VERIFIABLE AND MULTI-KEYWORD SEARCHABLE ATTRIBUTE-BASED ENCRYPTION SCHEME FOR CLOUD STORAGE

B. Santhosh Kumar[1], Sharon[2]

[1]*Dept of Computer Science, School of Arts and Science, Vinayaka Mission's Research Foundation, Chennai, India.*
Email: Santhoshbabu0411@gmail.com

[2]*Lecturer, Dept of Computer Science, School of Arts and Science, Vinayaka Mission's Research Foundation, Chennai, India.*
Email: sharon.csa@avit.ac.in

ABSTRACT

The advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in cloud, it is crucial for the search service to allow multi-keyword query and provide result similarity ranking to meet the effective data retrieval need. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely differentiate the search results. In this paper, for the first time, we define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. Among various multi-keyword semantics, we choose the efficient principle of "coordinate matching", i.e., as many matches as possible, to capture the similarity between search query and data documents, and further use "inner product similarity" to quantitatively formalize such principle for similarity measurement. We first propose a basic MRSE scheme using secure inner product computation, and then significantly improve it to meet different privacy requirements in two levels of threat models. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset further show proposed schemes indeed introduce low overhead on computation and communication.

Keywords: Multi-keyword, Attribute Authority, Central Authority.

1. INTRODUCTION

With the development of cloud computing, many of information can be shared through computer networks. The cloud server (CS) can provide users with a variety of services, such as outsourcing commission calculations and data storage. Users can store their large amounts of data to the CS and share data with other users. For the purpose of the security of storage data and user's privacy, data is usually stored in encrypted form in CS. However, under this environment users will encounter a difficulty problem of how to search keyword in ciphertext. Searchable Encryption (SE) is a cryptographic technology that has been developed for many years, which supports users' keyword search in ciphertext. In the meanwhile, it can save a lot of network and computational overhead for user, and take advantage of the huge computing power of CS. The SE technology mainly solves the problem of how to use the server to complete the search for interesting keywords when the data is encrypted and stored in

CS, but CS is not completely trusted. How to improve the efficiency of keyword search while reducing local computing load is still a problem to be solved. Most of existing schemes support single-keyword search. Single-keyword search waste network bandwidth and computing resources, as this search method returns a large number of results, this means that the search result is not accurate. That is, when a data user uses multi- keyword search, the cloud sever will return relatively few number of files containing these multi keyword, thus the search result is much more accurate than when a data user uses one keyword search. In order to solve this problem, multi-keyword search is proposed. Most of existing attribute-based encryption (ABE) schemes have high computational costs at user client. These problems greatly limit the applications of ABE schemes in practice. To solve the problems of network bandwidth waste and high computational cost, we propose a verifiable and multi-keyword searchable attribute-based encryption (VMKS-ABE) scheme for cloud storage, in which many computing tasks are outsourced to cloud proxy server to reduce local computing burden, the scheme also supports the verification of the correctness of outsourced private keys. In our new scheme multi-keyword can be searched and the search privacy is protected, which can greatly improve the accuracy of keyword search .

2. LITERATURE SURVEY

Title: Enabling Personalized Search over Encrypted Outsourced Data with Efficiency Improvement

Author: Z. Fu, K. Ren, J. Shu, X. Sun, and F. Huang

Abstract: In cloud computing, searchable encryption scheme over outsourced data is a hot research field. However, most existing works on encrypted search over outsourced cloud data follow the model of "one size fits all" and ignore personalized search intention. Moreover, most of them support only exact keyword search, which greatly affects data usability and user experience. So how to design a searchable encryption scheme that supports personalized search and improves user search experience remains a very challenging task. In this paper, for the first time, we study and solve the problem of personalized multi-keyword ranked search over encrypted data (PRSE) while preserving privacy in cloud computing. With the help of semantic ontology WordNet, we build a user interest model for individual user by analyzing the user's search history, and adopt a scoring mechanism to express user interest smartly. To address the limitations of the model of "one size fit all" and keyword exact search, we propose two PRSE schemes for different search intentions. Extensive experiments on real-world dataset validate our analysis and show that our proposed solution is very efficient and effective

Title: **Towards efficient content-aware search over encrypted Outsourced data in cloud**

