A secure and verifiable access control with integrity verification scheme in big data

Senthamil. K, PG Student, Department of CSE, Jayam College of Engineering and Technology, Dharmapuri, T.N, India

ABSTRACT

Due to complexity and volume, outsourcing cipher text to a cloud is one of the most effective approaches for big data storage and access. The cloud data storage service relieves the users from the burden of voluminous local data storage and their maintenance by outsourcing mass data to the cloud. 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 and the secret sharing mechanisms allow a secret to be shared and reconstructed by certain number of cooperative users. Trusted third party may lead to unauthorized exposure of stored data to other users. The Proposed work supports an encrypted data integrity checking scheme without a trusted of third party auditor. To investigate the security problems when a data owner outsources its data to multi cloud servers and to form attribute based access structure are updated dynamically. Integrity checking ensures the sensitive data which send by the data owner and if the data is insensitive it is directly stored in cloud service provider (CSP) or cloud server. Data integrity verification scheme can be assured by using asymmetric cryptosystem. Extensive security and performance analysis proves that the proposed scheme is highly secure and efficient.

Keyword: *Attributed-based encryption (ABE), Data integrity verification scheme, cloud Service provider (CSP)*

INTRODUCTION

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. Due to its complexity and large volume, managing big data using on hand database management tools is difficult. An effective solution is to outsource the data to a cloud server that has the capabilities of storing big data and processing users' access requests in an efficient manner. However, when a data owner outsources its data to a cloud, sensitive information may be disclosed because the cloud server is not trusted, therefore typically the ciphertext of the data is stored in the could. But how to update the ciphertext stored in a cloud when a new access policy is designated by the data owner and how to verify the legitimacy of a user who intends to access the data are still of great concerns. Most existing approaches for securing the outsourced big data in clouds are based on either attributed-based encryption (ABE).

There are several attacks that are carried out on data that is being stored in cloud especially like tampering, Data Forgery and there are threats like tag Forgery in circumstances where even a service Provider in cloud Environment tries to cheat the end users when they just try to verify the data instead of downloading and checking it. It is almost impossible to find out the attack in the latter case. Since Data Integrity becoming such a big threat it is mandatory for an efficient data verification Scheme in the Cloud Scenario and also while implementing such schemes it is important to check the amount of communication and computation cost and other overheads that occurs during implementation of Verification Process. The issues such as TPA becoming bottleneck, data leakage, introduction of new vulnerabilities, scalability, accountability, performance overhead, dynamic data support, extra hardware cost incurred etc. have motivated many researchers to address the data storage security problems without using a third party auditor, where the cloud user may be comprised of extra application or tool which helps it periodically checks the data integrity.

The proposed scheme achieves data storage correctness in cloud computing without making use of a third party auditor. The central design goal of the proposed scheme is to enhance user‘s control in managing the various aspects related to the secrecy of sensitive data.

EXISTING SYSTEM

Secret sharing mechanisms allow a secret to be shared and reconstructed by certain number of cooperative users but they typically employ asymmetric public key cryptograph 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 a data owner typically does not backup its data locally after outsourcing the data to a cloud, it cannot easily manage the data stored in the cloud. Besides, as more and more companies and organizations are using clouds to store their data, it becomes more challenging and critical to deal with the issue of access policy update for enhancing security and dealing with the dynamism caused by the users' join and leave activities. To the best of our knowledge, policy update for outsourced big data storage in clouds has never been considered by the existing research. Another challenging issue is how to verify the legitimacy of the users accessing the outsourced data in clouds. Existing schemes proposed in do not support user eligibility verification. On the other hand, verifiable secret sharing based schemes rely on RSA for access legitimacy verification. As multiple users need to mutually verify each other using multiple RSA operations, such a procedure has a high computational overhead. The NTRU cryptosystem is a type of lattice-based cryptography and its security is based on the shortest vector problem (SVP) in a lattice. The major advantages of NTRU are quantum computing attack resistance and lighting fast computation capability. However, NTRU suffers from the problem of decryption failures .The considerations mentioned above motivate 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.

