ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

Secure Accessibility for Big Data in Cloud

Surajkumar Singh[1], Niraj Chaudhary[2], Sreenu M[3], Manjunath B M[4]

U.G. Student, Department of Computer Engineering, Sri Krishna Institute of Technology, Chikkabanavara, Bangalore, India

U.G. Student, Department of Computer Engineering, Sri Krishna Institute of Technology, Chikkabanavara, Bangalore, India

U.G. Student, Department of Computer Engineering, BMS College of Engineering, Chikkabanavara, Bangalore, India

Programmer, Department of Computer Engineering, Sri Krishna Institute of Technology, Chikkabanavara, Bangalore, India

ABSTRACT: Big data are voluminous and complex for that retrieving cipher text to a cloud is deemed to be one Of most effective approaches for big data storage and access. The new policies are proposed in cloud where access legitimacy of user and updating cipher text security designated by data owner are two critical challenges to make cloud-based big data storage practical and effective. Existing approach are completely avoid access policy but in reality it is important to update the access policy amplify the security and dealing with different cause by user join and leave activity. In this paper, we plan a secure and verifiable access control scheme based on NTRU cryptosystem for big data storage in cloud. First the new NTRU decryption algorithm meet the decryption fault of the original NTRU, then details of itsanalyse its correctness, security strength and computational efficiency. When new access policy is specified by big data owner our system allows the cloud server to efficiently update the ciphertext. Which is able to update and validate policy against cheating scheme of cloud. It also authorize the end user to validate information by other user for data access and a user to validate the information provided by for recovery of plaintext.

1. INTRODUCTION

Cloud computing is revolution many of our ecosystems, including healthcare. Contrast with earlier methods of processing data, cloud platforms that handle sensitive information are required to deploy technical measures and organizational safeguards to avoid data protection breakdowns that might result in enormous and costly damages. Cloud computing environments provide significant benefits, such as the availability of automated tools to assemble, connect, configure and reconfigure virtualized resources on demand. These make it much convient to meet organizational goals as organizations can easily deploy cloud services. However, the shift in paradigm that accompanies the adoption of cloud computing is increasingly giving rise to security and privacy considerations relating to facets of cloud computing such as multi-tenancy, trust, loss of control and accountability.

BIG data is a high volume, and/or high velocity, high variety information asset, which requires new forms of processing to enable enhanced decision making, insight discovery, and process optimization. Most existing approaches for securing the outsourced big data in clouds are based on either attributed-based encryption (ABE) or secret sharing. ABE based approaches provide the flexibility for a data owner to predefine the set of users who are eligible for accessing the data but they suffer from the high complexity of efficiently updating the access policy and ciphertext. Secret sharing mechanisms allow a secret to be shared and reconstructed by certain number of cooperative users but they typically employ asymmetric public key cryptography such as RSA for users' legitimacy verification, which incur high computational overhead. Moreover, it is also a challenging issue to dynamically and efficiently update the access policies according to the new requirements of the data owners in secret sharing approaches.
As Big Data expands with the help of public clouds, traditional security solutions tailored to private computing infrastructures, confined to a well-defined security perimeter, such as firewalls and demilitarized zones (DMZs) are no more effective. Using Big Data, security functions are required to work over the heterogeneous composition of diverse

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology
An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018
2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"
13th-14th March 2018
Organized by
Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

hardware, operating systems, and network domains. In this puzzle-type computing environment, the abstraction capability of Software-Defined Networking (SDN) seems a very important characteristic that can enable the efficient deployment of Big Data secure services on-top of the heterogeneous infrastructure. SDN introduces abstraction because it separates the control (higher) plane from the underlying system infrastructure being supervised and controlled. Separating a network's control logic from the underlying physical routers and switches that forward traffic allows system administrators to write high-level control programs that specify the behavior of an entire network, in contrast to conventional networks, whereby administrators (if allowed to do it by the device manufacturers) must codify functionality in terms of low-level device configuration

The scheme motivates us to develop a verifiable access control scheme for securing the big data stored in clouds, tackling the challenges of the following security services:

• **Security.** The proposed scheme should be able to defend against various attacks such as the collusion attack. Meanwhile, access policy update should not break the security of the data storage, disclose sensitive information about the data owner, and cause any new security problem.

