A Survey on Searching of Keyword on Encrypted Data in Cloud Using Access Structure

Miss. Priyanka M. Abhale[1], Prof. M. B. Vaidya [2]

[1] P.G. Student, Department of Computer Engineering, Amrutvahini College of Engineering , Sangamner
[2] Associate Professor, Department of Computer Engineering ,Amrutvahini College of Engineering , Sangamner

---

Abstract - *The cloud computing is very popular as it gives more storage solution over the existing systems. In cloud data outsourcing is done using the encryption technique this gives the protection and secure data against unauthorized parties. Data in the cloud is in encrypted format and it is difficult to retrieve from cloud server. In this paper user request the data using access structure in the form of predicates are expressly in Boolean formulation, conjunctive, disjunctive and achieve the resultant document from server data storage. The main aim of this paper is searching of keyword from the cloud in the form encrypted data using cloud data storage this Multi-keyword ranked search is the beneficial technique for looking encrypted data inside the Cloud. The data duplication of documents is avoided and also it gives the top-K documents to the request users by analyzing this documents user find in which document more data is available for the requested structure then it downloads only this document. In this paper the new concept that is user interest model is designed this model is used for storing the history of searching data it works like cookies in the web. Different algorithms are RSBS is about to reduce the search time and AES is for encryption of data and also to avoid duplication specific duplication algorithms are used. This system gives the significant performance of the existing schemes.*

Key Words: **Cloud computing, Trapdoor, access structure, expressiveness, encrypted keywords.**

1. INTRODUCTION

Recently as a new industrial version, computers that do work for you, but that are stored somewhere else and maintained by other companies has attracted lots interest from both the world of college and industry. A major gain of cloud is that it useful things supplies completely and totally unlimited storage abilities and elastic aid provisioning. In order to reduce the capital and operational costs for hardware and software program, plenty of IT businesses and people are paying someone else to do something their statistics to cloud servers instead of building and keeping their own statistics facilities. Cloud computing has been thought about as a new version of large business IT basic equipment needed for a business or society to operate, that can organize big aid of calculating, storage and packages, and enable customers to enjoy existing everywhere, convenient and on demand network access to a shared pool of configurable calculating useful things supplies with excellent wasting very little while working or producing something and very little money-based overhead.

In the realistic programs, look for predicates (i.e., guidelines) should be communicating a lot of thought or emotion such that they can be expressed as not having a connection, conjunction, or any Boolean system of very important phrases. In the above cloud-based college system, to and the connection among department and stud name or class, person who works to find information may also problem to search question with an access structure (i.e. Predicate) (department = computer AND (stud name = Ram OR class = ME)). In order to help data use and sharing, it is incredibly clearly connected with or related to have a Searchable Encryption (SE) layout which permits the cloud provider company to look over unreadable data for the legal clients (which include scientific researchers or college authority) without studying statistics about the hidden plaintext[1].

The cloud provider providers (CSP) that preserve the statistics for users might also get speaking the truth about something bad to customers sensitive records on the not being there present of known approval. A famous approach to protect the facts confidentiality is to turn into secret code the statistics in advance than paying someone else to do something. However, this can purpose a big cost in phrases of statistics usability. Downloading all the statistics from the cloud and change secret codes into readable messages within a large area is obviously not having common sense. In this paper, we recommend a public-key based totally communicating a lot of thought or emotion SE layout in most important-order groups, that's specially good for key-word searching for over unreadable in situations of a couple of data owners and many statistics users which include the cloud-based college records system that hosts paid someone else to do something available data from many colleges or from different colleges.

2. RELATED WORK

Wang, N. Cao, K. Ren, and W. Lou [2] This paper define the problem of secure ranked key-word search over encrypted cloud facts and give effective protocol this will satisfy the secure ranked search operations using some piece of information over keyword The data owner outsources encrypted files and their index to the cloud server then this encrypted file converted in byte stream. The data user send search request to server after that server identifies the particular user and sends files using ranking to users.

