Multiauditing Based Cloud Storage Using a Dynamic Hash Table

OPEN ACCESS

Manuscript ID:
ASH-2020-08012763

Volume: 8

Issue: 1

Month: July

Year: 2020

P-ISSN: 2321-788X

E-ISSN: 2582-0397

Received: 13.04.2020

Accepted: 29.05.2020

Published: 02.07.2020

Citation:
Arun Kumar, K., et al. "Multiauditing Based Cloud Storage Using a Dynamic Hash Table." *Shanlax International Journal of Arts, Science and Humanities*, vol. 8, no. 1, 2020, pp. 143–148.

DOI:
https://doi.org/10.34293/sijash.v8i1.2763

K. Arun Kumar[1], R. Mohammed Harun Babu[2], S. Kalaivanan[3] and V. Kanimozhi[4]
[1,3]Senior Faculty and [2,4]Faculty
iNurture Education Solutions Pvt. Ltd., Bengaluru, Karnataka, India
https://orcid.org/0000-0001-7229-6662

Abstract
Cloud depository is one of the customary supplications of the cloud computing system, which offers on-demand offloading services for both individuals and institutions. Although utilizes do not have full faith in the cloud service providers (CSPs) within that, it is hard to decide either the CSPs meet their licit expectations for data security or not. Thence, it is evaluative to come up with productive auditing techniques to boost owners' trust and reliance in cloud storage. In this paper, we are presenting a scheme of auditing for assured cloud storage based on a 2-dimensional data structure called a dynamic hash table (DHT), used to record the data information for public auditing. This scheme emigrates the approved information from the CSP to the TPA and thereby a remarkable reduction in the computational cost and communication overhead. Also, the deduplication technology is utilized to lower the capacity and bandwidth prerequisites of the utilities by removing repetitious information and stockpiles, only an original replica of them. We enhance our design encourages privacy preservation by homomorphism authenticator constructed on the public key, and attend batch auditing by aggregate BLS signature technique. Experimental results indicate that our mechanism achieves secure deduplication and tag generation improvements.

Keywords: Public auditing, Cloud security, Data storage and Deduplication of data

1. Introduction

Storage depositary in the cloud is an influential field of cloud computing, whose intent is to supply on-demand data out - sourcing facility for end-users through distinctly virtualized infrastructures. Because of the outrageous performance and reasonable cost and of cloud depository, an increasing number of organizations and users are frequently outsource their data storage to Cloud Services Providers (CSP), However, as a cloud storage technology still encounters numerous challenges in security. A portion of the consternation is how to check whether a cloud container system and its contributor reach the constitutional expectations of users for the security of data. The major issues in data security include data privacy, data protection, data attainable, data placement, and secure data dissemination, hazard, data depletion, service disruption, unauthorized attacks, and the data malversation issues. Hence from the user's point of view, security, honesty, privacy, and confidentiality of the preserved data on the cloud must be considered essential requirements. To procure all of the above requirements, the latest methods or techniques should be developed and to be accomplished.

Need for this Work

Data auditing technique is initiated in Cloud computing that acts with solid data storage. Auditing is an operation of inspecting the user data, which is agreed by the data owner or a TPA. It assists in keeping the integrity of stored data on the cloud. The TPA is an entity that can act in favor of the client or owner of data, who has all the required expert knowledge, capabilities, and professional mastery that are required to handle the functions of integrity verification, which reduces the burden of the client. It must be crucial that TPA should and frequently systematically audit the data in the cloud on the request of the user.

Motivation

Distributed cache service is one of the significant facilities provided by the distributed computing, where the customers can easily arrange themselves as the clump and share the data among themselves. Nowadays, as many customers are sharing the data, cloud storage utility is associated with expanding the capacity of information cached at distant servers. Hence, one critical challenge of today's distributed depository utility is to manage the ever-evolving capacity of data. Instead of maintaining multiple duplicates of information with identical content, deduplication deletes monotonous records by maintaining only one physical version and pointing to that copy of the other redundant documents. The paper focuses on effective auditing and deduplication of information submitted by proprietors of information as well as tests for deduplication of the blocks of current customers.

