ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 10, Issue 3, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 6.551

+91 9940572462 +91 9940572462 ijarasem@gmail.com www.ijarasem.com

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

Deduplication of Secure Cloud Data with Effective Re-Encryption for A Bank Loan

V.P.Gladis Pushpa Rathi [1], Nuthakki.Abhiram Chowdary [2], Puvvula.Gokul [3], Akula.Sai Siva [4], Moturu.Hitesh [5]

[1]Professor, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[2] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[3] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[4] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[5] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

ABSTRACT*:* The data deduplication technique has been widely adopted by commercial cloud storage providers, which is both important and necessary in coping with the explosive growth of data. To further protect the security of users' sensitive data in the outsourced storage mode, many secure data deduplication schemes have been designed and applied in various scenarios. Among these schemes, secure and efficient re-encryption for encrypted data deduplication attracted the attention of many scholars, and many solutions have been designed to support dynamic ownership management. In this paper, we focus on the re-encryption deduplication storage system and show that the recently designed lightweight rekeying-aware encrypted deduplication scheme (REED) is vulnerable to an attack that we call it stub-reserved attack. Furthermore, we propose a secure data deduplication scheme with efficient re-encryption based on the convergent all-or-nothing transform (CAONT) and randomly sampled bits from the Bloom filter. Due to the intrinsic property of the one-way hash function, our scheme can resist the stub-reserved attack and guarantee the data privacy of data owners' sensitive data. Moreover, instead of re-encrypting the entire package, data owners are only required to re-encrypt a small part of it through the CAONT, thereby effectively reducing the computation overhead of the system. Finally, security analysis and experimental results show that our scheme is secure and efficient in re-encryption.

KEYWORDS: Deduplication, re-encryption, cloud computing.

I. INTRODUCTION

With the rapid development of cloud storage, more and more individuals and enterprises tend to outsource their sensitive data to remote cloud service providers in a pay-per-use manner.According to the study from Internet Data Center (IDC) [6] sponsored by Dell EMC, the digital universe is doubling in size every two years and the volume of the universe data is expected to reach 44 zettabytes (ZB) or 44 trillion gigabytes (GB) in 2020 (more than 5,200 gigabytes for each man, woman, and child). However, the growth of data puts heavy pressure on cloud service providers. To cope with it, a straightforward method is to require cloud service providers continuously increase the capacity of storage space, to meet users' requirements for high-quality storage services. However, cloud service providers may store plentiful and repetitive data [1] (such as movies, music, and genome data), which inevitably incurs a mass of redundant storage and backup space, consequently costing a vast amount of computing and management overhead during its whole life cycle. To solve this problem, Bolosky et al. [7] first proposed the technique of data deduplication, which decreases the redundant storage space and bandwidth by eliminating duplicate copies and only storing one copy of them.

II. LITERATURE REVIEW

Cross-client deduplication [11] is an arising procedure to facilitate the weight of distributed storage in an enormous information period by putting away just a single duplicate of copy information. Proficiency mirrors plausibility and is a fundamental thought while applying the deduplication plot. In this paper, we propose an effective cross-client deduplication plot for scrambled information, which starts to lead the pack to consider proficiency during deduplication and makes it a reality to accomplish secure deduplication in a lightweight manner. To guarantee proficiency, we use the proposed re-encryption method along with a non-intuitive united key age plan to dispose of the expense of clients and backing group checks of recuperated information. The reenactment results show that, with a similar degree of safety ensured, the proposed conspire is more effective contrasting and cutting edge.

Secret sharing is an important means to achieve confidentiality and data privacy. Secret sharing deals with splitting a piece of secret information with various players. The goal of secret sharing is the security of secrets, privacy, and hiding information. There are numerous techniques available for secret sharing e.g. polynomial, Chinese remainder theorem, vector space, and matrix projection. Techniques have characteristics like threshold, proactive, and verifiable. Proactive secret-sharing schemes allow the user to change shares in case of doubt or theft. In this work, we propose a proactive secret-sharing scheme based on homomorphic [14] techniques.