Drawback of Existing System

- Data owner outsources its data to a cloud, sensitive information may be disclosed because the cloud server is not trusted.
- Integrity checking doesn't form in existing system.

PROPOSED SYSTEM

The cloud data storage service relieves the users from the burden of voluminous local data storage and their maintenance by outsourcing mass data to the cloud. Trusted third party may lead to unauthorized exposure of stored data to other users. The Proposed work supports an encrypted data integrity checking scheme without a trusted of third party auditor. And to investigate the security problems when a data owner outsources its data to multi cloud servers and to form attribute based access structure are updated dynamically. Integrity checking ensures the sensitive data which send by the data owner and if the data is insensitive it is directly stored in cloud service provider (CSP) or cloud server. Data integrity verification scheme can be assured by using asymmetric cryptosystem. The proposed scheme is to enhance user's control in managing the various aspects related to the secrecy of sensitive data. To update the cipher text stored in clouds without increasing any risk when the access policy is dynamically changed by the data owner for various reasons. The correctness of the proposed scheme to investigate its efficiency and security strength. The cloud customer can flexibly control and manage the different privacy mechanisms necessary to protect sensitive data and achieve legal compliance. Users will be aware of all the operations carried out to secure the storage and processing of their sensitive information through a secure privacy.

COMPONENTS OF ARCHITECTURE

Cloud Service Provider

A cloud server provides spaces for data owners to store their outsourced cipher text data that can be retrieved by the users. It is also responsible for updating the cipher texts when the data owner changes its access policy. A cloud service provider which manages and operates a cloud infrastructure of storage and computing services.

Data Owner

A data owner that employs the cloud storage and computing resource facilities to remotely store and process data. The data owner designates the access policy for its data, encrypts the data based on 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. A message certificate for the data, and stores the encrypted data with the access policy in the cloud.

Authorized User

A user who access the data from cloud storage. When a user needs to use the data, it solicits help from other users to recover the data. The cloud server can update the encrypted data with a new policy is designated by the data owner.

Sensitive data and insensitive data

- Data marked with this attribute is stored encrypted by the data owner specific key.
- Data marked with this attribute is not sensitive and hence the provider is fully trusted to store it without any form of encryption.

Data Dynamics

After storing data in the cloud the data owner can dynamically update it. The data dynamics offer data insertion, deletion and modification. Data owner often changes the data to the cloud service provider and the complete data has to be updated, deleted, assign and stated are to be performed.

Integrity Checking

Data Integrity can be used to describe a state, a process or a function and is often used as a proxy for "data quality". Data with "integrity" is said to have a complete or whole structure. Data values are standardized according to a data model and/or data type. All characteristics of the data must be correct including business rules, relations, dates, definitions and lineage for data to be complete. Data integrity is imposed within a database when it is designed and is authenticated through the ongoing use of error checking and validation routines.

ADVANTAGES OF PROPOSED SYSTEM

- Cloud storage providing integrity protection of user's important data.
- It is proved to be secure against unauthorized users since it does not involve any trusted third party in data integrity checking operation.
- Authenticated users are allowed to check the integrity of data.
- It has very good efficiency in the aspects of communication, computation and storage costs.

