International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor (2018): 7.426

A Privacy Preserving Mechanism for Integrity Auditing and Data Sharing

R. Ramya Devi[1], R. Rangeela[2], S. Shalini[3], R. Sankara Narayanan[4]

[1]B. Tech, Department of Information Technology, Valliammai Engineering College, Chennai, Tamil Nadu, India

[4]Assistant Professor, Department of Information Technology, Valliammai Engineering College, Chennai, Tamil Nadu, India

Abstract - ***With cloud storage services, users can store huge amount of data and share their data with others. This leads to security and integrity problems. Data on the web can be manipulated and corrupted and therefore will lose its integrity. To avoid this, data verification was introduced, which is achieved by using a technique called integrity auditing, through which verification of multiple data of a user is performed at a time. This paper avails a shared environment in which the data owner uploads the data and the users in that environment access it. In order to provide security to the uploaded data, Advanced Encryption Standard (AES) is used by the data owner to store encrypted data in cloud server. AES uses a 16 bit key which is sent to users through Simple Mail Transfer Protocol (SMTP), using which the users can retrieve the required data. A public auditor challenges the cloud server to prove the integrity of data stored by the data owner and verifies the integrity of data by using the proof generated by the server. Message Digest (MD5) generates a hash value which is used by public auditor to verify the integrity of data. Thus, this paper aims for providing privacy preserving mechanism using cryptographic techniques for integrity auditing and data sharing, and the performance analysis shows the efficiency of the mechanism when auditing data integrity.***

Keywords: Cloud Storage, Integrity Auditing, Data Sharing, Privacy Preserving, Cryptographic Techniques

1. Introduction

Cloud refers to application, services, or resources made available to users on demand via the internet from a cloud computing providers server with high speed and accuracy. Cloud computing uses hardware and software to deliver a service over a network. Information data is stored on physical or virtual servers, which are maintained and controlled by a Cloud Service Providers(CSP). Cloud storage services maintains and manages user's data over a network, cloud storage providers also provide unlimited growth, the ability to increase and decrease storage capacity on demand and flexibility to access data any time from anywhere.

In cloud environment, the user data is being centralized or outsourced to the cloud. Even though, it provides various advantages, it also brings new security threats and challenges towards the outsourced data. Since the Cloud Service Providers (CSP) are separate entities, outsourcing the data creates a situation in which the users cannot control their data. As a result, the correctness of the data is at risk due to the following reasons. At first, the service provider may experience failures and may decide to hide the data errors from the users for their benefit. Second, although the cloud infrastructure is much more powerful and reliable than personal devices, it still faces various internal and security threats.

In order to verify the correctness of data, the auditing for data integrity was introduced. The general flow is, the users generate the signatures using the keys and they sign the blocks using that signature and upload it on the server. The auditor generates a challenge for the server to verify the integrity of data and the server sends the proof to the auditor. According to the role of auditor, auditing is divided into two types: Private Auditing and Public Auditing.

In private auditing, the users generate a private key for generating the signatures and the user will themselves challenge the server and verify the integrity of data stored in the cloud. The users may not be able to perform frequent integrity checks and if the user is going to perform verification, the user may require a copy of the outsourced data.

To overcome, the above efficiency problems, the public auditing is performed. In public auditing, the users generate a public key for generating the signatures and the user will delegate a Third Party Auditor (TPA), to evaluate the integrity of data stored in the cloud. The integrity auditing is used to verify multiple blocks of data simultaneously instead of verifying one by one and the batch auditing is used to verify multiple user data simultaneously at a time. To provide privacy for data cryptographic techniques can be used.

Cryptography can be used to provide message Confidentiality, Integrity and Sender Verification. The functions of cryptography are encryption, decryption and cryptographic hashing. Encryption converts plain text into cipher text. Decryption converts cipher text into plain text. In order to encrypt and decrypt messages, the sender and recipient need to share a secret key, which is used by the sender to encrypt the message and by the recipient to decrypt the message. Various algorithms such as Advanced Encryption Standard(AES) can be used for encryption. Cryptographic hashing such as Message Digest(MD5), is the process of generating a fixed length string from an arbitrary length of message.

