International Journal of Trend in Scientific Research and Development (IJTSRD)
International Open Access Journal
ISSN No: 2456 - 6470 | www.ijtsrd.com | Volume - 2 | Issue – 4

Secure Data Deduplication with Dynamic Ownership Management in Cloud Storage

K Archana[1], Mrs. Shubangini Patil[2]
[1]M.Tech Student, [2]Professor
Department of Computer Science and Engineering, AIET College, Karnataka, India

ABSTRACT
In this paper, we propose a novel server-side deduplication scheme for encrypted data. It allows the cloud server to control access to outsourced data even when the ownership changes dynamically by exploiting randomized convergent encryption and secure ownership group key distribution. This prevents data leakage not only to revoked users even though they previously owned that data, but also to an honest-but-curious cloud storage server. In addition, the proposed scheme guarantees data integrity against any tag inconsistency attack. Thus, security is enhanced in the proposed scheme. The efficiency analysis results demonstrate that the proposed scheme is almost as efficient as the previous schemes, while the additional computational overhead is negligible.

I. INTRODUCTION
In cloud storage services, deduplication technology is commonly used to reduce the space and bandwidth requirements of services by eliminating redundant data and storing only a single copy of them. Deduplication is most effective when multiple users outsource the same data to the cloud storage, but it raises issues relating to security and ownership. Proof of ownership schemes allow any owner of the same data to prove to the cloud storage server that he owns the data in a robust way. However, many users are likely to encrypt their data before outsourcing them to the cloud storage to preserve privacy, but this hampers deduplication because of the randomization property of encryption. Recently, several deduplication schemes have been proposed to solve this problem by allowing each owner to share the same encryption key for the same data. However, most of the schemes suffer from security flaws, since they do not consider the dynamic changes in the ownership of outsourced data that occur frequently in a practical cloud storage service.

Cloud computing provides scalable, low-cost, and location-independent online services ranging from simple backup services to cloud storage infrastructures. The fast growth of data volumes stored in the cloud storage has led to an increased demand for techniques for saving disk space and network bandwidth. To reduce resource consumption, many cloud storage services, such as Drop box, Wuala, Mozy, and Google Drive, employ a deduplication technique, where the cloud server stores only a single copy of redundant data and provides links to the copy instead of storing other actual copies of that data, regardless of how many clients ask to store the data. The savings are significant, and reportedly, business applications can achieve disk and bandwidth savings of more than 90%. However, from a security perspective, the shared usage of users' data raises a new challenge

II. LITERATURE SURVEY

A Dynamic Layering Scheme of Multicast Key Management Group key management is a difficult task in implementing large and dynamic secure multicast. In this paper, a new scheme is proposed in the basis of in-depth analysis of the requirements of the secure multicast and group key management. The scheme is based on the multicast group security architecture and multicast security group key management architecture proposed by IETF. This scheme constructs group key based on pairings and distributes the group key using HSAH function polynomial, and manages group key making use of

the dynamic layering GCKS. The scheme is better in security, lower in computation cost and communication cost. The analysis comparison proves that the scheme has strong scalability and efficiency.

Tree-based Group Key Agreement (2003): KIM, PERRIG AND TSUDIK [3]: In this paper it is said that fault-tolerant, scalable and reliable communication services have become critical in modern computing. An important and popular trend is to convert traditional centralized services (e.g., file sharing, authentication, web, and mail) into distributed services spread across multiple systems and networks. Many of these newly distributed and other inherently collaborative applications (e.g., conferencing, white-boards, shared instruments, and command-and-control systems) need secure communication. However, experience shows that security mechanisms for collaborative and dynamic peer groups tend to be both expensive and unexpectedly complex.

Energy and Communication Efficient Group Key Management Protocol for Hierarchical Sensor Networks(2006):PANJA, MADRIA AND BHARGAVE [4]: In this paper it is said that the security of sensor networks has become one of the most pressing issues in further development of these networks. Compared to the traditional wireless network, Wireless Sensor Network (WSN) provides a different computation and communication infrastructure. These differences originate not only from their physical characteristics, but also from their typical applications. For example, the physical characteristics include the large scale of deployment, limited computing capability, and constraints on power consumption. As a result, the requirements for the key management of a WSN are noticeably different from those for traditional networks.

An Efficient Hierarchical Group Key Management Protocol for Mobile Ad-Hoc Networks (2009):DAWOOD, MNENEY, AGHDASI AND DAWOUD [5]: In this paper it is said that a mobile ad hoc network (MANET) is a collection of autonomous nodes that communicate with each other, most frequently using a multi-hop wireless network. Nodes do not necessarily know each other and come together to form an ad hoc group for some specific purpose. Key distribution systems usually require a trusted third party that acts as a mediator between nodes of the network.

