A Well Structured Remote Data Possession Verification Protocol in Cloud Storage

[1]Sandesh R, [2]Shashi Rekha H
[1]Fourth Sem M.Tech, [2]Assistant Professor
[1]Department of studies in CSE, VTU PG Center, Mysuru,
[2]Department of studies in CSE, VTU PG Center, Mysuru

Abstract - As a major cloud computing application, cloud storage offers scalable, flexible and high-quality data storage and processing services. More and more data owners choose to outsource data files to the cloud. Because cloud storage servers are not completely reliable, data owners need reliable tools to verify that their files are outsourced to remote servers in the cloud. To solve this crucial problem, a Remote Data Possession Verification (RDPV) protocol was introduced. But many existing systems have vulnerabilities in the efficiency or dynamics of the data. In this article, we provide a new efficient RDPV protocol based on the homomorphic hash function. The new system is surely safe against counterfeiting attacks, replacing attack and copy attacks based on a typical security model. To support data dynamics, an operations log table (ORT) is introduced to track the operations of the file block. We also provide a new and optimized implementation for ORT that makes the ORT access costs almost constant. In addition, we perform a comprehensive performance analysis that shows that our system has advantages in terms of computing and communication costs. The implementation of the prototype and the experiments show that the scheme is feasible for real applications.

IndexTerms - Cloud computing, RDPV: Remote Data Possession Verification, ORT: Operation Record Table, Data Dynamics, Homomorphic Hash Functions.

I. INTRODUCTION

Cloud computing appears as a new paradigm of calculation after grid computing. By managing a large number of distributed computing resources over the Internet, it has enormous computing capacity and virtualized storage space. As a result, cloud computing is widely accepted and used in many real world applications. As a leading cloud service, the cloud service provider provides users with reliable, scalable and low-cost outsourced storage services. It provides users with a more flexible means called pay-as-you-go model for computing and on-demand storage resources. Based on this model, users can engage the necessary IT infrastructure for initial user investments that will be significantly reduced. In addition, it is convenient for them to adjust the capacity of the leased resource by changing the scale of their applications.

The cloud service provider tries to provide a promising service for data storage, saving users the investment and resource costs. However, cloud storage also involves various security issues for outsourced data. Although some security issues have been resolved, important issues related to data falsification and data loss are still present in cloud storage. On the one hand, the crash disk error or the cloud storage server (CSS) hardware failure can cause unexpected corruption of outsourced files. On the other hand, CSS is not completely reliable from the point of view of the data owner; It can actively delete or modify files for huge economic benefits. At the same time, CSS can hide the data owner's incorrect behavior and data loss incidents in order to maintain a good reputation. Therefore, it is crucial that the data owner uses an effective means of verifying the integrity of data outsourcing.

Remote Data Possession Verification Protocol (RDPV) is an effective technique for ensuring the integrity of data files stored on CSS. RDPV provides a data owner method to effectively check whether the cloud service provider is faithfully storing the original files without recovering them. In RDPV, the data owner can challenge the CSS on the integrity of the destination file. CSS can generate evidence to show that it maintains complete and incorrect data. The basic requirement is that the data owner can verify the integrity of the files without accessing the complete original file. Another desired requirement is that dynamic data operations must be supported by the protocol. In general, the data owner can add, insert, delete, or modify file blocks as needed. In addition, the complexity of the calculation and the communication overhead of the protocol must be taken into account for real applications.

1.1 Motivation

It is essential that data owners verify the integrity of the data stored on CSS before using it. For example, a large international trading company stores all import and export files on CSS. According to these files, the company can obtain key

information such as logistic quantity, trade volume, etc. If a registration file is discarded or falsified, the business will incur a significant loss that could affect its business and development. To avoid such circumstances, it is mandatory to check the integrity of externalized data files. In addition, since these files could refer to a trade secret, any exposure of information is unacceptable. Auto competitor of the company can check the integrity of the file, frequently checking the files, you can get useful information like when the file changes, the rate of growth of the file, etc., with which they can guess the development of the file company. So to avoid this situation, consider the type of private audit in our system, the data owner is the sole auditor In fact, the current direction of RDPV research focuses on public audit, in which everyone can perform the verify the integrity of the files with the public key of the system. Although CPD with public audit seems better than that with private audit, but is not suitable for the scenario mentioned above.

We present a well-structured RDPV protocol schema with data dynamics. The basic schema uses the homomorphic hash technique of the function, in which the hash value of the sum by two blocks is equal to the product for two hash values of the corresponding blocks. We introduce a linear table called ORT to record data operations to support data dynamics such as block modification, block insertion, and block cancellation. To improve the efficiency of access to ORT, we use the double-link list and matrix to present an optimized ORT list implementation that reduces the cost to an almost constant level. We show that the presented system is safe against counterfeit attacks, replay attacks and attack replacement based on a typical security model. Finally, we implement our system and make a thorough comparison with previous schemes. Experience shows that the new system has better performance and is useful for real applications.

