www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 03, Issue 03, March 2023, pp : 403-406

e-ISSN : 2583-1062

Impact Factor : 5.725

PRIVACY IN THE CLOUD: SOLUTION FOR CURRENT ISSUES AND CHALLENGES IN THE CLOUD

Nandagopal S[1], BharathKumar S[2], Lakshmanakumar V[3], NaveenKumar M[4]

[1]Professor, Department of Information Technology, Nandha College of Technology, Perundurai 638 052, Tamilnadu, India.

[2,3,4] UG Students – Final Year, Department of Information Technology, Nandha College of Technology, Perundurai 638 052, Tamilnadu, India.

ABSTRACT

Data owners are being encouraged to move their complex data management systems from local locations to commercial public clouds for greater flexibility and cost savings with the advent of cloud computing. However, sensitive data must be encrypted before being outsourced in order to protect privacy, which renders plaintext keyword search as a method of data utilization obsolete. Consequently, it is essential to implement an encrypted cloud data search service. For effective data retrieval, it is essential for the search service to support multi-keyword queries and result similarity ranking due to the large number of cloud data users and documents. Studies on searchable encryption tend to concentrate on Boolean keyword searches or searches with just one term, with few attempts made to differentiate the search results. In this project, for the first time, we define and solve the difficult problem of privacy-preserving multi-keyword ranked (EARM). Additionally, we establish stringent privacy requirements that must be met for such a secure cloud data utilization system to become a reality. We choose the effective principle of "Boolean keyword coordinate matching," or as many matches as possible, to capture the similarity between the search query and data documents. Additionally, we employ "inner product similarity" to quantitatively formalize this principle for measuring similarity. First, we present a straightforward EARM system that is based on safe inner product computing. We then significantly expand this system to meet distinct privacy requirements in two threat scenarios. The privacy and efficiency guarantee of the proposed strategies are thoroughly examined, and tests on real-world datasets show that the proposed strategies do not significantly increase computation or communication costs.

Keywords: Data management system, public cloud, multi key word ranked, EARM, privacy requirements, Boolean key word, Inner product similarity, encryption.

1. INTRODUCTION

1.1 The Generic Privacy Preserving Problem

It has become extremely crucial since data has begun to increase a million times quicker, as have the desires to preserve it for oneself privately. When the challenge was given, the web was still in its infancy; today, it is mature, vast, and spread to the farthest reaches of the globe. Back then, the authors noted that rapid advancements in networking, storage, and processing technology had resulted in the construction of ultra-large databases that recorded an unprecedented quantity of transactional data. We are more concerned with preserving privacy in the context of data mining methods. This is one place when privacy can be compromised. Data mining and data warehousing go hand in hand, according to the paper: Most tools work by collecting all data into a single location and then running an algorithm against it. However, privacy issues may impede the construction of a centralized warehouse data may be divided among numerous custodians, none of whom are permitted to move their data to another location.

It should be highlighted that data mining algorithms generate knowledge, and that data mining results seldom breach privacy since they often expose high-level information rather than revealing instances of data. However, privacy activists are right to be concerned, because bringing data together to allow data mining facilitates misuse. The issue is not data mining it, but rather how data mining is carried out.

1.2 Cloud Based Privacy Data Sharing Using Datamining

Because of the continuous increase of data, data owners are increasingly storing their data in the cloud to alleviate the burden of data storage and upkeep. However, because cloud customers and cloud servers are not in the same trusted domain, our outsourced data may be at danger. Thus, sensitive data must be encrypted before being transported to the cloud in order to safeguard data privacy and prevent unauthorized access. Unfortunately, typical plaintext search algorithms can no longer be used directly to encrypted cloud data. Traditional information retrieval (IR) has previously offered data users with multi-keyword ranking ontology keyword mapping and search. Similarly, the cloud server must give a comparable function to the data user while safeguarding data and search privacy. It is only when data can be quickly searched and used that putting it on the cloud server makes sense. According to the literature, searchable encryption approaches can enable users with secure search over encrypted material.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 03, Issue 03, March 2023, pp : 403-406

