International Journal of Technology Research and Management
ISSN (Online): 2348-9006
Vol 3 Issue 10 October 2016

A Study for Authentication and Integrity of Data Files in Cloud Computing

Utkarsh Pandey[1], Prof. Anurag Jain[2]
M. Tech., RITS, Bhopal[1]
HOD, CSE, RITS, Bhopal[2]
utkarsh.pandey581@gmail.com[1], anurag.akjain@gmail.com[2]

Abstract: *With the development of cloud computing, a large number of individuals and companies transfer their data and business to cloud computing. But the management of the data and services of the cloud may not be fully trustworthy. How to ensure the integrity of data stored in cloud attracts more and more attention.*

A number of data files authentication and integrity schemes have been conducted to recognize any modification in the exchange of data files between two entities within a cloud environment. Existing solutions are based on combining key-based hash function with traditional factors (steganography, smart-card, timestamp). However, none of the proposed schemes appear to be sufficiently designed as a secure scheme to prevent from attacks. In recent years, lots of efficient and safe solutions to ensure the outsourced data integrity are proposed. This paper provides a survey of the main research results of the previous studies.

Keywords: *Cloud Computing, Security, Privacy, Integrity Checking, Authentication.*

1. INTRODUCTION

The "Cloud Computing" is continuously evolving terminology and concepts in IT enterprises. Cloud computing is Internet based development and use of computer technology [1]. Cloud computing describes a new supplement, consumption and delivery model for IT services based on Internet, and it typically involves the provision of dynamically scalable [2] and often virtualized resources as a service over the Internet and details are abstracted from the users who no longer need knowledge.

Cloud computing can be confused with Grid computing or cluster computing [3] (a form of distributed computing and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks) and Utility computing [4] (packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity).

Cloud computing has a variety of characteristics:

Shared Infrastructure: Uses a virtualized shared resources or services [1,2].

Dynamic Provisioning: Allows for the provision of services based on current requirements [2].

Broad Network Access: Needs to be accessed across the internet from anywhere over a broad range of devices [3].

Managed Metering: Uses metering for managing and optimized service provided [4].

In short, cloud computing allows for the sharing and scalable deployment of services, as needed, from almost any location,

and for which the customer can be billed based on actual usage.

Once a cloud is established, how its cloud computing services are deployed is the issue. For this deployment three basic deployment model is designed i.e. Software as a Service (SaaS)- Clients obtaining the ability to access and use an application or service that is presented in the cloud, eg. Salesforce.com [5]. Platform as a Service (PaaS)-Users rent the accessing platforms, to deploy their own software and applications in the cloud [6]. Infrastructure as a Service (IaaS)-Users control and manage the resources such as operating systems, storage and network connectivity, but do not they control the cloud infrastructure [4,6].Deploying cloud computing can differ depending on requirements, and the following four deployment models have been identified i.e. Private cloud (cloud infrastructure is maintained and operated for a specific organization), Public cloud (cloud infrastructure is available to the public on a commercial basis), Community cloud (cloud infrastructure is shared among a number of organizations with similar interests) and Hybrid cloud(cloud infrastructure consists of a number of clouds of any type)[2-6].

