IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SECURE DOCUMENT STORAGE AND KEYWORD SEARCH SYSTEM USING CRYPTOGRAPHY

[1] HARDI AMRANIYA, [2] MAYUR ATTARDE, [3] DIVYA MACHHA, [4] AYUSH MEHARE,

[5] PROF.ISHWARI RASKAR(Guide)

DEPARTMENT OF INFORMATION TECHNOLOGY (BE IT)
TRINITY COLLEGE OF ENGINEERING AND RESEARCH, PUNE, INDIA

ABSTRACT

Now a day's cloud computing is used in many areas like industry, military colleges, healthcare, etc. to store huge amounts of data. We can retrieve data from cloud at request of a user. To store data on the cloud we have to face many issues. To provide the solution to these issues there are n number of ways. In Cloud Users can remotely store their data to the cloud & realize the data sharing with others. In Some Common cloud storage system such as the electronic health records system, the cloud file might contain some sensitive information. Encrypting the whole shared file can realize sensitive information hiding, but will make this shared file unable to be used by others. In cloud computing more Sensitive information hiding in cloud. This is very big problem that remote data integrity auditing scheme that realizes data sharing with sensitive information hiding in this cloud. In our System, our scheme makes the file stored in the cloud able to be shared and used by others on the condition that the sensitive information is hidden, while the remote data integrity auditing is still able to be efficiently executed. Meanwhile, the proposed scheme is based on identity-based cryptography, which simplifies complicated certificate management.

Keywords: *AWS, EC2, Cloud computing, data storage, security, data management.*

INTRODUCTION

Compared with the traditional information sharing and communication technology, cloud computing has attracted the interest of most researchers because a lot services are provided by the cloud service providers which helps to reduce costs needed for various resources. Cloud storage is one of the most vital service in cloud computing. Scalability is another attracting factor which allows user to scale up and scale down the resources as required. Cloud computing also provides convenient and flexible ways for data sharing. There are two ways to share data in cloud storage. The first case refers to the scenario where one client authorizes access to his/her data for many clients known as one-to-many pattern and the second case refers to a situation in which many clients in the same group authorize access to their data for many clients at the same time known as many-to-many pattern. As the data shared on the cloud is valuable, various security methods are provided by cloud. In current cloud applications various algorithms are used for data encryption and decryption. In encryption is based on ABE [Attribute Based Encryption. Symmetric-key cryptography is used in to enable efficient encryption. Practical group key management algorithm based on a proxy re-encryption technology. In existing system when a user is revoked from a group, he is still able to access files from his previous group which leads to collision attack. Another gap is that a user is not allowed to upload multiple files of same name. The process of

cryptography encrypts data so that it cannot be decoded. Strictly speaking, there are two types of cryptography: symmetric key and public key. This method makes data unreadable by encoding it with special keys. So only those who have been granted access to the cloud server are able to access the data. All people can see the cipher text data.

Symmetric key cryptography algorithms are AES, DES, 3DES, IDEA, BRA and blowfish. The main issue is delivering the key to receiver into multi user application. These algorithms require low delay for data encoding decode provides low security. The publicc key cryptography algorithm is RSA and ECC algorithm. Public and private keys are manipulated into public key cryptography algorithms. These algorithms accomplished high level of security but increase delay for encoding decoding code. Steganography hides the secret data existence into envelope. In this technique existence of data is not visible to all people. Only valid receiver knows about the data existence. Text steganography technique is used to produce high security for data.The secret data of user hide into text cover file. After adding text into text cover file it looks like normal text file. If text file found by illegitimate user then also cannot get sensitive data. If illegitimate user try to recover original data than large amount of time is essential.DES algorithm is used for text encode and decode. Advantage of text steganography technique is providing security to text.