1.2 Organizations

The remnant of the paper is organized as follows. Section II describes the Background or the related work of our scheme. Section III describes the construction of our new scheme. Section IV describes the Expected outcomes of the RDPV Protocol and the Section V describes the Conclusions and Future enhancement.

II. BACKGROUND

2.1 Purpose

The purpose of introducing a new RDPV protocol to ensure whether the data owners file is safe in cloud storage. Since the cloud service provider is not entirely trustworthy, the data owner may feel that his file is not safe. By the RDPV scheme, the data owner can request the cloud storage server if its files are not attacked by any external attacker. In turn, the cloud service provider can see the challenges that are made by data owners and must answer for the same by sending evidence to the data owner. Then the data owner can verify the evidence if the evidence is acceptable or not. Once the proof is verified by this system, the data owner can update, modify, or delete the contents of that particular file. This framework is therefore essential for the owner of the data to check the integrity of the file stored on the cloud server. Thus, the RDPV protocol becomes more efficient in the possession of data and to verify the integrity of the files in the cloud computing.

2.2 Existing Solutions

An increasing number of data owners are choosing to outsource data files to the cloud. Since cloud storage servers are not completely reliable, data owners need reliable ways to verify ownership of their outsourced files on remote cloud servers. To solve this crucial problem, some remote data possession verification protocols have been introduced. But many existing systems have vulnerabilities in terms of efficiency or data dynamics. Some of the main disadvantages are,

- It is essential that the data owner uses an effective means of verifying the integrity of outsourced data.
- Does not provide efficiency in verifying remote data integrity.
- More expensive.

The existing system offers less flexibility.

III. PROPOSED SOLUTIONS

In the proposed system, an efficient and flexible distributed schema with explicit dynamic data support, including updating, deleting and adding blocks. Adapting the proposed protocol with the distributed verification of erased coded data, the scheme performs public verifiability and data dynamics against third-party verifiers that shows the detection of data corruption during verification of storage accuracy on the distributed servers. The protocol design will provide the following security and performance assurance: Public auditability, Storage correction, Confidentiality preserving, Batch audit, Light. The model we propose aims to protect cloud data from unreliable service providers. This model involves data owners, cloud service providers, and data users. Data owners store data in the cloud and send each share of data entries to service providers.

We provide a new and efficient RDPV protocol. The new system is surely secure against falsification attacks, replaces attacks, and replays attacks based on a typical security model. We also give a new optimized implementation for the ORT which makes the cost of access to this process almost constant. We do the complete performance analysis that shows that our system has advantages in computing and communication costs. The implementation of the prototype and the experiments show that the scheme is feasible for real applications. Some of the main advantages of the proposed systems are,

- The results of experiments show that the new system has better performance and is practical for real applications.
- We show the advanced RDPV scheme supporting fully dynamic ORT-based block operations.
- Minimum calculation costs.
- The data owner can perform dynamic file operations.

3.1 Schema of RDPV Protocol

In this article, we study the cloud storage system that includes two participants: the CSS and the owner of the data. CSS has powerful storage and compute resources, accepts requests from data owners to store data files outsourced, and provides access service. The data owner benefits from the CSS service and puts a large amount of CSS files without local backup. Since it is not assumed that CSS is reliable and sometimes does not work, for example by modifying or deleting partial data files, the data owner can effectively verify the integrity of outsourced data in cloud computing.

The Remote Data Possession Verification Protocol (RDPV) is an effective technique for ensuring the integrity of data files stored on CSS. RDPV provides a data owner method to effectively check whether the cloud service provider is faithfully storing the original files without recovering them. In RDPV, the data owner can challenge the CSS on the integrity of the destination file. CSS can generate evidence to show that it maintains complete and incorrect data. The basic requirement is that the data owner can verify the integrity of the files without accessing the complete original file. Since it is not assumed that CSS is reliable and sometimes does not work, such as modifying or deleting partial data files, the data owner can effectively verify the integrity of the outsourced data. An RDPV schema includes the following process. They are KeyGen, TagGen, Challenge, ProofGen, Check, PrepareUpdate, ExecUpdate.