Author: Z. Fu, X. Sun, S. Ji, and G Xie

Abstract: With the increasing adoption of cloud computing, a growing number of users outsource their datasets into cloud. The datasets usually are encrypted before outsourcing to preserve the privacy. However, the common practice of encryption makes the effective utilization difficult, for example, search the given keywords in the encrypted datasets. Many schemes are proposed to make encrypted data searchable based on keywords. However, keyword-based search schemes ignore the semantic representation information of users retrieval, and cannot completely meet with users search intention. Therefore, how to design a content-based search scheme and make semantic search more effective and context-aware is a difficult challenge. In this paper, we proposed an innovative semantic search scheme based on the concept hierarchy and the semantic relationship between concepts in the encrypted datasets. More specifically, our scheme first indexes the documents and builds trapdoor based on the concept hierarchy. To further improve the search efficiency, we utilize a tree-based index structure to organize all the document index vectors. Our experiment results based on the real world datasets show the scheme is more efficient than previous scheme. We also study the threat model of our approach and prove it does not introduce any security risk.

Title: Towards A dynamic secure group sharing framework in public cloud computing

Author: Kaiping. Xue and Peilin. Hong

Abstract: With the popularity of group data sharing in public cloud computing, the privacy and security of group sharing data have become two major issues. The cloud provider cannot be treated as a trusted third party because of its semi-trust nature, and thus the traditional security models cannot be straightforwardly generalized into cloud based group sharing frameworks. In this paper, we propose a novel secure group sharing framework for public cloud, which can effectively take advantage of the cloud servers' help but have no sensitive data being exposed to attackers and the cloud provider. The framework combines proxy signature, enhanced TGDH and proxy re-encryption together into a protocol. By applying the proxy signature technique, the group leader can effectively grant the privilege of group management to one or more chosen group members. The enhanced TGDH scheme enables the group to negotiate and update the group key pairs with the help of cloud servers, which does not require all of the group members been online all the time. By adopting proxy re-encryption, most computationally intensive operations can be delegated to cloud servers without disclosing any private information. Extensive security and performance analysis shows that our proposed scheme is highly efficient and satisfies the security requirements for public cloud based secure group sharing.

Title: Attribute-based access to scalable media in cloud-assisted content sharing

Author: Y. Wu, Z. Wei, and H. Deng

Abstract: This paper presents a novel Multi-message Ciphertext Policy Attribute-Based Encryption (MCP-ABE) technique, and employs the MCP-ABE to design an access control scheme for sharing scalable media based on data consumers' attributes (e.g., age, nationality, or gender) rather than an explicit list of the consumers' names. The scheme is efficient and flexible because MCP-ABE allows a content provider to specify an access policy and encrypt multiple messages within one ciphertext such that only the users whose attributes satisfy the access policy can decrypt the ciphertext. Moreover, the paper shows how to support resource-limited mobile devices by offloading computational intensive operations to cloud servers while without compromising data privacy.

Title : Hierarchical attribute-based encryption for fine-grained access control in cloud storage services

Author : G. Wang, Q. Liu, and J. Wu,

Abstract: Cloud computing, as an emerging computing paradigm, enables users to remotely store their data into a cloud so as to enjoy scalable services on-demand. Especially for small and medium-sized enterprises with limited budgets, they can achieve cost savings and productivity enhancements by using cloud-based services to manage projects, to make collaborations, and the like. However, allowing cloud service providers (CSPs), which are not in the same trusted domains as enterprise users, to take care of confidential data, may raise potential security and privacy issues. To keep the sensitive user data confidential against untrusted CSPs, a natural way is to apply cryptographic approaches, by disclosing

decryption keys only to authorized users. However, when enterprise users outsource confidential data for sharing on cloud servers, the adopted encryption system should not only support fine-grained access control, but also provide high performance, full delegation, and scalability, so as to best serve the needs of accessing data anytime and anywhere, delegating within enterprises, and achieving a dynamic set of users. In this paper, we propose a scheme to help enterprises to efficiently share confidential data on cloud servers. We achieve this goal by first combining the hierarchical identity-based encryption (HIBE) system and the ciphertext-policy attribute-based encryption (CP-ABE) system, and then making a performance-expressivity tradeoff, finally applying proxy re-encryption and lazy re-encryption to our scheme.

