Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7, July 2021: 5776 - 5783

Research Article

Secure and Accurate Multi-Keyword Retrieval With Privacy Protection for Multiple Data Owners in Cloud Computing

Miss. Supriya Shete, Prof. Nagaraju Bogiri
(Computer Department, K J College of Engineering and Management Research Pune, India

Abstract

Cloud computing is a remarkable technology, due to the rapid rise in demand for infinite storage and high-quality retrieval services. Numerous research studies on privacy prob-lems have been conducted using a variety of keyword searches. Cloud storage is encrypted in compliance with the definition of multiple data owners. But most of these systems are vulnerable to threats and test attacks with keywords. Moreover, the top search results can be returned correctly From individual data owners. These inconveniences will obviously result in privacy-sensitive leakage Keywords and inaccurate of search results returned. in this paper, recovery system with various keyword confidentiality data owners that allow the cloud server to scan the cloud over multiple words and then return the relevant files to data customer search results without any keyword leakage. Moreover, we prove rigorous Safety analysis that our system is safe from attacks started by indoor and outdoor assailants. In conclusion, The performance assessment shows that our system has more efficient than the current graded keyword search system.

Index Terms—Cloud computing, Encryption, multi-keyword retrieval, multiple data owner model, privacy protection

I. INTRODUCTION

A. Overview

In cloud computing, individuals and companies will benefit from high-quality storage facilities by moving locally deployed software and providing a service to cloud servers, thus significantly alleviating the responsibility of local data processing and maintenance. However, given that data owners are not capable of entirely completing their outsourcing data in real time, the main challenges with data protection and privacy are the extensive use of cloud computing. The most common means of safeguarding data privacy is to send confidential data to the server as a cyphertext. Data protection does not allow data users to decrypt encrypted information on the server effectively. An ingenuous solution such as copying and local deciphering of all encrypted data for the search would significantly reduce and reduce cloud usefulness. As a result, a key issue that needs to be solved immediately is how to quickly and cost-effectively research encrypted target data.

For multiple data owners, we suggest a new, classified multi-keyword retrieval with security of privacy. The linear splitting technique is used to conceal the keyword behind any randomization to prevent keyword guessing attacks initiated by the server cloud. As a consequence, the cloud service can’t derive information on both ciphertext and trapdoor which is vulnerable to privacy. In addition, the cloud server

may avoid this splitting strategy from assessing if both trapdoors have the same collection of keywords. A updated keyword balanced binary tree (KBB-tree) is first built to allow the cloud server to search for encrypted data from various data owner systems and to provide relevant searches to data users.

II. LITERATURE SURVEY

Q in long Huang, Yixian Yang, Wei Yue and Yue He: In this paper [1], In this paper, authors propose a fast multi-keyword semantic ranked search scheme. Firstly, for the first time, the concept of weighted domain scoring is introduced to searchable encryption to calculate the document relevance scores. The keywords in different domains (title, abstract, etc.) are measured by different weighted domain score. Secondly, the retrieved keywords are semantically expanded to their synonym sets and the semantic similarities of the synonyms are calculated. Combining the semantic similarity, the weighted domain score and the relevance scores, authors construct the encrypted document index with higher accuracy. To improve the efficiency of MRSE (multi-keyword ranked search over encrypted cloud data), we partition the document index vector into several pieces and generate mark vector according to these pieces. Comparing the document mark vector and the query mark vector, authors effectively filter a large number of irrelevant documents. The time for calculating the relevance scores and ranking is greatly reduced. Finally, document index vectors are partitioned into several sub-vectors, which are encrypted by the matrices with smaller dimensions.

C. Guo, S. Su, K. K. R. Choo, and X. Tang:In the
paper, authors suggest a reliable and efficient scheme to locate the exact closest neighbour over encrypted medical images. Instead of calculating the Euclidean distance, authors reject candidates by computing the lower bound of Euclidean distance that is connected to the mean and standard deviation of results. Unlike other current schemes, our scheme will obtain the exact nearest neighbour rather than an estimated result. authors then test our proposed approach to show its effectiveness.

