Security for Encrypted Data in Cloud Server Search

S S Lakshmi Lavanya M [1], G Shruthi [2]
[1, 2] Sree Dattha Institute of Engineering and Science, Hyderabad, India.

Abstract – Cloud computing is service provider technology for software information domains. Accessing data and security of the data is one issue of data in cloud. Searching encrypted data in cloud is an open issue since search is not performed by the data users. The existing search techniques supports for only single or conjunctive keyword and expressive search which were computationally inefficient. In this paper we proposed an attribute keyword search which allows to perform search by the cloud server over the encrypted data in cloud. Security is one of the features of the proposed method and our scheme. The experimental result shows that our technique is better than the existed.

Index Terms – Prime order groups, expressive public key, charm, key word search, cloud, security.

1. INTRODUCTION

Rapid technology development made the information storage easy as per the comfort of the user's requirement. Internet technology became source of many information resources. Retrieving stored data from warehouses, data bases and clouds becoming an effort due to search time, security and extracting required patterns. Encryption is the good tool for information security but searching encrypted data in cloud is an interesting thing. Cloud computing provides SaaS, PaaS and IaaS.

Cloud computing provides utmost security for the stored data in cloud, but getting encrypted data in cloud with security while searching is difficult factor. Many existed methods used for this problem which proved computationally inefficient. These existed methods support single or conjunctive keyword search, expressive keyword only hence we propose an expressive public key search encryption method uses prime order groups predicates, access structures to be expressed in conjunctive, disjunctive or any monotonic Boolean formulas and achieves significant performance improvement over existing schemes.

2. LITERATURE SURVEY

This literature survey gives us an abundant inspiration to carry out our work. During the survey we found different works with various algorithms and their outcomes. We made a comparative study of different key word Searchable methods given below.

"Efficient Similarity Search over Encrypted Data" , Mehmet Kuzu, Mohammad Saiful Islam, Murat Kantarcioglu ,ACM publications, 2012. " Ginix: Generalized Inverted Index for Keyword Search ",Hao Wu , Guoliang Li, and Lizhu Zhou, IEEE Transcation, Volume 18, Number 1, February 2013 Cloud computing, a new terminology used to access data remotely from the centralized pool that can be rapidly deployed with great scalability and less computation overhead. Cloud computing comes with lots of advantages such as self-service on need, easy network access, accessing data independent of location, , rapid resource elasticity, low pricing, transference of risk, etc.

	Keyword Privacy	Expressiveness	Bilinear Group	Security	Unbounded keywords
BCOP04 [7]	keyword guessing attacks on trapdoors	AND	prime	full random oracle	yes
KSW13 [16]	keyword guessing attacks on trapdoors	AND, OR	composite	full standard model	no
LZDLC13 [8]	keyword guessing attacks on trapdoors	AND, OR	composite	full standard model	no
LHZF14 [14]	no keyword guessing attacks on trapdoors	AND, OR, NOT	composite	full standard model	no
Our scheme	keyword guessing attacks on trapdoors by designated server only	AND, OR	prime	selective standard model	yes

Table.1. Comparative study

"K-Gram Based Fuzzy Keyword Search over Encrypted Cloud Computing " Cong Wang1, Qian Wang1, Kui Ren1, and Wenjing Lou2 Presents a real approach having fuzzy logic as base to search data over ciphered cloud data. Here author said that it is the very first approach for the same. Here system returns the file in which exact match of predefined word is get found.

"Above the clouds: A Berkley view of cloud computing" Proposed a new technique known as Secured Multikeyword search (SMS) for searching the input query over ciphered data in cloud. To search files containing documents efficient rule of coordinate matching is used. Coordinate matching refers to a technique of finding the huge number of matches as many as possible so that it will become easy to find similarity between the input query and the list of records. To improve further accuracy they developed an alert system, this system will gives an alert whenever an unauthorized users are trying to access the system. This alert is given by the emails and messages to the related users.

