e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

ACHIEVING EFFICIENT SECURE DEDUPLICATION WITH USER-DEFINED ACCESS CONTROL IN CLOUD

S Pradeepa*1, Banupriya A*2

*1Master Of Engineering, Computer Science And Engineering, CSI College Of Engineering,
Near Ooty, Ketti, The Nilgiris, Tamil Nadu, India.

*2Assistant Professor, Department Of Information Technology, CSI College Of Engineering,
Near Ooty, Ketti, The Nilgiris, Tamil Nadu, India.

DOI : https://www.doi.org/10.56726/IRJMETS45056

ABSTRACT

Cloud storage as one of the most important services of cloud computing significantly facilitates cloud users to outsource their data to the cloud for storage and share them with authorized users. In cloud storage, secure deduplication has been widely investigated as it can eliminate redundancy over the encrypted data to reduce storage space and communication overhead. Regarding security and privacy, many existing secure deduplication schemes generally focus on achieving the following properties: data confidentiality, tag consistency, access control and resistance to brute-force attacks. However, as far as we know, none of them can achieve these four requirements at the same time. To overcome this shortcoming, in this paper, we propose an efficient secure deduplication scheme that supports user-defined access control. Specifically, by allowing only the cloud service provider to authorize data access on behalf of data owners, our scheme can maximally eliminate duplicates without violating the security and privacy of cloud users. Detailed security analysis shows that our authorized secure deduplication scheme achieves data confidentiality and tag consistency while resisting brute-force attacks. Furthermore, extensive simulations demonstrate that our scheme outperforms the existing competing schemes, in terms of computational, communication and storage overheads as well as the effectiveness of deduplication.

I. INTRODUCTION

With the great benefits of cloud computing, an increasing amount of data has been outsourced by data owners to the cloud and shared with authorized users. For example, the Cisco global cloud index shows that the data stored in cloud will nearly reach1.3 zettabytes by 2021. As a result, the management of the ever-increasing data becomes a critical challenge for cloud storage services. In fact, the study shows that about 75% of digital data are identical, and redundancy in backup and archival storage system is significantly more than 90%. In this situation, data deduplication technique has been widely developed in cloud storage because it can significantly reduce storage costs by storing only a single copy of redundant data. Indeed, data deduplication can reduce storage costs by more than 50% in standard file systems and by more than 90% for backup applications, and these savings are transformed into huge financial savings to cloud service providers and users. However, considering security and privacy concerns of outsourced data, users are likely to encrypt data with their keys before out- sourcing. As a result, data deduplication would be impeded as identical data will be encrypted into different ciphertexts. To make data deduplication feasible on encrypted data, convergent encryption and its implementations or variants have been developed in. However, convergent encryption suffers from brute-force attacks for predictable messages. To overcome this issue, some server-aided encryption schemes have been presented. they suffer from the duplicate faking attack that prevents legitimate users from obtaining correct data. Although some schemes try to resist this attack, the tag consistency is verified after downloading the ciphertext by users, which cannot exactly conclude whether the incorrect data is caused by duplicate faking attacks in data upload or is corrupted during data storage. The main reason for this attack is that the ciphertext and the corresponding tag are generated independently, which makes it impossible for the cloud service provider to check the tag consistency. Thus, a solution that computes the tag directly by hashing the ciphertext is presented, which obviously supports the tag consistency check by com- paring the hash of ciphertext to the received tag. Based on this idea, tag consistency and message authentication have been considered in, where the cloud service provider checks tag consistency and users conduct the message authentication after data download.

II. METHODOLOGY

Existing System:

Existing solutions for deduplication suffer from brute-force attack. They cannot flexibly support data access control and revocation at the same time. Most existing solutions cannot ensure reliability, security, and privacy with sound performance. Cloud storage service providers such as Dropbox, Google Drive, Mozy, and others perform deduplication to save space by only storing one copy of each file uploaded. However, if clients conventionally encrypt their data, storage savings by deduplication are totally lost. This is because the encrypted data are saved as different contents by applying different encryption keys. Existing industrial solutions fail in encrypted data deduplication. For example, deduplication system, but it cannot handle encrypted data.

Proposed Methodology:

In this paper, we propose a scheme based on data ownership challenge and Proxy Re-Encryption (PRE) to manage encrypted data storage with deduplication. We aim to solve the issue of deduplication in the situation while the data holder is uploading the data. Meanwhile, the performance of data deduplication in our scheme is not influenced by the size of data, thus applicable for cloud. Proxy-re encryption is performed to secure the data double time. The key generation for securing the data which is generated by the cloud server.

Software and Technology Description

Java Technology

Java technology is both a programming language and a platform.

