ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

Removal of Duplicate Storage of Encrypted Data in Cloud Computing Environment

[1] M.Maharasi, [2]S.Keerthiga, [3]P.Kiruthika, [4]R.Nivetha, [5]S.Priya
[1][2][3][4][5] Department of Computer Science and Engineering, V.S.B Engineering College,Karur,Tamilnadu,INDIA

***Abstract* – Duplication removal is the important aspect of data storage in cloud. CP-ABE ciphertext-policy attribute-based encryption (CP-ABE) scheme supports secure 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. Existing ABE encryption system only used for secure encryption and does not support secure deduplication, duplication removal from identical data is important in order to save storage space and network bandwidth. Proposed system achieves confidentiality in data sharing and also achieves standard notion of semantic security for data confidentiality. In addition, we put forth a methodology to modify a ciphertext over one access policy into ciphertext of the same plaintext but under other access policies without revealing the underlying plaintext. In proposed work of this project implement using AES with blowfish algorithm is used to encrypt the files before storing it on the cloud. Blowfish has high security when compared to CP-ABE. The combination of enhanced AES and blowfish is used for data confidentiality. Performance of the proposed system is evaluated based on time taken for encryption/decryption, throughput and memory usage for different data formats like text file, image file, audio file and video file. And also key sharing process is to modified in proposed system. A one valid key is sharing to the users to access the data.**

INTRODUCTION

Cloud computing greatly facilitates data providers who want to outsource their data to the cloud without disclosing their sensitive data to external parties and would like users with certain credentials to be able to access the data This requires data to be stored in encrypted forms with access control policies such that no one except users with attributes (or credentials) of specific forms can decrypt the encrypted data. Attribute-Based Encryption (ABE) scheme, where a user's private key is associated with an attribute set, a message is encrypted under an access policy (or access structure) over a set of attributes, and a user can decrypt a ciphertext with his/her private key if his/her set of attributes satisfies the access policy associated with this ciphertext. ABE and secure deduplication have been widely applied in cloud computing, it would be desirable to design a cloud storage system possessing both properties. In a typical storage system with secure deduplication, to store a file in the cloud, a data provider generates a tag and a ciphertext. The data provider uploads the tag and the ciphertext to the cloud. Upon receiving an outsourcing request from a data provider for uploading a ciphertext and an associated tag, the cloud runs a so-called equality checking algorithm, which checks if the tag in the incoming request is identical to any tags in the storage system. If there is a match, then the underlying plaintext of this incoming ciphertext has already been stored and the new ciphertext is discarded. In this system, a tag appended to the ciphertext does not provide the standard notion of semantic security for data confidentiality, because if the plaintexts can be predicated from their tags, an adversary can always make a correct guess by computing the tag of a plaintext and then testing it against the tag in the challenge phase in the semantic security game. To circumvent this obstacle, we bring in our system a hybrid cloud architecture, which consists of a private cloud responsible for tag checking and ciphertext regeneration (to be introduced later) and a public cloud storing the ciphertext. Thanks to this architecture, we manage to achieve semantic security with respect to the public cloud, whilst in terms of the private cloud, a weaker security notion called privacy under chosen distribution attacks (PRV-CDA security) is accomplished under the assumption that the message space is sufficiently large such that each message to be uploaded to the cloud is unpredictable.

RELATED WORK

1.daptable Ciphertext-Policy Attribute-Based Encryption
In this paper, we introduce a new cryptographic primitive, called adaptable ciphertext-policy attribute-based encryption (CP-ABE). Adaptable CP-ABE extends the traditional CP-ABE by allowing a semitrusted proxy to modify a ciphertext under one access policy into ciphertexts of the same plaintext under any other access policies; the proxy, however, learns nothing about the underlying plaintext. With such "adaptability" possessed

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

by the proxy, adaptable CP-ABE has many real world applications, such as handling policy changes in CP-ABE encryption of cloud data and outsourcing of CP-ABE encryption. Specifically, we first specify a formal model of adaptable CP-ABE; then, based on the CP-ABE scheme by Waters, we propose a concrete adaptable CP-ABE scheme and further prove its security under our security model.

2. Unbounded HIBE and Attribute-Based Encryption

In this work, we present HIBE and ABE schemes which are unbounded" in the sense that the public parameters do not impose additional limitations on the functionality of the systems. In all previous constructions of HIBE in the standard model, a maximum hierarchy depth had to be fixed at setup. In all previous constructions of ABE in the standard model, either a small universe size or a bound on the size of attribute sets had to be fixed at setup. Our constructions avoid these limitations. We use a nested dual system encryption argument to prove full security for our HIBE scheme and selective security for our ABE scheme, both in the standard model and relying on static assumptions. Our ABE scheme supports LSSS matrices as access structures and also provides delegation capabilities to users.

