Semantic Security For Data Confidentiality Using ABE With Identified Access Policies In Secure De-Duplication Of Encrypted Data In Cloud

Aamena muther[1] and MdAteeq Ur Rahman[2],
[1]Research Scholar, Dept. of Computer Science & Engineering,
SCET, Hyderabad, India
[2]Professor and Head, Dept. of Computer Science & Engineering,
SCET, Hyderabad, India

Abstract - Cloud computing, a convenient means of accessing services, resources and applications over the web, shifts the main target of industries and organizations aloof from the readying and day-to-day running of their IT facilities by providing associate degree on-demand, self-service, and pay-asyou-go business model. It is, therefore, expected that cloud computing has continued to increase in quality in recent times. While cloud computing provides numerous edges to users, there square measure underlying security and privacy risks. as an example, multi-tenancy, resource pooling and shareability options are often exploited by cybercriminals and anyone with a malicious intent, to the harm of each cloud users and cloud service suppliers. Attribute-based coding (ABE) has been wide utilized in cloud computing wherever an information supplier outsources his/herencrypted data to a cloud service supplier, and might share the information with users possessing specific credentials (or attributes). However,the standard ABE system doesn't support secure deduplication, that is crucial for eliminating duplicate copies of identical information so as to avoid wasting space for storing and network information measure. during this paper, we have a tendency to gift AN attribute-based storage system with secure deduplication in an exceedingly hybrid cloud setting, wherever a personal cloud is liable for duplicate detection and a public cloud manages the storage. Compared with the previous information deduplication systems, our system has 2 benefits. Firstly, it are often accustomed confidentially share information with users by specifying access policies instead of sharing cryptography keys. Secondly, it deliver the goodss the quality notion of linguistics security for information confidentiality whereas existing systems solely achieve it by process a weaker security notion. additionally, we have a tendency to place forth a technique to change a ciphertext over one access policy into ciphertexts of constant plaintext however underneath different access policies while not revealing the underlying plaintext.
it's expected, then, that cloud computing has
emerged as a salient space of inquiry for security researchers. as an example, once user knowledge (e.g. documents, videos and photos) square measure uploaded or keep in an exceedingly cloud computing service, the data owners square measure unlikely to grasp the trail of the transmitted knowledge or whether or not the info square measure being collected and analyzed by a 3rd party, as well as a administrative unit (see the revelations by Edward Snowden – European Parliament 2014). As posited by Choo and Sarre (2015), it is important to strike a balance between privacy, legitimate police investigation and lawful knowledge access, so as to confirm that the privacy of innocent people won't be compromised (e.g. that fine-grained aspects of associate degree individual's life can not be derived or inferred from the intelligence assortment and analysis).

Index Terms— ABE, Storage, Deduplication, Data integrity, homomorphic encryption and searchable encryption

I. INTRODUCTION

Cloud computing greatly facilitates knowledge suppliers United Nations agency need to source their knowledge to the cloud while not revealing their sensitive knowledge to external parties and would love

users with sure credentials to be able to access the info. this needs knowledge to be hold on in encrypted forms with access management policies specified nobody except users with attributes (or credentials) of specific forms will decode the encrypted knowledge. associate coding technique that meets this demand is named attribute-based coding (ABE), wherever a user's non-public secret is related to associate attribute set, a message is encrypted underneath associate access policy (or access structure) over a group of attributes, and a user will decode a ciphertext with his/her non-public key if his/her set of attributes satisfies the access policy related to this ciphertext. However, the quality ABE system fails to attain secure deduplication , that may be a technique to avoid wasting cupboard space and network information measure by eliminating redundant copies of the encrypted knowledge hold on within the cloud. On the opposite hand, to the simplest of our data, existing constructions for secure deduplication aren't designed on attribute-based coding.

