International Conference on New Horizons in Science Engineering Technology (NHSET-2018)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 5 | ISSN : 2456-3307

A Novel Data Auditing Approach to Achieve Data Privacy and Data Integrity in Cloud Computing

Soumya Shinde, Ramya V. Shinde, Priyanka Kamadhenu
Department of ISE, SKSVM Agadi College of Engg. & Tech, Laxmeshwar, India

Abstract
Cloud computing is a general term for the delivery of hosted services over the internet. Using the cloud services saves both users time and money. However, there are some security issues to be solved for personal users and enterprises to store data in the cloud. The fact is that users will not have any physical control over the outsourced data. The cloud user is concerned about the integrity of his data stored in the cloud as it can be attacked by attacker. The purpose of this paper is to suggest an efficient public auditing technique using Third Party Auditor (TPA) to verify the integrity of data stored in the cloud. The proposed auditing scheme makes use of AES algorithm for encryption and Secure Hash Algorithm (SHA-2) algorithm to generate verification metadata or message digest for data integrity check. The analysis shows that the proposed scheme is provably secure and TPA takes a constant time to audit files of different sizes.

Keywords : Cloud Computing, Public Auditing, Cloud Service Provider (CSP), Data Integrity, Third Party Auditor (TPA)

I. INTRODUCTION

Cloud computing is a new computing paradigm in which resources are shared as a service over the internet. Cloud data storage is one of the service provided by Cloud computing in which data is maintained, managed, backed up remotely and made available to users over a network (typically the Internet) [1]. Moving data into the cloud offers great convenience to users since they don't have to care about the complexities of direct hardware management. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) [2] are both well-known examples. These internet-based online services do provide huge amounts of storage space and customizable computing resources. This computing platform eliminates the responsibility of local machines and users can be relieved from the burden of local data storage and maintenance.

Even with these many benefits of cloud computing, previously mentioned, users are reluctant (hesitate) to adopt this technology because of security threats [3].There are many threats facing the cloud not only from an outsider but also from an insider which can utilize cloud vulnerabilities to do harm. These threats may harm data confidentiality, data integrity, and data availability. Some untrusted providers could hide data breaches to save their reputations or free some space by deleting the less used or less accessed data [4]. Hence for sensitive and confidential data, there should be some security mechanism to provide protection for cloud data. To keep away from the security risks, concept of cryptography and audit services are significant to make sure about the confidentiality and integrity of outsourced data.

Data auditing is introduced in Cloud computing to deal with secure data storage. Auditing is a process of verifying the users data carried out either by the client or by a Third Party Auditor (TPA). Auditing helps to maintain the integrity of client's data stored in the cloud. The auditing process can be categorized into two types: first one is private auditing where client or data owner is allowed to check the integrity of the stored data. But it increases verification overhead of the client. Second is public auditing, which allows anyone, to challenge the cloud server and performs data verification check with the help of TPA.

TPA is the third party auditor who will audit the data of data owner or client. The TPA has expertise and capabilities that users do not. It is necessary that TPA should efficiently audit the cloud data storage without requesting for the local copy of data.

II. RELATED WORK

Different factors such as confidentiality, data integrity etc. affects the performance of cloud data storage. There has been a lot of development in this field and lots of algorithms have been proposed by various researchers. Here we are presenting the related work that are found to be useful.

TABLE 1 represents the comparison done by considering different factors such as methods used, supports public auditing, supports privacy preserving, maintaining data integrity and data confidentiality.

TABLE I. COMPARISION OF EXISTING PUBLIC AUDITING SCHEMES

Research Work	**Methods Used**	**Supports Public Auditing**	**Supports Privacy Preserving**	**Maintaining Data Integrity**	**Maintaining Data Confidentiality**
Privacy Preserving Public Auditing for Secure Cloud Storage[5]	HLA with BLS Signature	Yes	Yes	Yes	No
Privacy-Preserving Public Auditing In Cloud Using HMAC Algorithm[6]	HMAC	Yes	Yes	Yes	No

Privacy-preserving Public Auditing for Data Storage Security in Cloud Computing [7]	HLA with Random Masking	Yes	Yes	Yes	No
Privacy Preserving and Public Auditing Service for Data Storage in Cloud Computing [8]	RSA and MHT	Yes	Yes	Yes	Yes
Towards Secure and Dependable Storage Services in Cloud Computing [9]	Homomorphic Tokens and Erasure code	Yes	Yes	Yes	No
Secure and Efficient Privacy Preserving Public Auditing Scheme for Cloud Storage [10]	HLA with BLS Signature	Yes	Yes	Yes	No

Swapnali Morea et al. [1] proposed a secure and efficient privacy preserving public auditing scheme. It achieves privacy preserving and public auditing for cloud by using a Third Party Auditor (TPA). The data owner or the user is responsible for splitting the file into blocks, encrypting those using AES algorithm, generating a SHA-2 hash value for each, concatenating the hashes and generates a RSA signature on it. In the Verification process, the signature generated by TPA and the one stored in the TPA which is provided by the data user are compared by the TPA. If both matches, then data is intact and if mismatch occurs then it indicates that the data integrity has been affected or tampered.