Zhangjie Fu, Xinle Wu, Chaowen Guan and Xingming Sun,[3] This paper propose an efficient multi keyword fuzzy ranked search scheme this is capable of address the above point out issues. First, we develop a brand new method of keyword transformation primarily based on the uni-gram, if you want to concurrently improve the accuracy and creates the capability to handle other spelling mistakes. The design goal are it help spelling errors like netward or netwrok , it take guarantee of privacy and maintain safety, the support of report and key-word updating, it generate the result in step with score. Stemming algorithm, Bloom Filter, Locality-Sensitive Hashing (LSH) this 3 essential strategies are used in this layout. The stemming, bloom and encryption is achieve by the usage of trapdoor technology center. The seek time and index production is important in trapdoor.

Zhihua Xia, Xinhui Wang, Xingming Sun, and Qian Wang[4] In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. The generation of query and also construction of index merge the vector model and TF-IDF model. In this Greedy Depth-first Search algorithm and KNN algorithm are used for tree based construction of keyword searching in multi keyword rank search and to encrypt the index and query. This paper describe the unencrypted dynamic multi-keyword ranked search (UDMRS) scheme based on two secure search schemes BDMRS and EDMRS schemes. The UDMRS scheme is constructed using the KNN algorithm and the privacy preserving is achieved by BDMRS scheme.

M. Kuzu, M. S. Islam, and M. Kantarcioglu [5] This paper offers an efficient scheme for similarity seek over encrypted facts. To achieve this, this utilize a state-of-threat algorithm for fast near neighbor seek in high dimensional spaces referred to as locality sensitive hashing. To make certain the confidentiality of the touchy information, This paper offer a rigorous safety definition and show the security of the scheme beneath the provided definition. In addition, this paper provide a actual international software of this scheme and verify the theoretical results with empirical observations on a actual dataset. Locality Sensitive Hashing (LSH) is an approximation algorithm for near neighbor search in high dimensional spaces . The basic idea of LSH is to use a set of hash functions to map objects into several buckets such that similar objects share a bucket with high probability.

3. ALGORITHMS

The Expressive keyword search scheme consists of five algorithm Setup, sKeyGen, Trapdoor, Encrypt , Test.

1. Setup(λ ,U) (PK,MSK). The setup algorithm takes the security parameter and the attribute universe description U as the input. It outputs the public parameters PK and a master secret key MSK.

2. Key Gen(MSK, S) SK. The key generation algorithm takes the master secret key MSK and a set of attributes S as input. It outputs a secret key SK.

3. Trapdoor(pars, pks, msk)TM. Taking the public parameter pars, the server public key pks and an access structure as the input, generates a trapdoor TM. This algorithm is run by the trapdoor centre.

4. Encrypt(PK,M,A) CT. The encryption algorithm takes the public parameters PK, a message M, and an access structure A as input. The algorithm will encrypt M and produce a ciphertext CT .

5. Test(pars, sks, CT, TM) → 1/0. Taking the public parameter pars, the server private key sks, a ciphertext CT associated with a keywords set W and a trapdoor TM for an access structure as the input, and generate outputs either 1 when the ciphertext satisfies the access structure of the trapdoor TM or 0 otherwise. This algorithm is run by the designated server.

A. Ranked Serial Binary Search (RSBS) algorithm:

Input : Noised trapdoor: t1
The number of document to return: k
Encrypted record indexes: E
Output : Document request: D

1. Create the scores as an N zeros
2. For i = 1 to N do
3. For n = 1 to E do
4. Find the keywords seems in any of the s slice of the report
5. End for
6. End for
7. Sorted, indices=sort, (scores)
8. Acquire the top-k files
9. Go back D

4. SYSTEM OVERVIEW

The system architecture of key-word search is proven in Figure, which is composed of 5 entities: a trusted trapdoor technology center who publishes the machine parameter and holds a master non-public key and is answerable for trapdoor era for the system, records owners who outsource encrypted information to a public cloud, data customers who're privileged to go looking and get right of entry to encrypted information, and a designated cloud server who executes the key-word seek operations for records users. To permit the cloud server to look over ciphertexts, the data proprietors append each encrypted file with encrypted key phrases. A records consumer issues a trapdoor request via sending a key-word get admission to structure to the trapdoor generation middle which generates and returns a trapdoor corresponding to the access structure.

