Privacy-Preserving Multi-Keyword Top-K Similarity Search Over Encrypted Data

Vaithianathan K[1], Mrs. J. Kalaivani[2]
[2] Asst. Prof.
[1, 2] Krishnasamy College Of Engineering And Technology

Abstract- *Identity-based encryption is used to build data sharing system. In order, access control is not static. It means that when authorization of some users is expired, the system should remove his (or) him. By that the removed user cannot access both forward and backward data. For this we use a concept called revocable-storage identity-based encryption (RS-IBE), which provide security of cipher text for forward/backward data by introducing the user revocation functionalities and simultaneous update of cipher text. We provide a detailed structure of RS-IBE, which certifies its secrecy in the described security model. The realistic and cost-effective system of data sharing is achieved by this RS-IBE scheme which has tremendous benefits of operability and capability. Certainly, we provide implementation outcome of this suggested scheme to define its feasibility. In this paper, we address this practical problem, which is largely neglected in the literature, by proposing the novel concept of key aggregate searchable encryption (KASE) and instantiating the concept through a concrete KASE scheme, in which a data owner only needs to distribute a single key to a user for sharing a large number of documents, and the user only needs to submit a single trapdoor to the cloud for querying the shared documents. The security analysis and performance evaluation both confirm that our proposed schemes are provably secure and practically efficient. This approach acquaints the utilities of user repudiation and cipher text update concurrently. Additionally, we provide a detailed structure of Searchable-IBE, which certifies its secrecy in the described security model. The realistic and cost-effective system of data sharing is achieved by this Searchable-IBE scheme which has tremendous benefits of operability and capability. Certainly, we provide implementation outcome of this suggested scheme to determine its feasibility.*

Keywords- key aggregate searchable encryption , revocable-storage identity-based encryption, trapdoor, data encryption.

I. INTRODUCTION

Cloud storage is nowadays very popular storage system. Cloud storage is storing of data off-site to the physical storage which is maintained by third party. Cloud storage is saving of digital data in logical pool and physical storage spans multiple servers which are manage by third party. Third party is responsible for keeping data available and accessible and physical environment should be protected and running at all time. Instead of storing data to the hard drive or any other local storage, we save data to remote storage which is accessible from anywhere and anytime. It reduces efforts of carrying physical storage to everywhere. By using cloud storage we can access information from any computer through internet which omitted limitation of accessing information from same computer where it is stored. While considering data privacy, we cannot rely on traditional technique of authentication, because unexpected privilege escalation will expose all data. Solution is to encrypt data before uploading to the server with user's own key. Data sharing is again important functionality of cloud storage, because user can share data from anywhere and anytime to anyone. For example, organization may grant permission to access part of sensitive data to their employees. But challenging task is that how to share encrypted data. Traditional way is user can download the encrypted data from storage, decrypt that data and send it to share with others, but it loses the importance of cloud storage. Cryptography technique can be applied in a two major ways- one is symmetric key encryption and other is asymmetric key encryption. In symmetric key encryption, same keys are used for encryption and decryption. By contrast, in asymmetric key encryption different keys are used, public key for encryption and private key for decryption. Using asymmetric key encryption is more flexible for our approach.

II. LITERATURE VIEW

1.Adaptively secure ciphertext-policy attribute-based encryption with dynamic policy updating(2016)- This paper is presented by Zuobin YING, Hui LI, Jianfeng MA, Junwei ZHANG & Jiangtao CUI and described as "Securely search the data on cloud server using symmetric algorithm and limitations are Less secure Easy to access by authorized one's".

2.Privacy-preserving multi-keyword top-k similarity search over encrypted data(2017) -This paper is presented by Ding et al. and described as " Using symmetric searchable scheme which is based on pursue-random function to improve the

security and limitations are More time used to complete the process and lessEfficient and secure".

3.Efficient searchable symmetric encryption for storing multiple source dynamic social data on cloud(2017) -This paper is presented by Chang Liu, Liehuang Zhu,Jinjun Chen and described as "To allow multiple parties to aggregate their data together while maintaining privacy and limitations are each participant communication and computational complexity to a small constant".

