Review of Privacy Preserving Public Auditing Techniques

Sarah Shaikh
M.E Information Technology
Vidyalankar Institute of Technology
University of Mumbai
Mumbai,India

Deepali Vora
Information Technology Dept
Vidyalankar Institute of Technology
University of Mumbai
Mumbai,India

ABSTRACT

Cloud computing provides services where users can store their data remotely on cloud and can access their data from anywhere, anytime by means of internet. Cloud computing enables users to use the data stored on cloud as if the data is local without worrying about its accuracy and reliability. But a major challenge in cloud computing is to ensure the integrity of user's outsourced data on cloud server. Public auditing service makes use of an independent third party to check data storage correctness on cloud on behalf of data owner itself. This paper analyzes various techniques employed in order to perform secure cloud auditing to verify the integrity of outsourced data on cloud. Also it discusses limitations associated with these protocols and lastly proposes a scheme to ensure privacy and security in public auditing for cloud storage

Keywords

Cloud computing; Data Integrity; ElGamal algorithm; SHA-256; Third party auditor.

1. INTRODUCTION

Cloud Computing is a rapidly growing technology due to advancement of internet. Cloud computing provides an easy and cost-effective way for companies and individuals to outsource their large volumes of data to cloud and focus only on their core business needs.

However, there are a number of challenges to be faced while placing the data on cloud. The data owner would worry that the data could be lost in cloud due to hardware malfunction or the service provider may intentionally delete the data to satisfy an immense number of customers. Also certain service providers may misuse user's outsourced data and can put data into risk. Many privacy attacks can occur from within the Cloud service provider (CSP) as they usually have direct access to stored data. They may steal the data and sell it to third parties to gain profit. The data owners are eventually not aware about their data being stolen as CSP will not inform about it to cloud users. Therefore, certain security measures must be taken to protect stored data both from malicious outside attackers as well as from service provider itself.

To assure cloud data integrity, an obvious way is to check on retrieval. However, it is impossible to assure integrity for whole data since large amount of data are outsourced to cloud but only few are frequently accessed which means checking the integrity of such data is bypassed by this technique. Another technique is to download the whole data and then perform integrity check, but this will eventually lead to heavy I/O overhead on cloud server and high communication cost for data owner.

To completely guarantee on data integrity and save cloud user's data, it is more significant to allow third party auditing service for cloud data storage. Third party auditors are skilled persons in a particular verification field. The owner can employ those third party auditors to verify their outsourced data on their behalf and ensure storage correctness of their data in cloud.

The rest of the paper is organized as follows: section 2 provides literature review of the various techniques employed to achieve data integrity and storage correctness of users outsourced data on cloud. Section 3 provides an efficient solution against problem definition and literature review using cryptography algorithms. Lastly, section 4 gives the concluding remark of whole paper and on the literature review.

2. TECHNIQUES FOR SECURE CLOUD AUDITING

Various Techniques used for verifying the integrity of remotely stored data on cloud and to assure their storage correctness are as follows:

2.1 Provable Data Possession (PDP) model

PDP model by Ateniese [1] was the first in considering the concept of public auditing to verify whether the server possess the original data files without retrieving it. This model uses RSA based homomorphic tags to verify data blocks stored on remote server. PDP generates probabilistic proofs of data possession by randomly sampling few blocks of files from server, which reduces I/O cost drastically. The server accesses only a small portion of file for generating proofs. The client can verify whether the server has retained the file data by using the metadata generated and maintained by them while data outsourcing. Due to reduced network communication, this model is highly suitable for large datasets in distributed systems. Instead of accessing the entire file, the server accesses only minor sections of data blocks to verify the data integrity of the entire file. Also, it requires linear combination of data blocks for verification. The data owner has to bear the overhead of generating the metadata. PDP does not consider the case of dynamic data and only static data is used for generating probabilistic proofs.