Utilizing cloud information clients might move information from their PC frameworks to cloud servers. Thus, the client won't have any weight of upkeep and he gets top-notch information capacity administrations. Numerous security issues are worried to distributed storage. Cloud specialist co-ops or capacity servers are not reliable. The outsider reviewer checks the client's information for rightness and gives the precision of the information that is put away in the cloud server. The correspondence and calculation above are diminished. The deduplication [12] procedure is utilized to check whether the document that the client needs to store in distributed storage is as of now exists on a cloud server or not.

III. METHODOLOGY

The systems architect establishes the basic structure of the system, we propose a Hash code Solomon algorithm and we can put a small part of data in the local machine and fog server to protect privacy. Moreover, based on computational intelligence, this algorithm can compute the distribution proportion stored in the cloud, fog, and local machines, respectively. Through the theoretical safety analysis and experimental evaluation, the feasibility of our scheme has been validated, which is a powerful supplement to the existing cloud storage scheme

To achieve data deduplication, our scheme enables the different cloud users with the same file always generate the same ciphertext. The process of encryption is described as follows. A cloud user first uses the MLE algorithm to encrypt the message and get the ciphertextC1. Then, the cloud user transforms the ciphertextC1 into the t andC2 packages by using CAONT. Second, the cloud user concatenates the randomly selected 256-bit of packageC2 by using the Bloom filter-based location selection method and the last 256-bit from the packageC2 asgatheredpackage.