• **Verification.** When a user needs to decrypt a stored ciphertext, its access legitimacy should be verified by other participating users and the secret shares obtained from other users must be validated for correct recovery.

• **Authorization.** To reduce the risk of information leakage, a user should obtain authorization from the data owner for accessing the encrypted data

II. RELATED WORK

Because big data contains a voluminous personal identifiable information, it is biggest challenges to provide access control policy securely, we mainly summarize the state of-the-art of securing the big data stored in clouds. retrieving to clouds is one of the most effictive approaches to securing the big data storage , in which the data get encrypted by owners based on cryptographic primitives and store the encrypted data to the clouds. To contract out aa secure mechanism should be created between a data owner and a cloud. In order for the cloud to perform operations over the encrypted data, "Fully Homomorphic Encryption" (FHE) was usually adopted, which allows direct addition and multiplication operations over the ciphertexts while preserving decryptability. Homomorphic encryption was also applied to guarantee the security of data storage . Nevertheless, it is an immature cryptosystem, and is extremely inefficient in practice, which renders it hardly applicable in real world applications. Securely outsourcing big data computations to the clouds was also extensively studied but this topic is out of the scope of the paper.

Adequate access control is key to protect the stored data. Access control has traditionally been provided by operating systems or applications restricting access to the information, which typically exposes all the information if the system or application is hacked . A better approach is to protect the information using encryption that only allows decryption by authorized entities. Attribute-Based Encryption(ABE) is one of the most powerful techniques for access control in cloud storage systems. In the past years, quite a few attribute-based access control schemes were proposed, in which the data owners define the access policies based on the attributes required by the data and encrypt the data based on the access policies. By this way the data owners are able to ensure that only the users meeting the access policies can decrypt the cipher texts. However, it is difficult to update the policies when these ABE based schemes are applied because the data owners do not store the data in their local systems once they outsource the data into the cloud servers. It is also difficult to verify the legitimacy of the downloaded data as the clouds housing the data are not trustworthy. Besides, the operations of encryption and decryption in ABE have a high computational overhead and incur a large energy consumption.

Exemplar based Inpainting technique is used for inpainting of text regions, which takes structure synthesis and texture synthesis together. The inpainting is done in such a manner, that it fills the damaged region or holes in an image, with surrounding colour and texture. The algorithm is based on patch based filling procedure. First find target region using mask image and then find boundary of target region. For all the boundary points it defined patch and find the priority of these patches. It starts filling the target region from the highest priority patch by finding the best match patch. This procedure is repeated until entire target region is inpainted.

The algorithm automatically generates mask image without user interaction that contains only text regions to be inpainted.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology
An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018
2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"
13th-14th March 2018
Organized by
Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

III.SYSTEM MODEL

We consider a cloud storage system that is applicable for both public and private clouds as shown in Fig. 3. It consists of the following three types of entities: cloud server, data owner (owners), and data user (users).

Cloud server. A cloud server provides spaces for data owners to store their outsourced ciphertext data that can be retrieved by

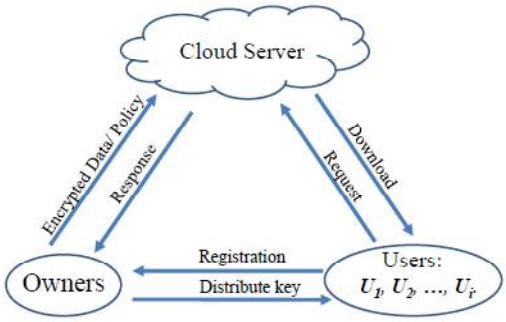

Fig. 3. System model

the users. It is also responsible for updating the cipher texts when the data owner changes its access policy. Owner .A data owner designates the access policy or its data, encrypts the database don the access policy before outsourcing the data to the cloud server, and requests the cloud server to update the encrypted data when a new access policy is adopted. It can also check whether the cipher text at the cloud server is correctly updated.