2. Objective

1) The objective of this system is to ensure the integrity of data outsourced in the cloud by data owner.
2) To perform integrity auditing of the stored data by a public auditor.
3) To provide a privacy preserving mechanism in which the auditor will audit the data without knowing about the data.

3. Related Work

To verify the integrity of the data outsourced in the cloud, many auditing schemes have been proposed. A Third Party Auditor(TPA) is introduced to verify the integrity of the data stored in the cloud.

In [1], this paper motivates the public auditing system of the data storage security in cloud computing and proposes Merkle Hash tree Construction for supporting fully dynamic data operations especially block Insertion. This paper has two goals a) Efficient data dynamics b) Batch auditing (i.e) Handling multiple auditing tasks simultaneously. To achieve efficient data dynamics, the paper improves the existing proof of storage models by manipulating the Merkle Hash Tree Construction. To support efficient handling of multiple auditing tasks, the technique of bilinear aggregate signature is performed. The proposed system is efficient as it provides an average of batch auditing of files as 97%, but the scheme has large communication costs during updating and verification process.

In [2], this paper motivates privacy preserving public auditing system for data storage security in cloud computing. This paper has two goals a) privacy Preserving Auditing b) Batch Auditing. This paper utilizes the homomorphic linear authenticator and random masking to guarantee that the TPA would not learn any knowledge about the data content stored on the cloud server during auditing process (i.e) Privacy preserving auditing: This is further extended to support batch auditing. The scheme is efficient for batch auditing, since when compared to individual auditing batch auditing reduces 15% of computational cost and hence it is more efficient.
In [3], this paper proposes a novel privacy-preserving mechanism that supports public auditing on shared data stored in the cloud. This scheme protects the identity privacy of user. This paper utilizes ring signature to construct homomorphic authenticators, so that an auditor is able to audit the shared data integrity in a group of users, without retrieving the entire data and by not distinguishing who is the signer on each blocks. But this scheme, does not support traceability (i.e) the ability of the group manager to reveal the identity of the signer based on verification metadata in some special situations, proof for data freshness (i.e) prove that the cloud possess the latest version of shared data and dynamic groups, where the old users leave and the new users join the group at any time.
In [4], this paper proposes a public auditing mechanism for shared data with efficient user revocation in the cloud. When a user in the group is revoked, the blocks which were previously signed by this revoked user must be resigned by an existing user. This scheme allows the semi-trusted cloud to re-sign blocks that were signed by the revoked user with proxy-resignatures. This paper provides efficient user revocation and also saves a significant amount of computation and communication resources for the existing users in the group during user revocation.

In [5], this paper proposes Global and Sampling Verification and Data Dynamics using Dynamic Structure. Guarantee of sampling verification makes data owners believe that their data is stored properly in the cloud. Global verification gives the cloud confidence against owners who are not always behaving appropriately and thus eliminating the fear of being wrongly accused. Data dynamics is efficiently done by using doubly linked info table and location array. This scheme provides efficient dynamics support and reduces overhead. In [6], this paper proposes a scheme in which the file stored in the cloud can be shared and used by others on the condition that the sensitive information is hidden. This scheme also uses user's identity information such as user's name or email address to replace the public key, in order to take less computational time. It uses a sanitizer to sanitize the data blocks corresponding to the sensitive information of the file. The time of signature verification and that of sensitive information sanitization are 2.318s and 0.041s respectively and it also shows that the computation overheads increases as the challenged data blocks increases during auditing.

4. Proposed System

The administrator is responsible for maintaining and managing the details of users, data owner and the auditor. In Fig – 1, the data owner uploads the data in the cloud for a group of users to access it. For security reasons, the data owner stores the data in the encrypted format using Advanced Encryption Standard (AES). Advanced Encryption Standard is a symmetric key block cipher for secure encryption and decryption. Users register by giving their required details and then gets added to the group of registered users list of data owner. The data owner sends the key to authorized users through Simple Mail Transfer Protocol (SMTP). Simple Mail Transfer Protocol is the standard protocol for email services on a network. SMTP provides the ability to send and receive email messages.