METHODOLOGY

User:

- In the proposed scheme, members are people with the same interests (e.g., bidder, doctors, and businessmen) and want to share data in the cloud.
- The most worrying problem when users store data in the cloud server is the confidentiality of the outsourced data.
- In this system, users of the same user conduct a key agreement. Subsequently, a common conference key can be used to encrypt the data that will be uploaded to the cloud to ensure the confidentiality of the outsourced data.

Admin:

- admin is responsible for generating system parameters, managing group members (i.e., uploading member's encrypted data, authorizing users), and for fault tolerance detection.
- The admin in our scheme is a fully trusted third party to both the cloud and users.
- If an external user tries to access files from a different group more than three times then the manager will remove that particular user from the applications.

Cloud Service Provider(CSP):

- CSP provides users with seemingly unlimited storage services.
- In addition to providing efficient and convenient storage services for users, the cloud can also provide data sharing services.
- However, the cloud has the characteristic of honest but curious.
- In other words, the cloud will not deliberately delete or modify the uploaded data of users.

LITERATURE SURVEY

1. Security challenges for the public cloud" Cloud computing is the newest term for the long-dreamed vision of computing as a utility. The cloud provides convenient, on-demand network access to a centralized pool of configurable computing resources that can be rapidly deployed with great efficiency and minimal management overhead.1 With its un-precedence advantages, cloud computing enables a fundamental paradigm shift in how we deploy and deliver computing services — that is, it makes possible computing outsourcing such that both individuals and enterprises can avoid committing large capital outlays when purchasing and managing software and hardware, as well as dealing with the operational overhead therein.

2. "Provable Data Possession at Untrusted Stores" We introduce a model for provable data possession (PDP) that allows a client that has stored

data at an untrusted server to verify that the server possesses the original data without retrieving it. The model generates probabilistic proofs of possession by sampling random sets of blocks from the server, which drastically reduces I/O costs. The client maintains a constant amount of metadata to verify the proof. The challenge/response protocol transmits a small, constant amount of data, which minimizes network communication. Thus, the PDP model for remote data checking supports large data sets in widely-distributed storage systems. We present two provably-secure PDP schemes that are more efficient than previous solutions, even when compared with schemes that achieve weaker guarantees. In particular, the overhead at the server is low (or even constant), as opposed to linear in the size of the data. Experiments using our implementation verify the practicality of PDP and reveal that the performance of PDP is bounded by disk I/O and not by cryptographic computation.

3. "POR S: Proofs of Retrievability for Large Files" In this paper, we define and explore proofs of retrievability (PORs). A POR scheme enables an archive or back-up service (prover) to produce a concise proof that a user (verifier) can retrieve a target file F, that is, that the archive retains and reliably transmits file data sufficient for the user to recover F in its entirety. A POR may be viewed as a kind of cryptographic proof of knowledge (POK), but one specially designed to handle a large file (or bit string) F. We explore POR protocols here in which the communication costs, number of memory accesses for the prover, and storage requirements of the user (verifier) are small parameters essentially independent of the length of F. In addition to proposing new, practical POR constructions, we explore implementation considerations and optimizations that bear on previously explored, related schemes. In a POR, unlike a POK, neither the prover nor the verifier need has knowledge of F. PORs give rise to a new and unusual security definition whose formulation is another contribution of our work. We view PORs as an important tool for semi-trusted online archives. Existing cryptographic techniques help users ensure the privacy and integrity of files they retrieve. It is also natural, however, for users to want to verify that archives do not delete or modify files before retrieval. The goal of a POR is to accomplish these checks without users having to download the files themselves. A POR can also provide quality-of-service guarantees, i.e., show that a file is retrievable within a certain time bound.