3. RELATED WORKS

In this section, we present a brief summary of related works dealing with the searchable encryption schemes. Searchable encryption scheme can be designed in either public cloud or private cloud. The first searchable encryption scheme in public cloud was proposed by Jin Li et al. [1][3]. Public cloud services on resources that are shared between many customers, managed off-premises and scalable homogeneous infrastructure.

Li, Yu, Zheng proposed a framework which enhance an existing MA-ABE scheme to handle efficient and on-demand user revocation, and prove its security. Wenhai Sun*,Hui proposed first attribute-based keyword search technique with user revocation (ABKS-UR) that enables fine-grained (i.e. filelevel) search authorization.

Ren, and W. Lou proposed a novel framework for access control to PHRs within computing environment. To enable fine-grained and scalable access control for PHRs;they used attribute based encryption (ABE) techniques to encrypt each patients' PHR data.

Goyal, Pandey, Sahai, developed a new cryptosystem for finegrained sharing of encrypted data that they call Key-Policy Attribute-Based Encryption (KP-ABE).Ravindra Changala, explained how health care data management can be done with security in hospitals patient health records using big data technologies HIVE, ETL tools in [5].

4. SYSTEM ARCHITECTURE

However, the solution in as well as other existing PEKS schemes which improve on only support equality queries As such, our scheme is not only capable of expressive multi-keyword search, but also significantly more efficient than existing schemes built in composite-order groups. Using a randomness splitting technique, our scheme achieves security against offline keyword dictionary guessing attacks to the cipher texts. Moreover, to preserve the privacy of keywords against offline keyword dictionary guessing attacks to trapdoors, we divide each keyword into keyword name and keyword value and assign a designated cloud server to conduct search operations in our construction. We formalize the security definition of expressive SE, and formally prove that our proposed expressive SE scheme is selectively secure in the standard model.

The approach based on set intersection leaks extra information to the cloud server beyond the results of the conjunctive query, whilst the approach using meta keywords require 2m meta keywords to accommodate all the possible conjunctive queries for m keywords. It is straightforward to see that compared to the existing ones, our construction make a good balance in that it allows unbounded keywords, supports expressive access structures, and is built in the prime-order groups.

We propose an expressive public-key searchable encryption scheme in the prime-order groups, which allows keyword search policies (i.e., predicates, access structures) to be expressed in conjunctive, disjunctive or any monotonic Boolean formulas and achieves significant performance improvement over existing schemes. We formally define its security, and prove that it is selectively secure in the standard model. In order to tackle the keyword search problem in the cloud-based healthcare information system scenario, we resort to public-key encryption with keyword search (PEKS) schemes, which is firstly proposed in . In a PEKS scheme, a ciphertext of the keywords called "PEKS ciphertext" is appended to an encrypted PHR. To retrieve all the encrypted PHRs containing a keyword, say "Diabetes", a user sends a "trapdoor" associated with a search query on the keyword "Diabetes" to the cloud service provider, which selects all the encrypted PHRs containing the keyword "Diabetes" and returns them to the user while without learning the underlying PHRs.

The expressive SE scheme inherits the advantages of the Rouselakis-Waters scheme . Thus, it is straightforward to see that in our SE scheme, the size of the public parameter is immutable with the number of keywords, and the number of the keywords allowed for the system is unlimited and can be freely set. According to the analysis in terms of the pairing-friendly elliptic curves, prime order groups have a clear advantage in the parameter sizes over composite order groups. The advantage for a polynomial time adversary that can distinguish between the games Game0 and Game1 is negligible.

The architecture of our keyword search system is shown in Fig. 1, which is composed of four entities: a trusted trapdoor generation centre who publishes the system parameter and holds a master private key and is responsible for trapdoor generation for the system, data owners who outsource encrypted data to a public cloud, data users who are privileged to search and access encrypted data, and a designated cloud server who executes the keyword search operations for data users. To enable the cloud server to search over ciphertexts, the data owners append every encrypted document with encrypted

keywords. A data user issues a trapdoor request by sending a keyword access structure to the trapdoor generation centre which generates and returns a trapdoor corresponding to the access structure.

We assume that the trapdoor generation centre has a separate authentication mechanism to verify each data user and then issue them the corresponding trapdoors. After obtaining a trapdoor, the data user sends the trapdoor and the corresponding partial hidden access structure (i.e., the access structure without keyword values) to the designated cloud server.

