ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 3, March 2018

Advanced Crypto Standard to Secure the Deduplication of Data in Cloud

R.Lavanya[1], Jayanthi.S, M.E., (Ph.D.).[2],

M.E. CSE, Department of Computer Science and Engineering, Agni College of Technology, Chennai, Tamil Nadu, India.

Assistant Professor, Department of Computer Science and Engineering, Agni College of Technology, Chennai, Tamil Nadu, India

ABSTRACT: The main objective of this system is to develop an experimental model to store the data securely in cloud computing environment and provides high level of trustworthy data maintenance scheme to its users with Advanced Cryptographic Standards (ACS) scheme. More and more clients would like to store their data to PCS (public cloud servers) along with the rapid development of cloud computing. New security problems have to be solved in order to help more clients process their data in public cloud. Security is the main constraint in cloud computing environment, which depicts the nature of the surveillance mechanisms in remote server based data maintenance scheme. Authorization and Authentication schemes provide the data owners and users to feel safe. In this system a novel proof based trustworthy data management scheme is introduced called Advanced Cryptographic Standards (ACS), which allows the data owners to upload the data into the server with proper cryptographic norms. With these systems the data owner and data users can successfully maintain and retrieve the data to and from the server.

KEYWORDS: Storage Security, Cloud Environment, Advanced Cryptographic Standard, ACS Algorithm, Deduplication, Remote Server Maintenance.

I. INTRODUCTION

As the cloud computing technology develops during the last decade, outsourcing data to cloud service for storage becomes an attractive trend, which benefits in sparing efforts on heavy data maintenance and management. Since the outsourced cloud storage is not fully trustworthy, it raises security concerns on how to realize data de-duplication in cloud while achieving integrity auditing. In this system, study the problem of integrity auditing and secure de-duplication on cloud data. Specifically, aiming at achieving both data integrity and de-duplication in cloud in the Cloudme secure system. To avoid redundancy in cloud storage and to achieve integrity auditing and secure de-duplication on cloud data.

Auditing entity with maintenance of a MapReduce cloud, which generate data tags before uploading as well as audit the integrity of data stored in cloud? The Internet has its roots in the 1960s, but it wasn't until the early 1990s that it had any relevance for businesses. The World Wide Web was born in 1991, and in 1993 a web browser called Mosaic was released that allowed users to view web pages that included graphics as well as text. Cloud computing means storing and accessing data and programs over the Internet instead of computer's hard drive. The cloud is just a metaphor for the Internet. It goes back to the days of flowcharts and presentations that would represent the gigantic server-farm infrastructure of the Internet as nothing but a puffy, white cumulonimbus cloud, accepting connections and doling out information as it floats.

DOI: 10.15680/IJIRCCE.2018.0603112

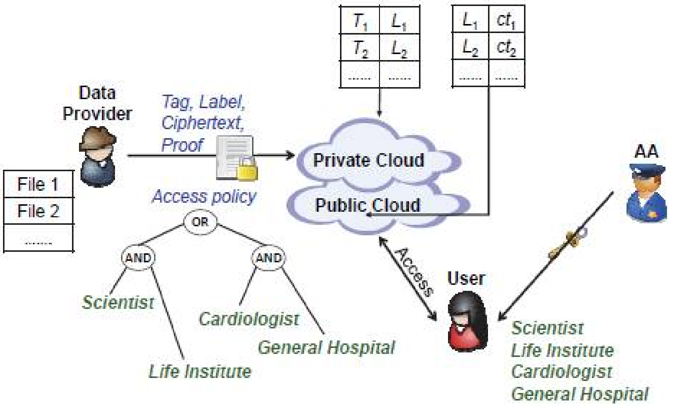


Fig.1 Proposed System Architecture

Cloud computing is a type of computing in which groups of remote servers are linked to allow centralized storage of data and online access to information technology services and resources. When storing data on or running programs from the hard drive, that's called local storage and computing. Accessing the data is fast and easy, for that one computer, or others on the local network. Working off hard drive is how the computer industry functioned for decades. The cloud is also not about having dedicated network attached storage (NAS) hardware or server in residence. Storing data on a home or office network does not count as utilizing the cloud. NAS allows remotely access things over the Internet. The massive data-processing is happening on the other end. The end result is the same, with an online connection, cloud computing can be done anywhere, anytime.