2. LITERATURE REVIEW

Nowadays, transmitted data has radically magnified and has been exponentially distributed [7]. Cloud computing is usually thought to be the computing infrastructure of future generation; it's an efficient method suggested that of enabling users to utilize giant volumes of resources and provides an efficiently and without delay accessible on-demand service [8]Consequently, information security requirements is increased as compared with other challenges [9]. Document integrity have become huge security issue. Different schemes based on key-based hash function to ensure document authentication and integrity have been proposed. In [10-12], the authors proposed schemes that combining steganography approach with hashed value to transfer it securely. Most of those schemes used sequence mapping in the least significant bits (LSBs) of a cover-image to hide such value, which lacked in hiding efficiency and brought up a lot of security problems. Meanwhile, simply combining traditional factors (smart-card, timestamp, and shared key) with cryptographic hash function [13-17] are the most widely used methods to overcome security challenge and support data integrity routinely. These schemes still suffer from drawbacks related to such factors, thus these techniques might be vulnerable to common form of attacks. The probability of common attacks can be reduced by combining cryptographic hash function with a strong factor that should be periodically changed. For this reason, one-time authentication scheme [18, 19] can ensure the integrity and authentication, used one-time key to achieve such a goal. In this paper, we report a scheme designed with a one-time bio-key. In details, the key used in the proposed scheme captures the advantages of a biometric technique that involves the use of the robust features extracted by the histogram of local binary pattern filter (LBP) [20] after combining the handwritten signature of the sender, the handwritten signature of the receiver, for generating one-time bio-key. This key is combining with MAC-SHA-1. The result of the combination is one-time message authentication code. Thereafter, the summation of such code called MACLESS is hidden in a cover image through LSBs [21] and discrete wavelet transformation (DWT) [21] based steganography anonymity. Concealing the anonymity of MACLESS depends on the one time random pixel sequences generated by Rivest Cipher 4 (RC4) [22] based on a one-time biometric stego-key. Although, a cloud service provider is needed in the proposed scheme, such provider in our scheme does not provide service in run time, but only in configuration phase. In [23] presents a new and efficient one-time MACLESS to ensure message/image document integrity and authenticity among users in a cloud environment. A robust method is developed by extracting handwritten signature features to generate a symmetric one-time bio-key and stegno-key. Furthermore, this technique can be employed to maintain the authenticity of the transferred document, and verify the integrity of the received document. Overall, the scheme is secure and simple to use. In [24], intended a privacy-preserving public auditing system for data storage security in Cloud Computing. We utilize 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 the efficient auditing process, which not only eliminates the burden of cloud user from the tedious and

possibly expensive auditing task, but also alleviates the users' fear of their outsourced data leakage.

3. REPRESENTATIVE APPROACHES AND ANALYSIS

Generally, the efficiency of the schemes supporting private verification is higher, but the data integrity verification can only be verified by the data owner himself. The schemes supporting public verification allow data owner himself or any other delegated trusted third party to complete the data integrity verification task. In the cloud computing environment, when clients store a large amount of data in CSS, in consideration of its own computing capacity constraints and communication overhead, etc. they can choose to mandate the data integrity verification to Third Party Auditor [25] which has the technical expertise and powerful computing capabilities. Thus, a tripartite relationship is formed in the scheme of data integrity audit. As is shown in the Figure 1.

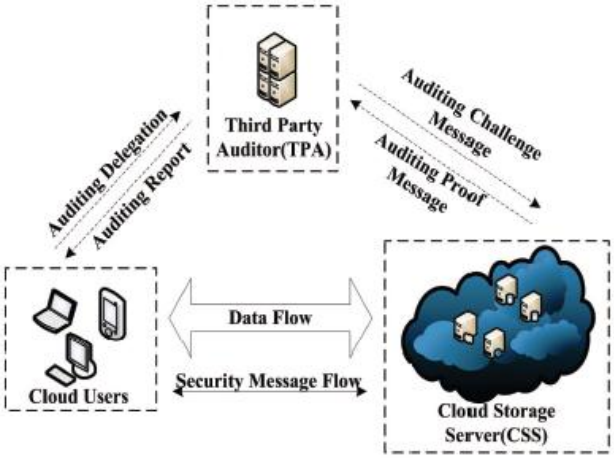

Figure 1: Third party auditor [25]

A public auditing system consists of two phases, Setup and Audit phase:

- Setup: The user initializes the public and secret parameters of the system and pre-processes the data file to generate verification metadata. The user then stores the data file and the verification metadata at the cloud server, and deletes its local copy. As part of pre-processing, the user may alter the data file by expanding it or including additional metadata to be stored at server.

- Audit: The TPA issues an audit message or challenge to the cloud server to make sure that the cloud server has retained the data file properly at the time of the audit. The cloud server will derive a response message from a function of the stored data file and its verification metadata. The TPA then verifies the response.