H. Cui, X. Yi, and S. Nepal: In this paper [3], propose a notion called a proxy-assisted encrypted text policy Attribute-based encryption (PA-CPABE), which outsources most decryption calculations to peripheral devices. Respect For the existing ABE with outsourced decryption schemes (ABE-OD), PA-CPABE has the advantage that the distribution of keys It does not require any secure channel. Author present a generic construction of PA-CPABE and therefore authors demonstrate its security. Moreover, implement an instance of the proposed PA-CPABE framework to evaluate its performance.

K. Xue, W. Chen, W. Li, J. Hong, and P. Hong: In this paper [4], propose a solution to secure encrypted cloud storages from EDoS attacks and provide resource consumption accountability. It uses CP-ABE schemes in a black-box manner and complies with arbitrary access policy of CP-ABE. Present two protocols for different settings, followed by performance and security analysis.

N. Paladi, C. Gehrmann, and A. Michalas: In this paper [5], describe a framework for data and operation security in IaaS, consisting of protocols for a trusted launch of virtual machines and domain-based storage protection. Author continue with an extensive theoretical analysis with proofs about protocol resistance against attacks in the defined threat model. The protocols allow trust to be established by remotely attesting host platform configuration prior to launching guest virtual machines and ensure confidentiality

of data in remote storage, with encryption keys maintained outside of the IaaS domain. Presented experimental results demonstrate the validity and efficiency of the proposed protocols. The framework prototype was implemented on a test bed operating a public electronic health record system, showing that the proposed protocols can be integrated into existing cloud environments.

Q. Huang, Y. Yang, and J. Fu: In this paper [6], propose an identity based data group sharing and dissemination scheme in public cloud, in which data owner could broadcast encrypted data to a group of receivers at one time by specifying these receivers' identities in a convenient and secure way. In order to achieve secure and flexible data group dissemination, author adopt attribute-based and timed-release conditional proxy re-encryption to guarantee that only data disseminators whose attributes satisfy the access policy of encrypted data can disseminate it to other groups after the releasing time by delegating a re-encryption key to cloud server. The re-encryption conditions are associated with attributes and releasing time, which allows data owner to enforce fine-grained and timed-release access control over disseminated cipher texts. The theoretical analysis and experimental results show our proposed scheme makes a tradeoff between computational overhead and expressive dissemination conditions.

L. Jiang, and D. Guo: In this paper [7], based on conditional proxy broadcast re-encryption technology, an encrypted data sharing scheme for secure cloud storage is proposed. The scheme not only achieves broadcast data sharing by taking advantage of broadcast encryption, but also achieves dynamic sharing that enables adding a user to and removing a user from sharing groups dynamically without the need to change encryption public keys. Moreover, by using proxy re-encryption technology, our scheme enables the proxy (cloud server) to directly share encrypted data to the target users without the intervention of data owner while keeping data privacy, so that greatly improves the sharing performance. Meanwhile, the correctness and security is proved, the performance is analyzed and the experimental results are shown to verify the feasibility and efficiency of the proposed scheme.

K. Xu, Y. Guo, L. Guo, Y. Fang, and X. Li: In this paper [8], they attempt to address this issue and study the scenario when a user shares a photo containing individuals other than himself/herself (termed co-photo for short). To prevent possible privacy leakage of a photo, Author design a mechanism to enable each individual in a photo be aware of the posting activity and participate in the decision making on the photo posting. For this purpose, need an efficient facial recognition (FR) system that can recognize everyone in the photo. However, more demanding privacy setting may limit the number of the photos publicly available to train the FR system. To deal with this dilemma, our mechanism attempts to utilize users' private photos to design a personalized FR system specifically trained to differentiate possible photo co-owners without leaking their privacy. Author also develop a distributed consensus based method to reduce the computational complexity and protect the private training set.

L. Fang, L. Yin, Y. Guo, Z. Wang, and Fenzhua Li : This paper [9], propose a bargain based incentive method to resolve the policy conflict problem. We propose a novel pricing system to achieve the balance between privacy loss and sharing benefit. Besides, author introduce a Clark-tax-based punishment mechanism to make sure that no co-owners would act maliciously. Game analysis and user studies are performed to illustrate the effectiveness of our proposed scheme.