Nevertheless, since ABE and secure deduplication are wide applied in cloud computing, it might be fascinating to style a cloud storage system possessing each properties. we have a tendency to take into account the subsequent state of affairs within the style of associate degree attribute-based storage system supporting secure deduplication of encrypted knowledge within the cloud, during which the cloud won't store a file quite once even if it should receive multiple copies of a similar file encrypted underneath completely different access policies. an information supplier, Bob, intends to transfer a fileM to the cloud, and shareM with users having sure credentials. so as to try to to thus, Bob encrypts M underneath associate degree access policy A over a collection of attributes, and uploads the corresponding ciphertext to the cloud, such solely users whose sets of attributes satisfying the access policy will rewrite the ciphertext. Later, another knowledge supplier, Alice, uploads a ciphertext for a similar underlying file M however ascribed to a distinct access policy A0. Since the file is uploaded in associate degree encrypted type, the cloud isn't ready to pick out that the plaintext love Alice's ciphertext is that the same as that love Bob's, and can store M doubly. Obviously, such duplicated storage wastes cupboard space and communication information measure. we gift associate degree attribute-based storage system that employs ciphertext-policy attribute-based secret writing (CP-ABE) and supports secure deduplication. Our main contributions will be summarized as follows. _ Firstly, the system is that the 1st that achieves the quality notion of linguistics security for knowledge confidentiality in attribute-based deduplication systems by resorting to the hybrid cloud design . Secondly, we have a tendency to place forth a technique to change a ciphertext over one access policy into ciphertexts of a similar plaintext however underneath the other access policies while not revealing the underlying plaintext. this system may well be of freelance interest additionally to the applying within the projected storage system. _ Thirdly, we have a tendency to propose associate degree approach supported 2 cryptanalytic primitives, as well as a zero-knowledge proof of information and a commitment theme, to realize knowledge consistency within the system.

II. Related Works

Cloud computing, a convenient manner of accessing services, resources and applications over the net, shifts the main focus of industries associate degreed organizations off from the preparation and regular running of their IT facilities by providing an on-demand, self-service, and pay-asyou-go business model. It is, therefore, unsurprising that cloud computing has continued to extend in quality in recent times. whereas cloud computing provides numerous edges to users, there square measure underlying security and privacy risks. for instance, multi-tenancy, resource pooling and shareability options may be exploited by cybercriminals and anyone with a malicious intent, to the impairment of each cloud users and cloud service suppliers. it's unsurprising , then, that cloud computing has emerged as a salient space of inquiry for security researchers.

For example, once user information (e.g. documents, videos and photos) square measure uploaded or hold on in a very cloud computing service, {the information|the info|the information} house owners square measure unlikely to grasp the trail of the transmitted data or whether or not the info square measure being collected and analyzed by a 3rd party, as well as a office (see the revelations by Edward Snowden – European Parliament 2014). As posited by Choo and Sarre (2015), it's vital to strike a balance between privacy, legitimate police investigation and lawful information access, so as to make sure that the privacy of

innocent people won't be compromised (e.g. that fine-grained aspects of associate degree individual's life can not be derived or inferred from the intelligence assortment and analysis). a very promising approach to realize security and privacy during this new computing paradigm is thru cryptography (Qin et al. 2013; Wang et al. 2014). for instance, as noted by rule et al. (2015), to make sure the safety and privacy of user information, specifically against associate degree untrusted cloud service supplier, one might inscribe the info before uploading and storing the info within the cloud. This special issue is devoted to providing each scientists and practitioners with a forum to gift their recent analysis on the employment of novel cryptanalytic techniques to boost the safety and privacy of the underlying cloud design or scheme, significantly analysis that integrates each theory and follow. for instance, however will we style associate degree economical cloud cryptography system that gives increased security and/or privacy while not compromising on usability and performance? within the sequel, we tend to in brief survey the content of papers during this special issue.

Cloud log forensics (CLF) mitigates the investigation method by distinguishing the malicious behavior of attackers through profound cloud log analysis. However, the accessibility attributes of cloud logs obstructs accomplishment of the goal to analyze cloud logs for numerous susceptibilities. Accessibility involves the problems of cloud log access, choice of correct cloud log file, cloud log information integrity, and trait of cloud logs. Therefore, rhetorical investigators of cloud log files area unit hooked in to cloud service suppliers (CSPs) to induce access of various cloud logs. Accessing cloud logs from outside the cloud while not betting on the CSP may be a difficult analysis area; whereas, the rise in cloud attacks has inflated the necessity for CLF to analyze the malicious activities of attackers. This paper reviews the progressive of CLF and highlights totally different challenges and problems concerned in work cloud log information. The work mode, the importance of CLF, and cloud log-as-a-service area unit introduced. Moreover, case studies associated with CLF area unit explained to focus on the sensible implementation of cloud log investigation for analyzing malicious behaviors. The CLF security necessities, vulnerability points, and challenges area unit known to tolerate totally different cloud log susceptibilities. we have a tendency to determine and introduce challenges and future directions to focus on open analysis areas of CLF for motivating investigators, academicians, and researchers to analyze on them.