Title : Identity-based cryptosystems and signature schemes

Author : A. Shamir

Abstract: In this paper we introduce a novel type of cryptographic scheme, which enables any pair of users to communicate securely and to verify each

other's signatures without exchanging private or public keys, without keeping key directories, and without using the services of a third party. The scheme assumes the existence of trusted key generation centers, whose sole purpose is to give each user a personalized smart card when he first joins the network. The information embedded in this card enables the user to sign and encrypt the messages he sends and to decrypt and verify the messages he receives in a totally independent way, regardless of the identity of the other party. Previously issued cards do not have to be updated when new users join the network, and the various centers do not have to coordinate their activities or even to keep a user list. The centers can be closed after all the cards are issued, and the network can continue to function in a completely decentralized way for an indefinite period.

Title : Public key encryption with keyword search

Author : D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano

Abstract: We study the problem of searching on data that is encrypted using a public key system. Consider user Bob who sends email to user Alice encrypted under Alice's public key. An email gateway wants to test whether the email contains the keyword "urgent" so that it could route the email accordingly. Alice, on the other hand does not wish to give the gateway the ability to decrypt all her messages. We define and construct a mechanism that enables Alice to provide a key to the gateway that enables the gateway to test whether the word "urgent" is a keyword in the email without learning anything else about the email. We refer to this mechanism as Public Key Encryption with keyword Search. As another example, consider a mail server that stores various messages publicly encrypted for Alice by others. Using our mechanism Alice can send the mail server a key that will enable the server to identify all messages containing some specific keyword, but learn nothing else. We define the concept of public key encryption with keyword search and give several constructions.

3. EXISTING SYSTEM

In an attribute-based searchable encryption (ABSE) scheme, data owners can encrypt their data with access policy for security consideration, and encrypt keywords to obtain keyword index for privacy keyword search, and data users can search interesting keyword on keyword indexes by keyword search trapdoor. However, many existing searchable encryption schemes only support single keyword search and most of existing attribute-based encryption (ABE) schemes have high computational costs at user client. These problems significantly limit the application of attribute-based searchable encryption schemes in practice.

4. PROPOSED SYSTEM

In this paper, we propose a user-defined privacy grid system called dynamic grid system (DGS) to provide privacy preserving snapshot and continuous. The main idea is to place a semi trusted third party, termed query server (QS), between the user and the service provider (SP). QS only needs to be semi-trusted because it will not collect/store or even have access to any user location information. In our scheme multi-keyword can be searched, and the search privacy is protected. That is, CS can search the multi-keyword with keyword search trapdoor but it does not know any information about the keywords searched. Considering that keyword search is indispensable for ABE in practice, and our scheme supports multiple keyword search, so our scheme is also a combination of ABE and SE.

5. ARCHITECTURE

The Structural design of the proposed system is shown in the figure-1:

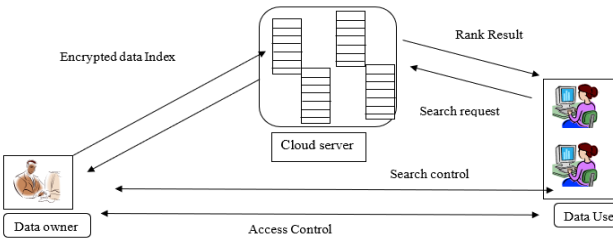

Fig 1: System Design

6. MODULE DESCRIPTION

6.1 USER

The data consumer (User) is assigned a global user identity Uid by CA. The user possesses a set of attributes and is equipped with a secret key associated with his/her attribute set. The user can freely get any interested encrypted data from the cloud server. However, the user can decrypt the encrypted data if and only if his/her attribute set satisfies the access policy embedded in the encrypted data.

6.2 OWNER

The data owner (Owner) defines the access policy about who can get access to each file, and encrypts the file under the defined policy. First of all, each owner encrypts his/her data with a symmetric encryption algorithm. Then, the owner formulates access policy over an attribute set and encrypts the symmetric key under the policy according to public keys obtained from CA. After that, the owner sends the whole encrypted data and the encrypted symmetric key (denoted as ciphertext CT) to the cloud server to be stored in the cloud.

6.3 ADMIN

Admin is a super user. they can view all the user and owner details. Admin can view the chart based on most number of word search, they can add related word, so user can easily mapping a related words for example Ambiguity level 2 refers to instances that most people think as ambiguous. These instances contain two or more unrelated senses, such as “apple” (fruit & company) and “jaguar” (animal & company). In this work, we only focus on disambiguation of instances.