e-ISSN : 2583-1062

Impact Factor : 5.725

2. LITERATURE REVIEW

2.1 Keyword Searches on Remote Encrypted Data

Maintaining the confidentiality of the keywords themselves without jeopardizing the data's security User U wants to store encrypted data on a far-off file server S. Later, User U wants to access encrypted files with keywords. For instance, a user might want to encrypt old e-mail messages and store them on a Yahoo or other major provider's server so that they can be accessed from a mobile device while traveling. solves this problem within clearly defined security constraints. The methods work well because no public-key cryptosystem is used. Indeed, the distant data's encryption technology has no effect on the procedure. They are also progressive. This can be accomplished by User U by submitting new files that are secure against previous queries but searchable against subsequent queries. The main idea that comes from this is that data can be stored remotely on other servers and accessed from anywhere using a mobile device, laptop, or other device.

2.2 Storage in the Cryptographic Cloud

Despite the obvious benefits of using a public cloud infrastructure, major security and privacy risks arise. Concerns regarding the integrity and security of data are the primary impediment to the adoption of cloud storage as well as cloud computing in general. Summarizes the advantages of a cryptographic storage service, such as ensuring regulatory compliance and lowering users' and cloud providers' legal liability. Aside from that, secure backups, archives, health record systems, safe data interchange, and e-discovery are a few examples of cloud services that could be built on top of a cryptographic storage service.

2.3 Effective and Safe Multi-Keyword Search on Cloud Data

On the one hand, users who do not necessarily have prior knowledge of the encrypted cloud data must post-process every file that is retrieved in order to locate the ones that are most relevant to their interests; On the other hand, in today's pay-as-you-go cloud model, retrieving all files that contain the query keyword requires additional network traffic, which is unacceptable. This study identifies and addresses the problem of performing ranked keyword searches over encrypted cloud data that are both efficient and safe. By delivering matched files in ranked order based on specified relevant criteria (such as keyword frequency), ranked ontology keyword mapping and search improves system usability and brings us one step closer to the practical implementation of privacy-preserving data storage services in Cloud Computing. For the first time, this work defines and solves the difficult problem of privacy-preserving multi-keyword ranked ontology keyword mapping and search over encrypted cloud data (BKCM). Additionally, it establishes a set of stringent privacy requirements that must be met for such a secure cloud data utilization system to be implemented. The documents that correspond to the entered search terms that are highly relevant are effectively returned by the suggested ranking algorithm. Our proposed system employs the proposed ranking method to enhance cloud service provider data security. Cloud Computing Offering Privacy Protection: In terms of user trust and legal compliance, privacy is a crucial concern for cloud computing, and it must be considered at every stage of design.

2.4 Easy Fuzzy Keyword Search Over Encrypted Data in Cloud Computing

The fundamental idea of this study is to formalize and address the difficulty of conducting efficient fuzzy keyword searches over encrypted cloud data while maintaining keyword privacy. Even though our proposed system is for multi-keyword raking search (the BKCM scheme), this fundamental idea is utilized. In, the author suggests developing a safe cloud storage service that addresses the issue of dependability while delivering nearly optimal overall performance. Cloud computing security, scalability, and fine-grained data access control: A problem that still needs to be solved in access control is how to simultaneously achieve data secrecy, scalability, and fine grainedness. The business locales this troublesome open issue by characterizing and upholding access strategies considering information ascribes from one viewpoint, and permitting the information proprietor to assign most of the calculation undertakings engaged with fine-grained information access control to untrusted cloud servers without uncovering the hidden information contents on the other. In this paper, the authors propose a public auditing mechanism for cloud computing data storage security that protects privacy. It uses random masking and the homomorphic linear authenticator to ensure that the TPA does not learn anything about the content of the data stored on the cloud server during the efficient auditing process. This removes the cloud user's fear of outsourced data leakage as well as the tedious and potentially costly auditing task.