Client Service Provider Port

This port is for CSP where CSP can approve data owner requests and view the in formations of data, but cannot view the contents of the uploaded data. It is encrypted and then stored in the cloud.CSP can also check the transactions down on the cloud.

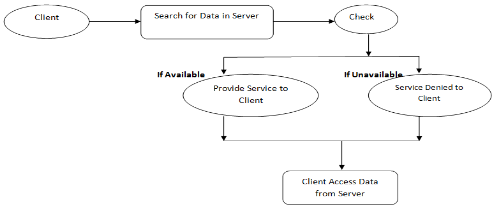

Client
Search for Data in Server
Check
If Available
If Unavailable
Provide Service to Client
Service Denied to Client
Client Access Data from Server

Fig.2 CSP Port

Data Owner

In this Data owner port upload his files to cloud. Before uploading each file is encrypted using ACS algorithm and a key for decrypting that file is created. Each file has its own key through this key only the file can be decrypted. The data owner can also download his files from cloud. For users downloading the data owner files the data owner has to give permission. Data owner can also check the transactions down on the cloud

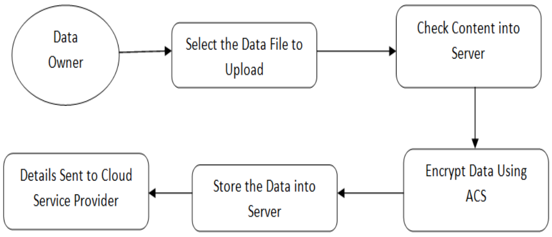

Data Owner
Select the Data File to Upload
Check Content into Server
Encrypt Data Using ACS
Store the Data into Server
Details Sent to Cloud Service Provider

Fig.3 Data Owner

End User

The end user can Search/Download files from the cloud. Before downloading he/she must get permission from the data owner. He can only download files uploaded by the data owners who approved his download permission. The search permission is granted by CSP.

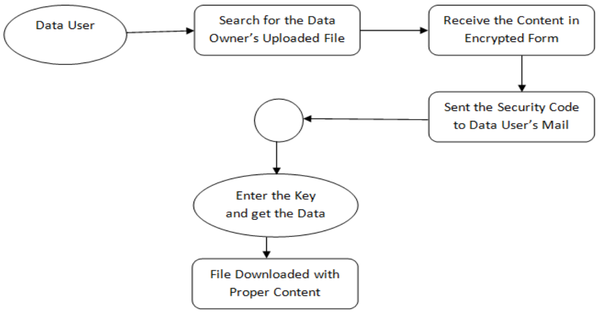

Data User
Search for the Data Owner's Uploaded File
Receive the Content in Encrypted Form
Sent the Security Code to Data User's Mail
Enter the Key and get the Data
File Downloaded with Proper Content

Fig.4 End User

Evaluator

The evaluator assigns roles to the end users. He can also have some view permissions.

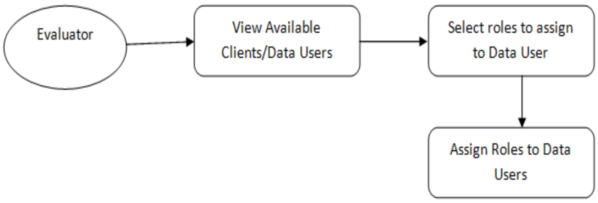

Evaluator
View Available Clients/Data Users
Select roles to assign to Data User
Assign Roles to Data Users