User. Each user is assigned with a sub-key for an encrypted data the user is eligible to access.In order to decrypt the ciphertext, the user's eligibility must be verified by at least t−1 other users that are also eligible to access the data. The information provided by the t − 1 verifiers must be validated by the user for correct message decryption based on the (t,n)-threshold secret sharing.

IV .PRELIMINARIES

we introduce the following two cryptographic primitives: **NTRU** and secret sharing.

The NTRU Cryptosystem

The NTRU cryptosystem is based on the shortest vector problem (SVP) in a lattice that makes it lightning fast and resistant to quantum computing attacks. It has been proved to be faster than RSA [28], [29]. In this subsection, we briefly introduce NTRU, and refer the interested readers to [27] and [26] for more details. NTRU consists of three integer parameters (N,p,q) and four sets Lf,Lg,Lφ and Lm of polynomials of degree N − 1 with integer coefficients. Note that p and q are not required to be prime, but they should simultaneously meet the following two criterions: gcd(p,q) = 1 and q > p. NTRU works in the ring R = Z[X]/(XN −1), and Lf,Lg,Lφ and Lm are defined in R whose selection criteria are detailed in [26]. Any element f ∈ Lf can be expressed as a polynomial or a vector as follows

f =N−1 X i=0

fixi = [f0,f1,...,fN−1]. (1) We use∗to denote the convolution multiplication of two polynomials f and g in R, which can be computed via (2).

f∗ g = h = (h0, h1, . . . , hN−1), with hk = X i+j≡k (mod N) figj .

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology
An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018
2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"
13th-14th March 2018
Organized by
Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

figN+k−i, (2) where k ∈ [0,N −1]. Note that when we do a multiplication of two polynomials modulo q, we mean to reduce the coefficients of the product modulo q.

NTRU implements the following three basic functions. For better presentation, we use Alice and Bob to represent the two entities for encrypting and decrypting a message.

• **Key Generation**:To create his public and private keys,Bob randomly chooses two polynomials g ∈Lg and f ∈ Lf, in which f is his private key and is required to have inverse modulo q and inverse modulo p.Letfq and fp respectively denote the inverse modulo q and the inverse modulo p of f, which are formally defined as follows: fq∗f = 1 (mod q) and fp∗f = 1 (mod p). (3) For any properly selected f, it is easy to compute fq and fp via the Euclidean algorithm. Then Bob computes his public key h according to (4), and publishes his public parameters{p,q,h}. h = fq∗g (mod q). (4)

• **Encryption:** Assume that Alice wants to send a message m to Bob, she needs to first convert m into a polynomial in Lm. For simplicity, we use m to denote the polynomial representation of m in Lm too. Then she randomly chooses a polynomial φ ∈Lφ, encrypts m using Bob's public key h according to (5), and sends the encrypted message e to Bob. e = pφ∗h + m (mod q). (5)

• **Decryption**: Bob receives the encrypted message e from Alice, and decrypts it using his private key f and the inverse of f module p, i.e., fp, via (6) and (7). a = e∗f (mod q). (6) m = a∗fp (mod p). (7)

V. CONCLUSION AND FUTURE WORKS

In this paper, we first propose an improved NTRU cryptosystem to overcome the decryption failures of the original NTRU and then present a secure and verifiable access control scheme based on the improved NTRU to protect the outsourced big data stored in a cloud. Our scheme allows the data owner to dynamically update the data access policy and the cloud server to successfully update the corresponding outsourced cipher text to enable efficient access control over the big data in the cloud. It also provides a verification process for a user to validate its legitimacy of accessing the data to both the data owner and t−1 other legitimate users and the correctness of the information provided by the t−1 other users for plaintext recovery. The security of our proposed scheme is guaranteed by those of the NTRU cryptosystem and the (t,n)-threshold secret sharing. We have rigorously analyzed the correctness, security strength, and computational complexity of our proposed scheme.