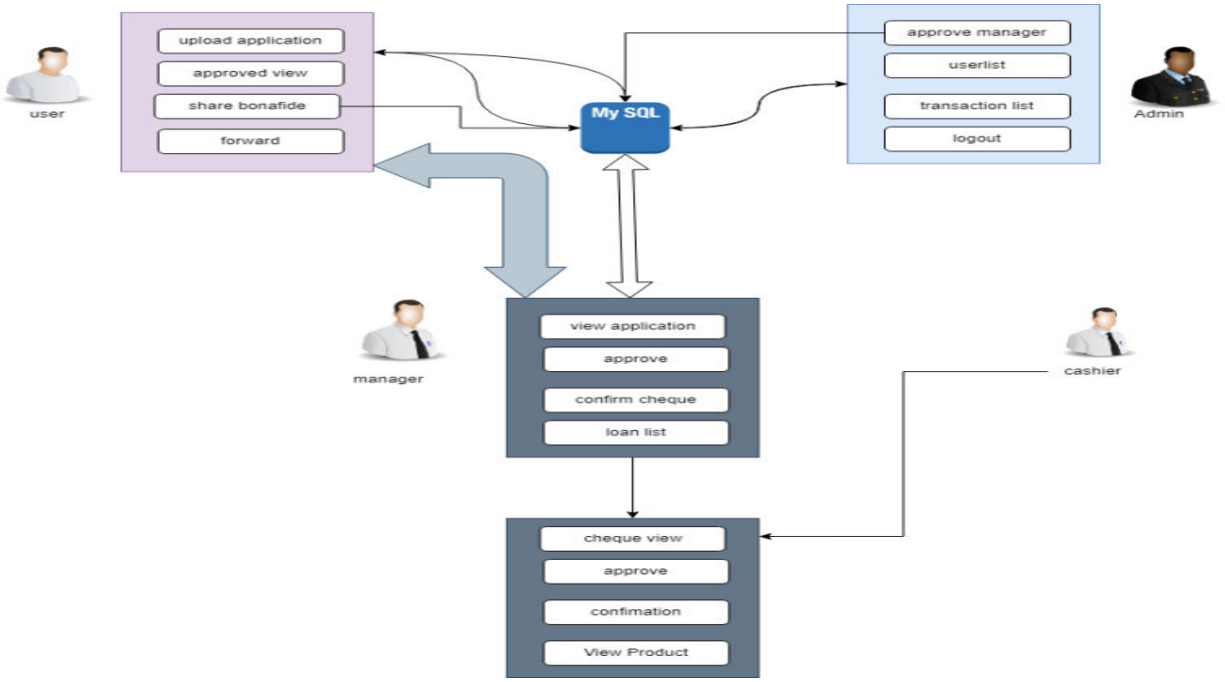


Fig1. The architecture of the proposed system

The package and a large part of the remained package C2 are concatenated as their main package. Finally, the cloud user uses the file key 2 to re-encrypt the gathered package and generate the gathered 0 packages. After that, the cloud user uploads the gathered 0 and the remaining packages to the CSP. In this way, our scheme allows different cloud users with the same data always to generate the same remainder and gathered0 packages. Since the different cloud users outsource the same remainder and gather 0 packages to the CSP, the CSP can remove the duplicate packages and only keep one copy of them. Thus, our scheme can achieve ciphertext deduplication Location Generation Performance. In this paper, we proposed the Bloom filter-based location selection method to choose random locations. Although we choose 256-bit random locations ass1 in our scheme. To achieve more efficiency or security, a cloud user also can select 128-bit or 512-bit random locations ass1. We measure the location selection time of different bit lengths. Gathered Package Generation Performance. We measure the performance of gathered package generation. The gathered package generation time consists of location generation, bits extraction, and concatenation time. We set the chunk size to 8 KB and 16 KB. The cost computation time for gathered package generation. The results show that the overhead of gathered package generation for 16 KB chunk is less than 8 KB chunk. This is because the large size of the chunk requires less processing overhead.

IV. RESULTS AND DISCUSSION

USER REGISTER:

The register module provides a conceptual framework for entering data on that user in a way that: eases data entry & accuracy by matching the user entry to the data source (usually paper files created at the point of care), ties easily back to individual user records to connect registers to user data, and collects data elements to enable better supervision of donation programs.

USER LOGIN:

This module in our project here symbolizes a unit of work performed within a database management system (or similar system) against a database and treated coherently and reliably independently of other transactions. A transaction generally represents any change in the database user will transfer the amount to the provider.

USER APPLY LOAN APPLICATION:

In this module, the user applies the application of a loan. The application details will be responsible for your file stored in the database.

USER UPLOAD BONAFIDE:

This module is used to help the manager approve the loan application after the user views the approved or reject the application if approved after the user uploads the bona fide certificate.

USER VIEW CHEQUE AND FORWARD:

In this module, the user view the cheque for the loan and changes to money, he shares the cheque with the cashier.

USER CHECK BALANCE:

This module in our project here symbolizes a unit of work performed within a database management system (or similar system) against a database, they check the loan credit balance in his bank account. A transaction generally represents any change in the database user will transfer the amount to the provider.

THE MANAGER APPROVES THE APPLICATION:

The manager who views the user loan application can then select a file to view the application and check the background verification of the user. After approving the application. They decide to accept or reject.

MANAGER APPROVES BONAFIDE AND CHEQUE:

This module helps the manager verify the bona fide certificate of the user and on the cashier side, they approve the cheque of the user for money transfer.

ADMIN LOGIN:

This module in our project here symbolizes a unit of work performed within a database management system (or similar system) against a database and treated coherently and reliably independently of other transactions. A transaction generally represents any change in the database user will transfer the amount to the provider.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

ADMIN APPROVE MANAGER:

The admin login module provides a conceptual framework for entering data on those team leaders in a way that: eases data entry & accuracy by matching the team leader entry to the data source. They approved the manager for verification of this application.

ADMIN VIEW USER LIST: In this module, the admin maintains the data source of the file and generates a key for the manager's request. Because of the key for security purposes. The admin maintains data of particular or total user lists.

