Applied Science and Biotechnology Journal for Advanced Research e-ISSN: 2583-553X
Peer Reviewed Journal
Volume-3, Issue-1 (January 2024), Page: 12-15 https://doi.org/10.5281/zenodo.10656146

Public Key Encryption and Keyword Search Mapping

Manoj Sharma[1] and Akshay Gupta[2]
[1]M.Tech Scholar, Department of Computer Science and Engineering, Patel College of Science & Technology, Bhopal, India
[2]Assistant Professor, Department of Computer Science and Engineering, Patel College of Science & Technology, Bhopal, India

[1]Corresponding Author: manojksharma.2011@gmail.com

Received: 27-12-2023 **Revised:** 15-01-2024 **Accepted:** 29-01-2024

ABSTRACT
This work addresses a difficult method for searching for crucial keywords for mixed cloud statistics (MRSE), which is a first step towards enabling secure cloud data processing. Among the different meanings of multiple terms, we choose the relevant concept of "relational coherence". Businesses can now more easily and affordably outsource a greater variety of products and services to community clouds thanks to cloud computing. To guarantee that sensitive personal data is kept enclosed before being supplied for their work. Given the enormous number of users of data and records in a search engine, it must be able to search for a specific phrase using numerous keyword searches and show a connection between them measure in order to successfully satisfy the demand of obtaining data. First, we suggest using the simple-to-use MRSE technique to protect a computer that contains internal objects. Real-world dataset experiments demonstrate that the suggested approaches do not significantly reduce the associated computing and communication expenses. One of the ways we effectively represent user information needs through optimization is by mapping feedback sessions to pseudo (fake) documents. Average precision (CAP) is a new metric used to assess the quality of the reconstructed online search results. Experiments with real data demonstrate that the suggested strategies have very little effect on transmission and computation. We extend these two methods to include more search semantics in order to enhance the search experience provided by the data search service.

Keywords: public key, encryption, search, mapping

I. INTRODUCTION

Performance, system usability, and scalability are all tough to achieve, making this a very difficult topic to solve. It's because of the high number of people relying on on-demand data, as well as the enormous amount of cloud-based data documents. Users should enter numerous keywords rather than just one in order to get the most relevant results. The search results can be further refined by including more keywords in the query. The plaintext information retrieval (IR) community has generally adopted "coordinate matching" as an effective similarity metric for such multi-keyword semantics. By providing customers with access to an on-demand pool of high-quality apps and services provided by programmable computer resources, "cloud computing" is finally bringing computing as a service to the masses. Records of one's own health care, picture folders, due financial transactions, documentation, and other sensitive data could require encryption by data holders before being sent to the commercial community cloud in order to ensure the data privacy and prevent unauthorised access in the cloud and elsewhere. A plaintext keyword search-based data consumption service will be replaced by this.

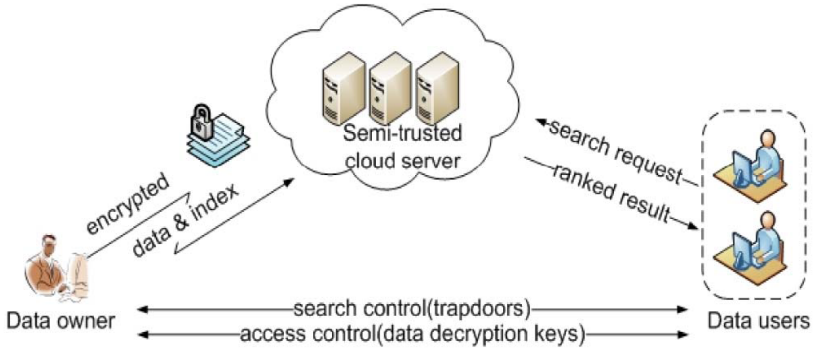

Figure1: Encrypted cloud data architecture

II. DATA MODELING

2.1 Threat Modeling

A well-known ciphertext model is It simply assumed that the cloud server knows the searchable value I and the encrypted data set C, both of which are externalized from the data owner, both of which are accessible to the cloud. A model that is well-known in the field An assumption is made that the cloud server has more knowledge than the known ciphertext model does. Two examples of this type of information are the link between specific search queries (trapdoors) and the statistical data relevant to the data set.

Furthermore, the cloud server adheres to the protocol specification and performs in a "honest" manner, as specified in the protocol. Curious about the data in its stowing and the missive flows that it receives for the duration of a protocol, however, is an inference and evaluation of this data (including the index). The based on the evidence the cloud server is aware of, we reflect two risk models with differing occurrence possibilities. Consistent with prior cloud security research, our model depicts a "honest but curious" cloud server.