Contribution

In this paper, we suggest Secure Deduplication & Auditing of data Shared in Cloud mechanism that supports secure document level & block-level deduplication. Our contributions compiled as follows:

- We present a public auditing scheme, which can thoroughly assist functions like dynamic auditing of data, batch auditing, and data deduplication.
- We design DHT a data structure to track data premises for auditing in the TPA to achieve efficient data updating and instant auditing
- We propose a Secure Deduplication and Audit of Shared Data in Cloud (STLDAS) scheme which supports secure document level deduplication and block level.
- The algorithm supports stable deduplication and has greatly reduced the time cost of creating tags. Experimental analysis manifests the adeptness and efficacy of Deduplication and Auditing of Shared Data in Cloud mechanism.

Organization

The paper list is structured as follows: In Section 2, we clarify the Relevant Works that include the pros and cons of current integrity audit and deduplication schemes. In Section 3, we discuss the earlier models and their drawbacks. In Section 4, we discuss several preliminaries. In Section 5, We describe the problem statement and model of the framework that illustrates the functioning of the architecture and the specifics relating to the design objectives. In Section 6, we explain scheme details of our Secure Two-Level Deduplication and Auditing of Data in Cloud. We clarify the Security Review in Section 7. In Section 8, we report the experimental test results. Conclusions are given in Section 9.

2. Related Work

As our work is incorporated with both dynamic Auditing and deduplication, we study the works in two of these areas in the following sections. Confirmable information ownership and Proofs of Retrievability (PoR) were originally suggested by Ateniese et al., and Juels et al. The homomorphic authentication method was incorporated in their techniques to minimize both the cost of transmission and the cost of reckoning. Subsequently, numerous alternatives of PDP and PoR strategies are constructed to increase the adeptness and upgrade the performance of fundamental strategies, such as permitting public validation and supporting information update. Jian Shen et al proposed scheme which comprises of public auditing with batch auditing

Pulse block less verification, where data dynamics are efficiently supported. The novel dynamic structure includes a location array, and doubly linked info table with this framework, computational and communication overheads can be considerably minimized. Analysis of security specifies that the scheme can conclude the given efficiency in practice with desired properties. But the entente consists of a two phases configuration or setup and verify phase, out of which only the verify phase commences cost of communication. Take-young Youn et al proposed a schema that performs both deduplication of data and public auditing of data. The scheme performs challenge-response protocols using the BLS signature-based homomorphic linear authenticator. The third-party auditor for conducting the public audit to clients. This scheme satisfies all the basic security requirements. But this scheme increases the computational overhead at the cloud storage server.

Deduplication in the cloud and other storage platforms is an operation in which frequent or duplicate data is removed from a data stream to minimize the amount of physical data stored in a setup or system. However, client-side deduplication is accompanied by the disclose of side-channel information. Halevi et al developed the proof of proprietorship mechanism that lets a consumer effectively prove to a server that the particular customer owns this document. Venugopal et al utilize soft computing methods for data mining applications. Geeta et al have performed an extensive review of the latest methods in information auditing and security in cloud computing. Y. Zhu et al suggested a schema that classifies the data premises for auditing using the IHT, and stores them in the TPA instead of the CSP. Thus, it can minimize computational costs and communication overhead. However, its updating operations (particularly, the insertion and deletion ones) are inefficacious, since they would induce the rearrangement of a mean of N/2 components in the IHT, where N indicates the number of blocks, due to the liner structure of the IHT. Moreover, the functions would automatically change the sequence numbers of a few blocks and eventually creates the recalculations of their labels, which would create extra computational costs of the CSP and unnecessary communication overhead.