Designing a secure, privacy preserving, and practical scheme for big data storage in a cloud is an extremely challenging problem. In our future research, we will further improve our scheme by combining the (t,n)-threshold secret sharing with attribute based access control, which involves an access structure that can place various requirements for a user to decryptan outsourced ciphertext data in the cloud. Meanwhile, we will investigate the security problems when a data owner outsources its data to multi cloud servers and consider an attribute-based access structure that can be dynamically updated, which is more applicable for practical scenarios in big data storage.

VII.ACKNOWLEDGMENT

This research was partially supported by the National Science Foundation of the US under grants CNS-1318872, CCF-1442642 and IIS-1343976, and the National Natural Science Foundation of China under grants 61672119, 61373027, 61472044, 61272475, 61571049, 61472403, and 61672321, and the Fundamental Research Funds for the Central Universities (No.2014KJJCB32).

REFERENCES

[1] M. A. Beyer and D. Laney, "The importance of big data: a definition," Stamford, CT: Gartner, 2012.
[2] V. Marx, "Biology: The big challenges of big data," Nature, vol. 498, no. 7453, pp. 255–260, 2013.
[3] G. P. Consortium et al., "A map of human genome variation from population-scale sequencing,"Nature,vol. 467,no. 7319,pp. 1061–1073, 2010.
[4] A. Sahai and B. Waters, "Fuzzy identity-based encryption," Advances in Cryptology–EUROCRYPT 2005, pp. 457–473, 2005.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization *Volume 7, Special Issue 6, May 2018*

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

[5] C. Hu, F. Zhang, X. Cheng, X. Liao, and D. Chen, "Securing communications between external users and wireless body area networks," in Proceedings of the 2nd ACM workshop on Hot topics on wireless network security and privacy. ACM, 2013, pp. 31–36.
[6] C. Hu, H. Li, Y. Huo, T. Xiang, and X. Liao, "Secure and efficient data communication protocol for wireless body area networks," IEEE Transactions on Multi-Scale Computing Systems, vol. 2, no. 2, pp. 94– 107, 2016.
[7] 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. ACM, 2006, pp. 89–98.
[8] B. Waters, "Ciphertext-policy attribute-based encryption: An expressive, efficient, and provably secure realization," Public Key Cryptography–PKC 2011, pp. 53–70, 2011.
[9] C. Hu, N. Zhang, H. Li, X. Cheng, and X. Liao, "Body area network security: a fuzzy attribute-based signcryption scheme," IEEE journal on selected areas in communications, vol. 31, no. 9, pp. 37–46, 2013.
[10] A. Lewko and B. Waters, "Decentralizing attribute-based encryption," Advances in Cryptology–EUROCRYPT 2011, pp. 568–588, 2011.
[11] C. Hu, X. Cheng, Z. Tian, J. Yu, K. Akkaya, and L. Sun, "An attribute based signcryption scheme to secure attribute-defined multicast communications," in Secure Comm 2015. Springer, 2015, pp. 418–435.
[12] A. Shamir, "Identity-based cryptosystems and signature schemes," in Advances in cryptology. Springer, 1985, pp. 47–53.
[13] M. Dehkordi and S. Mashhadi, "An efficient threshold verifiable multisecret sharing," Computer Standards & Interfaces, vol. 30, no. 3, pp. 187–190, 2008.
[14] Z. Eslami and J. Z. Ahmadabadi, "A verifiable multi-secret sharing scheme based on cellular automata," Information Sciences, vol. 180, no. 15, pp. 2889–2894, 2010.
[15] M. H. Dehkordi and S. Mashhadi, "New efficient and practical verifiable multi-secret sharing schemes," Information Sciences, vol. 178, no. 9, pp. 2262–2274, 2008. [16] J. Zhao, J. Zhang, and R. Zhao, "A practical verifiable multi-secret sharing scheme," Computer Standards & Interfaces, vol. 29, no. 1, pp. 138–141, 2007.