Ezhil Arasu et al. [6] has proposed a method that uses the keyed Hash Message Authentication Code (HMAC) with homomorphic tokens to enhance the security of TPA. It is a technique for verifying the integrity of a data transmitted between two parties that agree on a shared secret key. HMAC's are based on a key that is shared between the two parties, if either party's key is compromised, it will be possible for an attacker to create fraud messages.

Cong Wang et al. [7] proposed a secure cloud data storage system supporting privacy-preserving public auditing. In this paper, they utilized the homomorphic linear authenticator (HLA) and random masking technique to guarantee that the TPA would not learn any knowledge about the data content stored on the cloud server. They further extend their protocol into a multiuser setting, where the TPA can perform multiple auditing tasks in a batch manner for better efficiency. Even though this scheme provides efficient privacy preserving and public auditing, it lags in security concern. Less security is provided to data when compared to the systems that used encryption/ decryption algorithms.

Tejaswani et al. [8] has achieved integrity of data using a Merkle Hash Tree by Third Party Auditor and confidentiality of data is achieved using RSA based cryptographic algorithm. From the table 1, it is clearly seen that different methods have been implemented to check the integrity of the data. But each method has some issues connected with it. The existing methods succeeded in providing privacy preserving along with public auditing but failed to maintain the confidentiality of data. Therefore it is necessary to develop an efficient and secure auditing scheme which has to perform the public auditing effectively by maintaining both the integrity and confidentiality of cloud data.

III. PROPOSED SYSTEM

This paper mainly focuses on security issues of cloud data storage. We consider the cloud data storage service involving three different entities, as illustrated in Figure 1. The proposed scheme consists of three basic entities, they are cloud user, Third Party Auditor (TPA) and cloud server.

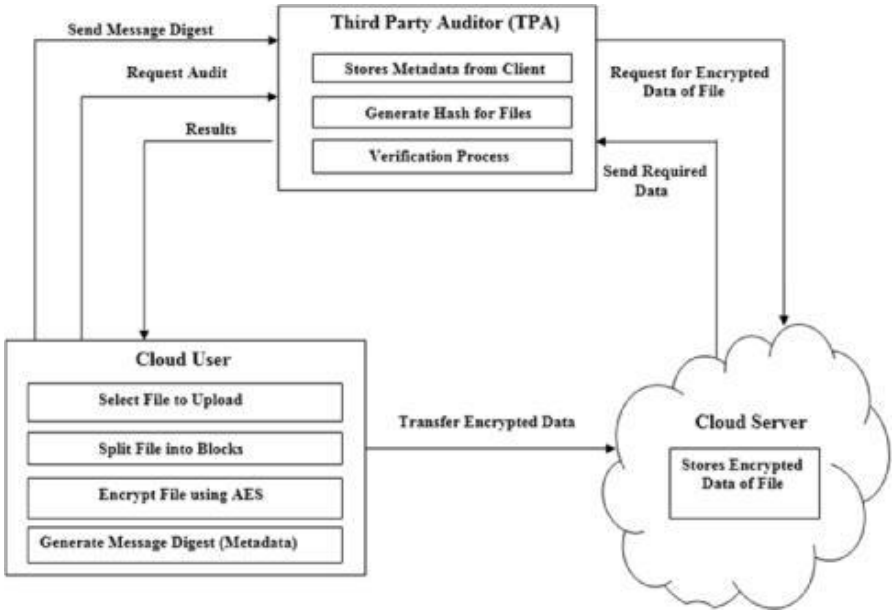


Fig. 1. Architecture of Proposed System

Once the data owner authenticates cloud server by providing his password, the data owner selects the file to be uploaded in cloud and splits the file in to blocks. The blocks are encrypted using AES algorithm followed by generating message digest using Secure Hash Algorithm (SHA-2). A copy of encrypted file is transferred to cloud for storage purpose. Later the message digest is sent to TPA. TPA uses this digest to check the integrity of data stored in the cloud server storage. Since metadata is sent to TPA, TPA will not get enough information about users actual data thus achieves user's data privacy.