The latter performs the testing operations between each ciphertext and the trapdoor using its private key, and forwards the matching ciphertexts to the data user. As mentioned earlier, a ciphertext created by a data owner consists of two parts: the encrypted document generated using an encryption scheme and the encrypted keywords generated using our SE scheme. From now on, we only consider the latter part of the encrypted document, and ignore the first part since it is out of the scope of this paper.

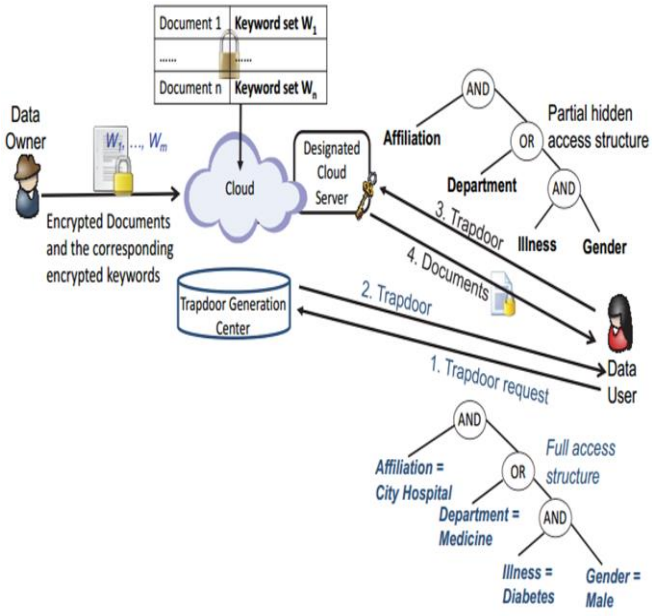

Fig.1. System architecture

In summary, the design goals of our expressive SE scheme are fourfold.

• **Expressiveness.** The proposed scheme should support keyword access structures expressed in any Boolean formula with AND and OR gates.

• **Efficiency.** The proposed scheme should be adequately efficient in terms of computation, communication and storage for practical applications.

• **Keyword privacy.** First, a ciphertext without its corresponding trapdoors should not disclose any information about the keyword values it contains to the cloud server and outsiders. Second, a trapdoor should not leak information on keyword values to any outside attackers without the private key of the designated cloud server. We capture this notion of security for the SE scheme in terms of semantic security to ensure that encrypted data does not reveal any information about the keyword values, which we call "selective indistinguishability against chosen keyword-set attack.

• **Provable security.** The security of the proposed scheme should be formally proved under the standard model rather than the informal analysis.

5. IMPLEMENTATION

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

The implantation can be done through the following steps.

- Searchable Encryption,
- Cloud Computing,
- Expressiveness Keyword search
- Attribute-Based Encryption

6. CLOUD COMPUTING

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services) which can be rapidly provisioned and released with minimal management effort. Cloud-based healthcare information system that hosts outsourced personal health records (PHRs) from various healthcare providers.

The PHRs are encrypted in order to comply with privacy regulations like HIPAA. In order to facilitate data use and sharing, it is highly desirable to have a searchable encryption (SE) scheme which allows the cloud service provider to search over encrypted PHRs on behalf of the authorized users (such as medical researchers or doctors) without learning information about the underlying plaintext. Note that the context we are

considering supports private data sharing among multiple data providers and multiple data users.

7. ATTRIBUTE-BASED ENCRYPTION

Attribute-based encryption is a type of public-key encryption in which the secret key of a user and the ciphertext are dependent upon attributes (e.g. the country in which he lives, or the kind of subscription he has). The basic idea of our scheme is to modify a key-policy attributed-based encryption (KP-ABE) scheme constructed from bilinear pairing over prime-order groups. Without loss of generality, we will use the large universe KP-ABE scheme selectively secure in the standard model proposed by Rouselakis and Waters in [to illustrate our construction during the rest of the paper.

