Quest Journals
Journal of Software Engineering and Simulation
Volume 7 ~ Issue 7 (2021) pp: 14-19
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

Over Encrypted Cloud Data, a Comparison of Privacy-Preserving Single-Keyword Search and Multi-Keyword Ranked Search Techniques

Dr. B. INDRANI
Assistant Professor
Department Of Computer Science
DirectorateOf Distance Education
Madurai Kamaraj University
Madurai, Tmail Nadu, India

HAIDER KAREEM MOHAMMED
Department Of Computer Science
Madurai Kamaraj University
Madurai, Tmail Nadu, India

ABSTRACT
On-demand computing is another term for cloud computing. It offers services through the internet. It allows users to store and access their data in and from a cloud server from any location and on any device. Because data is accessed through the internet and clients have no direct control over data after it is uploaded to a cloud server, there are numerous security risks. We start with a basic idea for a single keyword search over encrypted data, then go on to multi-keyword ranked searches. Secure inner product computation and efficient similarity measure of coordinate matching are used to search across encrypted cloud data (MRSE). We present two greatly enhanced MRSE strategies to achieve varied demanding privacy criteria in two separate threat models, i.e., as many matches as feasible, in order to capture the relevance of data documents to the search query. An anonymous ID is assigned to the user in order to increase the security of the data on the cloud server. To improve the data search service's search experience, the two approaches are being extended to enable more Search semantics. [5]

KEYWORDS: *Cloud computing, encryption, inner product similarity, single keyword search, multi-keyword search, and ranking are all terms that come to mind when thinking about cloud computing.*

Received 10 July, 2021; Revised: 24 July, 2021; Accepted 26 July, 2021 © The author(s) 2021.
Published with open access at www.questjournals.org

I. INTRODUCTION

Cloud computing is widely regarded as one of the most significant advancements in information technology in the recent epoch. Using resource virtualization, the cloud provides us with pay-as-you-go computing resources and services. Today's world is becoming increasingly digital, and cloud computing is the greatest notion for dealing with large datasets. Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS) are the three types of cloud computing services[13].

Cloud computing is a long-dreamed hallucination of computing as an efficacy, in which cloud users shakily store their data in the cloud in order to benefit from on-demand application and services from a public pool of programmable computing resources far above the ground. Because of its flexibility and cost benefits, both individuals and businesses are flocking to the cloud to outsource subcontract their local comprehensive data management system.

To protect data privacy and prevent unauthorised access in the cloud and beyond, sensitive data, such as e-mails, personal health records, photo albums, tax documents, and so on, may need to be encrypted by data owners before being outsourced to a commercial public cloud; however, this renders the traditional data employment service based on plaintext keyword invective obsolete.

Due to the large amount of bandwidth expense in cloud scale systems, the traditional way out of downloading all the data and decrypting locally is obviously unfeasible. Because images include a wealth of useful and important information, the projected system will also include image cataloguing in the MRSE scheme [1]. Furthermore, aside from eliminating local storage management, keeping data on the cloud provides no benefit unless it can be easily found and used. As a result, investigating a privacy-preserving and effective search service for encrypted cloud data is crucial. In the "pay-as-you-use" cloud paradigm, ranked search can also stylishly eliminate unnecessary network traffic by returning only the majority of germane material, which is quite appealing. Such a rating algorithm, however, should not provide any keyword-related information for privacy reasons. Furthermore, in order to increase search result accuracy as well as the user searching experience, such a ranking system must handle multiple keyword searches, as a single term search generally produces far too coarse results. As indicated by today's web search engines (for example, Google search), data users may want to supply a list of keywords rather than just one as an indicative of their search interest in order to get the most relevant data. Along with data privacy and effective penetration techniques, true privacy can only be achieved if the user's identity is kept secret from both the Cloud Service Provider (CSP) and the third-party user on the cloud server[6].