Q. Huang, W. Yue, Y. He, and Y. Yang: Data security [10] issue is one of the main obstacles to the wide application
of mobile healthcare social networks (MHSN), since health information is considered to be highly sensitive. In this paper, author introduce a secure data sharing and profile matching scheme for MHSN in cloud computing. The patients can outsource their encrypted health records to cloud storage with identity-based broadcast encryption (IBBE) technique, and share them with a group of doctors in a secure and efficient manner. Author then present an attribute-based conditional data re-encryption construction, which permits the doctors who satisfy the pre-defined conditions in the ciphertext to authorize the cloud platform to convert a ciphertext into a new ciphertext of an identity-based encryption scheme for specialist without leaking any sensitive information.

III. PROPOSED SYSTEM

A. Data user

1) In the proposed scheme, members are people with interests (e.g., bidder, doctors, and businessmen) and want to share data in the cloud.
2) The most worrying problem when users store data in the cloud server is the confidentiality of the outsourced data.
3) In this system, users of the same group conduct a key agreement. Subsequently, a common conference key can be used to encrypt the data that will be uploaded to the cloud to ensure the confidentiality of the outsourced data.
4) Attackers or the semi-trusted cloud server cannot learn any content of the outsourced data without the key.
5) Our scheme uses a technique called group signatures, which allows users in the same group to anonymously share data in the cloud.
6) Data user search files using multi-keyword search.

B. Data Owner

1) data owner is responsible for generating system parameters, managing group members (i.e., uploading member's encrypted data, authorizing group members) and for the fault tolerance detection.
2) The data owner in our scheme is a fully trusted third party to both the cloud and group members.
3) If an external user tries to access files from a different group more than three times then the manager will remove that particular user from the applications.

C. Cloud Service Provider (CSP)

1) CSP provides users with seemingly unlimited storage services.
2) In addition to providing efficient and convenient storage services for users, the cloud can also provide data sharing services.
3) However, the cloud has the characteristic of honest but curious.

4) In other words, the cloud will not deliberately delete or modify the uploaded data of users, but it will be curious to understand the contents of the stored data and the user's identity.

A. Architecture

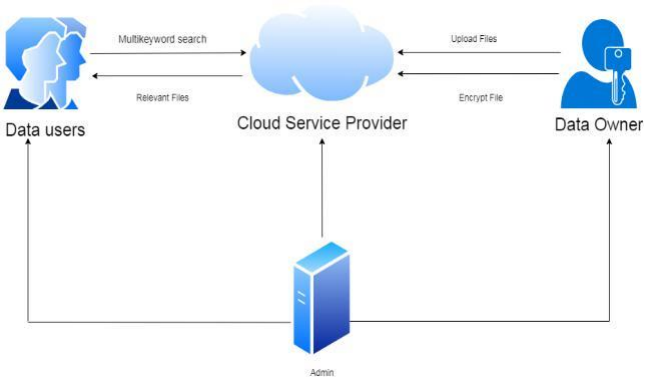

Fig. 1. Proposed System Architecture

B. Algorithms

1. **TFIDF Algorithm:** Terminology:

1. t — term (word)
2. d — document (set of words)
3. N — count of corpus
4. corpus — the total document set

• TF: Term Frequency, which measures how frequently a term occurs in a document. Since every document is different in length, it is possible that a term would appear much more times in long documents than shorter ones. Thus, the term frequency is often divided by the document length (aka. the total number of terms in the document) as a way of normalization:

tf(t; d) = count of t in d=number of words in d

• IDF: Inverse Document Frequency, which measures how important a term is. While computing TF, all terms are considered equally important. However it is known that certain terms, such as ”is”, ”of”, and ”that”, may appear a lot of times but have little importance. Thus we need to weigh down the frequent terms while scale up the rare ones, by computing the following:

idf(t) = log(N=(df + 1))

tf idf(I_i^j)log(tf(I_i^j ; d_j) + 1) log(D=1 + df(I_i^j ; D)

2. AES Algorithm for Encryption.

AES (advanced encryption standard).It is symmetric algorithm. It used to convert plain text into cipher text .The need for coming with this algo is weakness in DES. The 56 bit key of des is no longer safe against attacks based on exhaustive key searches and 64-bit block also consider asweak.AES was to be used128-bit block with128-bit keys.

Rijendeal was founder. In this drop we are using it to encrypt

the data owner file.
Input:
128 bit /192 bit/256 bit input (0, 1) Secret key (128 bit) +plain text (128 bit). Process:
10/12/14-rounds for-128 bit /192 bit/256 bit input
Xor state block (i/p)
Final round:10,12,14
Each round consists: sub byte, shift byte, mix columns, add
round key.
Output:
cipher text(128 bit)

IV. RESULT AND ANALYSIS

For experimental set up, use Windows 7 operating system, Intel i5 processor, 4 GB RAM, 200GB Hard disk, Eclipse oxygen JDK 8 tool and Tomcat server.