- TagGen: This algorithm is executed by the owner of the data to produce tags of the file. It enters the homomorphic key K, the private key sk and the file F, and outputs the set of variables T which is a sequential collection for the label of each block.
- Challenge: The data owner runs the algorithm to generate the challenge information. He takes into account the disputed blocks and grabs the challenge chal.
- ProofGen: The CSS executes this algorithm to generate the proof of integrity P. It introduces the file F, the set of labels T and the chal challenge and leaves the proof P.
- Verify: The data owner executes the algorithm to check the integrity of the file using the proof P returned from CSS. It takes homomorphism key K private key sk , challenge chal and proof P as inputs, and outputs 1 if P is correct, otherwise it outputs 0.
- PrepareUpdate: The data owner runs this algorithm to prepare dynamic data operations on data blocks. It takes as input the new file block Fi ', the block position i and the type of update UT and provides the update request information. The UT parameter has three optional elements: insert, modify, and delete.
- ExecUpdate: The CSS runs this algorithm to execute the update operation. It inputs request and outputs execution result. If the update operation is finished successfully, it returns Success, otherwise returns fail.

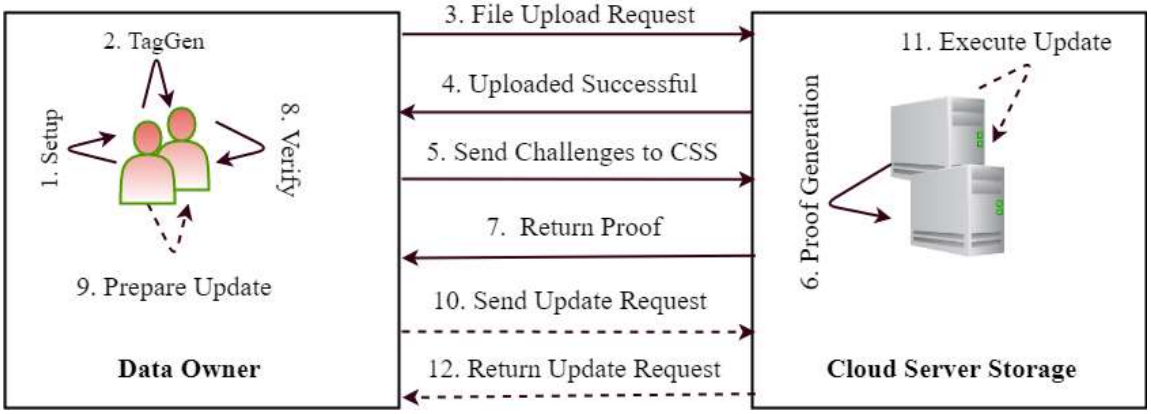

Figure 1: Schema and designing of new RDPV protocol

The complete workflow of our RDPV protocol is illustrated in Fig. 1, in which the solid lines and the dashed lines represent the processes of verification of the integrity of the data and the dynamic operations on the file

3.2 *System Framework*

In cloud computing, cloud storage offers scalable, flexible, and high-quality services for data storage and processing. An increasing number of data owners are choosing to outsource data files to the cloud. Because cloud storage servers are not completely reliable, data owners need reliable tools to verify that their files are outsourced to remote cloud servers. We provide a new and efficient RDPV protocol based on the homomorphic hashing function. The new system is surely secure against counterfeit attacks, replacing attacks of attack and reproduction based on a typical security model. To support data dynamics, an Operations Log Table (ORT) is introduced to track file block operations. The cloud storage system consisting of two participants: CSS and data owner. CSS has powerful storage and compute resources, accepts requests from data owners to store data files outsourced, and provides access service. The data owner benefits from the CSS service and puts a large amount of CSS files without local backup.

3.3 Homomorphic Hash Function and ORT:

The RDPV scheme uses the homomorphic hash function technique, in which the hash value of the sum for two blocks is equal to the product for two hash values of the corresponding blocks. We introduce a linear table called ORT to record data operations to support data dynamics such as block editing, block insertion, and block deletion. To improve the efficiency of access to ORT, the table is reserved by the data owner and is used to record any dynamic behavior on the file blocks. We use double-link lists and tables to present an optimized implementation of ORT that reduces the cost to an almost constant level. We show that the presented system is safe against counterfeit attacks, the attack of reproduction and the replacement of attacks based on a typical security model.

3.4 Security Requirement:

CSS is not completely reliable because it could lead to malicious behavior on outsourced data and hide data corruption occurrences from the data owner to maintain a good reputation. According to the dishonest, CSS can launch three types of attacks on the RDPV, namely a flogging attack, a replay attack and a substitution attack.

- Forge Attack: CSS forges a valid tag for the disputed block in order to fool the data owner.
- Repeated Attack: CSS chooses a valid test for possession of previous tests or other information, without access to the disputed block and tag.
- Replace Attack: CSS uses the other pair of valid blocks and tags as evidence of the disputed pair, which may have been corrupted or discarded.