The three major contributions to our planned effort are as follows:

1. Rank search with multiple keywords
2. Search for a single keyword
3. Data encryption with AES
4. Configuration of the cloud.
5. A comparison of MRES and single phrase search

II. LITERATURE SURVEY

2.1 Encrypted Secured Multi-Keyword Ranked Search:

Data owners that use cloud compute are irritated to move their multifarious data organisation systems from local sites to the marketable public cloud for better service and cost savings. It is necessary to encrypt data before storing it in the cloud to ensure data security. Cryptographic algorithms are used in these existing systems [1].

2.2 Remote Keyword Searches with Privacy Protection:

Consider the following scenario: a user U wants to save his data on a remote file server S in encrypted form. After then, user U wants to get some of the encrypted files that include exact keywords back in a professional manner, while keeping the keywords themselves secret and not jeopardising the security of the tenuously stored files. For example, a user may wish to save old e-mail messages in an encrypted format on a server managed by Yahoo or another large provider, and then retrieve certain messages while on the go with a mobile device. [2] shows how to solve this problem while adhering to well-defined safety constraints.

2.3 Searching Encrypted Cloud Data with Multiple Keywords in an Efficient and Secure Way:

On the one hand, users who do not have access to the encrypted cloud data must post-process each retrieved file in order to find the ones that best match their needs; on the other hand, retrieving all files that contain the query keyword generates unnecessary network traffic, which is unavoidable in today's pay-as-you-go cloud model. The problem of effective yet safe and sound rank keyword search over encrypted cloud data [2] has been identified and solved in this study. Ranked search considerably improves system usability by delivering matched files in a ranked order based on particular relevance criteria (e.g., keyword frequency), bringing Cloud Computing one step closer to rational use of privacy-preserving data hosting services. The study is the first to identify and solve the difficult problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE)[2], as well as to create a set of stringent privacy requirements for such a protected cloud data utilisation system to become a reality. The proposed ranking approach appears to be effective in retrieving highly relevant documents that correspond to the search terms entered. The proposed ranking technique will be implemented in our future system to improve the security of data on Cloud Service Providers.

2.4 Protecting Personal Information in Cloud Computing:

Privacy is a critical concern for cloud computing, both in terms of regulatory compliance and user confidence, and it should be taken into account at every stage of the planning process. The [5] paper discusses the need of ensuring individual privacy in cloud computing and discusses some privacy-preserving cloud computing methods. According to the paper, it is critical to consider privacy when building cloud services that include the collecting, processing, or exchange of personal data. The main focus of this study is the preservation of data privacy. This document solely discusses data privacy; it does not allow for indexed searches or the concealment of a user's identity. As a result, these two flaws are addressed in our ideal system.

2.5 Obligation to Protect Privacy When Sharing Data With Anonymous ID:

An algorithm for the mysterious exchange of private data across N parties is developed in this research. This approach is used to generate ID numbers for nodes ranging from 1 to N repeatedly. This assignment is anonymous in the sense that the identities given are unfamiliar to the rest of the group. In [6,] existing and novel

techniques for assigning anonymous IDs are compared in terms of communication and computational needs. These novel techniques are based on a safe sum data mining approach that employs Newton's identities and Sturm's theorem. The fundamental goal of this research is to give the consumer an unidentifiable ID on the cloud.

2.6 Cloud Encrypted Data Search with a Single Keyword:

These strategies only offer traditional Boolean keyword search, without capturing any relevance of the files in the search result. Obtainable searchable encryption schemes allow a user to confidently search for encrypted data using keywords without first decrypting it. When used in a large collaborative data outsourcing cloud environment, they have the following drawbacks.

The disadvantages of using a single keyword search strategy are as follows:

1. A single-keyword search that isn't ranked
2. Keyword search with Boolean operators but no ranking
3. Search with a single term and a ranking
4. No data is obtained that is meaningful.

Another unit, the Illegal User, is depicted in Figure 1. If an unauthorised user tries to access any data from the clod, alerts will be sent through email and text message. The alert is sent to the owner of the data who is authorised to receive it.