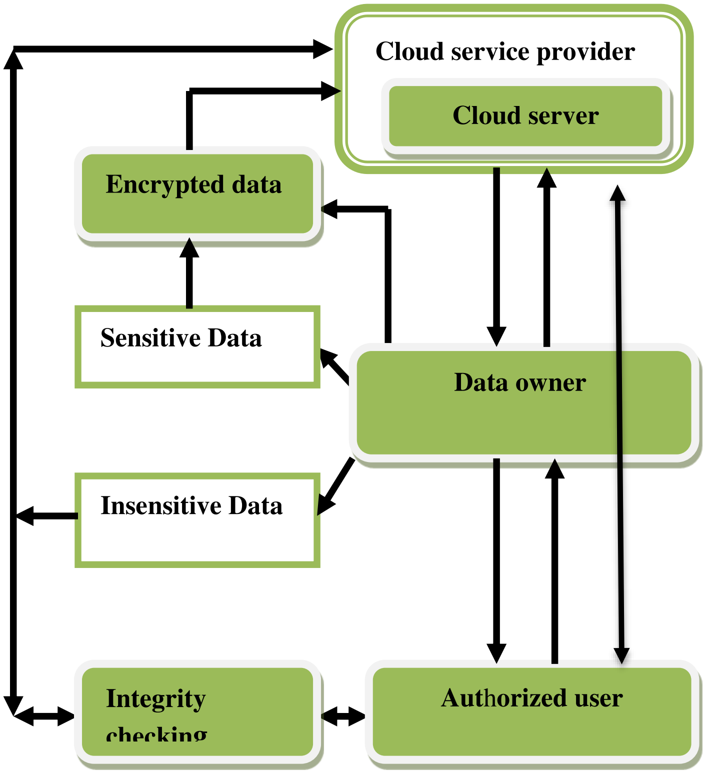

Fig1SYSTEM ARHITECTURE

SYSTEM MODULES

Provable Data Possession (PDP) (Outsourcing by the data Owner)
Authenticated Key Exchange Protocol (Encrypt the data)
Data Dynamics. Data Storage in Cloud Service Provider.
Decrypt the data (Verify the data).

Provable Data Possession (PDP)

Data outsourcing is a major component for data owners to distribute resources to external services for sharing with users and organizations. A crucial problem for owners is how to secure sensitive information accessed by legitimate users only using the trusted services. The problem with access control methods to enforce selective access to outsourced data without involving the owner in authorization. The basic idea is to combine cryptography with authorizations, and data owners assign keys to roles that will enforce access via encryption. Data owners would outsource their data in different formation of secret keys are to be assured.

Provable Data Possession (PDP) in the cloud is an efficient way for users to ensure the correctness of their data stored in the cloud. However, existing PDP solutions mainly focus on single-owner data, which cannot achieve anonymity and efficiency during the proving of the possession of multi-owner data, where data are shared and managed among multiple owners with cloud services. To propose a remote data possession checking protocol for critical information infrastructures. PDP protocol which is based on message authentication codes. This protocol supports block modification, deletion and append.

Authenticated Password Key Exchange

Authentications of the users are made secure by an Authenticated Password Key Exchange protocol. The proposed integrity verification scheme envisages a user centric approach in which the cloud data owner has full control on the privacy mechanisms to be applied on the cloud data. The security of the user's passwords is enhanced by the protocol. An Authenticated password key exchange is where based on the knowledge of their password, two or more parties establish a cryptographic key using message exchanging, such that an unauthorized party cannot involve in the process of brute force guessing the password. This method envisages that strong security can be obtained using weak passwords.

Diffie-Hellman key exchange typically assumes that parties use a fixed generator, and seem to require a public generator for their proofs of security. However, to require that no one know the discrete logarithms of any of the generators with respect to any other, and thus we need either a trusted party who generates the public information or else a source of randomness which can be used to publicly derive the information.

Data Dynamics

Storing data in the cloud and the data owner can dynamically update it. The data dynamics offer data insertion, deletion and modification. Data Owner can insert and delete the contents which the data are to be updated. And the modified data are to be updated in cloud service providers. CSP stores the data with the above mentioned techniques

Data Storage in Cloud Service Provider

A cloud service provider which manages and operates a cloud infrastructure of storage and computing services. The complete and large data are supposed to store in cloud service Provider where each data are to be encrypted before storing. And the encrypted data are again retrieved by the authorized user.

Decrypt the Data

The scheme proposed in, which independently envisages a RSA based methods for remote data integrity checking with the help of third party verifier. The proposed protocol can be viewed as an adaptation of to support more functionality. It inherits the support of data dynamics from, and supports integrity verification performed by anyone other than client and for providing security; it doesn‘t need to use a third-party verifier to check the integrity of data. It also uses asymmetric cryptosystem for providing integrity verification. PAKE protocol is used to provide security to the passwords provided for authentication.