III. EXISTING SYSTEM

In existing system, Cryptographic techniques were applied to access control for remote storage systems. The data owners encrypt files by using the symmetric encryption approach with content keys and then use every user's public key to encrypt the content keys. It requires each data owner to be online all the time. Some methods deliver the key management and distribution from the data owners to the remote server under the assumption that the server is trusted or semi-trusted.

Disadvantages:

- The key management is very complicated when there are a large number of data owners and users in the system.
- The key distribution is not convenient in the situation of user dynamically system.
- The server is cannot be trusted by the data owners in cloud storage systems.
- It cannot be applied to access control for cloud storage systems.

IV. PROPOSED SYSTEM

We propose an efficient group key management protocol in distributed group communication. This protocol is based on Elliptic Curve Cryptography and decreases the key length while providing securities at the same level as that of other cryptosystems provides. We provide the high level security and avoid the replication of file in the cloud service provider. In proposed system, we use hash function to generate key for the file .By using hash function to avoid the duplication in cloud. After that we are applying cryptographic technique for security purpose. We using ECC algorithm for encryption and decryption process.

Advantages:

- Avoid duplication in cloud.
- Increase the security level.
- High efficient.
- ECC algorithm provides high end security.

V. MODULES EXPLANATION

1. Registration and Login
2. Join Group and File Upload
3. File encrypt and store into Cloud
4. User request and Download

Registration and Login: In this process, new user registers the details and gets the username and password for further process. Using Username and Password user login into Group. Group generate key for the valid user and process inside the group under the valid key.

Join Group and File Upload: In file upload process, user choose the file from the system and generate hash key for each file. Hash key generation is provided to avoid duplication of file to the cloud. If the file is already in cloud

File Encrypt and Store into Cloud: After the validation of file from the user with cloud, we apply cryptographic technique to improve the security level in cloud. For cryptographic technique, we are using Elliptic Curve Cryptography (ECC) algorithm for encrypting the file. In Elliptic Curve Cryptography (ECC), it convert the file into binary format and store it in cloud.

User request and Download: User send request to the cloud, cloud service provider decrypt the file .For cryptographic technique, we using Elliptic Curve Cryptography (ECC) algorithm for decrypting the file. Send the requested file to the user after validate the user. Then file will be downloaded in user location.

VI. SYSTEM ARCHITECTURE

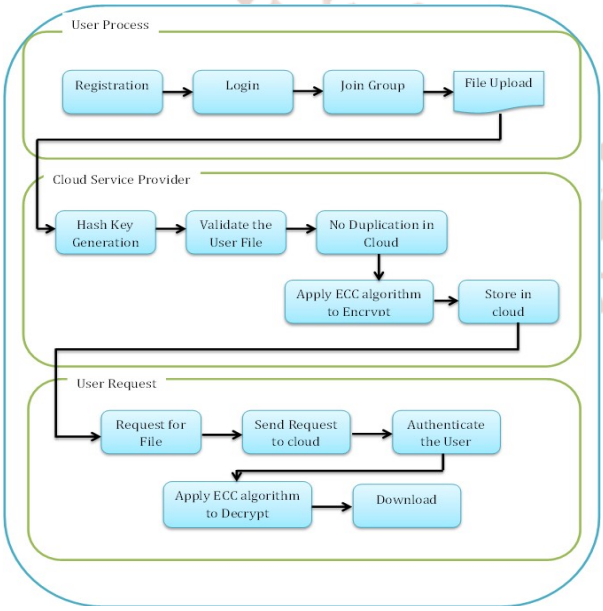

Figure-1: System Architecture

The above figure-1 depicts the System Architecture of the proposed system. Here initially the user has to undergo the registration, login procedure. Then the hash key is generated and then the validation of the user is checked and the user request is sent for processing regarding whether the file is need to be uploaded or not to upload based on whether the file is already existing or not.