In the proposed scheme, TPA is used to perform the task of data auditing. TPA performs data auditing on demand by the client. On receiving the auditing request from cloud user or data owner, the TPA challenges cloud server to send the encrypted data of files that are stored in cloud. After getting the encrypted data from cloud server the TPA follows the same process performed by data owner such as generating message digest for encrypted blocks of data using SHA-2 algorithm.

The working of the cloud user in our proposed system is illustrated in figure 2 and figure 3.

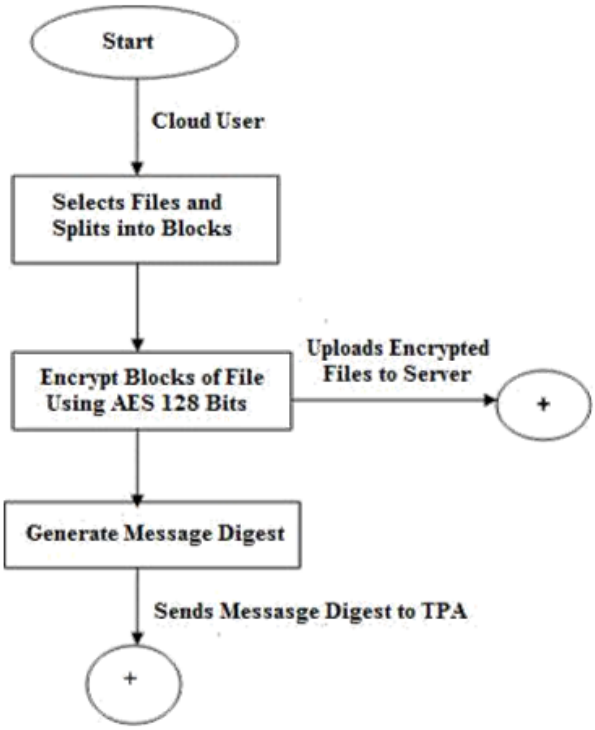


Fig. 2. Working of Cloud User

Later in verification process, it compares the newly generated message digest value with the earlier message digest sent by client. If both the values are matched then it indicates that the integrity of data is maintained. If there exists a mismatch, indicates that data is altered and integrity is not maintained. Finally the TPA will send auditing results to the data owner indicating the status of file. The figure 2 and figure 3 shows the working of the TPA in our proposed auditing scheme.

The cloud server is used to store the encrypted data of files. When it receives request from the TPA, the cloud server will send required encrypted blocks of data to TPA. In the proposed scheme cloud users can upload files to cloud server and can rely on TPA to check the integrity of data stored in cloud server.

IV. IMPLEMENTAION AND RESULTS

The proposed scheme is implemented using python programming on a system with Intel core i3 processor running at 2.9 GHz and 3GB RAM. HTML5 and CSS3 are used to develop front end. AES 128 bit encryption algorithm is implemented using python language and it is used for encrypting blocks of data. Secure Hash Algorith-2 (SHA-256 bit) is also implemented in python by using Hashlib module/library. The Hashlib module is used to implement a common interface to various secure hash algorithms and message digest algorithms.