Fig.5 Evaluator Port

II. EXISTING APPROACHES – A SUMMARY

When a user uploads data that already exist in the cloud storage, the user should be deterred from accessing the data that were stored before they obtained the ownership by uploading it (backward secrecy). These dynamic ownership changes may occur very frequently in a practical cloud system, and thus, it should be properly managed in order to avoid the security degradation of the cloud service. In the former approach, most of the existing schemes have been proposed in order to perform a process in an efficient and robust manner, since the hash of the file, which is treated as a “proof” for the entire file, is vulnerable to being leaked to outside adversaries because of its relatively small size. A data owner uploads data that do not already exist in the cloud storage, he is called an initial uploader; if the data already exist, called a subsequent uploader since this implies that other owners may have uploaded the same data previously, he is called a subsequent uploader. The existing system has several disadvantages; some of them are described below: (a) User deduplication on the client-side cannot generate a new tag when they update the file. In this situation, the dynamic Ownerships would fail. As a summary, existing dynamic Ownerships cannot be extended to the multi-user environment, (b) Whenever data is transformed, concerns arise about potential loss of data. By definition, data deduplication systems store data differently from how it was written. As a result, users are concerned with the integrity of their data, (c) One method for deduplicating data relies on the use of cryptographic hash functions to identify duplicate segments of data. If two different pieces of information generate the same hash value, this is known as a collision and (d) The probability of a collision depends upon the hash function used, and although the probabilities are small, they are always non zero.

III. PROPOSED SYSTEM SUMMARY

This Project the goal of saving storage space for cloud storage services also is used for secure deduplication, but several process have been this same concept for deduplication. However this project flow some different modules in there. In this case, if two users upload the same file, the cloud server can discern the equal ciphertexts and store only one copy of them and blocks the other or same uploader to upload the duplicate file. In this process some authentication schemes are available for security purpose. For security establishment in this system, a new security algorithm is implemented called “Advanced Cryptographic Standards (ACS)”. Through this process data owners and users can ensure for secured deduplication policies. An owner wants to outsource their data to the cloud and share it with users possessing certain credentials. The proposed system has several advantages; some of them are described below: (a) Firstly, it can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys and (b) Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion.

IV. LITERATURE SURVEY

In the year of 2013, the authors "IBM T. J. Watson Research Center, 2 IBM Haifa Research Lab, 3Bar Ilan University" proposed a paper titled "Proofs of Ownership in Remote Storage Systems", in that they described such as: we consider a few different formal formulations of the intuitive requirement above and associated protocols that realize them. The first solution is the most stringent in terms of security, but lacks in efficiency (mainly of computation time and memory requirements).

The next solutions two solutions each improve on these aspects at the cost of relaxations on the security guarantees. Our performance measurements indicate that the scheme incurs only a small overhead compared to naive client-side de-duplication. We remark that the new attacks that we consider and our solutions to them are more relevant for file-level de-duplication than for block-level de-duplication.We gave three definitions for security in this setting and three matching protocols, the last of which is very practical. Our streaming protocol allows the designer of the scheme to set a threshold for how “short” a summary can a file has (e.g., 64MBytes in our implementation). This seems suitable for the attack scenarios of common hash functions, malicious software, or accidental leakage that were described in the introduction.

In the year of 2013, the consortium "Advances in Cryptology - EUROCRYPT" proposed a paper titled "Message-Locked Encryption and Secure Deduplication", in that they described such as: It is natural to aim to build Message-Locked Encryption(MLE) from a D-PKE scheme or a CI-H function because the latter primitives already provide privacy on unpredictable messages. However, in attempting to build MLE from these primitives, several problems arise. One is that neither of the base primitives derives the decryption key from the message.

Indeed, in both, keys must be generated upfront and independently of the data. A related problem is that it is not clear how an MLE scheme might decrypt. CI-H functions are not required to be e-client invertible. D-PKE does provide decryption, but it requires the secret key, and it is not clear how this can yield message-based decryption. Numerical analysis and experimental results on Amazon AWS show that our scheme is efficient and scalable.