2.2 System's Model

Ruminate a cloud hosting service consisting of a documents titleholder, a data consumer, and an cloud server, as shown graphically in below figure. As a result, the cloud server will receive a number of encrypted data records F. Process of search controls, such as broadcast encryption, allowing an authorized user to obtain the corresponding trapdoor T in order for the purpose of conducting a search through the document collection the defined keywords t. After the cloud server has received T from the data user, it will search index I and then return the appropriate collection of cyphers. The search results ought to be arranged on the cloud server in accordance with a variety of categorization standards in order to make the text retrieval process more effective. To Communication costs will be minimized. There will be a searchable encrypted index, I from F first, which will be exported to the cloud server for data efficiency, and then the corresponding cipher text collection C will be exported as well. It is also possible to add new credentials and delete old ones within the framework of this access control system, which manages access to decryption and data collection capabilities.

2.3 Design Objectives

Protection of privacy the data set and index must be encrypted in order to meet the specified privacy requirements and prevent the cloud server from learning more from them. Effectiveness To meet the privacy and performance requirements described above, minimal communication and processing should be used. Security and performance assurance must be accomplished simultaneously in order to enable positioned search to make the best use of outsourced cloud data, according to our system architecture. a search ranking system that includes several words and phrases. Search algorithms should be developed that handle a variety of keyword queries and provide a similarity measure for efficient information retrieval rather than returning similar results.

2.4 Nomenclature

- A plaintext collection of m data documents is represented as a set by the expression F = (F1,F2,....Fm).
- As shown as (I1,I2,...,Im), the searchable index related to C is for which each subindex Ii is constructed.
- "C" stands for the cloud server's encrypted document collection.
- Words in the dictionary are represented by the W symbol (W=n), which is written as W=n (W1,W2,...,Wn).

2.5 Prepare for Coordinate Matching, Section

Searches using boolean operators work best when the user knows exactly what subset of the data they want returned. A "relevance match" is a combination of link and cross searches that count the document's total number of search phrases to determine how well they match the query.

III. PRIVACY AND FRAMEWORK

Using encrypted cloud data, we provide the multi keyword categorized examine framework secure cloud data must meet the MRSE which is a set of stringent concealment criteria at the system level utilisation.

3.1 MRSE Framework, Section

These algorithms are part of the MRSE system, which focuses on indexing and searching. Because the data owner can simply encrypt the data using ordinary symmetric key encryption before outsourcing it, the framework does not show any activities in the data documents for ease of presentation.

To begin, you need to create an a number in an index (F,SK). The owner of the data creates a searchable index I using the data set F, which is then scrambled using the symmetric key SK of the cloud server.

Establishment (1l). In response to being given a safety factor (l), A symmetric key is generated by the database owner (SK) and sends it out in the form of SK.

3.2 MRSE Privacy Requirements

Finally, the cloud server will be unable to access the outsourced data. According to the conventional privacy assurance in the relevant literature, which incorporates searchable encryption, the server should only learn the results of searches. Using this wide privacy description, we analyse and establish severe privacy standards for the MRSE architecture. The trapdoor generation function should be random rather than deterministic.

Particularly the any connection between the cloud server and the outside world should be undetectable to it. between two trapdoors, such as whether they were produced by the same search request from their relationship. Method of gaining access. Rank-ordered searches use an access pattern that determines the order in which search results are shown. FW keyword set, and it provides an all documents rated depending on how useful they are to fW in order of importance. In the next step, the user's search patterns are established. A catchphrase Private. In order to protect users' privacy, issue terms recommended through the corresponding trapdoor.

IV. PRIVACY AND EFFICIENT MRSE

First, we discuss the basic concept for MRSE using a secure internal internal computer, extend it more to be secure protection for various MRSE danger scenarios system by following the procedure of step by step. We also discuss how to make the system more dynamic and how to improve search semantics. Using "internal feature similarity" to objectively evaluate the best similarity measure "relevant match" to multiple keyword ranked searches is recommended. If the corresponding keyword Wj occurs in a given document, Di is the binary data vector, and Q is the binary query vector specifying the bits corresponding to those keywords.

4.1 I-Scheme of MRSE

Randomness in search results should be adequately calibrated to disguise document frequency and limit the possibility of keywords being reidentified, as indicated in the criterion for keyword privacy. Instead of eliminating as we had originally planned, fresh random number t is assigned to the expanded dimension in each query vector in our more advanced technique. The cloud server will have a harder time figuring out how the received trapdoors connect to one another with this new randomness.

