A Survey of Data Security and Privacy Protection in Cloud Computing

Zhongbing Yuan[1, a] and Xiaojun Liu[1, b *]

[1]School of Electronic & Information, Huanggang Normal University, Hubei Huanggang, China

[a]18623582@qq.com, [b]whutliuxiaojun@126.com

*The corresponding author

Keywords: Data security; Privacy protection; Fully homomorphic encryption; Cloud computing

Abstract. This paper introduces the data security and privacy requirements of cloud computing, including the research background, solutions to problems and research status. Emphatically introduces the Fully Homomorphic Encryption (FHE) in the cloud computing research progress in the field of data security. FHE mature technology will greatly promote the development of networking and cloud computing.

Introduction

In the cloud computing model, the storage and computation of user data are processed by the cloud service provider users lost the data of physical security protection, such as cloud computing data privacy is more dependent on cloud computing service provider. Therefore, in order to development of cloud computing technology and service platform, allowing users to safely will their data to the cloud service provider management, it is necessary to fully understand and solve the problem of data security and user privacy protection faced by cloud computing. This paper will elaborate on the security issues, including the research background, problem solving scheme and research status.

Security Requirements in Cloud Computing

Security in the cloud mainly includes: (1) data confidentiality: ensure that authorized users can use; (2) data integrity: any unauthorized users cannot modify, delete data in the cloud server to the user that the encrypted data hold is complete, also called provable data possession (3); availability: to ensure that the legitimate user at any time, any place can get the necessary services, such as encrypted data can not affect the data security and privacy case can be retrieved; (4) authentication and access control: the validity of the user's judgment, to prevent unauthorized users from accessing; (5) controllability and audit: especially multi-tenant environment data security, at the same time provide audit event list all the evidence and evidence, and the evidence will not be disclosed to other users' information. In practical application, the problem Is not isolated, but interdependent, the lack of any part of its security cannot be guaranteed.

Data Security. With the rapid development of computer and Internet, more and more people and enterprises to choose their own data and a large number of business into the cloud for processing, in order to reduce the hardware cost and maintenance cost. The local system in the private cloud, the cloud data storage device by the user self-management, data under the control of a user is authorized only trusted users access to ensure the security in the public cloud environment, because the outer characteristics of cloud physical device user data stored in the cloud service providers, users of data lost physical security protection. It is possible for users and attackers to share a storage device or physical host, the attacker through the sharing of equipment resources for side channel attack to obtain user data and privacy information. If the data security cannot be guaranteed, especially the user enterprise and government is not willing to own data Information is transferred to the cloud, which has hindered the rapid development of cloud computing. Therefore, how to ensure the security of user data and privacy information is the most important issue of cloud computing.

In order to ensure the confidentiality of cloud data, a basic idea is the first user data encryption, and then calculated the transmission of encrypted data to the cloud server, cloud service providers and non-authorized users can only access data ciphertext, and do not know the original data. If there are other users want to access the original data. Need to have a private key or certificate of authorization decryption. But this simple encryption storage can only guarantee the confidentiality of data, and data integrity is generally guaranteed by the hash function based signature. However, in the cloud computing environment should not only ensure the confidentiality and integrity of data storage, but also must meet the requirements of data sharing, effective search and availability service features. So the simple encryption and signature system cannot meet the requirements of data stored on the cloud. Here we will detail how to protect In the case of security data storage, data search or other operations can be carried out.

Search for Data. As mentioned above, the data encryption is to protect the cloud computing environment security of user data, the most basic method. However, under normal circumstances the encrypted data stored in the cloud is difficult to realize the documents fast retrieval and search. Users of cloud to the operation of the data, you need to first download solution close to the operation, which is contrary to the cloud computing service tenet. Ciphertext retrieval technology is an important technology to realize sharing of privacy and data security. This technique usually requires that the data owner before the data transmission of the ciphertext to the cloud server, first extract the data of keywords and encryption, the encrypted keywords and encryption of data as the ciphertext transmission to the cloud server.