4.Easy multi-keyword search over encrypted cloud data for multiple data for multiple data holders(2018)-This paper is presented by Peng et al. and described as "Encrypted the sensitive data before outsourcing on cloud server using the symmetric algorithm and limitations are Less secure and More system disruption".

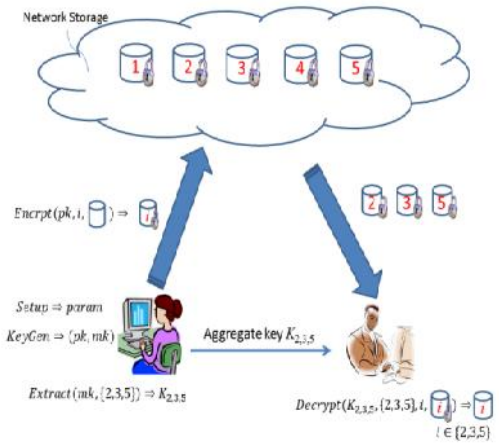
III. PROPOSED SYSTEM

In this paper, address this challenge by proposing the novel concept of key-aggregate searchable encryption (KASE), and instantiating the concept through a concrete KASE scheme. The proposed KASE scheme applies to any cloud storage that supports the searchable group data sharing functionality, which means any user may selectively share a group of selected files with a group of selected users, while allowing the latter to perform keyword search over the former. To support searchable group data sharing the main requirements for efficient key management are twofold. First, a data owner only needs to distribute a single aggregate key (instead of a group of keys) to a user for sharing any number of files. Second, the user only needs to submit a single aggregate trapdoor (instead of a group of trapdoors) to the cloud for performing keyword search over any number of shared files. In Addition with that we are implementing the Efficient Cryptography algorithm changing the file extension for the security purpose. And also we are implementing the user rights into account that user can modify the content after the uploading of file also. He can change the content of the file.

ADVANTAGES

- In a KASE scheme, the owner only needs to distribute a single key to a user when sharing lots of documents with the user, and the user only needs to submit a single trapdoor when he queries over all documents shared by the same owner In a practical data sharing system based on cloud storage, the user can retrieve data by any possible device and the mobile devices are widely used now.
- The security analysis and performance evaluation both confirm that our proposed schemes are provably secure and practically efficient.

SYSTEM ARCHITECTURE



IV. MODULE DESCRIPTION

1. **Data uploading**
2. **Data sharing**
3. **Keyword Search**
4. **Data Retrieving**

Data uploading

To upload a document, the owner runs KAE. Encrypt to encrypt the data and KASE. Encrypt to encrypt the keyword cipher texts, then uploads them to the cloud. The cloud assigns a docID for this document and stores the encrypted data in the path file Path, then inserts a record into the table docs. In addition, the owner can encrypt the keys using his/her private key and store them into the table docs.

Data sharing

To share a group of documents with a target member, the owner runs KAE. Extract and KASE. Extract to generate the aggregate keys, and distributes them to this member, then inserts/updates a record in table sharedDocs. If the shared documents for this member are changed, the owner must re-extract the keys and update the elddocIDSet in table sharedDocs.

Keyword Search

To retrieve the documents containing an expected keyword, a member runs KASE. Trapdoor to generate the

keyword trapdoor for documents shared by each owner, then submits each trapdoor and the related owners identity OwnerID to the cloud. After receiving the request, for each trapdoor, the cloud will run KASE. Adjust the trapdoor for each document in the docIDSet and run KASE. Test to perform keyword search. Then, the cloud will return the encrypted documents which contains the expected keyword to the member .

Data Retrieving

After receiving the encrypted document, the member will run KAE. Decrypt to decrypt the document using the aggregate key distributed by the documents owner.

V. FUTURE WORK

In future, we make the whole search process verifiable and data user can be assured of the authenticity of the returned search Result. We also formally prove the proposed scheme Semantically secure in the selective model.