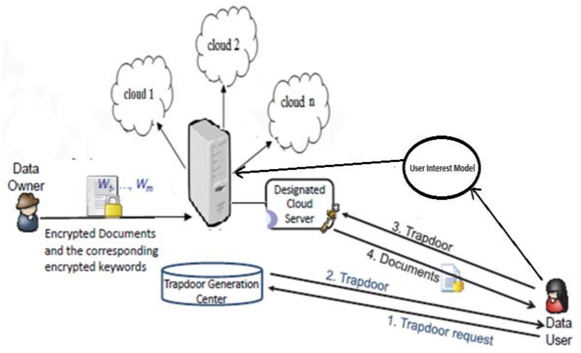

Fig -1: Architecture of the System

The trapdoor generation center has a separate authentication mechanism to confirm every facts person and then issues them the corresponding trapdoors. Obtaining a trapdoor, the facts consumer sends the trapdoor and the corresponding partial hidden access structure (i.e., the access structure without keyword values) to the targeted cloud server. The latter plays the testing operations among every ciphertext and the trapdoor the usage of its non-public key, and forwards the matching ciphertexts to the facts user. Data proprietor includes two elements: the encrypted record generated the usage of an encryption scheme and the encrypted key phrases generated the use of our SE scheme. [1]

Working:

In the proposed work the first expressive SE scheme in the public-key putting from bilinear pairings in prime order corporations. As such, our scheme isn't most effective able to expressive multi-key-word search, but additionally significantly more efficient than existing schemes constructed in composite-order companies. Using a randomness splitting method, our scheme achieves security against offline key-word dictionary guessing attacks to the cipher texts. Moreover, to maintain the privacy of keywords in opposition to offline keyword dictionary guessing attack to trapdoors, we divide each keyword into keyword call and key-word cost and assign a chosen cloud server to behavior seek operations in our creation [1]. A trusted trapdoor generation center who publishes the device parameter and holds a master non-public key and is answerable for trapdoor era for the system, records proprietors who outsource encrypted facts to a public cloud, records customers who are privileged to search and get entry to encrypted statistics, and a designated cloud server who executes the key-word seek operations for information customers.

5. SYSTEM REQUIREMENTS

A. Software Requirement

1) Operating System : Windows family
2) Application Server : NetBeans IDE 8.2, Glassfish
3) Front End : HTML, Java, Jsp
4) Scripts : JavaScript
5) Server side Script : Java Server Pages.
6) Database : My sql 5.5
7) Database Connectivity : JDBC.

B. Hardware Requirement

1) Processor : Intel
2) CPU Speed : 1.1 GHz or Higher
3) RAM : 2 GB or Higher
4) Hard Disk : 20 GB or Higher

6. IMPLIMENTATION DETAILS

We implement our scheme in java based on the java programming language. For testing we have hosted each entity on different machines. The cloud TPA has core i-3 processor with 4 gb RAM. Client system has i3 processor with 2 gb ram. On every system java runtime environment JRE-1.7 is installed. For development we have used jdk 1.7 and NetBeans IDE are used. For Database storage we have used mysql 5.3 database and also we have done the JDBC database connectivity.

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 "Department", "Position", "Affiliation" and a keyword value such as "Computer", "HOD", and "AVCOE". For the simple implementation, we use integers to denote keyword values, e.g., a keyword as "Department = 3" is expressed by "Department = Computer". In this way we generate a random set of keywords containing 10 to 20 keywords and use them to encrypt 2,000 documents. We then remove the keyword values in the ciphertexts such that they contain only generic names of keywords like "Department", "Position", as specified in our concrete construction.

We implement expressive SE in the prime-order group, which is a programming environment for rapid primitives.