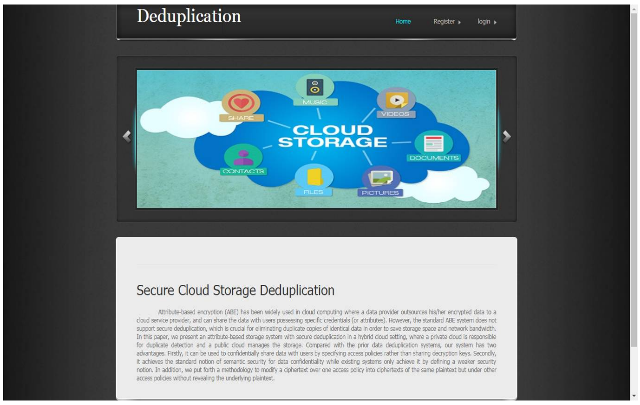
The proposed scheme should accept all valid secret keys and public keys, all valid authentication tags, all valid proof information generated based on valid public keys and all valid data blocks. Our Sample-Extract-Encrypt (SXE) construction builds an MLE scheme for certain classes of block sources where a random subset of the bits of each message remains unpredictable even given the rest of the bits and previous messages. For example, if a message has some subset of uniform bits embedded within it. The scheme then uses a random subset of the message bits as a key, applies an extractor, and then symmetrically encrypts the rest of the message.

In the year of 2017, the consortium "IEEE TRANSACTIONS ON BIG DATA" issued a paper titled "Attribute-Based Storage Supporting Secure Deduplication of Encrypted Data in Cloud", in that they described such as: Attribute-based encryption (ABE) has been widely used in cloud computing where a data provider outsources his/her encrypted data to a cloud service provider, and can share the data with users possessing specific credentials (or attributes).

However, the standard ABE system does not support secure deduplication, which is crucial for eliminating duplicate copies of identical data in order to save storage space and network bandwidth. In this paper, we present an attribute-based storage system with secure deduplication in a hybrid cloud setting, where a private cloud is responsible for duplicate detection and a public cloud manages the storage. Compared with the prior data deduplication systems, our system has two advantages. Firstly, it can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys. Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion. In addition, we put forth a methodology to modify a ciphertext over one access policy into ciphertexts of the same plaintext but under other access policies without revealing the underlying plaintext.

V. EXPERIMENTAL RESULTS

The following figure shows the Home Page of the proposed system.

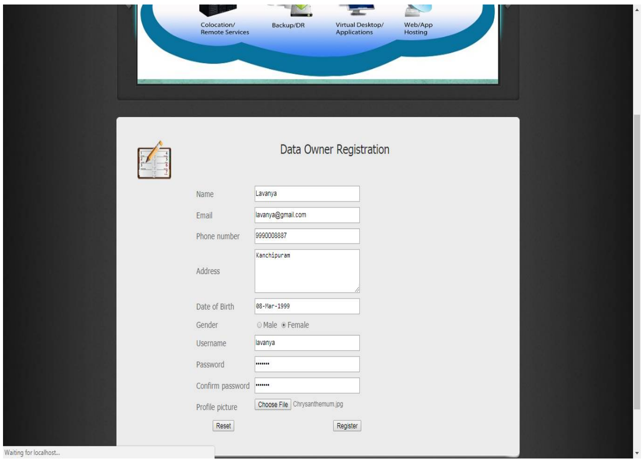


Secure Cloud Storage Deduplication

Attribute-based encryption (ABE) has been widely used in cloud computing where a data provider outsources his/her encrypted data to a cloud service provider, and can share the data with users possessing specific credentials (or attributes). However, the standard ABE system does not support secure deduplication, which is crucial for eliminating duplicate copies of identical data in order to save storage space and network bandwidth. In this paper, we present an attribute-based storage system with secure deduplication in a hybrid cloud setting, where a private cloud is responsible for duplicate detection and a public cloud manages the storage. Compared with the prior data deduplication systems, our system has two advantages. Firstly, it can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys. Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion. In addition, we put forth a methodology to modify a ciphertext over one access policy into ciphertexts of the same plaintext but under other access policies without revealing the underlying plaintext.

Fig.5 Home Page

The following figure illustrates the Registration Page of the Data Owner.

