ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Robust Data Integrity Mechanism for Outsourced Cloud Data

Suganya S[1] , Roshni Thanka M[2]

[1]Post-Graduate Student, Department of Computer Science and Engineering, Karunya University, India
[2]Assistant Professor, Department of Computer Science and Engineering, Karunya University, India

Abstract- Cloud computing is used to store data remotely and access high quality applications and services from shared pool of configurable computing resources. The user can only be able to store their data. But, the user will not have any confirmation about whether the data stored in cloud is intact. The public auditability enables an external auditor to audit the outsourced data without downloading the entire data. The dynamic data support allows the user to make changes at any time. All the existing technique can take more time to audit the outsourced data. But, the robust data integrity mechanism for outsourced cloud data mainly focuses on how to reduce the auditing time and improve the security.

Keywords– Public auditability, dynamic data, data storage, cloud computing , integrity

I INTRODUCTION

Cloud computing is referred as delivery of computing resources over the internet [19]. Cloud services allow individuals and organizations to use resources that are managed by external parties [19]. Some of the examples of cloud online file storage, social networking sites, webmail and online business applications [19]. The cloud computing provides resources like networks, computer processing power, data storage space and user applications [19]. The important characteristics of cloud computing are resource pooling, rapid elasticity, on demand self service and measured service [19]. The Software as a Service (SaaS) model provides application along with any required software, operating system hardware and networks [19]. The Platform as a Service (PaaS) model provides operating system, hardware and network. The Infrastructure as a Service (IaaS) model provides hardware and network [19].

The cloud provides benefits to user which includes scalability, reliability and efficiency. Even though cloud computing have some benefits, it also have security concerns too. The user will be sending their data over the internet and is stored in remote locations. But the user will not have any confirmation about the integrity of outsourced data. The public auditability concept allows an external auditor to audit outsourced data on behalf of user. The problem is that the third party auditor will know the exact content which is stored in cloud. Due to its lack of security, we are using a concept called homomorphic linear authentication with random masking technique. This technique allows the Third Party Auditor to audit users cloud data without learning the data content. And also this technique not only allows the auditing for single user case, but also for multi user setting. The dynamic data support allows the user to do any kind of modifications on the outsourced data. In existing system, all these kind of support will be achieved by HMAC algorithm. Our approach uses SHA 512 algorithm which is more secure and takes less time to perform auditing than HMAC algorithm.

The contribution of this work is summarized as the following aspects:

1. Compared to other techniques, the proposed scheme achieves storage correctness assurance and data error localization.
2. This scheme not only supports for static operation but also supports for dynamic operations like update, delete and append.
3. The experiment result shows that the proposed scheme is highly efficient.
4. The SHA 512 algorithm which is used in this project takes only less time to perform auditing than HMAC algorithm. This approach is highly secure due to its larger key size.

ISSN: 2321-9653

II RELATED WORKS

The Provable Data Possession(PDP) [2] technique allows the client to verify the data which is stored in cloud server. The server may delete some part of data or it may not store all data in cloud storage. The Proof Of Retrievability(POR) [2] uses special blocks called sentinals to verify storage correctness of data. These two techniques used symmetric key cryptography which is unsuitable for public verification.

The probabilistic proof [1] technique allows the client to verify whether the server possesses the original data without retrieving it. The client maintains the constant amount of meta data to verify the proof by comparing it. But this technique will not provide dynamic data support and public verifiability.

Homomorphic verifiable responses and hash index hierarchy [17] uses multiple cloud service providers which can cooperatively store and maintain the users data. In this technique the response collected from multiple cloud service providers can be combined into a single response as the final result. But it is very difficult to check the data stored in each cloud one by one.

The sentinel based Proof Of Retrievability protocol [7] encrypts the file and embeds a set of random check blocks into it which is called sentinals. The verifier sends the challenge request to the prover by specifying the position of sentinel block. After thet the prover will return the particular sentinel block to the verifier for integrity verification. If the prover did any modification in some portion of file then the sentinel will also gets changed. This technique will also have some drawback which is computational overhead.

The challenge response protocol for multiple replica provable data possession [14] allows the client to store N number of replicas in cloud storage. If any of these replicas gets deleted then the data still be available in other replicas. The main drawback of this technique is sometimes all the replicas may fail simultaneously. At that time the replication will not help to verify the correctness of data.