At present, the ciphertext retrieval method is mainly divided into two types: Symmetric Searchable Encryption and Asymmetric Searchable Encryption. symmetric encryption retrieval thought was first used by Song et al. Proposed by [1], is mainly used for encryption of data retrieval, can also be used to achieve symmetric encryption. Then search keywords can be many achievements [2-5] in terms of safety and efficacy have been significantly improved. The advantages of symmetric encryption scheme retrieval is strong security, because the encryption algorithm is a symmetric encryption algorithm (grouping algorithm and pseudo random number generating function), encryption and search efficiency is very high, but the disadvantage is less function. For example, Curtmola et al [2] retrieval scheme the method is optimal, but the data index update is very complicated. And the scheme of [3] Goh index update quickly and effectively, but the server search data is relatively slow. In addition, an important issue, the current has not been able to simultaneously search for several key words, including continuous and non-continuous situation.

Another search retrieval method is asymmetric encryption algorithm, suitable for different users can access the data. The main advantages of strong function, but because of the general operation of the elliptic curve of assignment, compared to the hash function and block cipher operation efficiency is not high, originally proposed by. Boneh public key encryption key based on the [6] algorithm, then, Abdalla et al. Anonymous encryption scheme improved [7] scheme based on Boneh, and gives the definition of security. Non symmetric encryption algorithm can achieve continuous retrieval keyword retrieval and retrieval. The interval query ciphertext interval, mainly through the Order-Preserving Encryption algorithm to realize, is put forward by Agrawal et al [8]. on the one hand this encryption algorithm will reveal the plaintext ciphertext sequence information, can not guarantee complete security expressly, on the other hand, this Encryption algorithm can well achieve interval search. The algorithm has the drawbacks of is: (1) only support exact match, for small input errors and format is not consistent with the lack of robustness; II does not support returns ranked results; (3) does not support multi keyword search.

To solve the above three problems, there are many improved algorithms have been proposed by [9-13].Li et al [9] proposed a fuzzy search scheme to solve the data encryption, only support exact match search the drawbacks of. Wang et al [10] proposed the support search results ranking search encrypted data is defined, and the structure of a support return result encrypted data sorted search scheme.Cao et al [11] proposed search and sort the returned results of a multi keyword support encryption scheme of.Li et al [12] proposed a search keyword search authorization scheme, through the improvement and use of hierarchical predicate encryption, the multidimensional search keywords, so as to protect user privacy and data indexing and query privacy purposes. So far at the same time, can realize the fuzzy search, support the return result sorting, data encryption scheme supporting multi keyword search is not text.

Fully Homomorphic Ecryption (FHE) and the Security of Cloud Computing

The user data is usually in the cloud computing is not only the search operation, also need statistical analysis and other more complex operation, and the emergence of homomorphic encryption technology to solve this problem. FHE is an encryption mode, allowing people to get specific algebraic operations of the ciphertext, and perform the same operation the results of the plaintext encryption It is because of homomorphic characteristics of FHE, fundamentally solves the data and operation outsourcing to third party security issues, to ensure data security and privacy information, provides a new opportunity for the development of cloud computing. In the cloud computing environment, users use the homomorphic encryption technology will store data encryption, then the ciphertext data to the cloud server, other users can directly access and manipulate the ciphertext data, number of users can not get the original According to. Subsequently, the user can from the server gets encrypted data processing results and decrypt, has been operating the plaintext data. Similarly, using FHE methods for encrypting user privacy information and storage to the cloud server, other users can the encrypted user privacy information such as retrieval, statistical analysis, such as handling, and can not be informed of the original privacy information.