Any event occurring in a corporation IT system or network is recorded with numerous entries during a log file. the method of recording log files is understood as work [Chuvakin et al. 2013]. The log file provides helpful info concerning previous events occurring within the system and network throughout a specified time span. as an example, a network administrator will conclude concerning the network information measure usage during a measure by analyzing the network logs. Similarly, application developers use application logs to spot and fix bugs within a program code.

2.1 Existing System

In a typical storage system with secure deduplication, to store a move into the cloud, an information supplier generates a tag and a ciphertext. the info supplier uploads the tag and also the ciphertext to the cloud. Upon receiving associate outsourcing request from an information supplier for uploading a ciphertext associated an associated tag, the cloud runs a questionable equality checking algorithmic program, that checks if the tag within the incoming request is a twin of any tags within the storage system. If there's a match, then the underlying plaintext of this incoming ciphertext has already been keep and also the new ciphertext is discarded. it's apparent that such a system with a tag appended to the ciphertext doesn't offer the quality notion of linguistics security for information confidentiality.

However, endowing such a tag checking ability to theprivate cloud isn't sufficient to attain deduplication in theattribute-based storage system that employs CP-ABE fordata cryptography. within the projected attributed-based system,the same file might be encrypted to totally different ciphertextsassociated with different access policies, storing solely oneciphertext of the file means users whose attributessatisfy the access policy of a discarded ciphertext (but notthat of the keep ciphertext) are denied to access thedata that they're entitled to.

2.2 Disadvantages:

Existing constructions for secure deduplication aren't engineered on attribute-based cryptography.

If the plaintexts is predicated from theirtags, associate human will forever create an accurate guess bycomputing the tag of a plaintext so testing it againstthe tag within the challenge introduce the linguistics security game.

III. PROPOSED SYSTEM

In this paper, we tend to gift associate degree attribute-based storage systemwhich employs ciphertext-policy attribute-based encryption(CP-ABE) and supports secure deduplication. Our maincontributions may be summarized as follows.

Firstly, the system is that the 1st that achieves the standardnotion of linguistics security for information confidentialityin attribute-based deduplication systems byresorting to the hybrid cloud design .

Secondly, we tend to place forth a strategy to modifya ciphertext over one access policy into ciphertextsof a similar plaintext however underneath the other accesspolicies while not revealing the underlying plaintext.This technique could be of freelance interest inaddition to the applying within the projected storagesystem.

3.1 advantages:

we conferred a completely unique approach to understand associate degree attribute-based storage system supporting secure deduplication.

Our storage system is constructed underneath a hybrid cloud design, wherever a non-public cloud manipulates the computation and a public cloud manages the storage.

it accomplishs the quality notion of linguistics security whereas existing deduplication schemes solely achieve it underneath a weaker security notion