Ateniese et al. [2] provided a model based on symmetric key cryptography in order to provide data dynamics support. It makes use of less storage space and less bandwidth. However, this method requires fixing priori the number of queries which can be answered. Also it lacked support for block insertion operation, thus not providing fully dynamic environment

2.2 Proofs of Retrievability (PoR) model

This method [3] suggests BLS method for data privacy. Spot checking and error correcting codes are used to verify remote data possession. Some special blocks known as "sentinels" are embedded into the file and this file is further encrypted to protect the position of these sentinel blocks. POR scheme cannot be used for public databases; it is suitable only for confidential data. Also it supports only static data, dynamic data operations are not allowed. This protocol supports only two party auditing i.e. the client has to perform the integrity check.

An improved PoR scheme was further defined in [4]. This protocol makes use of homomorphic linear authenticators (HLAs).These HLAs can be aggregated for authenticating a linear combination of individual data blocks. Still, this protocol considers only static Data files.

2.3 Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds

This method [5] supports updates to remotely stored data by using probabilistic query and periodic verification. The Third party auditor (TPA) must be allowed to manage the outsourced data instead of data owner itself. This technique employs fragment structure, random sampling and index-hash table.TPA checks data integrity and availability at regular time intervals.

Figure 1 shows the cloud storage architecture with Cloud service provider, Cloud users and TPA

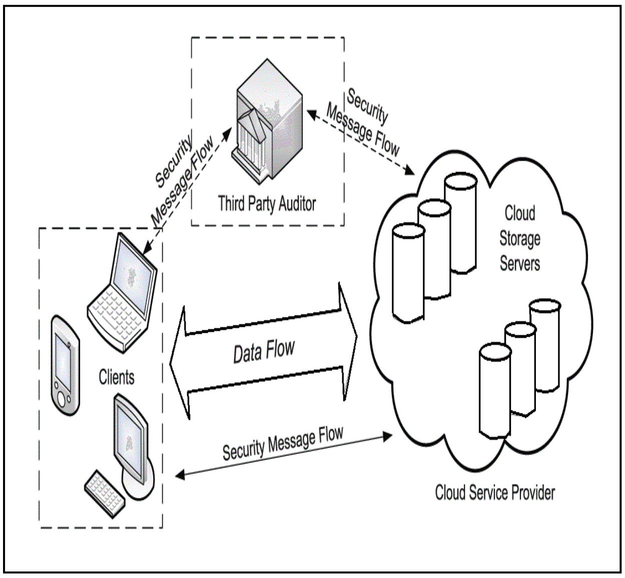

Fig 1: The architecture of cloud data storage service[7]

2.4 ORUTA: privacy preserving public auditing for shared data on cloud

This method [6] is specially employed for ensuring the privacy and security of the users sharing their data on cloud with other users. It provides identity privacy i.e. the identity of the signer is kept secret from the third party auditor (TPA). Each cloud user can access and modify the shared data. The shared data is signed using public and private keys of the data owner. The TPA checks the integrity of stored data by sending auditing challenge to cloud server. TPA verifies the correctness of stored data by checking the auditing proof given by server.

2.5 Privacy Preserving Public Auditing for Cloud Storage

This technique [7] [8] [9] employs homomorphic linear authenticator with arbitrary masking to achieve storage correctness and integrity assurance of data stored on cloud. It uses linear combination of sampled data blocks along with arbitrary masking because of which the TPA cannot have enough information to make up the correct group of linear equations to recover the original content. This system uses two algorithms which are as follows:

- MAC Based:

Message authenticator code (MAC) is used to validate the data blocks. The client transfers the data blocks and MAC to the cloud server. The cloud server provides the secret key to TPA. The TPA then recovers the data blocks and MAC uses secret key to verify the integrity of outsourced data. However, in this process the TPA obtains priori information of data blocks. Also it is suitable for static data, dynamic auditing is not supported.

- HLA Based:

The homomorphic linear authenticator system does not retrieve the data blocks itself. It validates a linear combination of individual data blocks, but this further requires user data information to TPA thereby violating the privacy of user's data.