4. "Privacy Preserving Public Auditing for Secure Cloud Storage" —Using Cloud Storage, users can remotely store their data and enjoy the on-demand high-quality applications and services from a shared pool of configurable computing resources, without the burden of local data storage and maintenance. However, the fact that users no longer have physical possession of the outsourced data makes data integrity protection in Cloud Computing a formidable task, especially for users with constrained computing resources. Moreover, users should be able to just use the cloud storage as if it is local, without worrying about the need to verify its integrity. Thus, enabling public audit ability for cloud storage is of critical importance so that users can resort to a third-party auditor (TPA) to check the integrity of outsourced data and be worry-free. To securely introduce an effective TPA, the auditing process should bring in no new vulnerabilities towards user data privacy, and introduce no additional online burden to users. In this paper, we propose a secure cloud storage system supporting privacy-preserving public auditing. We further extend our result to enable the TPA to perform audits for multiple users simultaneously and efficiently. Extensive security and performance analysis shows the proposed schemes are provably secure and highly efficient.

ALGORITHM(EQUATIONS):

AES (advanced encryption standard). It is a symmetric algorithm. It is used to convert plain text into cipher text. The need for coming up with this algo is a weakness in DES. The 56-bit key of des is no longer safe against attacks based on exhaustive key searches and the 64-bit block also consider weak.AES was to be used128-bit block with 128-bit keys.

Input:

128_bit /192 bit/256-bit input (0, 1)

Secret key (128_bit) +plain text (128_bit).

Process:

10/12/14-rounds for-128_bit /192 bit/256 bit input

Xor state block (i/p)

Final round:10,12,14

Each round consists of sub byte, shift-byte, mix columns, and add round key.

Output:

Cipher text (128 bit)

PROPOSED SYSTEM:

In Our System proposed an identity-based data integrity auditing scheme for secure cloud storage, which supports data sharing with sensitive information hiding. In our Application user uploads the data into the cloud with the user and researcher when Doctor Shares the data with the User that file goes to Admin and the admin converts it into Binary format after that binary format file again converts into Homomorphic encryption and Stored into Block Level.

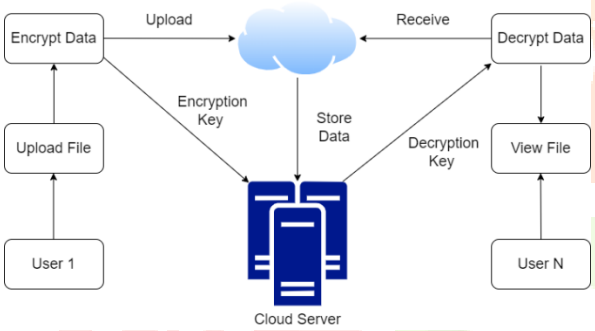
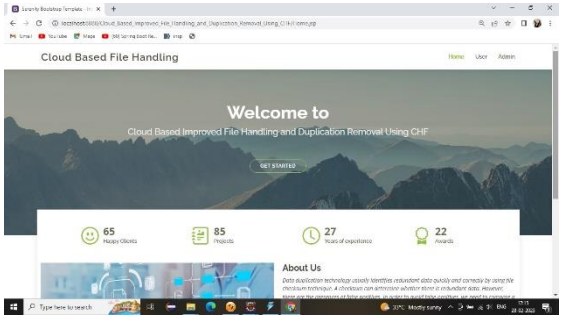


Figure 1. System Architecture

- In our system cloud then sensitive information is hidden with the help of hidden data identity Auditing called identity-based shared data.
- In Our System the file stored in the cloud can be shared and used by others on the condition that the sensitive information is protected while the remote data integrity auditing is still able to be efficiently executed.
- In Cloud Stored using Block Level concepts.

RESULT

The result of the project is that a secure system for storing and searching documents has been created using cryptography. This means that documents can be stored safely without the risk of unauthorized access. The system also allows for easy and efficient searching of stored documents using specific keywords.



CONCLUSION

In this paper, we proposed an identity-based data integrity auditing scheme for secure cloud storage, which supports data sharing with sensitive information hiding. In our scheme, the file stored in the cloud can be shared and used by others on the condition that the sensitive information of the file is protected. Besides, remote data integrity auditing is still able to be efficiently executed. The security proof and the experimental analysis demonstrate that the proposed scheme achieves desirable security and efficiency.