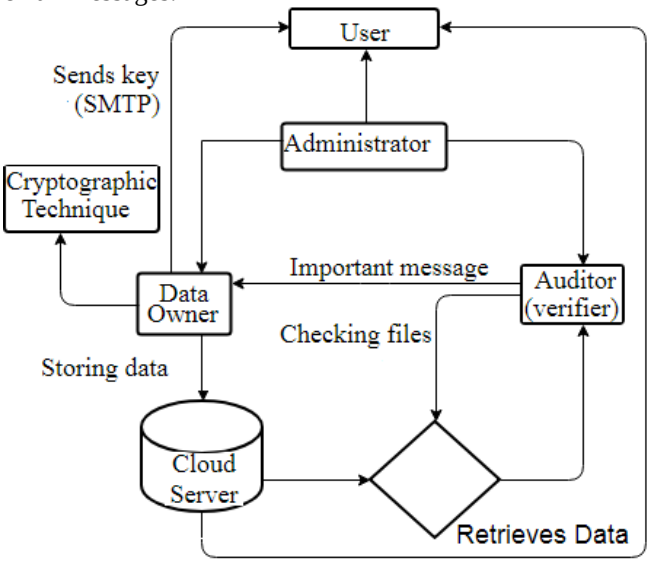

Figure 1: Architecture Diagram

Then the users use this key to download the file. The auditor is responsible for checking the integrity of data. For this purpose, the auditor uses Message Digest (MD5) algorithm, which is a cryptographic hash function. It takes arbitrary length message as input and generates 128 bits hash value. The auditor challenges the cloud to prove the integrity of the data stored by data owner. The cloud sends the metadata and the auditor generates the hash value to verify the integrity. If the data is found to be corrupted, then the auditor sends an important message to the data owner stating the corruption of data. The auditor will inform the admin about the user who corrupted

the data. Then the admin will remove the adversary user from the group.

5. Methodology

5.1 Registration

This module involves the registration of data owner, user and auditor. Admin is responsible for maintaining and managing the details of data owner, user and auditor. In Fig – 2, the data owner fills his/her details along with the product key in order to register. Then the data owner will login to upload the file in encrypted format. For encrypting the file, we use Advanced Encryption Standard (AES), which has a fixed block size of 128 bits and a key size of 128, 192, or 256 bits depending on the number of rounds. The block can take a maximum size of 256 bits, but the key size has no restrictions. AES operates on a 4×4 column-major order matrix of bytes. The encryption process is performed using four sub processes namely adding round keys, substituting bytes, shifting rows and mixing columns. The process of decryption is similar to encryption process but in the reverse order i.e., adding round key, mixing columns, shifting rows and substituting bytes. In this paper, AES is used for sending the 16 byte key to the users. Users register by giving the required details and selects the group of the data owner. The user will be added to the selected group. Finally, auditor register by giving the required details.

5.2 Data Sharing

In Fig – 3, the data owner shares data in a confidential and selective way by encrypting the file using the product key and uploads the file in the cloud server. The data owner then uses Message Digest (MD5) to generate the hash value which will be sent to the auditor for verification purposes. Message Digest (MD5) is a hash function used in cryptography. It takes an input of arbitrary length and produces a message digest that is 128 bits long. The digest is also called as the hash of the input. The steps involved in MD5 are appending of padding bits, appending representation of padded message to the original message, initialization of message digest buffer, processing of message in 16-word blocks and finally output the result. In this paper, Message Digest (MD5) is used by public auditor to verify the integrity of data. The user will receive the product key through email by using SMTP protocol and will use that key to retrieve and decrypt the required file from the cloud.