The Java platform has two components:

1) The Java Virtual Machine (Java VM)

2) The Java Application Programming Interface (Java API)

Java API:

The Java API is a large collection of ready-made software components that provide many useful capabilities, such as graphical user interface (GUI) widgets. The Java API is grouped into libraries of related classes and interfaces; these libraries are known as packages.

Every full implementation of the Java platform gives you the following features:

The essentials, Applets, Networking, Internationalization, Security, Software components, Object serialization, Java Database Connectivity (JDBC^{TM}).

URL:

The URL provides a reasonably intelligible form to uniquely identify or address information on the Internet. URLs are ubiquitous; every browser uses them to identify information on the Web. Within Java's network class library, the URL class provides a simple, concise API to access information across the Internet using URLs.

ODBC:

Microsoft Open Database Connectivity (ODBC) is a standard programming interface for application developers and database systems providers.

My-SQL:

MySQL is a simple, yet powerful Open-Source Software relational database management system that uses SQL. MySQL is a true multi-user, multithreaded SQL database server.

III. MODELING AND ANALYSIS

In propose an effective approach to verify data ownership and check duplicate storage with secure challenge and cloud support and we integrate cloud data deduplication with data access control in a simple way, thus reconciling data deduplication and encryption. The proposed encryption has protected the security of users' sensitive data effectively.

Data Owner and User Authentication

User Authentication & Data Owner authentication Verified in this module. If the User is a new user, they should submit their details are given to the data owner. Users Information is valid Then their registration will accept and providing data access Permission when user enter their correct id.

Data Selection and encryption

The data owner selects the Any files from Different source and to be encrypted. The data owner should have the encryption of their own key for authority. Then the file is encrypted forward to CSP

Data Upload in CSP and Proxy Re-Encryption

When the encrypted file Forward to CSP contains cloud Server. Then Its Forward The encrypted file Data to cloud Server In server will Re-Encrypt the specified file using symmetric Encryption algorithm. And store a encrypted data in server.

User Data Request and Download

In This Module when the user entered their access mode will select the file send request to the Data Owner. Data Owner will redirect The Request to the CSP. CSP has Forward the user request to Proxy-Server and send the user requested files encrypted data to user with a decrypted Key data from Proxy-Server.

Get Key & Decryption

In this Module User download the encrypted data with encrypt key from Proxy Using this Key The decrypt the Data and get The encrypted file format and again send the download file encrypted decrypt key requested to the data Owner and get the downloaded file encryption key for Decryption using This key user will get the plain text data.

System Design

The goal of designing input data is to make data entry as easy, logical and free from errors. In the system design phase input data are collected and organized into groups of similar data. The input forms in this project are login, customer details, purchase details, item details and sales details.

Output Design

Output design is a very important phase because the output will be in an interactive manner. The reports generated in this project are sales report, product report, purchase report.

Database Design

In the project, login table is designed to be unique in accepting the username and the length of the username and password should be greater than zero. The different users view the data in different formats according to the privileges given.

Testing

Unit Testing

This is the first level of testing. The different modules are tested against the specifications produced during the integration.

Integration Testing

The goal of integration testing to check whether the modules can be integrated properly emphasizing on the interfaces between modules.

Validation Testing

The objective of the validation test is to tell the user about the validity and reliability of the system. It verifies whether the system operates as specified and the integrity of important data is maintained. User motivation is very important for the successful performance of the system.

System Testing

The objective of system testing is to check if the software meets its requirements. System testing is done to uncover errors that were not found in earlier tests.