8. EXPRESSIVENESS KEYWORD

The proposed scheme should support keyword access structures expressed in any Boolean formula with AND and OR gates. Efficiency. The proposed scheme should be adequately efficient in terms of computation, communication and storage for practical applications.

A ciphertext without its corresponding trapdoors should not disclose any information about the keyword values it contains to the cloud server and outsiders. Second, a trapdoor should not leak information on keyword values to any outside attackers without the private key of the designated cloud server.

We capture this notion of security for the SE scheme in terms of semantic security to ensure that encrypted data does not reveal any information about the keyword values, which we call "selective indistinguishability against chosen keyword-set.

9. RESUTLS

We implement our scheme in Charm which is a framework developed to facilitate rapid prototyping of cryptographic schemes and protocols. Based on the Python programming language, Charm enables one to implement a cryptographic scheme with very few lines of code, significantly reducing development time. Meanwhile, computationally intensive mathematical operations are implemented with native modules, so the overhead due to Python in Charm is less than 1%. Since all Charm routines are designed under the asymmetric groups, our construction is transformed to the asymmetric setting before the implementation.

The computational costs of the Setup and sKeyGen algorithms are straightforward, and we focus on the computational costs of the Trapdoor, Encrypt and Test algorithms. In our experiments, a set of keywords is generated, of which every keyword contains a generic name such as "Illness", "Position", "Affiliation" and a keyword value such as "Diabetes", "Doctor", and "City Hospital". For the sake of simple implementation, we use integers to denote keyword values, e.g., a keyword as "Illness = 6" is expressed by "Illness = Diabetes". In this way, we generate a random set of keywords containing 10 to 50 keywords, and use them to encrypt 5,000 documents. We then remove the keyword values in the ciphertexts such that they contain only generic names of keywords like "Illness", "Position", as specified in our concrete construction.

10. CONCLUSION

PEKS has given good performance in searching encrypted data in cloud. We used other searchable encryption techniques due to communication overhead, searching and security. Due to computational inability of the existed searchable techniques we proposed public key searching scheme focused on design and analysis of the objective. Attribute based encryption performed expressive searchable encryption system in the primeorder group which supports expressive access structures expressed in any monotonic Boolean formulas. We proved its security with efficiency by analyzing different simulations.

REFERENCES

[1] M .Li, S.Yu, Yao Zheng, Kui Ren, Wenjing Lou ,‖Salable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute Based Encryption‖, IEEE Trans. Parallel and Distributed Systems, Vol.24,No.1 Jan 2013.

[2] Wenhai Sun*,Hui Li*,Shucheng Yu,Thomas Hou,Wenjing Lou,‖Protecting Your Right: Attribute-based Keyword Search with Finegrained Owner-enforced Search Authorization in the Cloud‖, Proc.IEEE 978-1-4799-3360-0,2014.

[3] Ravindra Changala, Automated Health Care Management System Using Big Data Technology, at Journal of Network Communications and Emerging Technologies (JNCET), Volume 6, Issue 4, April (2016), 2016, pp.37-40,ISSN: 2395-5317, ©EverScience Publications.

[4] M. Li, S. Yu, K. Ren, and W. Lou, ―Securing Personal Health Records in Cloud Computing: Patient-Centric and Fine-Grained Data Access Control in Multi-Owner Settings‖, Proc. Sixth Intl ICST Conf. Security and Privacy in Comm. Networks (SecureComm 10), pp. 89-106, Sept. 2010.

[5] avindra Changala, "Data Mining Techniques for Cloud Technology" in International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE),Volume 4, Issue 8, Pages 2319-5940, ISSN: 2278-1021, Agust 2015 by Tejass Publisheers.

[6] Ravindra Changala, "Retrieval of Valid Information from Clustered and Distributed Databases" in Journal of innovations in computer science and engineering (JICSE), Volume 6, Issue 1,Pages 21-25, Septemebr 2016.ISSN: 2455-3506.

[7] V.Goyal, O. Pandey,A. Sahai, and B.Waters, Attribute-based encryption for fine-grained access control of encrypted data‖, in Proc. ACM Conf.Computer and Communications Security, 2006, pp. 8998.