3. DupLESS: Server-Aided Encryption for Deduplicated Storage

Cloud storage service providers such as Dropbox, Mozy, and others perform deduplication to save space by only storing one copy of each file uploaded. Should clients conventionally encrypt their files, however, savings are lost. Message-locked encryption (the most prominent manifestation of which is convergent encryption) resolves this tension. However it is inherently subject to brute-force attacks that can recover files falling into a known set. We propose an architecture that provides secure deduplicated storage resisting brute-force attacks, and realize it in a system called DupLESS. In DupLESS, clients encrypt under message-based keys obtained from a key-server via an oblivious PRF protocol. It enables clients to store encrypted data with an existing service, have the service perform deduplication on their behalf, and yet achieves strong confidentiality guarantees. We show that encryption for deduplicated storage can achieve performance and space savings close to that of using the storage service with plaintext data.

4. New Constructions and Proof Methods for Large Universe Attribute-Based Encryption

Traditionally, public key encryption schemes provided any user with the ability to share data with another specific user in a private manner. However, in many applications we would like to have the additional capability to encrypt data for a set of users according to a specific policy on their credentials. For example, one might want to store data in a public server such that only parties with credentials of specific forms are able to decrypt. This encryption notion, called Attribute-Based Encryption (ABE), was introduced by Waters. In this setting, each user possesses a set of attributes/credentials and a secret key that corresponds to these credentials. The encrypting party can define any Boolean formula on the possible attributes and a user can decrypt if and only if his attribute set satisfies the Boolean formula. Several Attribute-Based constructions have been presented since then (see related work below). A common classification property is whether a system is a \small universe" or \large universe" constructions. In \small universe" constructions the size of the attribute space is polynomially bounded in the security parameter and the attributes were fixed at setup. Moreover, the size of the public parameters grew linearly with the number of attributes.

II. EXISTING SYSTEM

In a typical storage system with secure deduplication, to store a file in the cloud, a data provider generates a tag and a ciphertext. The data provider uploads the tag and the ciphertext to the cloud. Upon receiving an outsourcing request from a data provider for uploading a ciphertext and an associated tag, the cloud runs a so-called equality checking algorithm, which checks if the tag in the incoming request is identical to any tags in the storage system. If there is a match, then the underlying plaintext of this incoming ciphertext has already been stored and the new ciphertext is discarded. It is apparent that such a system with a tag appended to the ciphertext does not provide the standard notion of semantic security for data confidentiality, because if the plaintexts can be predicated from their tags, an adversary can always make a correct guess by computing the tag of a plaintext and then testing it against the tag in the challenge phase in the semantic security game. To circumvent this obstacle, we bring in our system a hybrid cloud architecture, which consists of a private cloud responsible for tag checking and ciphertext regeneration (to be introduced later) and a public cloud storing the ciphertexts. In this proposed architecture, we achieve security with respect to the public cloud. A weaker security method called privacy

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

under chosen distribution attacks (PRV-CDA security) is accomplished under the assumption that the message space is sufficiently large such that each message to be uploaded to the cloud is unpredictable. However, proposing such a tag checking ability for private cloud and CP-ABE for data encryption is not sufficient to achieve deduplication in the attribute-based storage system. In the proposed attributed-based system, the same file could be encrypted to different ciphertexts associated with different access policies, storing only one ciphertext of the file means that users whose attributes satisfy the access policy of a discarded ciphertext (but not that of the stored ciphertext) will be denied to access the data that they are entitled to. Another key challenge in secure deduplication is to make it secure against duplicate faking attacks in which a legally generated message is unnoticeably replaced by a fake one. In such an attack, a malicious user may intercept an outsourcing request and tamper with the ciphertext, and then sending the modified ciphertext but the original tag to the cloud. Later, an honest data provider wants to upload a ciphertext for an identical file.

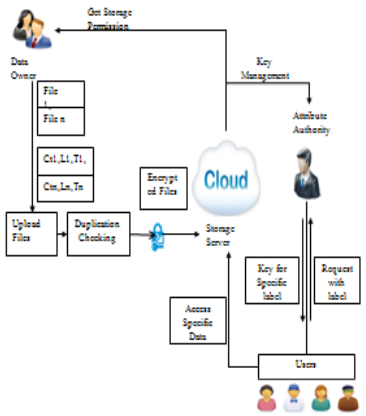
Disadvantages

•It violates data integrity, when a user downloads the ciphertext.
•Large number of tags making this system difficult to retrieve the data.
•CP-ABE does not support deduplication process.

III. PROPOSED SYSTEM

The proposed architecture of our attribute-based storage system with secure deduplication is shown in the figure. In which four entities are involved: data providers, attribute authority (AA), cloud and users. A data provider wants to outsource his/her data to the cloud and share it with users possessing certain credentials. The AA issues every user a decryption key associated with his/her set of attributes. The cloud consists of a public cloud which is in charge of data storage and a private cloud which performs certain computation such as tag checking. When sending a file storage request, each data provider firstly creates a tag T and a label L associated with the data, and then encrypt the data under an access structure over a set of attributes. Also, each data provider generates a proof pf on the relationship of the tag T, the label L and the encrypted message ct3, but this proof will not be stored anywhere in the cloud and is only used during the checking phase for any newly generated storage request. After receiving a storage request, the private cloud first checks the validity of the proof pf, and then tests the equality of the new tag T with existing tags in the system. If there is no match for this new tag T, the private cloud adds the tag T and the label L to a tag-label list, and forwards the label and the encrypted data, (L, ct) to the public cloud for storage. At the user side, all users are allowed to download the file, and decrypt the ciphertext with the attribute-based private key generated by the AA if this user's attribute set satisfies the access structure. Each user checks the correctness of the decrypted message using the label, and accepts the message if it is consistent with the label. Proposed system of attribute-based storage with deduplication achieves the security for data confidentiality in systems by resorting to the hybrid cloud architecture.