IV. System Architecture

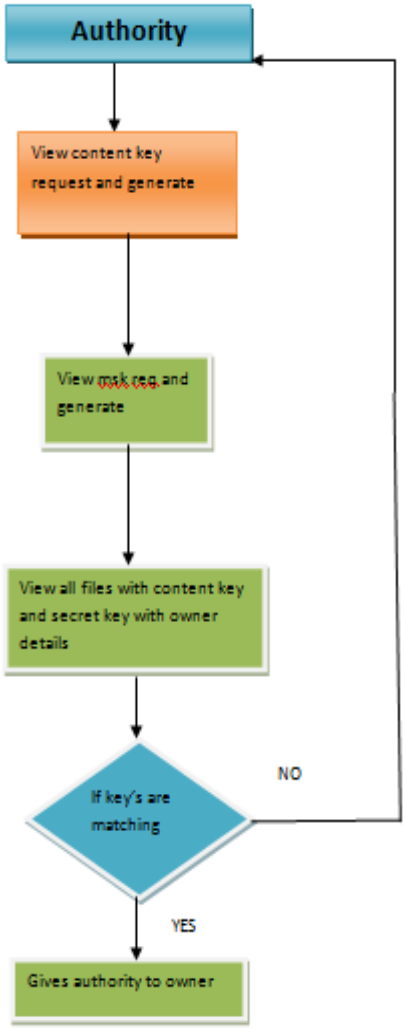

Figure 1: System Architecture[authority] of the Proposed System

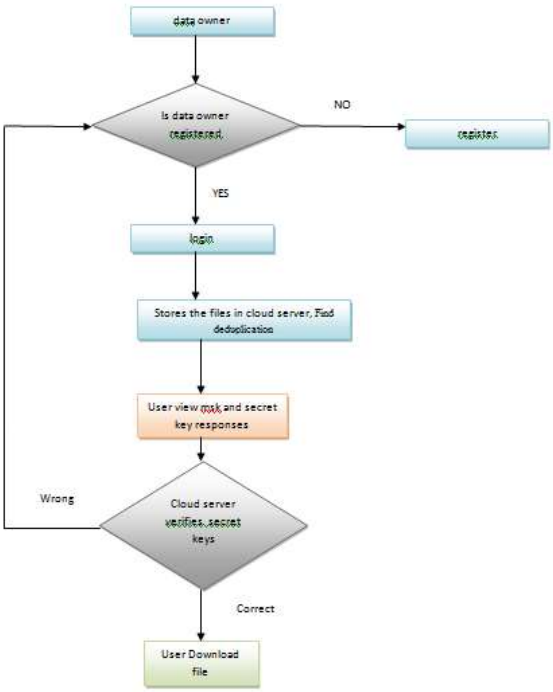

Figure 2: System Architecture[Data Owner] of the Proposed System

3.1 Module Description:

In this project, we have three modules.

- DATA OWNER:
- CLOUD SERVER
- AUTHORITY
- END USER

DATA OWNER:

In this module, initially the data owner has to register to the cloud server and get authorized. After the authorization from cloud data owner will encrypt and add file to the cloud server where in after the addition of file data owner requests the content key and the master secret key to the authority for the file he uploaded and finds Find deduplication ,only after the keys generated the file is uploaded to the cloud server. After the uploading of the file the data owner will have to provide download and the search permission for individual file for the users to perform search and download.

CLOUD SERVER

The cloud server manages a cloud to provide data storage service. Data owners encrypt their data files and store them in the cloud for sharing with cloud End users. To access the shared data files users will request the permission of content key and the MSK master secret key. And the cloud will provide the permission .and also views all the transactions and attackers related to the files.

AUTHORITY

Authority generates the content key and the secret key requested by the end user.
Authority can view all files with the content key and master secret key generated with the corresponding data owner details of the particular file.

END USER

User has to register and login for accessing the files in the cloud. User is authorized by the cloud to verify the registration. User has to request for the MSK master secret key and content key to download the file. User can only download and serach the file if the data owner of the particular file has provided the permissions.

V. Conclusion

In this paper, we tend to propose a novel
Attribute-based coding (ABE) has been wide used in cloud computing wherever information suppliers source their encrypted information to the cloud and may share the information with users possessing specified c redentials. On t he o ther hand, deduplication is a very important technique to save lots of the storage space and network information measure, that eliminates duplicate copies of identical information. However, the quality ABE systems do not support secure deduplication, that makes them costly to be applied in some business storage services. In this paper, we have a tendency to given a completely unique approach to understand associate degree attribute-based storage system supporting secure deduplication. Our storage system is constructed underneath a hybrid cloudarchitecture, wherever a personal cloud manipulates the computation and a public cloud manages the storage. The non-public cloud is given a trapdoor key related to the corresponding ciphertext, with that it will transfer the ciphertext over one access

policy into ciphertexts of constant plaintext underneath the other access policies while not being aware of the underlying plaintext. once receiving a storage request, the non-public cloud initial c hecks t he v alidity o f the uploaded item through the connected proof. If the proof is valid, the non-public cloud runs a tag matching algorithmic program to see whether or not constant information underlying the ciphertext has been keep. If so, whenever it's necessary, it regenerates the ciphertext into a ciphertext of constant plaintext over associate degree access policy that is that the union set of each access policies.

The projected storage system enjoys 2 major benefits. Firstly, it are often accustomed confidentially share information with alternative users by specifying associate degree access policy instead of sharing the decryption key. Secondly, it achieves the quality notion of linguistics security whereas existing deduplication schemes only come through it underneath a weaker security notion.

References

[1] D. Quick, B. Martini, and K. R. Choo, Cloud Storage Forensics. Syngress Publishing / Elsevier, 2014. [Online]. Available: http://www.elsevier.com/books/cloud-storageforensics/quick/978-0-12-419970-5

[2] K. R. Choo, J. Domingo-Ferrer, and L. Zhang, "Cloud cryptography: Theory, practice and future research directions," Future Generation Comp. Syst., vol. 62, pp. 51–53, 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 22-26, 2005, Proceedings, ser. Lecture Notes in Computer Science, vol. 3494. Springer, 2005, pp. 457–473.

[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 26-30, 2013. Proceedings, ser. Lecture Notes in Computer Science, vol. 7881. Springer, 2013, pp. 296–312.

[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 18-22, 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: Serveraided 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. N¨ urnberger, 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.