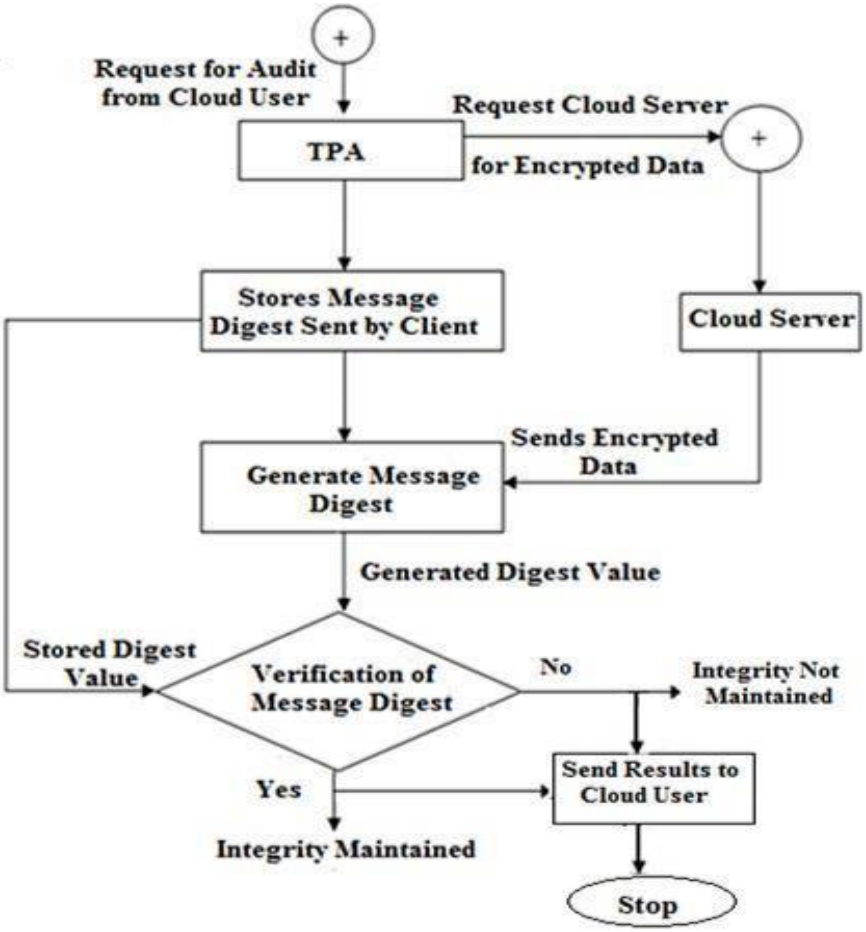


Fig. 3. Working of TPA

SHA 2 is used to generate verification metadata or a digest for a given data. It generates an almost-unique 256-bit (32-byte) cryptographic hash or digest for a text. The following are some of the 32 byte (256 bits) digest we obtained during the analysis of our proposed scheme.

1. 837c8ab3afaa0d10c1a58902d58d46a41085e03650

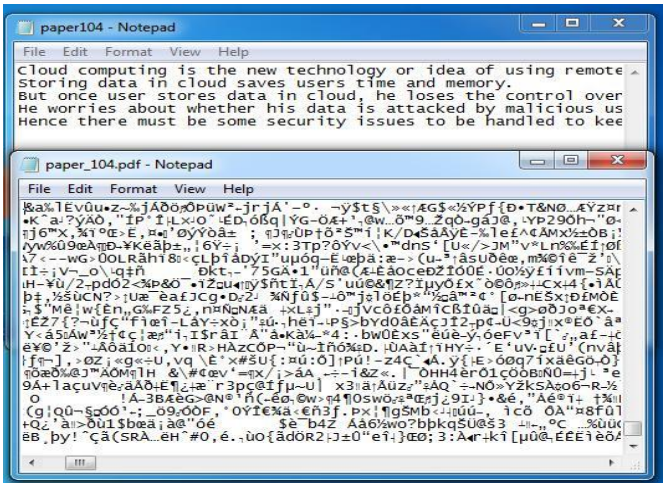
6092134d3fffb97c4507bf.

2. d3bdb4bb29c8e29be3f64f23d4df4673fd2a2f29c39

986933502e8f0465daee3.

3. 410486d8537970d628271a02776c0fdd5dbd14259 581fa53a53677f6ee9cdff5.

The snapshot of encrypted data file stored in the server is shown in figure 4



The figure 6 represents the time taken by TPA to audit the files ranging from 100KB to 1000KB. In our observation we find that TPA will take almost constant time to audit the files of different file sizes.