3. Background Work

Hui Tian et al proposed a public auditing scheme for cloud storage rest on the dynamic hash table (DHT), which is a 2-dimensional data structure placed at a third parity auditor (TPA) to track the data property information for dynamic auditing. The proposed scheme emigrates the authorized information from the CSP to the TPA, also achieves higher updating efficiency, and encourages privacy preservation by integrating the homomorphic authenticator found on the public key with the random masking created by the TPA, and perform batch auditing by performing the aggregate BLS signature technique. However, the search operation on the DHT during the verification may cost more time than the IHT. And this schema does not support the deduplication of data where storage cost of data will be more.

4. Preliminaries

Bilinear Maps

Bilinear maps are the tool of pairing-based crypto; let consider cyclic groups G_a, G_b and G_t be groups of the same order. A bilinear map from $G_a \times G_b$ to G_t is a function $e : G_a \times G_b \rightarrow G_t$ such that for all $u \in G_a$, $v \in G_b$, y, $z \in Z$, $e(u^y,v^z) = e(u,v)^{yz}$. Bilinear maps are called pairings because they relate pairs of elements from G_a and G_b with elements in G_t.

Computational Diffie-Hellman (CDH) Problem: The Computational Diffie-Hellman (CDH) problem is that, given g, g^m, $g^n \in G$ for unknown m, $n \in Z_p$, to estimate g^{mn}.

Homomorphic Verifiable Authenticator (HVA)

HVA is globally employed as a basic construction block for auditing, which allows a public auditor to authenticate the integrity of data stored in the cloud without retrieving or downloading the actual data. Typically, digital signatures (such as RSA-based signature and BLS-based signature) are used to induce HVAs.

5. Problem Definition and System Model

Problem Definition

Given the Cloud storage Model, the owner of the data outsources the document to the distributed server, group of customers distributes this document, the main objectives are:

- Public auditing Scheme that supports dynamic data auditing
- Data structure named Dynamic Hash Table(DHT) is designed to track data properties for auditing in the TPA
- To perform a secure document level or hunk level deduplication of data

System Model

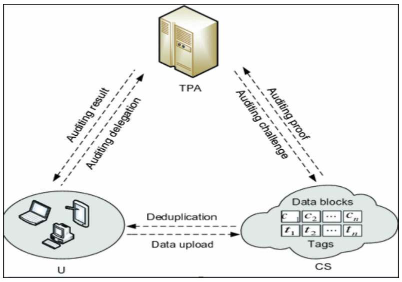

Fig 1 System Model

In this work, we focus on the model of a beneficial public auditing scheme based on the DHT illustrated in Fig. 1, which presumes the subsequent three entities: User, who stores a considerable volume of data records in the cloud, can be an individual or an organization; Cloud Service Provider, who controls and coordinates several cloud servers to proffers ascendible and on-demand outsourcing data facilities for users; and Third-Party Auditor, who can justify the reliability of the cloud storage utilities(CSS) tenable and devoted on behalf of the users after request. Users restful to the burden of storage and computation while enjoying the storage and prolongation service by externalization of their data into the CSP. Original customer or data owner Shared data are divided into blocks and sign with the secret key_k and upload to the CSP. The CSP deduplicates; once the file is saved, the CSP intimates the original consumer that the file already exists. If the file is not a duplicate, the CSP will save the file.

6. The Algorithm

System Setup

Consider two multiplicative groups G_1, G_2 of order p, and e: G_1 G_1 G_2 be a bilinear map. H is a hash function with H: (0,1) G_1; assuming that document is divided into n blocks i,e F = (b_1, $b_{2,}$.....b_n) and outsourced to the CSP. Let u be the user or customer of the cloud.