VII. DATA FLOW DIAGRAM

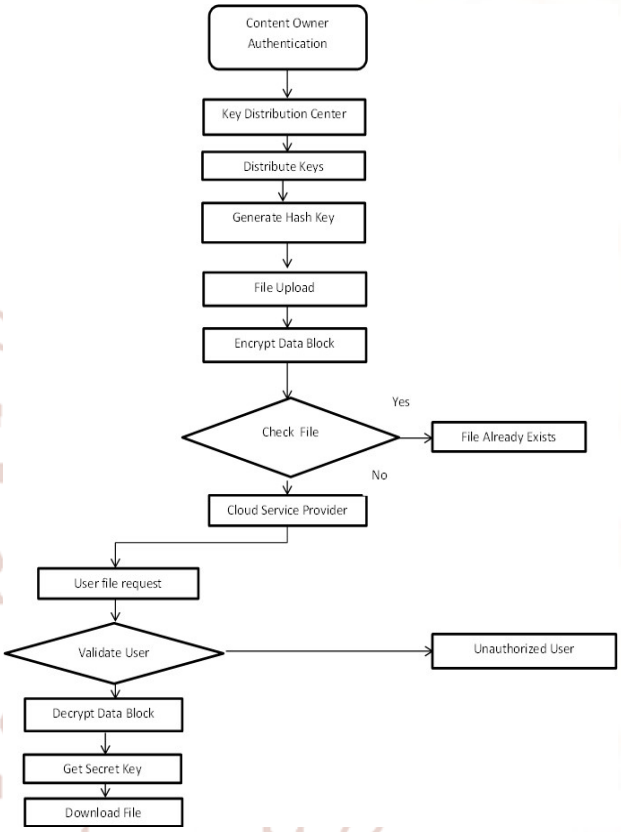

Figure 2 Data Flow Diagram

The above figure-2 shows the flow of the project in a step by step manner. Here in the above figure where ever the is need of some checking for the questions types whose answer may be yes or no are represented in the rhombus shape . Here the condition that is needed to be checked is return in this polygon and then based on the consequence the further steps are carried out. Normal statements will be return in the rectangular forms in the flow diagrams.

VIII. RESULTS

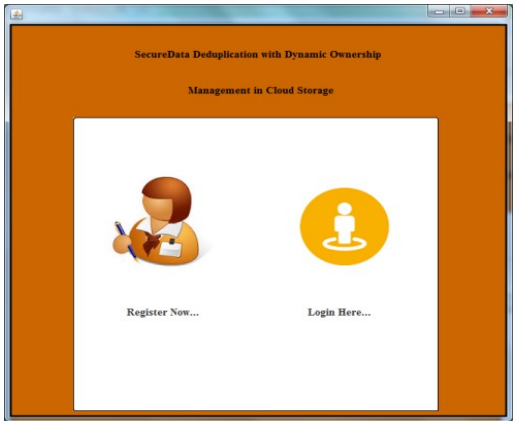

Figure-3: Beginning page of the project.

The figure-3 shows the beginning page of the project where the user is provided with choice of either registration for the new user or the login of the already previously registered user.

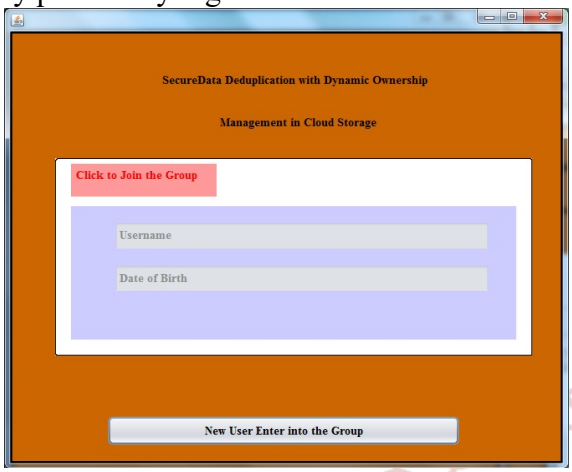

Figure-4: Form to join the group.

The figure-4 shows the form to join the group of users where the user is need to click the "click to join the group".

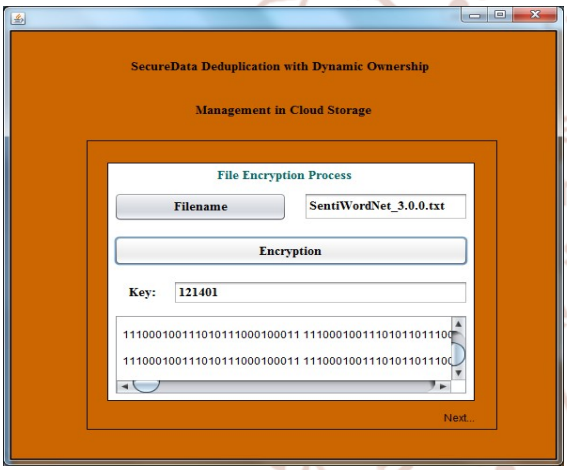

Figure-5: Encryption page for the newly uploaded file.

The figure-5 shows the encryption page for the fie where the file that is need to be encrypted is chosen and then the encrypting of the file will be carried out when the key is known and then the file will be encrypted.