APPLICATIONS

- Group sharing applications
- Security Applications
- Searching applications

REFERENCES

[1] J. Shen, J. Shen, X. Chen, X. Huang, and W. Susilo, "An efficient public auditing protocol with the novel dynamic structure for cloud data," IEEE Trans. Inf. Forensics Security, vol. 12, no. 10, pp. 2402–2415, Oct. 2017.

[2] Y. Li, Y. Yu, G. Min, W. Susilo, J. Ni, and K.-K. R. Choo, "Fuzzy identity-based data integrity auditing for reliable cloud storage systems," IEEE Trans. Depend. Sec. Comput., to be published

[3] J. Hur, D. Koo, Y. Shin, and K. Kang, "Secure data deduplication with dynamic ownership management in cloud storage," IEEE Trans. Knowl. Data Eng., vol. 28, no. 11, pp. 3113–3125, Nov. 2016.

[4] J. Li, J. Li, D. Xie, and Z. Cai, "Secure auditing and deduplicating data in the cloud," IEEE Trans. Comput., vol. 65, no. 8, pp. 2386–2396, Aug. 2016.

[5] Y. Zhang, J. Yu, R. Hao, C. Wang, and K. Ren, "Enabling efficient user revocation in identity-based cloud storage auditing for shared big

data," IEEE Trans. Depend. Sec. Comput., to be published.

[6] H. Wang, D. He, J. Yu, and Z. Wang, "Incentive and unconditionally anonymous identity-based public provable data possession," IEEE Trans. Serv. Comput., to be published.

[7] Y. Yu et al., "Identity-based remote data integrity checking with perfect data privacy-preserving for cloud storage," IEEE Trans. Inf. Forensics Security, vol. 12, no. 4, pp. 767–778, Apr. 2017.

[8] P.-J. Maenhaut et al., "A Dynamic Tenant-Defined Storage System for Efficient Resource Management in Cloud Applications," Journal of Network and Computer Applications, 2017.

[9] L. Rupprecht et al., "SwiftAnalytics: Optimizing Object Storage for Big Data Analytics," in IEEE IC2E, 2017.

[10] D. Espling et al., "Modeling and Placement of Cloud Services with Internal Structure," IEEE Transactions on Cloud Computing, vol. 4, no. 4, 2014.

[11] M. Bellare, S. Keelveedhi, and T. Ristenpart, "DupLESS: Server aided encryption for deduplicated storage," in Proc. 22nd USENIX Conf. Secur., 2013, pp. 179–194.

[12] J. Li, Y. K. Li, X. F. Chen, P. P. C. Lee, and W. J. Lou, "A hybrid cloud approach for secure authorized deduplication," IEEE Trans. Parallel Distrib. Syst., vol. 26, no. 5, pp. 1206–1216, May 2015, doi:10.1109/TPDS.2014.2318320.

[13] P. Meye, P. Raipin, F. Tronel, and E. Anceaume, "A secure two phase data deduplication scheme," in Proc. HPCC/CSS/ICESS, 2014, pp. 802–809, doi:10.1109/HPCC.2014.134.

[14] Z. Yan, X. Y. Li, M. J. Wang, and A. V. Vasilakos, "Flexible data access control based on trust and reputation in cloud computing," IEEE Trans. Cloud Comput., vol. PP, no. 99, Aug. 2015, doi:10.1109/ TCC.2015.2469662, Art. no. 1.

[15] P. Puzio, R. Molva, M. Onen, and S. Loureiro, "ClouDedup: Secure deduplication with encrypted data for cloud storage," in Proc. IEEE Int. Cof. Cloud Comput. Technol. Sci., 2013, pp. 363–370, doi:10.1109/CloudCom.2013.54.