A secure RDPV protocol should be able to withstand all previous attacks, which ensures that anyone who can build a valid test by passing the check should have the entire file.

IV. **EXPECTED OUTCOMES**

In the proposed system, an efficient and flexible distributed system with dynamic support for explicit data, including updating, deleting and adding blocks. Our schema uses a homomorphic hash function to check the integrity of the files stored on the remote server and reduces the storage costs and computational costs of the data owner. The model we propose is designed to protect cloud data from unreliable service providers. And the presented system is proven safe in the existing security model.

V. CONCLUSIONS

In this article, we examine the problem of controlling the integrity of remote-processed data files to the remote server, and we propose an efficient and secure RDPV protocol with dynamic data. Our schema uses a homomorphic hash function to check the integrity of the files stored on the remote server and reduces the storage costs and computational costs of the data owner. We are designing a new lightweight RDPV protocol to support dynamic block operations that result in minimal computational costs by reducing the number of node moves. Thanks to our new data structure, the data owner can perform file insertion, modification or deletion operations with great efficiency. The presented schema is safe in the existing security model. The results of the experiment will indicate that our system is convenient in cloud storage.

ACKNOWLEDGEMENT

The author would like to thank Dr. K Thippeswamy, Professor and chairman, Dept. of studies in computer science and engineering, VTU Regional office, Mysuru and anonymous reviewers encouragement and constructive piece of advice that of prompted us for new round of rethinking of our research, additional experiments and clearer presentation of technical content.

REFERENCES

[1]. R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic,"Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Gener. Comp. Sy., vol. 25, no. 6, pp. 599 – 616, 2009.

[2]. H. Qian, J. Li, Y. Zhang and J. Han, "Privacy preserving personal health record using multi-authority attribute-based encryption with revocation," Int. J. Inf. Secur., vol. 14, no. 6, pp. 487-497, 2015.

[3]. J. Li, W. Yao, Y. Zhang, H. Qian and J. Han, "Flexible and fine-grained attribute-based data storage in cloud computing," IEEE Trans. Service Comput., DOI: 10.1109/TSC.2016.2520932.

[4]. J. Li, X. Lin, Y. Zhang and J. Han, "KSF-OABE: outsourced attribute-based encryption with keyword search function for cloud storage," IEEE Trans. Service Comput., DOI: 10.1109/TSC.2016. 2542813.

[5]. J. Li, Y. Shi and Y. Zhang, "Searchable ciphertext-policy attribute-based encryption with revocation in cloud storage," Int. J. Commun. Syst., DOI: 10.1002/dac.2942.

[6]. J.G. Han, W. Susilo, Y. Mu and J. Yan, "Privacy-Preserving Decentralized Key-Policy Attribute-Based Encryption," IEEE Transactions on Parallel and Distributed Systems, vol. 23, no.11, pp. 2150-2162, 2012

[7]. Z. J. Fu, X. M. Sun, Q. Liu, L. Zhou, and J. G. Shu, "Achieving efficient cloud search services: multi-keyword ranked search over encryptedn cloud data supporting parallel computing," IEICE Transactions on Communications, vol. E98-B, no. 1, pp.190-200, 2015.

[8]. Z. J. Fu, K. Ren, J. G. Shu, X. M. Sun, and F. X. Huang, "Enabling personalized search over encrypted outsourced data with efficiency improvement," IEEE Transactions on Parallel and Distributed Systems, DOI: 10.1109/TPDS.2015.2506573, 2015.

[9]. Z. H. Xia, X. H. Wang, X. M. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," IEEE Transactions on Parallel and Distributed Systems, vol. 27, no. 2, pp. 340-352, 2015.

[10]. Y. J. Ren, J. Shen, J. Wang, J. Han and S. Y. Lee, "Mutual verifiable provable data auditing in public cloud storage," Journal of Internet Technology, vol. 16, no. 2, pp. 317-323, 2015.

BIOGRAPHIES

Sandesh R is presently pursuing his M.Tech degree in department of studies in CSE at Visvesvaraya Technological University, PG Center, Mysuru 570029. He has completed B.E in CSE branch at Vidyavardhaka College of Engineering, Mysuru, Karnataka in the year 2016. His M.Tech project area is on Cloud Computing. This paper is a survey paper of his M.Tech project.

Shashi Rekha H presently working as Asst. Professor in DOS in CS & E , VTU-PG Center , Mysuru since 2013. Her qualification is B.E, M.tech, (Ph.D). She has 11 years of teaching experience. She is currently pursuing research in the area of Big Data analytics. Her research interests are Image classification, Pattern recognition, Data Mining in E- healthcare. She has presented many papers in various journals and few conferences. She is a member of CSI, Research Gate Forum.