FHE also called Privacy Encryption, Rivest, Adleman and Dertouzos [13]. proposed the earliest homomorphic encryption scheme in 1978 is the factorization of the difficulty of RSA system based on [14] 1977 is proposed, at any time to meet the homomorphic multiplication characteristics, proposed by ElGamal[15] in 1984 based on the discrete logarithm problem of the cryptosystem is also satisfying multiplicative homomorphic characteristics, probability Goldwasser and Micali encryption scheme [16] belongs to addition modulo 2 homomorphism. In addition, there are a lot of homomorphic encryption system to meet some kind of operation in [17-19]. 2005 Boneh, Goh and Nissim proposed the BGN encryption system [20], support arbitrary additive homomorphism and a multiplicative homomorphism this is the problem of homomorphism, since the past 30 years, the most close to a result of FHE.

For decades, scientists have been exploring research on password homomorphic encryption, but do not always find FHE, to achieve a perfect solution until 2009, Gentry announced the first design scheme of [21]. encryption scheme is based on the fully homomorphic ideal lattice, its security is based on the bounded distance problem and sparse encoding subset of the problem and ideal lattices. Subsequently, many of the specific scheme of FHE have been proposed, including the ideal lattice Gentry scheme based on [22-24], DGHV and integer optimization scheme based on [25-27], proposed by Brakerski et al [28-29]. LWE problem which fully homomorphic scheme proposed by Dijk et al DGHV encryption scheme based on the basic operation modular integer arithmetic, to achieve efficiency compared to ideal lattice scheme of Gentry is more effective, its security is the approximate GCD problem based on synchronization.

Conclusion

FHE for cloud computing data security and privacy protection provides a key technology, the making of the encrypted data stored in the cloud server for complex arithmetic operations and possible, in order to ensure the safety and greatly reduces the computation overhead between the user and the cloud server. The FHE algorithm complex degree is large, difficult in the existing computing fast implementation. How to improve FHE scheme for encryption and decryption efficiency, reduce storage space of the key, cryptography research hotspot and difficulty. If these problems can be solved well, greatly promote the popularization and application of cloud computing and networking.

Acknowledgment

The paper was supported by the project of excellent ICT engineer training program in Huanggang Normal University (Grant No. 2014zy04); and the project of zxfz2016A014 in Huanggang Normal University

References

[1] Song D, Wagner D, Perrig A. Practical techniques for searches on encrypted data[C]// Proceedings of the IEEE Symposium on security and privacy(S&P’00). Piscataway, NJ, USA: IEEE, 2000: 44-55.

[2] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky. Searchable symmetric encryption: Improved definitions and efficient constructions. In A. Juels, R. Wright, and S. De Capitanidi Vimercati, editors, ACM Conference on Computer and Communications Security (CCS'06), pages 79-88. ACM, 2006.

[3] E-J. Goh. Secure indexes. Technical Report 2003/216, IACR ePrint Cryptography Arch- ive, 2003. See http://eprint.iacr.org/2003/216.

[4] Boldyreva, N. Chenette, Y. Lee, and A.O’Neill, Order-preserving symmetric encryption

[5] EUROCRYPT 2009, LNCS 5479, pp. 224–241, 2009.

[6] A.Boldyreva, N. Chenette, A. O'Neill: Order-Preserving Encryption Revisited: Improved Security Analysis and Alternative Solutions. CRYPTO 2011, LNCS 6841, pp.578-595, 2011

[7] Boneh D, Crescenzo G, et al. Public encryption with keyword search[C]//Advances in cryptology, Proceedings of the 23rd Annual international conference on the theory and applications of cryptographic techniques(Eurocrypt’04). Berlin, Germany: Springer-Verlag, 2004: 506-522.

[8] M. Abdalla, M. Bellare, D, et al. Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions. In: Shoup, Y(ed.) CRYPTO 2005. LNCS, vol. 3621, Heidelberg, Germany, 2005:205 - 222.

[9] Agrawal R, Kiernan J, et al. Order preserving encryption for numeric data[C]//Proceedings of 2004 ACM sigmod int conf on management of data(Sigmod’04). New York:ACM,2004:563

[10] -574.