System Architecture



Algorithm

ABFE Algorithm for Encryption

AES with Blowfish is a symmetric key encryption algorithm. It is symmetric because the same key is used for both encryption and decryption; the key has to be kept secret from all others except the sender and receiver. Blowfish is able to create a much longer key so that it is much more difficult to try to hack the key value.

Algorithm for Duplication Checking

Cyclic Redundancy Check-32

Cyclic redundancy checking is defined as a methodology that used for checking errors occurring in the data transmission between sender and receiver over a communications medium. A device of sender applies a 16- or 32-bit polynomial to a block of data that's to be

transmitted and appends the ensuing cyclic redundancy code (CRC) to the block. The receiving end additionally applies identical polynomial to the information that employed in sender aspect and compares its result with the result appended by the sender. If they agree, the information has been received with success. If not, the sender may be notified to resend the block of data.

IV.CONCLUSION

Attribute-based Encryption (ABE) has been wide employed in cloud computing for encrypting information that data owner outsource their encrypted data to the cloud and might share the information with users possessing given credentials. On the opposite hand, deduplication is a very important technique to save lots of the space for storing and network information measure, which eliminates duplicate copies of identical information. However, the quality ABE systems don't support secure deduplication that makes them expensive to be applied in some business storage services. In this paper, we tend to given a unique approach to understand associate attribute-based storage system supporting secure deduplication. Blowfish used with ABE provide high security compared with existing security schemes.

V.FUTURE WORK

In future the secure cloud storage architecture will enhanced to secure multimedia files like image, audio, video and graphical format files.

VI. REFERENCES

[1] D. Quick, B. Martini, and K. R. Choo, Cloud Storage Forensics. Syngress Publishing / Elsevier, 2014. [Online].

[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, Aarhus, Denmark, May 22-26, 2005, 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," 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.

[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.

[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.

[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.

[13] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge complexity of interactive proof-systems (extended abstract)," in Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA.

[14] M. Fischlin and R. Fischlin, "Efficient non-malleable commitment schemes," in Advances in Cryptology - CRYPTO 2000.

[15] S. Goldwasser and S. Micali, "Probabilistic encryption," J. Comput. Syst. Sci., vol. 28, no. 2, pp. 270–299, 1984.

[16] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in

Proceedings of the 13th ACM Conference on Computer and Communications Security, CCS 2006, Alexandria, VA, USA, Ioctober 30 - November 3, 2006, ser. Lecture Notes in Computer Science, vol. 5126. Springer, 2006, pp. 89–98.

[17] R. Ostrovsky, A. Sahai, and B.Waters, "Attribute-based encryption with non-monotonic access structures," in Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, Alexandria, Virginia, USA, October 28-31, 2007. ACM, 2007, pp. 195–203.

[18] A. B. Lewko and B. Waters, "Unbounded HIBE and attribute based encryption," in Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings, ser. Lecture Notes in Computer Science, vol. 6632. Springer, 2011, pp. 547–567.

[19] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in 2007 IEEE Symposium on Security and Privacy (S&P 2007), 20-23 May 2007, Oakland, California, USA. IEEE Computer Society, 2007, pp. 321–334.

[20] L. Cheung and C. C. Newport, "Provably secure ciphertext policy ABE," in Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, Alexandria, Virginia, USA, October 28-31, 2007. ACM, 2007, pp. 456–465.

[21] V. Goyal, A. Jain, O. Pandey, and A. Sahai, "Bounded ciphertext policy attribute based encryption," in Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part II - Track B: Logic, Semantics, and Theory of Programming & Track C: Security and Cryptography Foundations, ser. Lecture Notes in Computer Science, vol. 5126. Springer, 2008, pp. 579–591.

[22] Y. Rouselakis and B. Waters, "Practical constructions and new proof methods for large universe attribute-based encryption," in 2013 ACM SIGSAC Conference on Computer and Communications Security, CCS'13, Berlin, Germany, November 4-8, 2013. ACM, 2013, pp. 463–474.

[23] J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system," in ICDCS, 2002, pp. 617–624.

[24] M. W. Storer, K. M. Greenan, D. D. E. Long, and E. L. Miller, "Secure data deduplication," in Proceedings of the 2008 ACM Workshop On Storage Security And Survivability, StorageSS 2008, Alexandria, VA, USA, October 31, 2008. ACM, 2008, pp. 1–10.

[25] P. Anderson and L. Zhang, "Fast and secure laptop backups with encrypted de-duplication," in Uncovering the Secrets of System Administration: Proceedings of the 24th Large Installation System Administration Conference, LISA 2010, San Jose, CA, USA, November 7-12, 2010. USENIX Association, 2010.