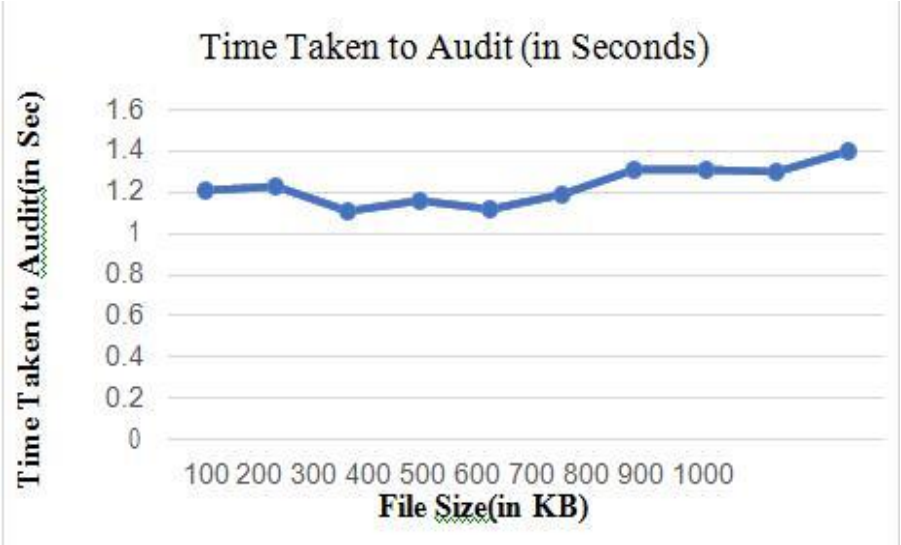


Figure 4 : Sample of Encrypted File

The figure 5 represents the time taken by our system to generate message digest for the encrypted data of different file sizes. From the observations it is seen that our system is approximately 60 percent better than the paper proposed by Abhishek Mohta et al [5].

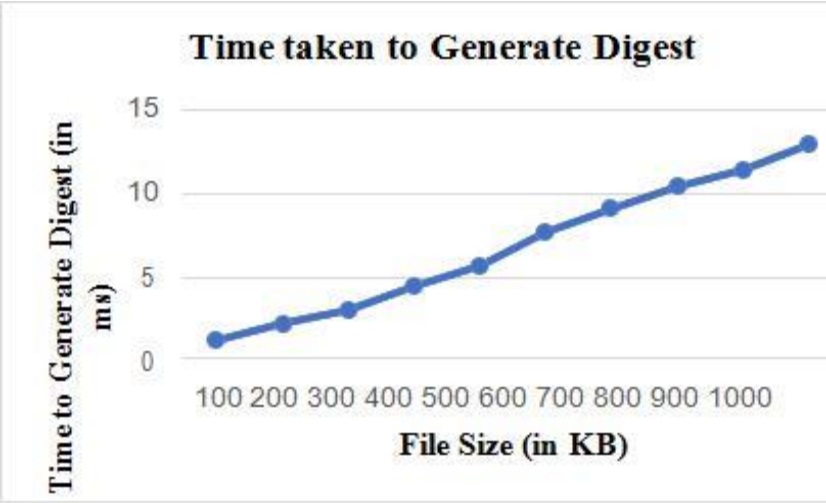


Figure 5: File Size vs. Time to Generate Digest

V. CONCLUSION AND FUTURE WORK

In the cloud storage, users place their data in the cloud and no longer retain the data locally. One of the key issues is to detect the integrity of cloud data. In this paper, we proposed a privacy preserving public auditing system for data storage security in cloud. TPA will perform auditing task without retrieving the data copy of a cloud user, thus achieves privacy preserving. Before uploading any data in cloud, the user's data is encrypted first and then stored in the cloud storage in encrypted form, thus achieving the confidentiality of data. The integrity of data is evaluated by TPA by verifying both the message digest. In the proposed system, TPA checks whether the data stored in cloud is tampered or altered and later intimates the same to the cloud user. All the modules in the system are implemented to develop an effective auditing scheme. In future, we would like to perform data dynamic operations such as updating, deletion and insertion of data.

VI. REFERENCES

[1] Swapnali Morea, Sangita Chaudhari , "Third Party Public Auditing Scheme for Cloud Storage", International Journal of Procedia Computer Science, Volume 79, pp. 69-76, 2016.
[2] Amazon.com, "Amazon Web Services (AWS)," Online at http://aws. amazon.com, 2008.
[3] W. A. Sultan Aldossary, "Data Security, Privacy, Availability and Integrity in Cloud Computing", International Journal of Advanced Computer Science and Applications,Volume 7, Issue 4, pp. 485-498, 2016.
[4] N. K. Yang and X. Jia, "Data Storage Auditing Service in Cloud Computing: Challenges, Methods and Opportunities", World Wide Web, Volume 15, Issue 4, pp. 409–428, 2012.

[5] Abhishek Mohta, Ravikant Sahu, Lalit Kumar. " Robust Data Security for Cloud While Using Third Party Auditor". International journal of Advanced research in CSE (IJARCSE), Volume 2, Issue 2, February 2012.

[6] S Ezhil Arasu, B Gowri, and S Ananthi. "Privacy-Preserving Public Auditing in cloud using HMAC Algorithm". International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277, 3878, Volume 2, Issue 1, pp. 149-152, March 2013.

[7] Pratiksha M, Roshani T, Rajesh B. "A System of Privacy Preserving Public Auditing for Secure Cloud Storage System". International Journal of Engineering Research and Technology (IJERT), Volume 3, Issue 8, pp. 1031-1035, August 2014.