ADMIN VIEW TRANSACTION LIST:

In this module, the admin views the total list of loan transactions and personal details. admin will respond to the data file fully analyzed data in category-wise view Bank will be responsible for your file stored in the database.

ALGORITHM

The **RSA algorithm** is an asymmetric cryptography algorithm; this means that it uses a *public* key and a *private* key (i.e two different, mathematically linked keys). As their names suggest, a public key is shared publicly, while a private key is secret and must not be shared with anyone.

The RSA algorithm is named after those who invented it in 1978: Ron Rivest, Adi Shamir, and Leonard Adleman.

The following illustration highlights how asymmetric cryptography works:

How it works

The RSA algorithm ensures that the keys, in the above illustration, are as secure as possible. The following steps highlight how it works:

1. Generating the keys

1. Select two large prime numbers, x*x* and y*y*. The prime numbers need to be large so that they will be difficult for someone to figure out.
2. Calculate n = x*n*=*x* x y*y*.
3. Calculate the ***totient*** function; \phi(n) = (x-1)(y-1)$\phi(n)$=(x−1)(y−1).
4. Select an integer e*e*, such that e*e* is ***co-prime*** to \phi(n)$\phi(n)$ and 1 < e < \phi(n)1<e<$\phi(n)$. The pair of numbers (n,e)(n,e) makes up the public key.

Note: Two integers are co-prime if the only positive integer that divides them is 1.

5. Calculate d*d* such that e.d = 1e.d=1 mod *mod* \phi(n)$\phi(n)$.
6. d*d* can be found using the ***extended euclidean algorithm***. The pair (n,d)(n,d) makes up the private key.

2. Encryption

Given a plaintext P*P*, represented as a number, the ciphertext C*C* is calculated as:

C = P^{e}*C*=*Pe* mod*mod* n*n*.

3. Decryption

Using the private key (n,d)(n,d), the plaintext can be found using:

P = C^{d}*P*=*Cd* mod*mod* n*n*.

Pseudocode

Consider an example of the RSA algorithm through the following pseudocode:

$$h_i(K_{file}) = \mathsf{H}(K_{file}||i) : \{0,1\}^* \rightarrow [1, n - 256].$$

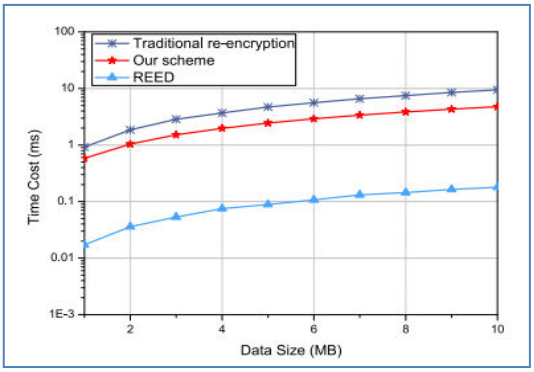
V. EXPERIMENT RESULTS

We measure the performance ofgatheredpackage generation. The gathered package generation time consists of location generation, bits extraction, and concatenation time. According to the paper, we set the chunk size to 8 KB and 16 KB. The cost computation time for gathered package generation. The results show that the overhead of gathered package generation for 16 KB chunk is less than 8 KB chunk. This is because the large size of the chunk requires less processing overhead.

Fig2. Computation cost time

Chunk Size	Traditional re-encryption		Our Scheme	
	Re-encryption	Re-decryption	Re-encryption	Re-decryption
4 KB	100%Enc	100%Dec	Gen + 1.56%Enc	Rec + 1.56%Dec
8 KB	100%Enc	100%Dec	Gen + 0.78%Enc	Rec + 0.78%Dec
16 KB	100%Enc	100%Dec	Gen + 0.39%Enc	Rec + 0.39%Dec