Table -1: Performance analysis of file node allocation

Node id	Allocated size	Free size	Total size
Node 1	11067	8933	20000
Node 2	0	20000	20000
Node 3	87	19913	20000
Node 4	18040	1960	20000

In this section, we show the expected result of comparing the computational cost, communication and storage overhead of our scheme with other existing schemes.

Table -2: Keyword searching in different scheme

Keyword	Index Keyword	Our scheme	Wang's scheme
Talking	Talk	Yes	No
achievement	Achieve	Yes	No
Thing	Night	No	Yes
TPA	TPA	Yes	No

After complete system implementation we will evaluate the system performance with

- Upload download time for different files
- File share and key allotment times

7. CONCLUSIONS

In this paper, we focus on improving the efficiency and the security of multi-keyword top-k similarity search over encrypted data. Then, in order to improve the search efficiency, we design the group multi-keyword top-k search scheme, which divides the dictionary into multiple groups and only needs to store the top-k documents of each word group when building index. . In future this proposed key-word searching technique proves efficient and return top k relevant documents or files corresponding to submitted search terms. This proposed system reduces the searching time the usage of Ranked Serial Binary Search (RSBS) set of rules.

ACKNOWLEDGEMENT

I would like to thank my project guide Prof M. B. Vaidya and project coordinator prof. S. K. Sonkar, Department of Computer Engineering, AVCOE, Sangamner for the valuable advice, support and the interest shown in this project by timely suggestions in this work. I would also like to thank my teachers for their encouragement, guidance, understanding and support.

REFERENCES

[1] Hui Cui, Zhiguo Wan, Robert H. Deng, Guilin Wang, and Yingjiu Li, "Effcient and Expressive Keyword Search Over Encrypted Data in Cloud", Transactions on Dependable and Secure Computing Journal Of , Vol. , No., 2016.

[2] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data", IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 8, pp. 14671479, Aug. 2012.

[3] Zhangjie Fu, Member, IEEE, Xinle Wu, Chaowen Guan, Xingming Sun, Senior Member, IEEE, and Kui Ren, Fellow, IEEE, "Toward Efcient MultiKeyword Fuzzy Search Over Encrypted Outsourced Data With Accuracy Improvement", IEEE Transactions On Information Forensics And Security, Vol. 11, No. 12, December 2016.

[4] Zhihua Xia, Member, IEEE, Xinhui Wang, Xingming Sun, Senior Member, IEEE, and Qian Wang, Member, IEEE, "A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data", IEEE Transactions On Parallel And Distributed Systems, Vol. 27, No. 2, February 2016.

[5] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data", in Proc. IEEE 28th Int. Conf. Data Eng., 2012, pp. 11561167.

[6] Priya S , Ambika R, "A Multi-keyword Ranked Search Scheme that is Dynamic and Secure over Cloud Data", International Journal of Innovative Research in Science, Engineering and Technology Vol. 5, Special Issue 10, May 2016.

[7] W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, "Secure ranked multi-keyword search for multiple data owners in cloud computing", in Dependable Syst. Networks (DSN), IEEE 44th Annu. IEEE/IFIP Int. Conf., 2014, pp. 276286.

[8] Xiaofeng Ding, Member, IEEE, Peng Liu and Hai Jin, Senior Member, IEEE, "Privacy-Preserving Multi-keyword Top-k Similarity Search Over Encrypted Data", IEEE Transactions on Dependable and Secure Computing,2016.

[9] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search", in Advances in Cryptology Euro crypt 2004. Springer, 2004, pp. 506522.

[10] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved denitions and effcient constructions", in Proceedings of the 13th ACM Conference on Computer and Communications Security. ACM, 2006, pp. 7988.

[11] E. Shen, E. Shi, and B. Waters, "Predicate privacy in encryption systems", in Theory of Cryptography. Springer, 2009, pp. 457473.

[12] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi keyword ranked search over encrypted cloud data", IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 1, pp. 222233, 2014.

[13] J. Baek, R. Safavi-Naini, and W. Susilo, "Public key encryption with keyword search revisited", in Computational Science and Its Applications. Springer, 2008, pp. 12491259.