6.4 ATTRIBUTE AUTHORITY

The attribute authorities (AAs) are responsible for performing user legitimacy verification and generating intermediate keys for legitimacy verified users. Unlike most of the existing multi-authority schemes where each AA manages a disjoint attribute set respectively, our proposed scheme involves multiple authorities to share the responsibility of user legitimacy verification and each AA can perform this process for any user independently. When an AA is selected, it will verify the users’ legitimate attributes by manual labor or authentication protocols, and generate an intermediate key associated with the attributes that it has legitimacy-verified. Intermediate key is a new concept to assist CA to generate keys.

6.5 CENTRAL AUTHORITY

The central authority (CA) is the administrator of the entire system. It is responsible for the system construction by setting up the system parameters and generating public key for each attribute of the universal attribute set. In the system initialization phase, it assigns each user a unique Uid and each attribute authority a unique Aid. For a key request from a user, CA is responsible for generating secret keys for the user on the basis of the received intermediate key associated with the user’s legitimate attributes verified by an AA. As an administrator of the entire system, CA has the capacity to trace which AA has incorrectly or maliciously verified a user and has granted illegitimate attribute sets.The cloud server provides a public platform for owners to store and share their encrypted data. The cloud server doesn’t conduct data access control for owners. The encrypted data stored in the cloud server can be downloaded freely by any user.