2.6 Use of cryptographic schemes

This method uses cryptographic algorithms to ensure data storage security in cloud environment. The paper [10] proposes RSA based storage security to ensure data storage correctness and also attempts to identify the misbehaving server. It also supports dynamic data operation using RSA signature and Merkle Hash Tree construction. The protocol proposed in [11] uses AES algorithm instead of RSA to verify data integrity at untrusted server. Since AES is faster than RSA in encryption-decryption time, this method attempts to improve the overall performance of cloud based storage security system. The paper [12] proposes yet another protocol to provide secure cloud auditing over encrypted data by making use of RSA algorithm and MD5 hash algorithm to ensure data integrity at cloud server.

However, using RSA resulted in longer encryption time and higher computational overhead even for small file size. Thus the proposed system attempts to use another public key cryptographic algorithm called ElGamal in place of RSA in order to improve cloud storage system security

3. RESEARCH WORK

The general idea of proposed system can be explained in following steps:

1. The proposed system considers the server being untrusted entity. Hence the data is not stored directly on the untrusted server, instead we encrypt it using ElGamal encryption algorithm before uploading the data files on cloud so that the cloud service provider or any other Third Party cannot directly view the contents of the file.

2. Now to check the integrity of data on cloud, the data owner employs an auditor (Third Party Auditor) to perform audit operation on stored data. The auditor takes the responsibility of verifying the data and performs integrity checking of encrypted data. Here SHA-256 hash algorithm is used for file authentication and integrity verification.

3. After performing integrity checking, the data owner (i.e. the client) is notified about the status of its data; whether the data is authentic or whether its integrity is lost. Initially the auditor poses the challenge to server for initial authentication of the entire data. After authentication and integrity checking, the result is maintained by the auditor and copy is sent to the data owner.

4. When there is a need to modify the data, the data owner updates the data on cloud. Also the TPA is notified of the modifications done to the data files. TPA then performs periodic auditing of the modified data to verify the integrity of outsourced data on the basis of information provided by data owner.

4. CONCLUSION

After reviewing the literature survey, it can be observed that public cloud auditing system possess problem statement like static data support, data leakage to third parties, unsafe data transmission violates the privacy preservation assurance and secure access to cloud data content. To overcome these limitations, the research system proposes a public auditing mechanism which makes use of encryption and Hash signatures to ensure data integrity at untrusted servers. The proposed system employs ElGamal algorithm to secure uploaded data on cloud and then verify its integrity by making use of SHA-256 hash algorithm.

5. REFERENCES

[1] G.Ateniese, “Provable Data Possession at Untrusted Stores”, Proc. 14th ACM Conf. Computer and Comm. Security (CCS’ 07), 2007.

[2] G.Ateniese, “Scalable and Efficient Provable Data Possession”, Proc.Fourth Int’l Conf. Security and Privacy in Comm. Networks (SecureComm ’08), 2008.

[3] Juels, “Pors: Proofs of Retrievability for Large Files”, Proc. 14th ACM Conf. Computer and Comm. Security (CCS ’07), pp. 584-597, 2007.

[4] H. Shacham and B. Waters, “Compact Proofs of Retrievability,” Proc. 14th Int’l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (ASIACRYPT ’08), pp. 90-107, 2008.

[5] Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu and S.S Yau,“Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds,” Proc. ACM Symp. Applied Computing (SAC’11), pp. 1550-1557, 2011.

[6] Wang, B. Li, and H. Li, “Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud,” Proc. IEEE Fifth Int’l Conf. Cloud Computing, pp. 43-56, 2014.

[7] 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.

[8] Q. Wang, C. Wang, K.Ren, W. Lou and Jin Li “Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing”, IEEE Transaction on Parallel and Distributed System, vol. 22, no. 5, pp. 847 –859, 2011.

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

[10] Trushna S Khatri, G B Jethava “Improving Dynamic Data Integrity Verification in Cloud Computing”, 4th ICCCNT 2013.

[11] Poonam M. Pardeshi,”Improving Data Integrity for Data Storage Security in Cloud Computing” ,International Journal of Computer Science and Information Technologies, Vol. 5(5), 2014.

[12] Sawan V. Baghel, Deepti P. Theng,” A Survey for Secure Communication of Cloud Third Party Authenticator “, IEEE 2nd International Conference on Electronics and Communication Systems, ICECS ‘2015.