There are two class of schemes that is used for integrity checking using TPA as follows:

MAC-based Solution: There are two possible ways to make use of MAC to authenticate the data. A trivial way is just uploading the data blocks with their MACs to the server, and sends the corresponding secret key to the TPA [25]. Later, the TPA can randomly retrieve blocks with their MACs and check the correctness via secret key. Apart from the high (linear in the sampled data size) communication and computation complexities, the TPA requires the knowledge of the data blocks for verification. To circumvent the requirement of the data in TPA verification, one may restrict the verification to just consist of equality checking.

HLA-based Solution: To effectively support public audit ability without having to retrieve the data blocks themselves, the HLA technique [25] can be used. HLAs, like MACs, are also some unforgivable verification metadata that authenticate the integrity of a data block. The difference is that HLAs can be aggregated. It is possible to compute an aggregated HLA which authenticates a linear combination of the individual data blocks.

PDP: PDP (provable data possession) [26] can provide efficient verification over their outsourced data and the corresponding security proof is given in the random oracle model. It is believed to be the first scheme to provide block less verification and public verifiability at the same time. The scheme proposed is based on RSA signature.

DPDP: If in the integrity verification scheme of the file, (for example, PDP), the index of file blocks is included in the metadata of the file block verification [27]. While the scheme

becomes infeasible when the acts of insertion, deletion and modification of a block are operated on the scheme. In order to effectively support these dynamic updates, based on the structure of rank-based authenticated skip list, DPDP (Dynamic PDP) is proposed.

Public Auditing of Dynamic Data: This scheme supports both dynamic data and public audit ability which uses the Merkle hash tree [28]to manage the authentication information in the scheme. Its leaf nodes are the hashes of the ith block of the file. The leaf nodes of the MHT are ordered. The MHT of this scheme can authenticate both the values and the positions of data blocks. Root is the authentication metadata, and the client gives the Hash function using the private key.

4. CONCLUSION

In order to verify the integrity of outsourced data, in different circumstances, an analysis is carried out of different effective and secure models and algorithms. In this paper only some representative model of scheme were analyzed. Although the existing research results have already achieved some satisfying goals, the topic of integrity verification of outsourced data in cloud computing has been attracting more and more interest from researchers and there is still a long way to go. Especially with the advent of the era of Big Data, the scale of data and applications grow exponentially, and this will bring more heavy computation to algorithm based on cryptography. So, huge challenges will be brought to data integrity auditing.