CONCLUSION

The proposed work forms a new data integrity verification protocol for cloud storage providing integrity protection of user's important data. It is proved to be secure against unauthorized users since it does not involve any trusted third party in data integrity checking operation. It has very good efficiency in the aspects of communication, computation and storage costs. To exploit the strengths of this technology and to overcome the drawbacks in order to ensure data integrity and consequently a big data security on the cloud. Enabling the different integrity proofs to keep the data in secure manner. So that sensitive data are encrypt and stored in cloud service provider. Insensitive data stored directly without any encryption.

FUTURE WORK

In future work, is focused on protocol support for data dynamics using Embedded Merkel Hash Tree(EMHT) and a hardware support for data integrity checking. And to upgrade Dynamic data updating is provided by Merkle B-tree algorithm which provide aspects of good efficiency in the communication, computation and storage costs less.

REFERENCE:

[1] Zhuo Hao, Sheng Zhong, Neng-Hai Yu, ―A privacy preserving data integrity checking protocol with data dynamics and public verifiability ‖ Knowledge and Data Engineering, IEEE transactions on, vol 23, 2011,pp. 1432 - 1437

[2] Z. Hao and N. Yu, ―A multiple-replica remote data possession checking protocol with public verifiability,‖ in Data, Privacy, and E-Commerce, 2010.

[3]C. Wang, Q. Wang, K. Ren, and W. Lou ‖Privacy-preserving public auditing for data storage security in cloud computing ‖ in *InfoCom2010, IEEE.*

[4] Feng Hao, Peter Ryan ,‖ J-PAKE: authenticated key exchange without PKI‖in Transactions on computational science XI, 2010 Pages 192-206.

[5]Li Xiao-fei, Shen Xuan-jing, Chen Hai-peng. ―An Improved ElGamal Digital Signature Algorithm Based on Adding a Random Number‖. In 978-0-7695-4011-5/10 $26.00 © 2010 IEEE.

[6] Wassim Itani, Ayman Kayssi, Ali Chehab ― Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures‖ in Proceedings of the 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, Pages 711-716 .

[7] C. Wang, S. S.-M. Chow, Q. Wang, K. Ren, and W. Lou,―Privacy preserving public auditing for secure cloud storage.‖ Cryptology ePrint Archive, Report 2009/579.

[8] C. Erway, A. K¨upc¸ ¨u, C. Papamanthou, and R. Tamassia, ―Dynamic provable data possession,‖ in *CCS '09*: Proceedings of the 16th ACM conference on computer and communications security , 2009, pp. 213–222.

[9] Wang C, Wang Q, Ren K, and Lou W, ―Ensuring data storage security in cloud computing,‖ in Proc. of IWQoS'09, 2009.

[10]Kyrakos Mouratidis, Dimitris Sacharidis‖Partially materialized digest scheme: an efficient verification method for outsourced databases‖ in the international journal on very large data bases archieve,Vol 18 Issue 1, 2009,pages 363-381.

[11]. Marco Bodrato, ―Public key cryptography. ElGamal, hints on implementation‖, Tokyo University of Science,2008.

[12] G. Ateniese, R. Di Pietro, L. V. Mancini, and G. sudik, ―Scalable and efficient provable data possession,‖ in SecureComm '08:proceedings of 4th international conference on Security and privacy in communication networks,
2008, pp. 1–10.

[13] Qingji Zheng, Shouhuai Xu, Giuseppe Ateniese ‖Efficient Query integrity of outsourced dynamic databases‖in Proceeding of the 2012 ACM Workshop on Cloud computing security workshop, 2012, Pages 71-82

[14] F.Sebe,J.Domigo-Ferrer,A.Martinez-Balleste et.al ―Efficient remote data possession checking in critical information infrastructures‖, IEEE Transactions on, vol. 20, 2008,pp. 1034-1038.