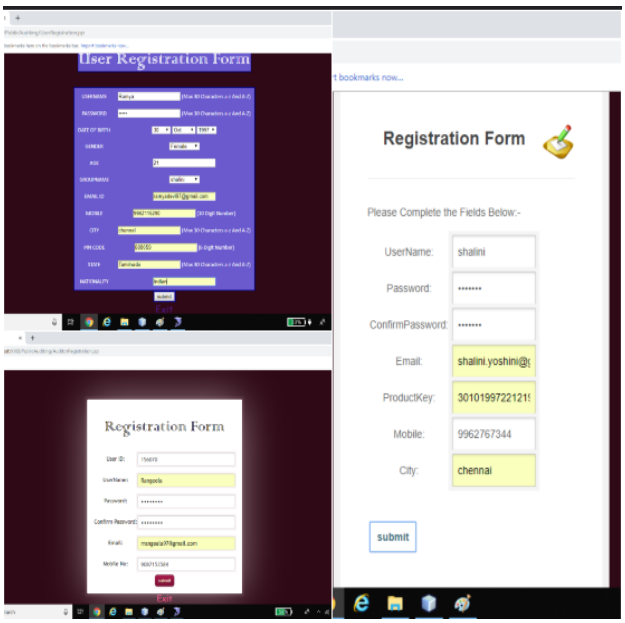

Figure 2: Registration form

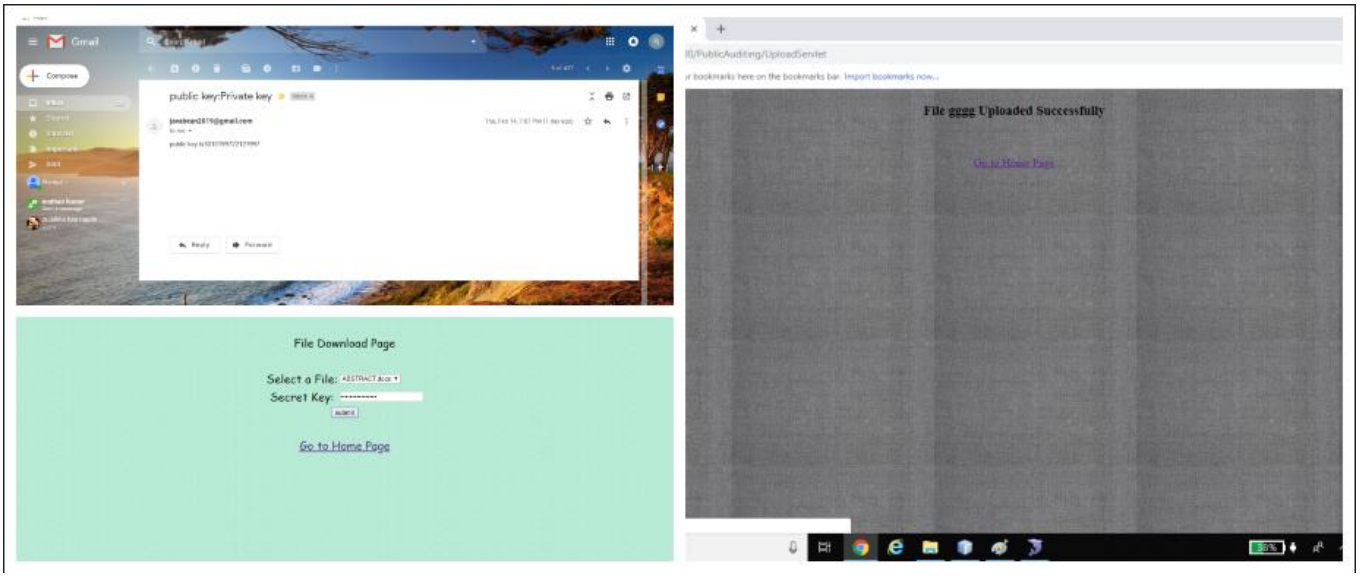

Figure 3: File upload and Download

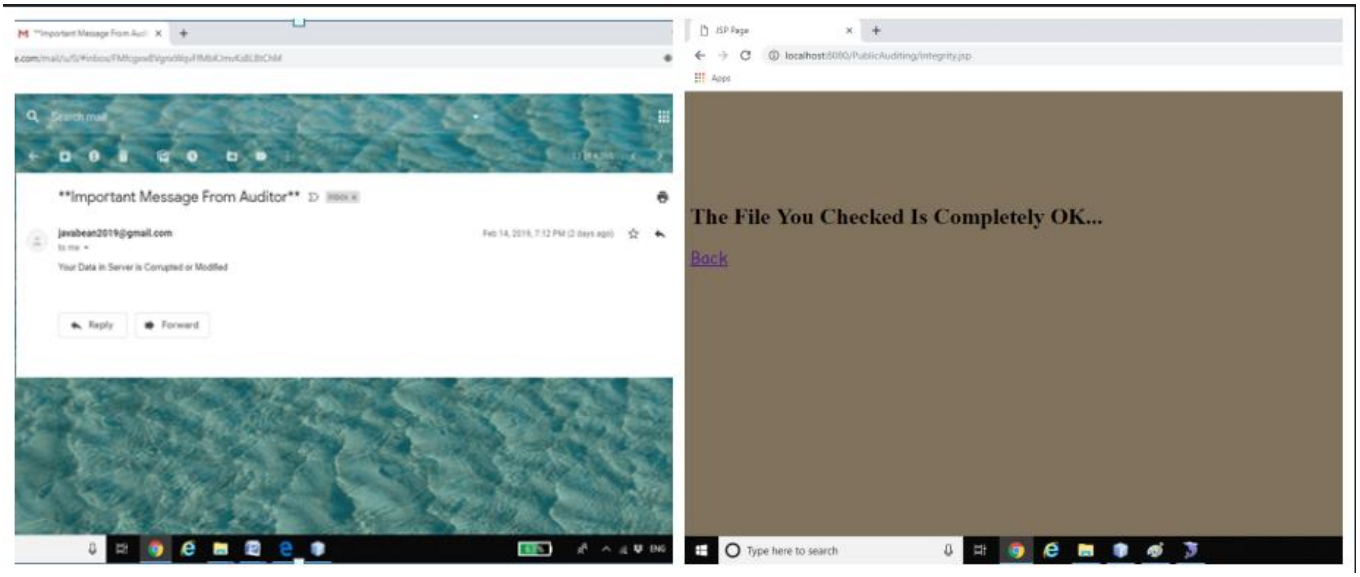

Figure 4: Public Auditing

5.3 Public Auditing

To perform public auditing the auditor will challenge the server to prove the data integrity of the file stored in the cloud. The server will send the metadata of the encrypted file to the auditor and the auditor will generate the hash value for the received file. Then the auditor will compare the generated hash value with the hash value sent by the server, to check the integrity of the file stored in the cloud. If the auditor senses corruption in data, an error message will be sent to the data owner through SMTP. Simple Mail Transfer Protocol (SMTP) is a protocol used for email services on a network. SMTP is used to send and receive email messages. SMTP is generally integrated with in an