Fig3. The computational time for re-encryption



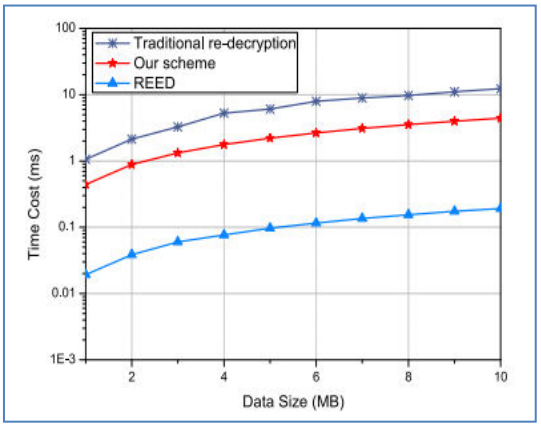


Fig 4. The computational time for re-decryption

VI. CONCLUSION

In this paper, we propose a Bloom filter-based location selection method and a secure data deduplication scheme with efficient re-encryption. By using symmetric encryption and the new location selection method, the revoked cloud user cannot obtain sensitive data from the data owner. In addition, instead of re-encrypting the entire package, data owners are only required to re-encrypt a small part of it through the CAONT, which saves excessive computation overhead. We also prove that our scheme can achieve the desired security goals and provide detailed simulation tests. The experimental results show that our scheme is efficient in re-encryption.

REFERENCES

[1] X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, "Verifiable computation over the large database with incremental updates, "IEEE Trans. Comput., vol. 65, no. 10, pp. 3184–3195, Oct. 2016.
[2] M. Gerla, J. Weng, and G. Pau, “Pics-on-wheels: Photo surveillance in the vehicular cloud,” in Proc. Int. Conf. Comput. Netw. Commun., 2013, pp. 1123–1127.
[3] X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, “New algorithms for secure outsourcing of modular exponentiations, "IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 9, pp. 2386–2396, Sep. 2014.
[4] H. Yuan, X. Chen, T. Jiang, X. Zhang, Z. Yan, and Y. Xiang, “DedupDUM: Secure and scalable data deduplication with dynamic user management, "Inf. Sci., vol. 456, pp. 159–173, 2018.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

[5] H. Huang, X. Chen, Q. Wu, X. Huang, and J. Shen, "Bitcoin-based fair payments for outsourcing computations of fog devices," Future Gener. Comput. Syst., vol. 78, pp. 850–858, 2018.
[6] IDC, "The digital universe of opportunities: Rich data and the increasing value of the Internet of Things," 2014. [Online]. Available: https://www.emc.com/leadership/digital-universe/2014iview/index.htm
[7] W. J. Bolosky, S. Corbin, D. Goebel, and J. R. Douceur, "Single instance storage in windows 2000," in Proc. Conf. Usenix Windows Syst. Symp., 2000, pp. 2–2.
[8] Dropbox, 2007. [Online]. Available: http://www.dropbox.com
[9] GoogleDrive, 2012. [Online]. Available: http://drive.google.com
[10] Memopal, 2018. [Online]. Available: http://www.memopal.com
[11] Netapp, "Netapp deduplication helps duke institute for genome sciences and policy reduce storage requirements for genomic information by 83 percent," 2008. [Online]. Available: http://www.netapp.com
[12] M. Dutch, "Understanding data deduplication ratios, "SNIA Data Manage. Forum, pp. 1–13, June 2008. [Online]. Available: http://www.snia.org
[13] T. Jiang, X. Chen, J. Li, D. S. Wong, J. Ma, and J. K. Liu, "TIMER: Secure and reliable cloud storage against data re-outsourcing," in Proc. 10th Int. Conf. Inf. Security Practice Experience, 2014, pp. 346–358.
[14] X. Chen, B. Lee, and K. Kim, "Receipt-free electronic auction schemes using homomorphic encryption," inProc. 6th Int. Conf. Inf. Security Cryptology, 2003, pp. 259–273.

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |

www.ijarasem.com