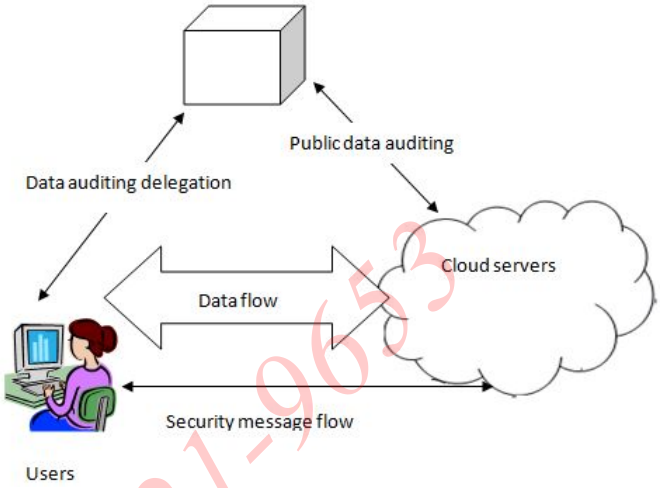

Fig. 1. Architecture of cloud data storage

In digital signature technique [6], the user and external auditor will generate their own public and secret keys with respect to RSA algorithm. The message is initially encrypted and signed using respective keys and this encrypted message is then stored in cloud. This technique provides storage security to cloud users. But it will not provide support for dynamic data auditing. The architecture diagram of cloud data storage is shown in fig. 1.

III PROPOSED SCHEME

In our proposed scheme SHA 512 algorithm is used to perform auditing more faster. The security will be high due to its larger key size. All the existing technique uses HMAC algorithm which takes more time to perform auditing.

A. HMAC Algorithm

The message authentication code is a widely used technique for performing message authentication to verify the integrity of data. The MAC algorithms involve the use of secret key to generate a small block of data which is known as Message Authentication Code. The HMAC refers Hash based Message Authentication Code [4]. In prior work, the auditing is performed with the help of MD5 with HMAC. The MD5 is a widely used cryptographic hash function producing a 128 bit hash value and is used to verify data integrity. The HMAC algorithm may generate the hash code which consist of more alphabets. The SHA 512 algorithm converts any stream of data

into 160 bit value. The key size of SHA 512 algorithm is lesser than HMAC algorithm. And the SHA 512 algorithm is more secure and it contains unbreakable hash code.

B. SHA 512 Algorithm

In this work, SHA 512 algorithm is implemented in order to perform auditing faster. The hash code generated by SHA 512 contains hexadecimal code rather than alphabets. So the hackers cannot easily hack the data which is stored in cloud. The computation overhead is minimized in this work.

IV AUDITING PROCESS

These are all the steps which is involved in auditing:

A. Data block computation and homomorphic key generation

The client will be storing their files in cloud server. Initially the client file will be splited into several blocks for easy integrity verification. The files are splitted based on its file size and is stored in cloud server. The key generation is performed for each data block and it is the process of generating public key parameters and hash code in the form of verification meta tag. The simplest method to read encrypted data is a brute force attack. Therefore it is sufficient to use longer key length. Because longer key size takes exponentially longer time to attack the data.

B. Homomorphic Linear Authenticator(HLA) based MAC signature

The files will be encrypted using their corresponding keys. After that, store the keys and data in a hash table. Because we cannot do search the entire data, just we are going to search the index of data. So the process will be very speed. The authentication will be processed in both TPA and cloud server for data security and data storage correctness.

C. Data storage on cloud server

The encrypted files will be stored in a different location of the cloud server. The Third Party Auditor will use their corresponding keys to perform data verification. The TPA cannot see the original data. The TPA can only check the validation using signature scheme in cryptography.

D. TPA integrity verification

The TPA will audit the data by comparing the MAC code of already stored data with the retrieved one. If both the MAC code are same then the TPA will confirm that the data has not modified.

E. Data dynamics

The user can not only store the data. But they can perform any kind of modifications like insert, delete and append on the outsourced data.

