweight deduplication and auditing protocol for secure data storage in cloud environment", Cluster Comput 22, 1247–1258 (2019). https://doi.org/10.1007/s10586-017-1382-6.
[11] G. Atenises, R. Burns, R. Curtmola, et al., "Provable Data Possession at untrusted Stores", In: Proceedings of CCS, pp. 598–609 (2007).
[12] Juels, A., Kaliski, B., Pors.: Proofs of retrievability for large files. In: Proceedings of CCS, pp. 584–597 (2007).
[13] Zhang, J., Yang, Y., Chen, Y., et al.: A general framework to design secure cloud storage protocol using homomorphic encryption scheme. Comput. Netw. 129, 37–50 (2017).
[14] Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC'2009, pp. 169–178.
[15] Erway, C., Küpçü, A., Papamanthou, C., Tamassia, R. : "Dynamic Provable Data Possession," in CCS'2009, pp. 213–222 (2009).
[16] Wang, Q., Wang, C., Ren, K., et al.: Enabling public auditability and data dynamics for storage security in cloud computing. IEEE Trans. Parallel Distrib. Syst. 22(5), 847–859 (2011).
[17] Zhu, Y., Ahn, G., Hu, H., et al.: Dynamic audit services for outsourced storage in clouds. IEEE Trans. Serv. Comput. 6(2), 227–238 (2013)
[18] Chen, Y., Liu, J.: Dynamic-hash-table based public auditing for secure cloud storage. IEEE Trans. Serv. Comput. 10(5), 701–714 (2017)
[19] Shen, J., Shen, J., Chen, X. et al.: An efficient public auditing protocol with novel dynamic sturcture for cloud data. In: IEEE Transactions on Information Forensics and Security, vol. 12(10), (2017)
[20] Chen, F., Xiang, T., Yang, Y., et al.: Secure Cloud Storage Meets with Secure Network Coding. IEEE Trans. Comput. 65(6), 1936–1948 (2016)
[21] Chang, J., Shao, B., Ji, Y., et al.: Secure network coding from secure proof of retrievability. SCI. CHINA Inf. Sci. (2020). https://doi.org/10.1007/s11432-020-2997-0.