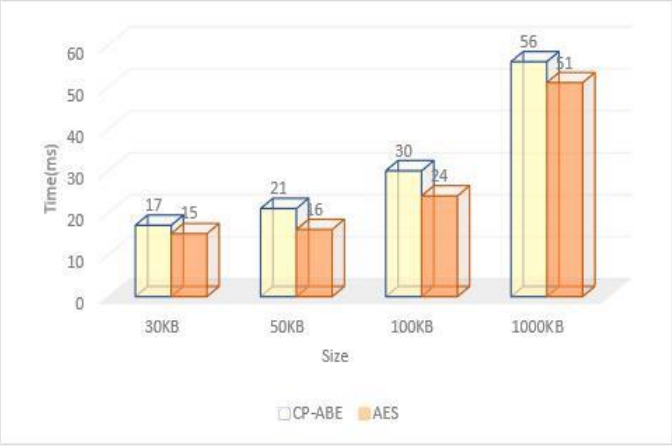

Fig. 2. Encryption and Searching time

In this paper,we compare system execution time with Y.

Yang et.al [1].

Fig. 3. overall system execution

V. CONCLUSION

This article proposes for the first time a new powerful multi-keyword search system with the security of privacy for various cloud data owners. Through delegate its search-ability to cloud repositories for high-ranking multi-keyword extraction, data consumers may obtain relevance results from the encrypted data of various data holders. The anonymity and trapdoor confidentiality of keywords and other malicious attackers are avoided during the whole search process. Later, the robust safety evidence indicates that our suggested scheme is safe regardless of internal and external assaults. Finally, the assessment of results shows that our scheme is more flexible, making our scheme more realistic and practical.

REFERENCES

[1] Y. Yang, J. Liu, S. Cai, and S. Yang, “Fast Multi-keyword Semantic Ranked Search in Cloud Computing,” Chinese Journal of Computers, vol. 41, no. 6, pp. 1126-1139, 2018.

[2] C. Guo, S. Su, K. K. R. Choo, and X. Tang, “A fast nearest neighbor search scheme over outsourced encrypted medical images,” IEEE Trans-actions on Industrial Infor matics, 2018, DOI: 10.1109/TII.2018.2883680.

[3] H. Cui, X. Yi, and S. Nepal, “Achieving scalable access control over encrypted data for edge computing networks,” IEEE Access, vol. 6, pp. 30049–30059, 2018.

[4] K. Xue, W. Chen, W. Li, J. Hong, and P. Hong, “Combining data owner-side and cloud-side access control for encrypted cloud storage,” IEEE Transactions on Information Forensics and Security, vol. 13, no. 8, pp. 2062–2074, 2018.

[5] N. Paladi, C. Gehrmann, and A. Michalas, “Providing user security guarantees in public infrastructure clouds,” IEEE Transactions on Cloud Computing, vol. 5, no. 3, pp. 405-419, 2017.

[6] Q. Huang, Y. Yang, and J. Fu, “Secure data group sharing and dis-semination with attribute and time conditions in Public Clouds,” IEEE Transactions on Services Computing, 2018.

[7] L. Jiang, and D. Guo “Dynamic encrypted data sharing scheme based on conditional proxy broadcast re-encryption for cloud storage,” IEEE Access, vol. 5, pp. 13336 – 13345, 2017.

[8] K. Xu, Y. Guo, L. Guo, Y. Fang, and X. Li, "My privacy my decision: control of photo sharing on online social networks," IEEE Trans. On Dependable and Secure Computing, vol. 14, no. 2, pp. 199-210, 2017.

[9] L. Fang, L. Yin, Y. Guo, Z. Wang, and Fenzhua Li, "Resolving access conflicts: an auction-based incentive approach," Proc. IEEE Military Communications Conference (MILCOM), pp. 1-6, 2018.

[10] Q. Huang, W. Yue, Y. He, and Y. Yang, "Secure identity-based data sharing and profile matching for mobile healthcare social networks in cloud computing," IEEE Access, vol. 6, pp. 36584–36594, 2018.