2.5 Semantic Multi-Keyword Ranked Ontology Keyword Mapping and Search Over Encrypted Cloud Data in an Efficient and Privacy-Protecting Manner

This paper discusses the increasing number of data owners who are storing their sensitive data in the cloud due to the numerous benefits of cloud computing. With countless information records put away on the cloud server, it is basic to

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
www.ijprems.com
editor@ijprems.com
Vol. 03, Issue 03, March 2023, pp : 403-406
e-ISSN : 2583-1062
Impact Factor : 5.725

furnish information clients with watchword-based search administrations. However, plaintext search technologies are rendered useless because sensitive data is frequently encrypted prior to being sent to a cloud server in order to protect privacy. A semantic multi-keyword ranking ontology keyword mapping and search strategy for encrypted cloud data that complies with several stringent privacy constraints is presented in this study. To get started, we will use "Latent Semantic Analysis" to find connections between texts and words. The latent semantic analysis makes use of implicit higher-order structure in the relationship between terms and texts ("semantic structure") to represent words and documents in a vector space with reduced dimensions.

Consequently, the relationship between words is automatically recorded. Second, we use secure "k-nearest neighbor (k-NN)" to provide secure search capabilities. The system that was suggested might not only return files that exactly match the query keyword, but it might also return files that have latent semantically linked phrases with it. The result of the experiment. The proposed system is fully demonstrated in the preceding section, except for the KeyGen method. In our method, we create the inverse matrix by employing Gauss-Jordan. The time required to generate a key is determined by the scale of the matrix. Additionally, the SVD algorithm will take longer to process than the suggested method. In terms of time consumption, our other proposed algorithms, such as index construction, trapdoor generation, and query, are compatible with the original BKCM.

3. EXISTING SYSTEM

Because of the enormous number of data users and documents in the cloud, it is critical for the search service to support multi-keyword queries and give result similarity ranking to satisfy the effective data retrieval need. The searchable encryption relies on single term or Boolean keyword searches and seldom separates search results.

4. PROPOSED SYSTEM

For a secure cloud data utilization system to become a reality, we define and solve the challenging problem of privacy-preserving multi-keyword ranked ontology keyword mapping and search over encrypted cloud data (EARM). Additionally, we establish a set of stringent privacy requirements. We select the effective principle of "coordinate matching" from a variety of multi-keyword semantics. We propose the issue of Secured Multi keyword Search (SMS) over encrypted cloud data (ECD) and develop a collection of privacy policies for a secure cloud data utilization system. To identify the similarity between the search query and the data, we choose the highly efficient rule of coordinate matching from several multi-keyword semantics. For further matching, we use inner data correspondence to quantitatively formalize this principle for measuring similarity. Using secure inner product computation, we first propose and then modify a fundamental Secured multi keyword ranked ontology keyword mapping and search scheme to meet various privacy requirements. The top k retrieval results can be found in the Ranked result. Additionally, we propose an alert system that will send out email and text messages whenever an unauthorized user attempts to access cloud data.

4.1 Cloud Setup Module

Instead of giving undifferentiated results, this module improves the methods that allow multi-keyword queries and give result similarity rating for optimal data retrieval. Maintaining privacy by preventing the edge cloud server from learning new information from the dataset and index. Efficiency: The above functionality and privacy criteria should be met with minimal communication and processing overhead.

4.2 BKCM Coordinate Matching

"Coordinate matching" is an intermediate similarity metric that quantifies the relevance of a document to a query by counting the number of query terms that occurs in it. When users select the exact portion of the dataset to be recovered, Boolean searches perform well with the user's stated search requirement. Users may more easily identify a list of keywords suggesting their concern and retrieve the most relevant publications in a rank order.