VI. CONCLUSION

Considering the practical problem of privacy preserving data sharing system based on public cloud storage which requires a data owner to distribute a large number of keys to users to enable them to access his/her documents, we for the first time propose the concept of key-aggregate searchable encryption(KASE) and construct a concrete KASE scheme. Both analysis and evaluation results confirm that our work can provide an effective solution to building practical data sharing system based on public cloud storage. In a KASE scheme, the owner only needs to distribute a single key to a user when sharing lots of documents with the user, and the user only needs to submit a single trapdoor when he queries overall documents shared by the same owner. However ,if a user wants to query over documents shared by multiple owners, he must generate multiple trapdoors to the cloud. How to reduce the number of trapdoors under multi-owners setting is a future work. More over, federated clouds have attracted a lot of attention now a days, but our KASE cannot be applied in this case directly. It is also a future work to provide thesolution for KASE in the case of federated clouds.

REFERENCES

[1] W. Sun, S. Yu, W. Lou, Y. T. Hou, and H. Li, “Protecting Your Right: Attribute-based Keyword Search with Fine-grained Owner-enforced Search Authorization in the Cloud,” in IEEE INFOCOM, pp. 226-234, 2014.

[2] Zuobin YING, Hui LI, Jianfeng MA, Junwei ZHANG & Jiangtao CUI, ”Adaptively secure ciphertext-policy attribute-based encryption with dynamic policy updating,” in Science China information sciences 59(4),1-16,2016.

[3] Ding et al. “Privacy-preserving multi-keyword top-k similarity search over encrypted data,” in ieexplore.iee.org, 2017.

[4] Chang Liu, Liehuang Zhu, Jinjun Chen, ”Efficient searchable symmetric encryption for storing multiple source dynamic social data on cloud”, in Journal of Network and Computer Applications, volume 86,pages 3-14,2017.

[5] S. Yu, C. Wang, K. Ren, and W. Lou, “Achieving secure, scalable,and fine-grained data access control in cloud computing,” in Proc.of IEEE INFOCOM, pp. 1-9, 2010.

[6] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, “Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption,” IEEE TPDS, vol. 24, no. 1, pp. 131-143, 2013.

[7] S. Kamara and K. Lauter, “Cryptographic cloud storage,” in Finan-cial Cryptography and Data Security, pp. 136–149, 2010.

[8] D. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Proc. of IEEE S&P, pp. 44-55, 2000.

[9] Y. Huang, D. Evans, J. Katz, and L. Malka, “Faster secure two-party computation using garbled circuits,” in USENIX Security Symposium, vol. 201, no. 1, 2011.

[10] C. Gentry, “A fully homomorphic encryption scheme,” Ph.D. dis-sertation, Stanford University, 2009.

[11] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, “Searchable symmetric encryption: improved definitions and efficient constructions,” in Proc. of ACM CCS, pp. 79-88, 2006.

[12] S. Kamara, C. Papamanthou, and T. Roeder, “Dynamic searchable symmetric encryption,” in Proc. of ACM CCS, pp. 965-976, 2012.

[13] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, “Privacy-preserving multi-keyword ranked search over encrypted cloud data,” in Proc. of IEEE INFOCOM, pp. 829-837, 2011.

[14] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H.Li, “Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking,” in Proc. of ACM ASIACCS, pp. 71-82, 2013.

[15] F. Bao, R. H. Deng, X. Ding, and Y. Yang, “Private query on en-crypted data in multi-user settings,” in Information Security Practice and Experience, Springer, pp. 71-85, 2008.

[16] Y. Yang, H. Lu, and J. Weng, “Multi-user private keyword search for cloud computing,” in Proc. of IEEE CloudCom, pp. 264-271, 2011.

[17] Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in Proc. of Pairing, pp. 2-22, 2007.
[18] M. Blaze, G. Bleumer, and M. Strauss, "Divertible protocols and atomic proxy cryptography," in Proc. of EUROCRYPT, pp. 127-144, 1998.
[19] M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu, "Plutus: Scalable secure file sharing on untrusted storage," in Proc. of FAST, vol. 42, pp. 29-42, 2003.
[20] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. of EUROCRYPT, pp. 506-522, 2004.
[21] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. of ACNS, pp. 31-45, 2004.
[22] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Theory of Cryptography, pp. 535-554, 2007.
[23] J. Katz, A. Sahai, and B. Waters, "Predicate encryption supporting disjunctions, polynomial equations, and inner products," in Proc. of EUROCRYPT, pp. 146-162, 2008.