IV. RESULTS AND DISCUSSION

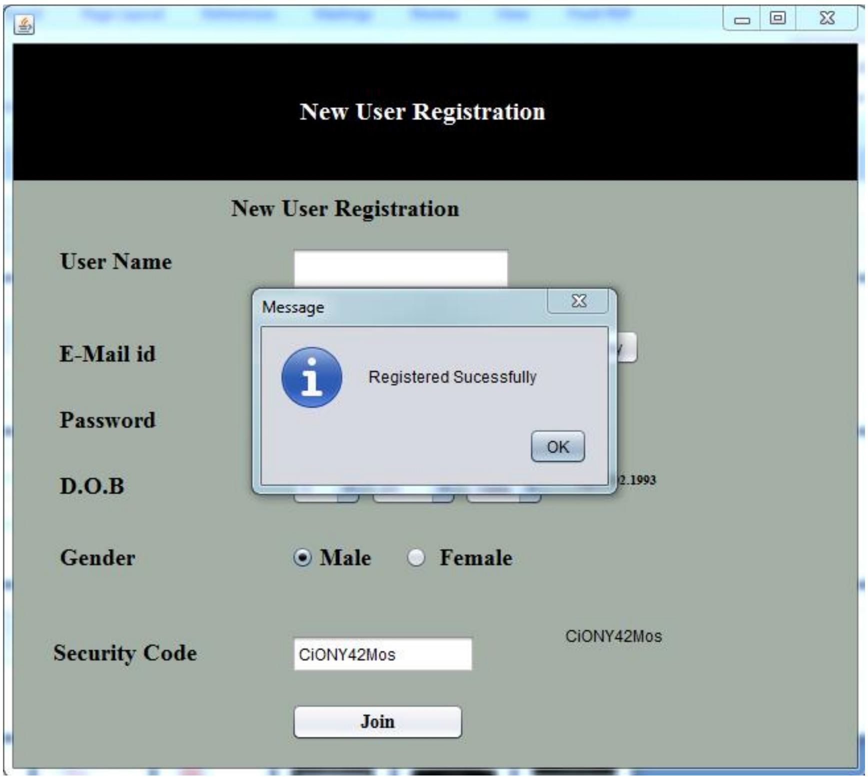
Home Page:

The homepage serves as the entry point where users can click to access the registration page.

Framework Using Proxy Re-Encryption and Secure Sharing in Cloud

Registration:

This page functions as the login portal where users can input their client details for the purpose of registration.



New User Registration

New User Registration

User Name:

E-Mail id:

Password:

D.O.B: ...2.1993

Gender: ● Male ○ Female

Security Code: CiONY42Mos

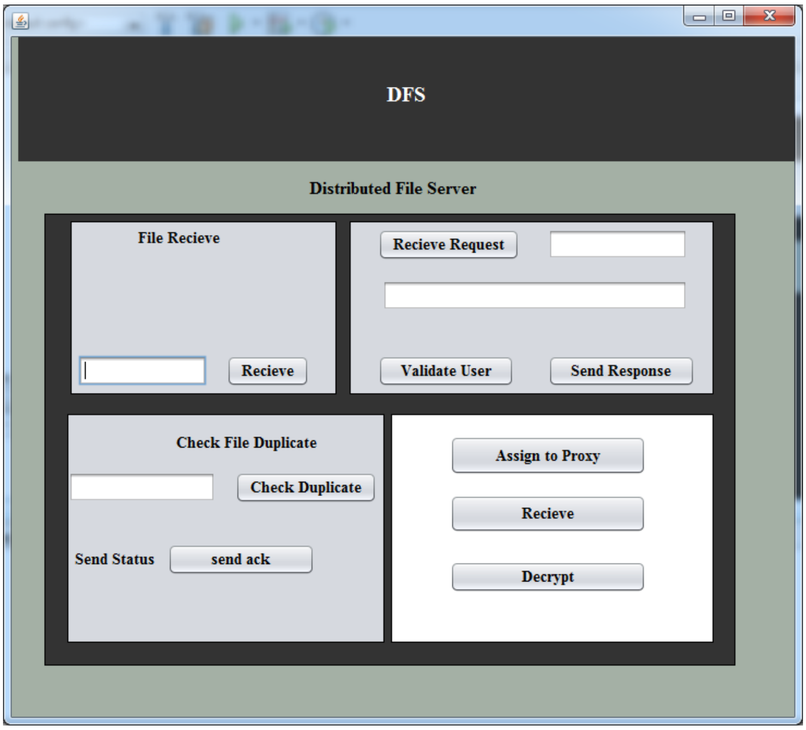
CiONY42Mos

Message

Registered Sucessfully

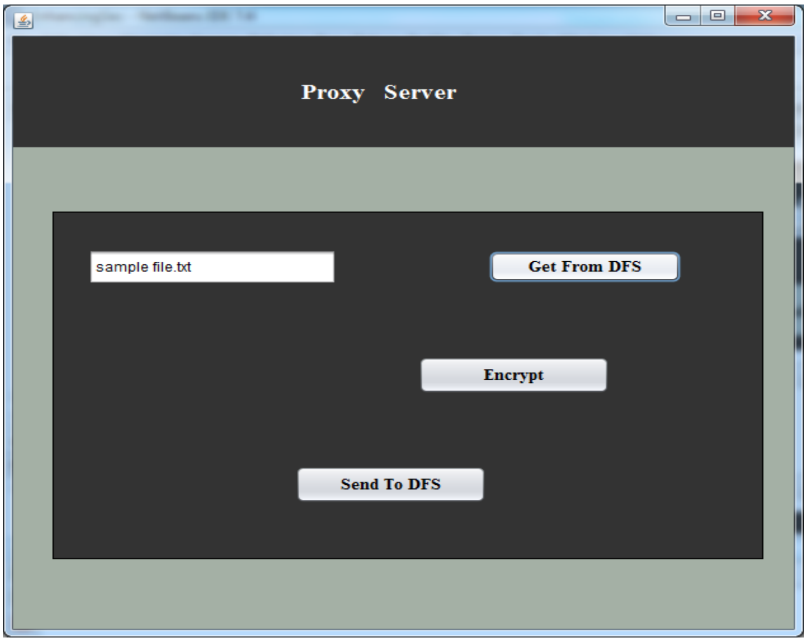
DFS:

The uploaded file will be received in the Distributed File System (DFS), where it will undergo duplication checks facilitated by the utilization of a hash key.



Proxy Server Page:

If the file is determined to be a duplicate within the DFS, the system will automatically identify it. However, in the case of a new file, the data will be encrypted and subsequently assigned to the proxy server, where it will be stored.



sample file.txt

V. CONCLUSION

Managing encrypted data with deduplication is important and significant in practice for achieving a successful cloud storage service, especially for big data storage. In this paper, we proposed a practical scheme to manage the encrypted data in cloud with deduplication based on ownership challenge and PRE. Our scheme can flexibly support data update and sharing with deduplication even when the data holders are offline. Encrypted data can be securely accessed because only authorized data holders can obtain the symmetric keys used for data decryption. Extensive performance analysis and test showed that our scheme is secure and efficient under the described security model and very suitable for big data deduplication. The results of our computer simulations further showed the practicability of our scheme.

VI. REFERENCES

[1] "Cisco global cloud index: Forecast and methodology, 2016-2021 white paper," https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html.

[2] J. Gantz and D. Reinsel, "The digital universe decade-are you ready," https: //hk.emc.com/collateral/analyst-reports/idc-digital-universe-are-you-ready. pdf.

[3] H. Biggar, "Experiencing data de-duplication: Improving efficiency and reducing capacity requirements," The Enterprise Strategy Group., 2007.

[4] D. T. Meyer and W. J. Bolosky, "A study of practical deduplication," TOS, vol. 7, no. 4, pp. 14:1–14:20, 2012.

[5] D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Side channels in cloud services: Deduplication in cloud storage," IEEE Security & Privacy, vol. 8, no. 6, pp. 40–47, 2010.

[6] J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer, "Re- claiming space from duplicate files in a serverless distributed file system," in ICDCS, 2002, pp. 617–624.

[7] J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl, "A secure data deduplication scheme for cloud storage," in Financial Cryptography and Data Security - 18th International Conference, FC 2014, Christ Church, Barbados, March 3-7, 2014, Revised Selected Papers, 2014, pp. 99–118.

[8] P. Puzio, R. Molva, M. O¨ nen, and S. Loureiro, "Cloudedup: Secure dedu- plication with encrypted data for cloud storage," in IEEE 5th International Conference on Cloud Computing Technology and Science, CloudCom 2013, Bristol, United Kingdom, December 2-5, 2013, Volume 1, 2013, pp. 363–370.

[9] S. Keelveedhi, M. Bellare, and T. Ristenpart, "Dupless: Server-aided encryp- tion for deduplicated storage," in Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013, 2013, pp. 179–194.

[10] M. Miao, J. Wang, H. Li, and X. Chen, "Secure multi-server-aided data dedu- plication in cloud computing," Pervasive and Mobile Computing, vol. 24, pp. 129–137, 2015.

[11] Y. Shin, D. Koo, J. Yun, and J. Hur, "Decentralized server-aided encryption for secure deduplication in cloud storage," IEEE Trans. Services Computing, vol. PP, no. 99, pp. 1–1, 2017.

[12] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-locked encryption and secure deduplication," in Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings, 2013, pp. 296–312.

[13] X. Liu, W. Sun, W. Lou, Q. Pei, and Y. Zhang, "One-tag checker: Message- locked integrity auditing on encrypted cloud deduplication storage," in 2017 IEEE Conference on Computer Communications, INFOCOM 2017, Atlanta, GA, USA, May 1-4, 2017, 2017, pp. 1–9.

[14] R. Chen, Y. Mu, G. Yang, and F. Guo, "BL-MLE: block-level message- locked encryption for secure large file deduplication," IEEE Trans. Information Forensics and Security, vol. 10, no. 12, pp. 2643–2652, 2015.

[15] S. Jiang, T. Jiang, and L. Wang, "Secure and efficient cloud data deduplication with ownership management," IEEE Trans. Services Computing, vol. PP, no. 99, pp. 1–1, 2017.