PRIVACY AWARE DATA DEDUPLICATION FOR SIDE CHANNEL IN CLOUD STORAGE

Mrs.N.Antony Sophia[1], Sanapala Upendra Rao[2], TamilSelvan G[3]

[1] Assistant Professor, Computer Science And Engineering, Jeppiaar SRR Engineering College, Tamil Nadu, India
[2]UG Student, Computer Science And Engineering, Jeppiaar SRR Engineering College, Tamil Nadu, India
[3]UG Student, Computer Science And Engineering, Jeppiaar SRR Engineering College, Tamil Nadu, India

ABSTRACT

Data de duplication technique has been widely adopted by commercial cloud storage providers, which is both important and necessary in dealing with the explosive growth of information. To further protect the protection of users' sensitive data within the outsourced storage mode, many secure data de duplication schemes are designed and applied in various scenarios. Among these schemes, secure and efficient re-encryption for encrypted data de duplication attracted the eye of the many scholars, and lots of solutions are designed to support dynamic ownership management. during this paper, we target the re-encryption de duplication storage system and show that the recently designed lightweight rekeying-aware encrypted de duplication scheme is liable to an attack which we call it stub-reserved attack. Furthermore, we propose a secure data de duplication scheme with efficient re-encryption supported the convergent all-or-nothing transform and randomly sampled bits from the Bloom filter. because of the intrinsic property of one-way hash function, our scheme can resist the stub-reserved attack and guarantee the info privacy of information owners' sensitive data. Moreover, rather than re-encrypting the complete package, data owners are only required to re-encrypt a tiny low a 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.

Keyword: *-Sparse Social Dimension, Line Graph Partition*

1. 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) 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 5200 gigabytes for each man, woman, and child). However, the growth of data puts heavy pressures on cloud service providers. To cope with it, a straightforward method is to require cloud service providers continuously increasing the capacity of storage space, so as to meet users' requirements for high-quality storage services. However, cloud service providers may store plentiful and repetitive data (such as movies, music and genome data), which inevitably incurs a mass of redundant storage and backup space, consequently to cost a vast amount of computing and management overhead during its whole life cycle.

2. EXISTING SYSTEM

In existing, the unquestionable SSE plans supporting information dynamic update are altogether founded on deviated key cryptography confirmation, which includes tedious activities. The overhead of check may turn into a critical weight because of the sheer measure of cloud information.

3. PROPOSED SYSTEM

In the proposed framework, we explore achieving watchword search over interesting encoded cloud data with symmetric-key based affirmation and propose a reasonable plot in this paper. In order to help the capable check of dynamic data, we plan a novel Accumulative Authentication Label reliant on the symmetric-key cryptography to deliver an affirmation tag for each catchphrase.

3.1. Architecture Diagram

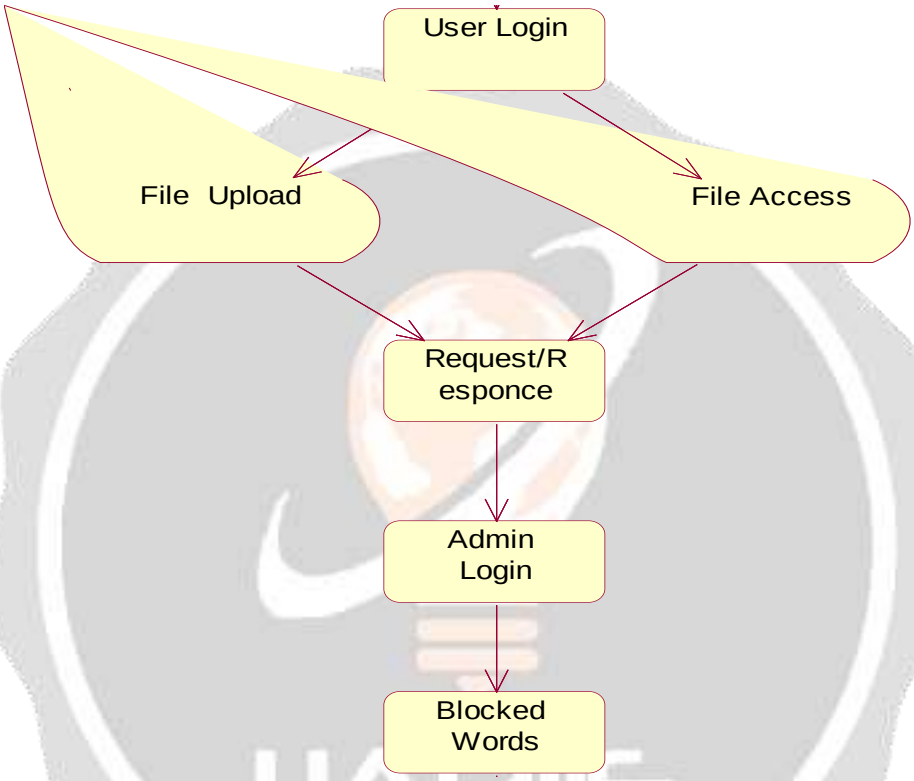

Fig-1: Architecture Diagram

4. SYSTEM HARDWARE AND SOFTWARE

Hardware Used

The hardware requirements may serve as the basis for a contract for the implementation of the system and should therefore be a complete and consistent specification of the whole system. They are used by software engineers as the starting point for the system design. It shows what the system does and not how it should be implemented.