REFERENCES

[1] Fang Liu, Jin Tong, Jian Mao, Robert Bohn, John Messina, Lee Badger and Dawn Leaf, "NIST Cloud Computing Reference Architecture", US Department of Commerce, Gaithersburg, MD, 2011.
[2] P. Mell and T. Grance, "The nist definition of cloud computing", special publication 800-145, US Department of Commerce, Gaithersburg, MD, 2011.
[3] Bhaskar Prashad Rimal, Eunmi choi, Ian Lumb, "A Taxonomy and Survey of Cloud Computing System", International Joint Conference on INC, IMS and IDC, IEEE, 2009.
[4] Armbrust M, Fox A, Griffith R, Joseph D A, Katz H R, Konwinski A, Lee Gunho, Patterson A D, RabkinA, Stoica A, Zaharia M, "Above the clouds: A Berkeley view of Cloud Computing", UC Berkeley, EECS, 2010.
[5] Bhaskar Prasad Rimal, Admela Jukan, Dimitrios Katsaros, Yves Goeleven, "Architectural Requirements for Cloud Computing Systems: An Enterprise Cloud Approach", Journal of Grid Computing, Springer, 2010.
[6] Rajkumar Buyya, Karthik Sukumar, "Platforms for Building and Deploying Applications for Cloud Computing", CSI Communications, 2011.
[7] T. Rethika, I. Prathap, R. Anitha, and S. V. Raghavan, "A novel approach to watermark text documents based on Eigen values", Proceedings of the Ninth International Conference on Network and Service Security (N2S'09), France, IEEE,pp.1-5, 2009.
[8] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches", Wireless Communications and Mobile Computing, John Wiley, vol. 13, no. 19,, pp. 1587-1611, Dec. 2012.
[9] A. T. Velte, T. J. Velte, and R. Elsenpeter, "Cloud Computing: A Practical Approach", McGraw-Hill Companies, 1st Edition, 2010.
[10] J. Shen and K. Liu, "A Novel Approach by Applying Image Authentication Technique on a Digital Document", Proceedings of International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, pp. 119-122, June, 2014.
[11] J. Qiu and P. Wang, "An Image Encryption And Authentication Scheme", Proceedings of Seventh International Conf. on Computational Intelligence and Security (CIS), China, pp. 784-787, Dec, 2011.
[12] N. Jamil and A. Aziz, "A Unified Approach to Secure and Robust Hashing Scheme for Image and Video Authentication", Proceedings of 3rd IEEE International Congress on Image and Signal Processing(CISP), Yantai, China, Oct., pp. 274-278, 2010.
[13] Z. Liu, H. S. Lallie, L. Liu, Y. Zhan, and K. Wu, "A hash-based secure interface on plain connection", Proceedings of the sixth International Conference on Communications and Networking in China(ChinaCom'11), Harbin, China, pp. 1236-1239, IEEE, 2011.
[14] N. Rabadi and S. Mahmud, "Drivers anonymity with a short message length for vehicle-to-vehicle communications network", Proceedings of the fifth IEEE Consumer Communications and Networking Conference(CCNC'08), Las Vegas, NV, USA, IEEE, pp. 132-133, Jan. 2008.
[15] S. I. Naqvi and A. Akram, "Pseudo-random key generation for secure HMAC-MD5", Proceedings of the 3rd IEEE International Conference on Communication Software and Networks (ICCSN), Xi'an, China, pp.573-577, May, 2011.
[16] C. Chaisri, N. Mettripun, and T. Amornraksa, "Facsimile Authentication Based on MAC", IT Convergence and Services, Lecture Notes in Electrical Engineering, vol. 107, pp. 613-620, 2012.

[17] K. Alla, G. G. Shankar, and G. B. Subrahmanyam, "Secure Transmission of Authenticated Messages using New Encoding Scheme and Steganography", Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology, Coimbatore UNK, India, pp. 749-752, 2012.
[18] J. Song and S. Han, "One-time key authentication protocol for PMIPv6", Proceedings of the Third International Conference on Convergence and Hybrid Information Technology (ICCIT'08), South Korea, IEEE, pp. 1150-1153. 2008.
[19] J. Y. Park, D. Lee and H. H. Lee, "Data Protection in Mobile Agents; one-time key based approach", Proceedings of the 5th International Symposium on Autonomous Decentralized Systems (ISADA'05), USA, IEEE, pp.411-418. 2001.
[20] M. A. Ferrer, F. Vargas, A. Morales, and A. Ordonez, "Robustness of offline signature verification based on gray level features", IEEE Trans. Inform. Forensics Security, vol. 7, no. 3, pp. 966-977, Jun. 2012.
[21] P. Wayner, "Disappearing Cryptography: Information Hiding: Steganography & Watermarking", Morgan Kaufmann, 3th Edition, 2009.
[22] W. Stallings, "Cryptography and Network Security: Principles and Practice", Prentice Hall, 6th Edition, 2013.
[23] Zaid Ameen Abduljabbar, Hai Jin, Ali A.Yassin, Zaid Alaa Hussien, "Robust Scheme to Protect Authentication Code of Message/Image Documents in Cloud Computing", IEEE, 2016.
[24] Solomon Guadie Worku, Chunxiang Xu, Jining Zhao, Xiaohu He, "Secure and efficient privacy-preserving public auditing scheme for cloud storage", Computers and Electrical Engineering, Elsevier, 2013.
[25] 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, 2013.
[26] G. Ateniese, R. B. Johns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores", in Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS '07), pp. 598-609, 2007.
[27] C. Erway, A. Küpçü, C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession", in Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS□09),Chicago, pp. 213-222, USA, 2009.
[28] Q. Wang, C. Wang, K. Ren, W. Lou and J. 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, 2011.