email client application and is composed of four key components namely Mail User Agent (MUA), Mail Submission Agent (MSA), Mail Transfer Agent (MTA), and Mail Delivery Agent (MDA). Fig – 4, shows the successful auditing of the file and the auditing of the corrupt file.

5.4 User Revocation

In this module, auditor will inform the data owner and the admin about the change of data in the files and the auditor will find the user who tried to change the data. The auditor will inform the admin and then the admin will remove the malicious user from the cloud environment. Fig – 5, shows the successful revocation of the adversary user.

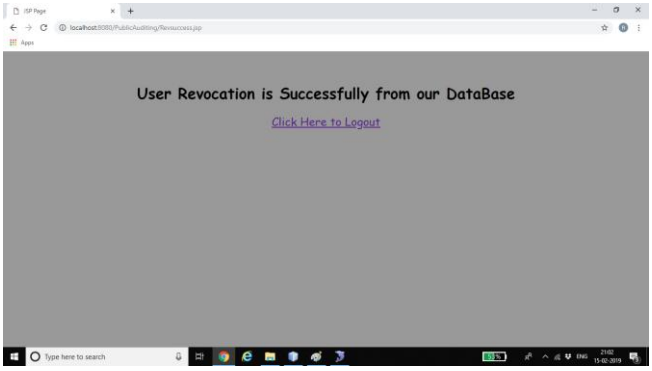

Figure 5: User Revocation

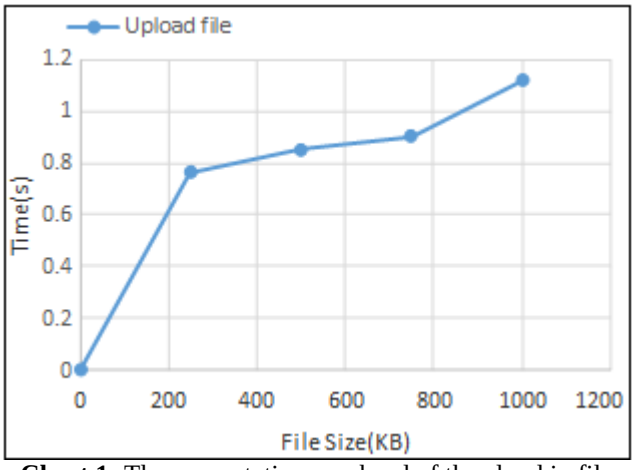

Chart 1: The computation overhead of the cloud in file upload process

6. Performance Analysis

The experiment is run on a Windows machine with an Intel Pentium 2 GHz processor and 80GB memory.