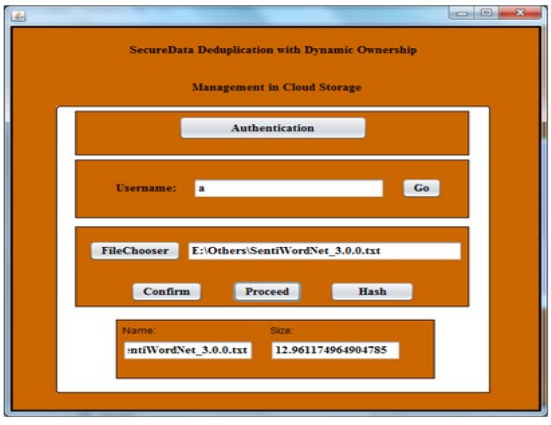

Figure-6: Uploading page for the file.

The figure-6 shows the uploading dialog box. Here the user has to choose the file that he wants to upload and also select the path and then confirm and then click on the hash.

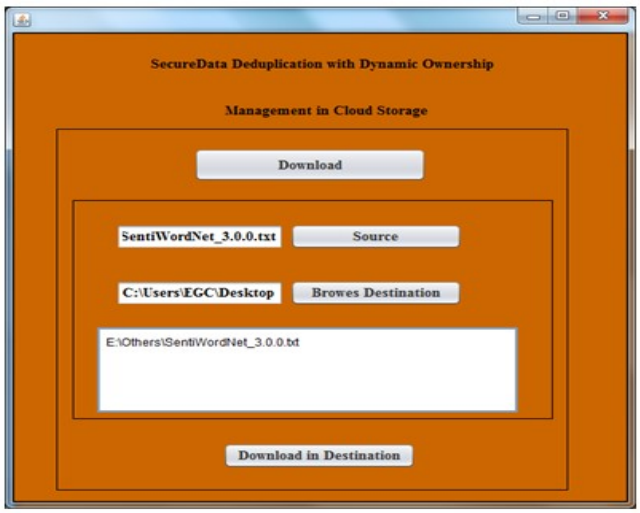

Figure-7: Downloading page for the encrypted file.

The figure-7 shows the downloading page for file. Here the file that is to be downloaded has to be selected and also we need to mention the destination folder of where the file is need to be downloaded so that once the file is downloaded than the file can be seen whenever the user requires the file.

CONCLUSION

In this paper, we proposed an efficient and privacy preserving big data deduplication in cloud storage. We then analyzed the security of our proposed scheme and demonstrated that it achieves improved privacy preserving, accountability and data availability, while resisting brute-force attacks. Future research includes extending the proposed scheme to fully protect the duplicate information from disclosure, even by malicious data providers, without affecting the capability to perform data deduplication.

REFERENCES

1. J. Xu, E. Jhang, and hu, "Leakage-resilient client-side deduplication of encrypted data in cloud storage," ePrint, IACR, http://eprint.iacr.org/2011/538.
2. A Dynamic Layering Scheme of Multicast Key Management : FAN, PING, KUAN AND MING
3. Tree-based Group Key Agreement (2003): KIM, PERRIG AND TSUDIK
4. Energy and Communication Efficient Group Key Management Protocol for Hierarchical Sensor

Networks(2006):PANJA, MADRIA AND BHARGAVE

5. An Efficient Hierarchical Group Key Management Protocol for Mobile Ad-Hoc Networks (2009):DAWOOD, MNENEY, AGHDASI AND DAWOUD
6. 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. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-38348-9 18
7. S. Keelveedhi, M. Bellare, and T. Ristenpart, "Dupless: Server-aided encryption for deduplicated storage," in Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013, 2013, pp. 179–194. [Online]. Available: https://www.usenix.org/conference/usenixsecurity13/technical-sessions/presentation/bellare
8. M. Abadi, D. Boneh, I. Mironov, A. Raghunathan, and G. Segev "Message-locked encryption for lock-dependent messages," in Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I, 2013, pp. 374–391. [Online]. Available: http://dx.doi.org/10.1007/ 978-3-642-40041-4 21
9. Y. Zhou, D. Feng, W. Xia, M. Fu, F. Huang, Y. Zhang, and C. Li, "Secdep: A user-aware efficient fine-grained secure deduplication scheme with multi-level key management," in IEEE 31st Symposium on Mass Storage Systems and Technologies, MSST 2015, Santa Clara, CA, USA, May 30 - June 5, 2015, 2015, pp. 1–14. [Online]. Available: http://dx.doi.org/10.1109/MSST.2015.7208297
10. Dropbox, http://www.dropbox.com/.
11. Wuala, http://www.wuala.com/.
12. Mozy, http://www.mozy.com/.
13. Google Drive, http://drive.google.com.
14. IDC, "Executive summary: Data growth, business opportunities, and the it imperatives,," http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm, 2014.