[11] Li, J, Wang Q, et al. Fuzzy key-word search over encrypted data in cloud computing[A].In IEEE INFOCOM’10,Mini-Conference[C].NJ:IEEE Press,Piscataway,2010:441-445

[12] Wang C, Cao N,et al. Secure ranked keyword search over encrypted cloud data[A].In ICDCS 2010[C].Washington, DC:IEEE Computer Society,2010:253-262

[13] Cao N,Wang C, et al. Privacy-Preserving multi-keyword ranked search over encrypted cloud data[A].31st International conference on distributed computing systems (ICDCS)[C].2011:

[14] 393-402

[15] Li M,Yu SC,et al. Authorized private keyword search over encrypted data in cloud computing[A].In ICDCS),2011[C].USA:IEEE Press,2011:383-39.

[16] R. L. Rivest, L. Adleman, and M. L. Dertouzos. On data banks and privacy homomorphisms. In Foundations of Secure Computation, 1978.

[17] Ronald L. Rivest, Adi Shamir, Leonard M. Adleman: A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. Commun. ACM 21(2): 120-126 (1978)

[18] Taher El-Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In CRYPTO, pages 10-18, 1984.

[19] Sha Goldwasser and Silvio Micali. Probabilistic encryption and how to play mental poker keeping secret all partial information. In STOC, pages365-377. ACM, 1982.

[20] David Naccache, Jacques Stern: A New Public Key Cryptosystem Based on Higher Residues. ACM Conference on Computer and Communications Security 1998:59-66

[21] Tatsuaki Okamoto, Shigenori Uchiyama: A New Public-Key Cryptosystem as Secure as Factoring. EUROCRYPT 1998: 308-318

[22] Pascal Paillier. Public-key cryptosystems based oncomposite degree residuosity classes. In EUROCRYPT, pages 223-238, 1999.

[23] Dan Boneh, Eu-Jin Goh, and Kobbi Nissim. Evaluating 2-DNF formulas on ciphertexts. In Theory of Cryptography - TCC'05, volume 3378 of Lecture Notes in Computer Science, pages 325{341. Springer, 2005.

[24] Craig Gentry. FHE using ideal lattices. In Michael Mitzenmacher, editor, STOC, pages 169{178. ACM, 2009.

[25] Gentry, C., Halevi, S.: Implementing Gentry’s Fully-Homomorphic EncryptionScheme. In: Paterson, K.G. (ed.) EUROCRYPT 2011. LNCS, vol. 6632, pp. 129–148. Springer, Heidelberg (2011)

[26] Smart, N.P., Vercauteren, F.: FHE with Relatively Small Key and Ciphertext Sizes. In: Nguyen, P.Q., Pointcheval, D. (eds.) PKC 2010. LNCS, vol. 6056, pp. 420–443. Springer, Heidelberg (2010)

[27] Damien Stehl_e and Ron Steinfeld. Faster FHE. In Masayuki Abe, editor, ASIACRYPT, volume 6477 of Lecture Notes in Computer Science, pages 377-394. Springer, 2010.

[28] van Dijk, M., Gentry, C., Halevi, S., Vaikuntanathan, V.: FHE over the Integers. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 24-43. Springer, Heidelberg (2010)

[29] Coron, J.S., Naccache, D., Tibouchi, M.: Public-key Compression and Modulus Switching for FHE over the Integers. Full version of this paper. Cryptology ePrint Archive, Report 2011/440

[30] Coron, J.-S., Mandal, A., Naccache, D., Tibouchi, M.: FHE over the Integers with Shorter PublicKeys. In:Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 487–504. Springer, Heidelberg (2011)

[31] Brakerski, Z., Vaikuntanathan, V.: Efficient FHE from (Standard) LWE. In: Proceedings of FOCS 2011 (2011); Full version available atIACR eprint

[32] Brakerski, Z., Vaikuntanathan, V.: FHE from Ring-LWE and Security for Key Dependent Messages. In: Rogaway, P CRYPTO 2011.LNCS 6841, pp. 505- 524. Springer, Heidelberg (2011)