Function: Key generation

1) Generates the key pairs of public and private keys.
2) Input: u, u1, global parameter (g, Zp)
3) Output: pki, ski
4) for each i up to n
5) Generate random number x from Zp
6) Assign Private key ski= xi
7) Compute Public key pki=g^x
8) The user creates data information that contains id's of all blocks in the document.
9) For each block, user creates the signature
10) End

File Uploading

User u1 is considered as the information proprietor of the cluster. The information proprietor produces private key ski and public-key pki for all the blocks, as shown in Function Key Generation. The information proprietor executes the deduplication test by transmitting the hash value of the document Hash F1 to the server (see Algorithm 1, Phase 1). If there is an identical document, the cloud user executes proof of proprietorship convention with the distributed server. If it is passed, the client is certified to retrieve this cached document without uploading the document. Otherwise, the CSP divides the file F1 into blocks, creates a tag for each block generated dynamically using Pairing Based Cryptography, where the tags are represented in the form of b(x, y) where b blocks and (x, y) is a vector. The CSP verifies with the respective clients for the deduplication of the piece. If it is the modified chunk, then CSP allows uploading; otherwise, CSP executes the proof ownership convention; if it is a duplicate, then CSP allows the respective customers to retrieve the chunk as shown in Algorithm 1, Phase 2. A summary of the Notations used in the Algorithm is, as shown in Table 1.

Algorithm 1: Deduplication and dynamic Auditing

Input: F1 = (m_1, m_2,.....m_n) ge, mi Z_p, id_i where k [1,n]

Output: σi

(1) For every outsourcing document by the user, the following tasks are implemented:

(2) CSP examines for the deduplication of the document. If it is a current document, then it moves to step 4. If the document exists, then PoW convention is performed between CSP and the user.

(3) After the validation that there is no duplicate copy of the document, then divides the document into chunks F1 = (m_1, m_2,.....m_n) and outsources to the CSP.

(4) CSP produces id and signature for every block that is created actively utilizing Pairing Based Cryptography.
(5) for each bk with idk
(6) Estimate σi = (H(idi), $g^{m}i$) x
(7) end for
(8) the owner then outsources blocks to CSP and sends data information to TPA, TPA stores this info in DHT.
(9) CSP validates for the deduplication of the block. If it is an update block, then it moves to step 10. If the block is present, then PoW convention is executed between CSP and the prevailing user.
(10) If the block does not exist in the cloud, then the prevailing user uploads the modified block to the cloud.

Table 1: Summary of the Notations used in the Algorithm

Notation	Description
G1, G2	Groups of order p
g, x	Generator polynomial of G1
H	Hash function with H:(0,1)* → G1
PK	Public key
Sk	Secret key
N	number of blocks in document
g, x	Generator polynomial of G1

7. Security Analysis

We will analyze the certainty of the suggested strategy by evaluating the effectiveness of attack prevention polices in this section.

Theorem: For some opponent, it is computationally unreasonable to found an HVA under BLS signature strategy, if the computational Diffe-Hellman (CDH) supposition in bilinear groups holds Proof. This proposition comes from Wang's work, where it demonstrates the HVA scheme is empirically unforgettable in that BLS short signature procedure is secure with the presumptions that the CDH is a frozen problem in bilinear groups. Hence, we exclude the featured proof here. 2) Secure Deduplication: Let us assume that an adversary tries to upload his hunks of the record to the server. He sends these hunks as challenges to the CSP. After receiving these hunks, CSP runs the proof of ownership protocol and identifies that the .challenger is an attacker and informs the information proprietor. Therefore the CSP safely performs deduplication and efficiently protects the shared data from the opponents.

8. Performance Analysis

In this section, we present an experimental analysis of our scheme. We exploit Pairing Based Cryptography (PBC) Library to perform cryptographic operations in our convention. We have used Intel(R) Core(TM) i3-3217U, CPU @1.80GHz, 2GB RAM. To accomplish 80-bit security, the prime order p of the bilinear groups G and GT are respectively chosen as 160 bits in length.