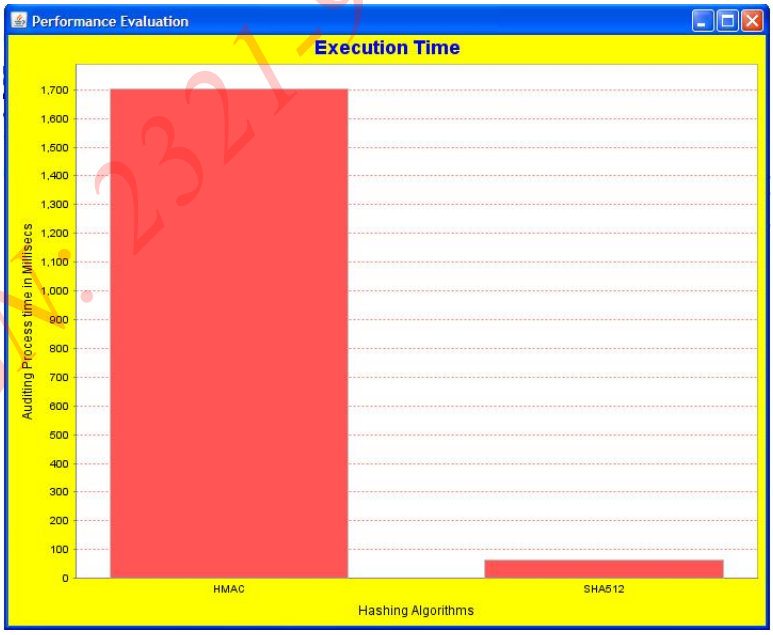

Fig. 2. Performance evaluation

V RESULT ANALYSIS

The TPA can audit only one task at a time using individual auditing case. But in batch auditing the TPA can audit N number of task from different users simultaneously. The auditing time will be reduced in batch auditing than individual auditing. The average per task auditing time is calculated by dividing the total auditing time by the number of task involved. The performance evaluation of HMAC and SHA 512 algorithm is shown in fig. 2.

VI CONCLUSION

To ensure cloud data security, it is very important to enable an external auditor to audit client data without cheating them. In this paper, we explored the problem of providing simultaneous public auditability and data dynamics for remote data integrity check in cloud computing. In order to perform auditing faster SHA 512 algorithm is implemented which has longer key size. Because longer key size needs more time to attack the system. The security and performance analysis shows that the proposed scheme is highly efficient and provably secure.

REFERENCES

[1] Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 598-609, 2007.

[2] Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. Int'l Conf. Security and Privacy in Comm. Networks (SecureComm '08), pp. 1-10, 2008.

[3] Boneh, B. Lynn, and H. Shacham, "Short Signatures from the Weil Pairing," J. Cryptology, vol. 17, no. 4, pp. 297-319, 2004.

[4] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Storage Security in Cloud Computing," Proc.IEEE INFOCOM '10, Mar. 2010.

[5] Erway, A. Kupcu, C. Papamanthou, and R. Tamassia,"Dynamic Provable Data Possession," Proc. ACM Conf. Computer and Comm. Security (CCS '09), pp. 213-222, 2009.

[6] Govinda, V. Gurunathaprasath, H. Sathishkumar, "Third Party Auditing for secure data storage in cloud through digital signature using RSA", International Journal of Advanced Scientific and Technical Research, (ISSUE 2, VOLUME 4-August 2012).

[7] Juels and J. Burton, S. Kaliski, "PORs: Proofs of Retrievability for Large Files," Proc. ACM Conf. Computer and Comm. Security (CCS '07), pp. 584-597, Oct. 2007.

[8] K.D. Bowers, A. Juels, and A. Oprea, "Proofs of Retrievability:Theory and Implementation," Proc. ACM Workshop Cloud Computing Security (CCSW '09), pp. 43-54, 2009.

[9] M.A. Shah, M. Baker, J.C. Mogul, and R. Swaminathan, "Auditing to Keep Online Storage Services Honest," Proc. 11th USENIX Workshop Hot Topics in Operating Systems (HotOS '07), pp. 1-6, 2007.

[10] M.A. Shah, R. Swaminathan, and M. Baker, "Privacy-Preserving Audit and Extraction of Digital Contents," Cryptology ePrint Archive, Report 2008/186, 2008.

[11] Makhija, V. Gupta, I. Rajput, " Enhanced Data Security in Cloud Computing with Third Party Auditor", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 3, Issue 2, February 2013.

[12] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," IEEE Trans. Parallel and Distributed Systems, vol. 22, no. 5, pp. 847-859, May 2011.

[13] R. Curtmola, O. Khan, and R. Burns, "Robust Remote Data Checking," Proc. Fourth ACM Int'l Workshop Storage Security and Survivability (StorageSS '08), pp. 63-68, 2008.

[14] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession," Proc. IEEE Int'l Conf.Distributed Computing Systems (ICDCS '08), pp. 411-420, 2008.

[15] Sebe, J. Domingo-Ferrer, A. Martı´nez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient Remote Data Possession Checking in Critical Information Infrastructures," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 8, pp. 1034-1038, Aug.2008.

[16] Shacham and B. Waters, "Compact Proofs of Retrievability,"Proc. Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (Asiacrypt), vol. 5350, pp. 90-107,Dec. 2008.

[17] Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. Yau, "Efficient Provable Data Possession for Hybrid Clouds," Cryptology ePrint Archive, Report 2010/234, 2010.

[18] P. Mell and T. Grance, "Draft NIST Working Definition of Cloud Computing," http://csrc.nist.gov/groups/SNS/cloudcomputing/ index.html, June 2009.

[19] http://www.priv.gc.ca/resource/fs-fi/02_05_d_51_cc_e.pdf