1)Performance of File Upload: To evaluate the performance of encrypting the file using Advanced Encryption Standard(AES) and then uploading the file in cloud, the number of file sizes varies from 0KB to 1000KB increased by an interval of 250 in the experiment. As shown in Chart - 1, the time cost of uploading the files increases with the size of the file being uploaded. The time of uploading files ranges from 0s to 1.118s

2) Performance of Auditing: With different number of challenged data files, the computation overhead of the auditor and the cloud is shown in the Chart – 2. The number of challenged file sizes in the experiment varies from 0KB to 1000KB. In Chart-2, the computation overheads of challenge generation and the proof verification on the auditor side increases with the number of challenged files size. The computation overhead of the auditing process varies from 0s to 1.201s.

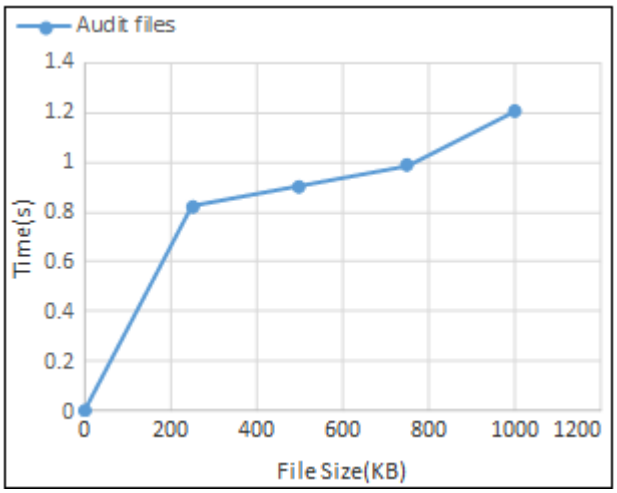

Chart 2: The computation overhead of the cloud in auditing process

7. Conclusion

In this paper, a privacy preserving mechanism for integrity auditing and data sharing in cloud environment is proposed. In this scheme, the file is stored in the cloud securely and by using privacy preserving mechanism, the auditor is able to audit the integrity of shared data without knowing about the data in the file successfully. This scheme can also be extended to perform data dynamics operations in cloud, which can be used for future work. The performance analysis demonstrate that the proposed scheme achieves desirable security and efficiency.

References

[1] Qian Wang, Cong Wang, Kui Ren, Wenjing Lou and Jin Li,"Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", IEEE Transactions On Parallel And Distributed Systems, Vol. 22, No. 5,pp.847-859,May 2011.

[2] Cong Wang, Sherman S.M. Chow, Qian Wang, Kui Ren and Wenjing Lou," Privacy-Preserving Public Auditing for Secure Cloud Storage", IEEE Transactions On Computers, Vol. 62, No. 2, February 2013.

[3] Boyang Wang, Baochun Li and Hui Li," Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud", IEEE Transactions On Cloud Computing, Vol. 2, No. 1,pp.43-56, January-March 2014.

[4] Boyang Wang, Baochun Li and Hui Li, "Panda: Public Auditing for Shared Data with Efficient User Revocation in the Cloud", IEEE Transactions On Services Computing, Vol. 8, No. 1,pp.92-106, January/February 2015.

[5] Jian Shen, Jun Shen, Xiaofeng Chen, Xinyi Huang, and Willy Susilo," An Efficient Public Auditing Protocol With Novel Dynamic Structure for Cloud Data", IEEE Transactions On Information Forensics And Security, Vol. 12, No. 10,pp.2402-2415,October 2017.

[6] Wenting Shen, Jing Qin, Jia Yu , Rong Hao, and Jiankun Hu , "Enabling Identity-Based Integrity Auditing and Data Sharing With Sensitive Information Hiding for Secure Cloud Storage", IEEE Transactions On Information Forensics And Security, Vol. 14, No. 2,pp.331-346, February 2019.