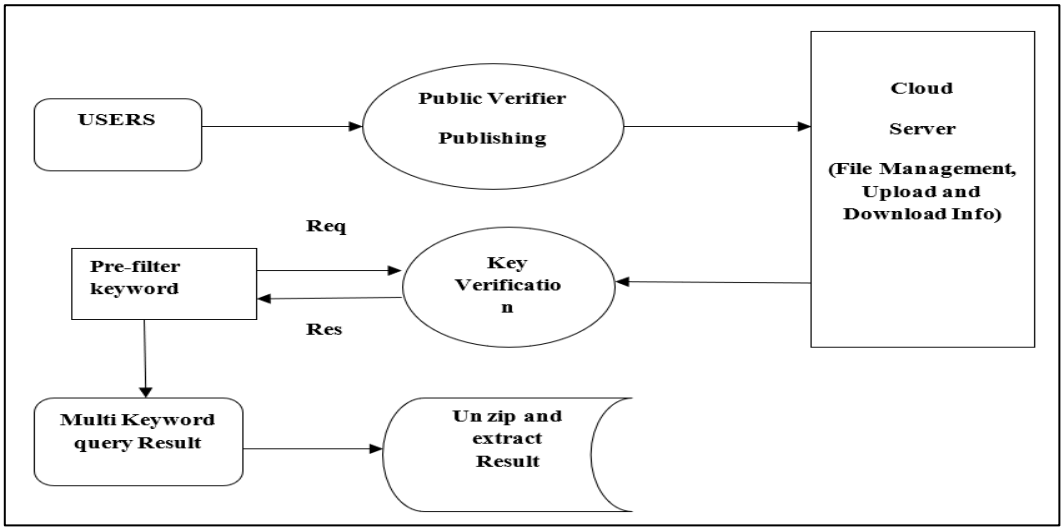
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 03, March 2023, pp : 403-406

e-ISSN : 2583-1062

Impact Factor : 5.725



4.3 Prefiltering and Security Management Module

The data owner can use typical symmetric key cryptography to encrypt the data prior to outsourcing, essentially preventing the cloud server from accessing the outsourced data. Index privacy is compromised if the cloud server deduces any link between keywords and encrypted documents from the index. As a result, a searchable index should be created to prevent the cloud server from engaging in such an association assault.

5. CONCLUSION

In this work, we describe and solve the problem of multi-keyword ranking ontology keyword mapping and search over encrypted cloud data for the first time, as well as develop a variety of privacy constraints. We choose the efficient principle of "coordinate matching," i.e., as many matches as possible, among various multi-keyword semantics to effectively capture similarity between query keywords and outsourced documents, and we use "inner product similarity" to quantitatively formalize such a principle for similarity measurement. To address the difficulty of providing multi-keyword semantic without compromising privacy, we first present a simple BKCM approach based on safe inner product computation, and then dramatically improve it to fulfill privacy requirements in two threat scenarios. A thorough investigation of the suggested schemes' privacy and efficiency guarantees is provided, and tests on real-world datasets indicate that our proposed methods impose negligible overhead on both computation and communication. We will investigate supporting alternative multi-keyword semantics (e.g., weighted query) over encrypted data, integrity verification of rank order in search result, and privacy assurances in a more robust threat model in the future.

6. REFERENCES

[1] "A break in the clouds:" by L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner. toward a definition of the cloud," ACM SIGCOMM Comput.

[2] "Cryptographic cloud storage," by S. Kamara and K. Lauter, published in January 2010 in RLCPS, LNCS, Springer, Heidelberg.

[3] Contemporary information retrieval: A brief synopsis," volume IEEE Data Engineering Bulletin 24, no. 4, 2001, pp. 35-43.

[4] "Gibbytes management: Compressing and indexing photographs and documents," by I. H. Witten, A. Moffat, and T. C. Bell, Morgan Kaufmann Publishing, San Francisco, May 1999.

[5] "Practical strategies for searching encrypted data," by D. Song, D. Wagner, and A. Perrig, in S&P Proc., 2000.