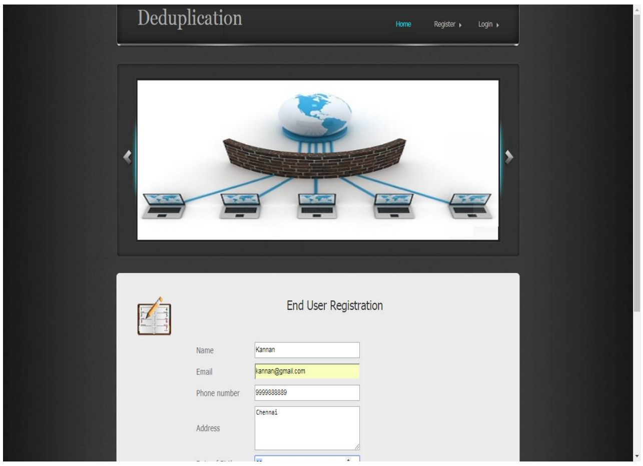


Data Owner Registration

Name	Lavanya
Email	lavanya@gmail.com
Phone number	9990008887
Address	Kanchipuram
Date of Birth	08-Mar-1999
Gender	○ Male ● Female
Username	lavanya
Password	•••••••
Confirm password	•••••••
Profile picture	Choose File Chrysanthemum.jpg

Fig.6 Data Owner Registration

The following figure illustrates the End User Registration Page.



End User Registration

Name: Kannan
Email: kannan@gmail.com
Phone number: 9999888889
Address: Chennai

Fig.7 End User Registration

VI. CONCLUSION

Secured Cloud Maintenance with crypto policies has been widely used in all places, where a data provider outsources his/her encrypted data to a cloud service provider, and can share the data with users possessing specific credentials. However, the standard crypto system does not support secure deduplication, which is crucial for eliminating duplicate copies of identical data in order to save storage space and network bandwidth. In this system, a new scheme called Advanced Cryptographic Standard (ACS) is introduced with secure deduplication in a hybrid cloud setting, where a private cloud is responsible for duplicate detection and a public cloud manages the storage. That can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys. In addition, we put forth a methodology to modify a ciphertext over one access policy into ciphertexts of the same plaintext but under other access policies without revealing the underlying plaintext.

REFERENCES

[1] D. Quick, B. Martini, and K. R. Choo, Cloud Storage Forensics. Syngress Publishing/Elsevier, 2014. http://www.elsevier.com/books/cloud-storageforensics/quick/978-0-12-41970-5

[2] K. R. Choo, J. Domingo-Ferrer, and L. Zhang, “Cloud cryptography: Theory, practice and future research directions,” Future Generation Comp. Syst., 2016.

[3] K. R. Choo, M. Herman, M. Iorga, and B. Martini, “Cloud forensics: State-of-the-art and future directions,” Digital Investigation, vol. 18, pp. 77–78, 2016.

[4] Y. Yang, H. Zhu, H. Lu, J.Weng,Y. Zhang, and K. R. Choo, “Cloud based data sharing with fine-grained proxy re-encryption,” Pervasive and Mobile Computing, vol. 28, pp. 122–134, 2016.

[5] D. Quick and K. R. Choo, “Google drive: Forensic analysis of data remnants,” J. Network and Computer Applications, vol. 40, pp. 179–193, 2014.

[6] A. Sahai and B. Waters, “Fuzzy identity-based encryption,” in Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 2005, Proceedings, ser. Lecture Notes in Computer Science, vol. 3494. Springer, 2005,

[7] B. Zhu, K. Li, and R. H. Patterson, “Avoiding the disk bottleneck in the data domain deduplication file system,” in 6th USENIX Conference on File and Storage Technologies, FAST 2008, February 26-29, 2008, San Jose, CA, USA. USENIX, 2008, pp. 269–282.

[8] 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, 2013.

[9] 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, 2013. Proceedings, Part I, ser. Lecture Notes in Computer Science, vol. 8042. Springer, 2013, pp. 374–391.

[10] 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. USENIX Association, 2013, pp. 179–194.

[11] M. Bellare and S. Keelveedhi, "Interactive message-locked encryption and secure deduplication," in Public-Key Cryptography - PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 - April 1, 2015, Proceedings, ser. Lecture Notes in Computer Science, vol. 9020. Springer, 2015, pp. 516–538.

[12] S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider, "Twin clouds: Secure cloud computing with low latency - (full version)," in Communications and Multimedia Security, 12th IFIP TC 6 / TC 11 International Conference, CMS 2011, Ghent, Belgium, October 19-21,2011. Proceedings, ser. Lecture Notes in Computer Science, vol. 7025. Springer, 2011, pp. 32–44.