Fig 2: Graph of Data Deduplication

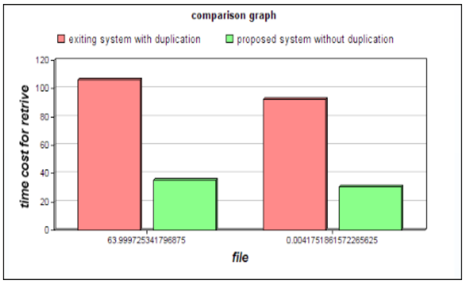

Fig. 2: Shows the experimental comparison results of data with deduplication. Without deduplication stored in the cloud, the x-axis represents the file size or file length; the y-axis represents time and cost.

9. Conclusions and Future Work

With simulation to present a public auditing scheme for secure cloud storage using a dynamic hash table used to the target of performing information probity. We were introducing Deduplication and Data auditing in the Cloud system. To list the data property information for auditing dynamically. DHT, our venture, can also reach preferable performance than other schemes in the updating phase. Additionally, our scheme further exploits the aggregate BLS signature approach from bilinear maps to enact multiple auditing jobs simultaneously, of which the principle is to compound all the signatures on varying data blocks into an isolated one and ratify it for only one time to truncate the communication cost in the verification process.

References

Ateniese, G., et al. "Provable Data Possession at Untrusted Stores." *Proceedings of the 14th ACM Conference on Computer and Communications Security*, 2007, pp. 598-609.

Dewan, H and R.C. Hansdah. "A Survey of Cloud Storage Facilities." *IEEE World Congress on Services*, 2011, pp. 224-231.

Erway, C.C., et al. "Dynamic Provable Data Possession," *ACM Transactions on Information and System Security*, vol. 17, no. 4, 2015.

Geeta, C.M. et al. "Data Auditing and Security in Cloud Computing: Issues, Challenges and Future Directions." *International Journal of Computer (IJC)*, vol. 28, no. 1, 2018, pp. 8-57.

Juels, Ari, and Burton S. Kaliski Jr. "PORs: Proofs of Retrievability for Large Files." 2013.

Liu, Jian, et al. "Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage." *IEEE Transactions on Information Forensics and Security,* vol. 7, no. 7, 2015, pp. 1513-1528.

Ren, Kui, et al. "Security Challenges for the Public Cloud." *IEEE Internet Computing*, vol. 16, no. 1, 2012, pp. 69-73.

Ryoo, Jungwoo, et al. "Cloud Security Auditing: Challenges and Emerging Approaches." *IEEE Security Privacy*, vol. 12, no. 6, 2014, pp. 68-74.

Sookhak, M. et al. "Towards Dynamic Remote Data Auditing in Computational Clouds." *The Scientific World Journal*, 2014.

Tian, Hui, et al. "Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage." *IEEE Transactions on Services Computing*, vol. 10, no. 5, 2017, pp. 701-714.

Wang, Cong, et al. "Toward Secure and Dependable Storage Services in Cloud Computing." *IEEE Transactions on Services Computing*, vol. 5, no. 2, 2012, pp. 220-232.

Yang, G., et al. "Enabling Public Auditing for Shared Data in Cloud Storage Supporting Identity Privacy and Traceability," *Journal of Systems and Software*, vol. 113, 2016, pp. 130-139.

Youn, Taek-Young, et al. "Efficient Client-Side Deduplication of Encrypted Data With Public Auditing in Cloud Storage." *IEEE Access,* vol. 6, 2018.

Zhu, Yan. "Dynamic Audit Services for Outsourced Storage in Clouds." *IEEE Transactions on Services Computing*, vol. 6, no. 2, 2013, pp. 227-238.

Author Details

K. Arun Kumar, *Senior Faculty, iNurture Education Solutions Pvt. Ltd., Bengaluru, Karnataka, India,*
Email ID*: arunkumar.k@inurture.co.in*

R. Mohammed Harun Babu, *Faculty, iNurture Education Solutions Pvt. Ltd., Bengaluru, Karnataka, India*

S. Kalaivanan, *Senior Faculty, iNurture Education Solutions Pvt. Ltd., Bengaluru, Karnataka, India*

V Kanimozhi, *Faculty, iNurture Education Solutions Pvt. Ltd., Bengaluru, Karnataka, India*