7. RESULTS

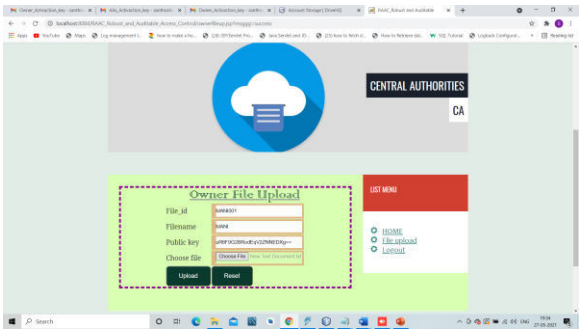

Fig.2: Owner File Upload

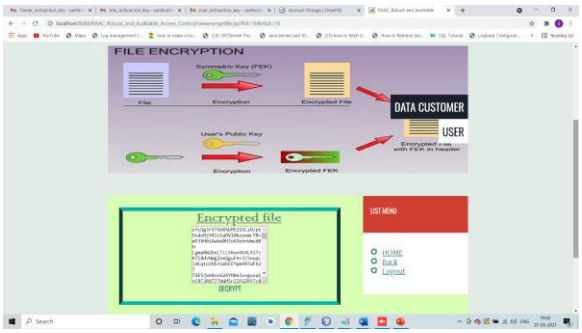

Fig 6: Encrypted File

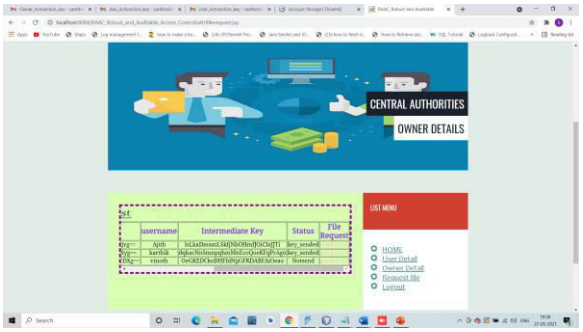

Fig 3: User request send to CA from AA

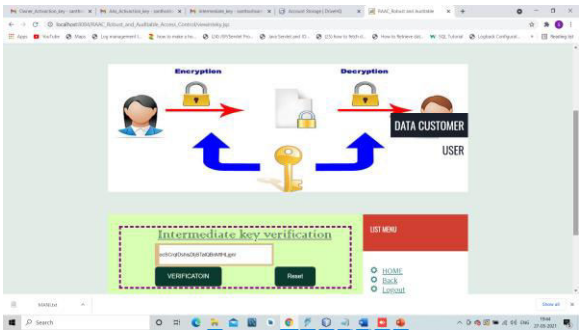

Fig 7: Intermediate key verification

Fig 4: Requested file send to user from CA

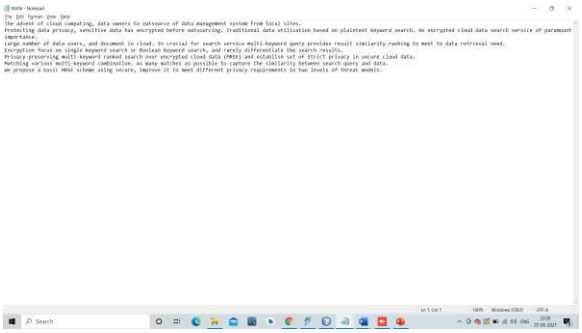

Fig 8: Decrypted File

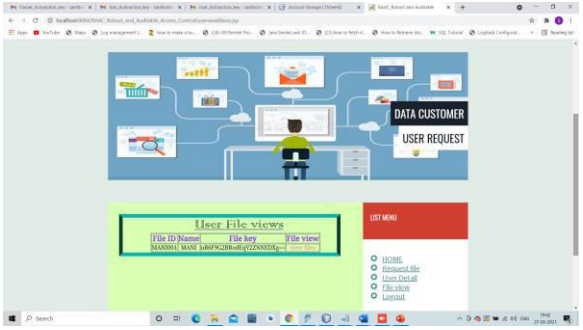

Fig 5: User File Views

8. CONCLUSION

We define and solve the problem of multi-keyword ranked search over encrypted cloud data, and begin a variety of privacy requirements. Among different multi-keyword semantics, we choose the efficient principle of "coordinate matching", as many matches as possible, to effectively capture similarity between query keywords and outsourced documents, and use "inner product similarity" to quantitatively formalize such a principle for similarity measurement. For meeting the challenge of supporting multi-keyword semantic without privacy breaches, the propose a basic MRSE scheme using secure inner product computation, and significantly

improve it to achieve privacy requirements in two levels of threat models. By effectively reformulating CPABE cryptographic technique into our novel framework, our proposed scheme provides a fine-grained, robust and efficient access control with one-CA/multi-AAs for public cloud storage. Our scheme employs multiple AAs to share the load of the time-consuming legitimacy verification and standby for serving new arrivals of users' requests. We also proposed an auditing method to trace an attribute authority's potential misbehavior. We conducted detailed security and performance analysis to verify that our scheme is secure and efficient.

REFERENCES

1. D.X. Song, D. Wanger, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," IEEE Symposium on Security & Privacy, Washington, DC, USA: IEEE Computer Society, May 2000, pp. 44-55.

2. C. Dong, G. Russello, N. Dulay, "Shared and Searchable Encrypted Data for Untrusted Servers," Lecture Notes in Computer Science, Berlin, Germany: Springer, Jul. 2008, pp.127-143.

3. S. Li and M. Xu, "Attribute-Based Public Encryption with Keyword Search," Chinese Journal of Computers, vol. 37, no. 5, 1018-1024, Jun. 2014, doi: 10.3724/SP.J.1016.2014.01017.

4. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," Journal of Computer Security, vol. 19, no. 5, 2011, pp. 895-93.

5. M. Bellare, A. Boldyreva, and A. O'Neill, "Deterministic and Efficiently Searchable Encryption," the 27th annual international cryptology conference on Advances in cryptology, Berlin, Germany: Springer, Aug. 2007, pp. 535-552.

6. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," International Journal of Engineering Research and Applications, vol. 4, no. 7, pp. 441-445, May 2014, doi: 10.1109/INFCOM.2010.5462196.

7. W. Sun, S. Yu, W. Lou, Y. Hou, and H. Li, "Protecting Your Right: Verifiable Attribute-Based Keyword Search with Fine-Grained OwnerEnforced Search Authorization in the Cloud," IEEE INFOCOM, vol. 27, no. 4, pp. 226-234, Jul. 2014, doi: 10.1109/INFOCOM.2014.6847943.

8. Q. Dong, Z. Guan, and Z. Chen, "Attribute-Based Keyword Search Efficiency Enhancement via an Online/Offline Approach," the 21th IEEE International Conference on Parallel and Distributed Systems, Dec. 2015, pp. 298-305.

9. Y. Ye, J. Han, W. Susilo, T. H. Yuen, and J. Li, "ABKS-CSC: attributebased keyword search with constant-size ciphertexts," Security & Communication Networks, vol. 9, no. 18, pp. 5003-5015, Oct. 2016, doi: org/10.1002/sec.1671.

10. Q. Chai and G. Gong, "Verifiable symmetric searchable encryption for semi-honest-but-curious cloud servers," IEEE International Conference on Communications, Jun. 2012, pp. 917-922.