4.2 Review and Assessment

In light of Section's three design objectives, we assess this MRSE scheme. Because of the high purity of the documents retrieved, it is expected to be less effective in terms of efficiency. user-friendliness and productivity Consider that a document has a certain number of search phrases.

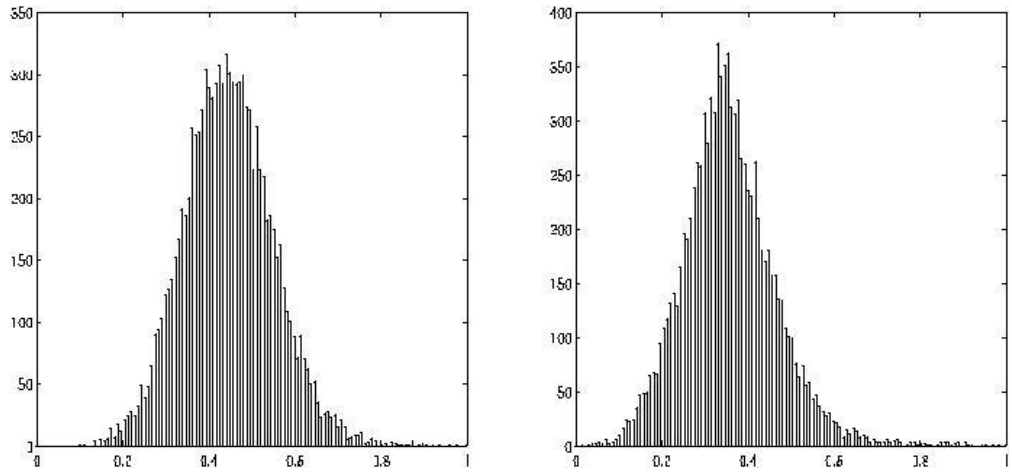

Figure 2: Functionality and Efficiency of 11k documents, 11 search terms, and several SSE were used to construct the standard deviation of the final similarity score distribution. TFS 4.0

However, there are many additional factors that can influence the usefulness of a search. Keywords that appear often in large numbers of documents are less important to the query than those that appear on fewer pages. The "coordinate matching" algorithm assigns a value of 1 to each term in a document or query.

V. PERFORMANCE ANALYSIS

With the use of our own method, we compare the performance of four suggested MRSE systems to our own method. In-depth testing of the proposed technique is based on the Enron Email Data Set, which has undergone extensive analysis. In order to collect data, a random sample of emails is chosen at random. The Linux server and an Intel Xeon processor clocked at 2.93 GHz are used to power the experiment system written in C. The inverse of a matrix is computed using open-source algorithms from Numerical Recipes.

VI. CONCLUSION

In this work, we integrated new search semantics into our old search algorithm, making it much better. We introduce the fundamental idea of MRSE utilizing safe inner product computing to address the problem of revealing the meaning of several keywords without breaching confidentiality. We'll discuss two upgraded MRSE solutions later on that will better safeguard your privacy in two distinct threat scenarios. We may now impose further privacy limitations after defining and resolving the multiple keyword syntax searches for encrypted cloud data. Utilizing the efficient "contact overlap," To put it another way, the more matches that are conceivable, the more pertinent the material will be on externalized sites to query phrases. We use a concept known as "internal feature similarity" to measure the degree of likeness.

REFERENCES

Xiaojun Zhang, Yao Tang, Huaxiong Wang, Chunxiang Xu, Yinbin Miao, & Hang Cheng. (2019). Lattice-based Proxy-oriented Identity-based encryption with keyword search for cloud storage. *Inf. Sci., 494*, 193–207.

Joël Alwen, & Chris Peikert. (2011). Generating shorter bases for hard random lattices. *Theory Comput. Syst., 48*(3), 535–553.

Rouzbeh Behnia, Muslum Ozgur Ozmen, & Attila Altay Yavuz. (2020). Latticebased public key searchable encryption from experimental perspectives. *IEEE Trans. Dependable Secur. Comput., 17*(6), 1269–1282.

Mahnaz Noroozi, & Ziba Eslami. (2019). Public key authenticated encryption with keyword search: Revisited. *IET Inf. Secur., 13*(4), 336–342.

Oded Regev. (2009). On lattices, learning with errors, random linear codes, and cryptography. *J. ACM, 56*(6).

Xiaojun Zhang, Chunxiang Xu, Huaxiong Wang, Yuan Zhang, & Shixiong Wang. (2021). FS-PEKS: Lattice-based forward secure public-key encryption with keyword search for cloud-assisted industrial internet of things. *IEEE Trans. Dependable Secur. Comput., 18*(3), 1019–1032.