PROCESSOR	**:**	PENTIUM IV 2.6 GHz, Intel Core 2 Duo.
RAM	**:**	4 GB DD RAM
MONITOR	**:**	15” COLOR
HARD DISK	**:**	40 GB

Software Used

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost,

planning team activities, performing tasks and tracking the teams and tracking the team's progress throughout the development activity.

FRONT END	**:**	J2EE (JSP, SERVLETS) JAVASCRIPT
BACK END	**:**	MY SQL 5.5
OPERATING SYSTEM	**:**	Windows 07
IDE	**:**	Eclipse

.

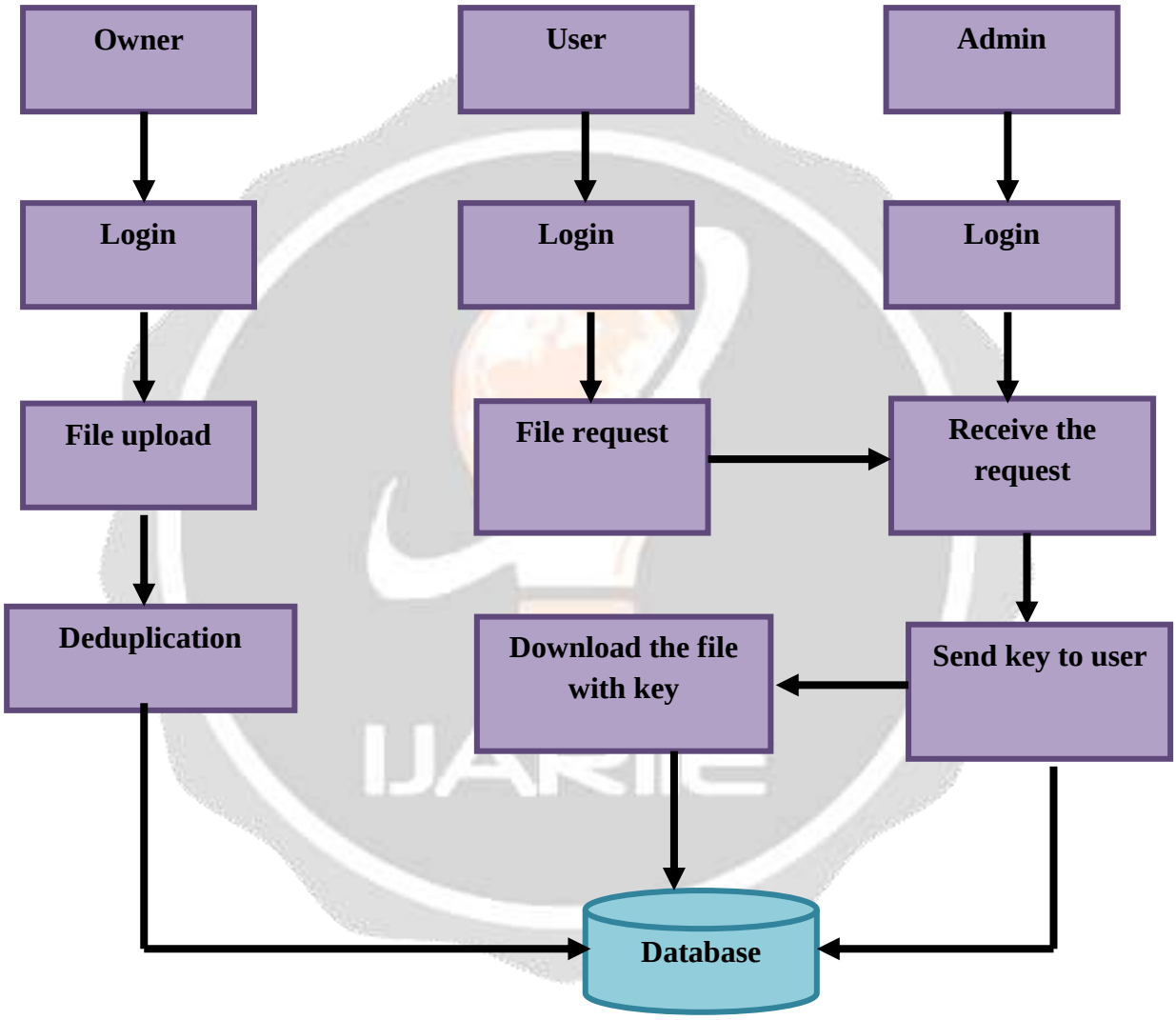

Fig-2: Detailed Diagram

5. WORKING

MODULES:

- **User interface design**
- **File upload**

- **De duplication process**
- **Request file to user**
- **Admin response by sending the key**
- **Download the file**

- **USER INTERFACE DESIGN:**

This is the first module of our project. The important role for the user is to maneuver login window to user window. This module has created for the security purpose. In this login page we've to enter login user id and password. It will check username and password is match or not (valid user id and valid password). If we enter any invalid username or password we can't enter into login window to user window it'll shows error message. So we are preventing from unauthorized user getting into the login window to user window. It will provide an honest security for our project. So server contain user id and password server also check the authentication of the user. It well improves the safety and preventing from unauthorized user enters into the network. In our project we are using JSP for creating design. Here we validate the login user and server authentication.

- **FILE UPLOAD:**

In this module, after login the owner will upload the file which can be related to science, technology or anything else and while storing in the database it will be getting encrypted.

- **DE DUPLICATION PROCESS :**

In this module, once the owner uploaded the file it can't be uploaded one more time. Because if the owner uploading the same for more than 5 times means, there will be duplicate data stored in the cloud and after that there won't be storage for storing those data's.

- **REQUEST FILE TO USER:**

In this module, after the deduplication process, the user will be sending the request the file to the admin for access.

- **ADMIN RESPONSE BY SENDING THE KEY:**

In this module, the admin will be receiving the file request from user. After that, the admin will be accepting the request and key will be send to user.

- **DOWNLOAD THE FILE:**

In this module, after receiving the key from admin the user will be downloading the file by using the key allocated for that file.

6. EXPERIMENTAL RESULTS

The result involves a simple yet efficient model, called dual sentiment analysis (DSA), to affect the polarity shift problem in sentiment classification. By using the property that sentiment classification has two opposite class labels (i.e., positive and negative), we first propose a knowledge expansion technique by creating sentiment reversed reviews. The original and reversed reviews are constructed during a one-to-one correspondence.

7. APPLICATIONS

- In Hospital Based System
- In Banking Sector for security Purpose
- Airlines and transport based system
- In Educational Institute
- Computerized Library System

8. CONCLUSION

In this paper, we investigate acknowledging catchphrase search over dynamic encoded cloud information with symmetric-key based check. So as to help the productive confirmation of dynamic information, we structure a novel Accumulative Authentication Tag in view of symmetric-key cryptography to create an aggregate confirmation tag for every watchword. Besides, another protected record dependent on the symmetrical rundown and the single connected list is intended to improve the refreshed productivity. The security investigation and the presentation assessment show that the proposed conspire are secure and productive.

9. REFERENCES

[1] X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, "Verifiable computation over large database with incremental updates," IEEE Trans. Computers, vol. 65, no. 10, pp. 3184–3195, 2016.
[2] M. Gerla, J. Weng, and G. Pau, "Pics-on-wheels: Photo surveillance in the vehicular cloud," International Conference on Computing, Networking and Communications, pp. 1123–1127, 2013.
[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, 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.
[5] H. Huang, X. Chen, Q. Wu, X. Huang, and J. Shen, "Bitcoinbased fair payments for outsourcing computations of fog devices," Future Generation Comp. Syst., vol. 78, pp. 850–858, 2018.
[6] X. Chen, B. Lee, and K. Kim, "Receipt-free electronic auction schemes using homomorphic encryption," Information Security and Cryptology - ICISC 2003, 6th International Conference, Seoul, Korea, November 27-28, 2003, Revised Papers, pp. 259–273, 2003.
[7] J. Wang, X. Chen, J. Li, K. Kluczniak, and M. Kutylowski, "Trdup: enhancing secure data deduplication with user traceability in cloud computing," IJWGS, vol. 13, no. 3, pp. 270–289, 2017.
[8] X. Zhang, X. Chen, J. Wang, Z. Zhan, and J. Li, "Verifiable privacypreserving single-layer perceptron training scheme in cloud computing," Soft Comput., vol. 22, no. 23, pp. 7719–7732, 2018.
[9] M. Dutch, "Understanding data deduplication ratios," in SNIA Data Management Forum, 2008, pp. 1–13.
[10] 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," Information Security Practice and Experience - 10th International Conference, pp. 346–358, 2014.