III. IMPLEMENTATION DETAILS

3.1 MRES System

To recognise the relationship between search query and data credentials, we choose the attitude of harmonise matching for our organism. In coordinate matching concept, we employ internal data correspondence, i.e., the number of query keywords present in a document, to assess the document's closeness to the search query. Each page is linked to a binary vector as a sub index, with each bit representing whether the document contains a comparable keyword [6]. The search reservation can alternatively be described as a dual vector, with each bit indicating whether or not the associated phrase exists in this search request, allowing the similarity to be quantified precisely by the inner product of the query vector and the information vector. Direct outsourcing of data vectors or query vectors, on the other hand, will compromise index or search privacy.

To improve document retrieval accuracy, cloud servers should rank search results based on specific ranking parameters. Only the top-k pages that are most relevant to the search query are returned by the cloud server.

The stream in the wished-for organism begins with the user. To enjoy the amenities, the user must first register with CSP. Once user data is stored in CSS, it is no longer under the control of the user. The user must hire a TPA auditor to review the user data in CSS[on a regular basis. The user should be able to use the TPA to check the integrity of a specified piece of data for an unambiguous period of time without having access to the actual data. Below is an algorithm that explains how the TPA does the audit.

The AES technique is used to encrypt the data in the cloud server. As a result, when TPA conducts an audit, it simply receives a false representation of original files. The hash value of an encrypted file calculated jointly is the value on which TPA calculates or checks the integrity. The only thing the TPA is permitted to do is check for uprightness. It first examines for integrity, then determines whether uprightness has been maintained, and last determines whether integrity has been maintained or lost. TPA can give or cancel the concession at any time. The ability to upload, download, and change data is granted to the user. Instead of retrieving the entire file, the user's edit request is also served for a specific portion of the file[6].

Notation and Definitions for the AES Algorithm

AES(K,W)	Encrypt W using the AES codebook with keyK
AES-1(K,W)	Decrypt W using the AES codebook with keyK
MSB(j, W)	Return themost significantjbitsofW
LSB(j,W)	Return theleastsignificant jbitsofW
B1 |B2	ConcatenateB1andB2
K	Thekey-encryptionkeyK
s	Thenumberofsteps inthewrappingprocess, =6n
P[i]	Theith plaintext keydatablock
C[i]	Theith ciphertextdatablock
A	The64-bitintegritycheckregister

R[i]	Anarrayof64-bitregisterswherei=0,1,2, ...,n

Algorithm
1. A series of round keys derived from a cypher key.
2. Add the initial round key to the starting state array and initialise the state array.
3. Execute the Usual Round (rounds 1 to 9).
1. Carry out the last round.
5. Final Round Step's corresponding cypher text chunk output
ii. The Standard Round Execute the procedures outlined in the preceding section.
Sub Bytes, No. 1

2. Rows with a Shift 3.MixColumns
4.Use K to add a Round Key (round)
iii. The Grand Finale:
Execute the procedures outlined in the preceding section. 1.Secondary Bytes
2. Rows with a Shift
3. Use K to add Round Key (10)
The concatenation occupation will be employed in the key wrap technique to concatenate 64-bit numbers to generate the 128-bit input to the AES codebook. The AES codebook's 128-bit amount will be broken into two 64-bit quantities using the pulling out functions.
The AES codebook [AES] is required for the setup of the key enfold algorithm. The key envelop algorithm, key unwrap method, and inherent data integrity check will be discussed in the following sections.
Wrapping an algorithm's key:
1) Sub Bytes: The encryption site employs the first transformation, Sub Bytes. We interpret a byte as two hexadecimal digits to substitute it.
2) Shift Rows: Shift Rows is a transformation used in encryption.
3) Mix Columns: The Mix Columns transformation alters each column individually.
to a new column the state column
4) Add Round Key: Add Round Key goes through each column one by one.
The KEY and the plaintext to be wrapped are the inputs to the key wrapping technique. The plaintext is made up of n 64-bit blocks that include the key data for the life form. The key covering procedure is outlined below.
The key wrap method can also be explained in terms of indexing rather than shifting. This method avoids the rotation in the previous account by calculating the wrap key in place. This yields the same outcomes and is easier to programme in software.

IV. IMPLEMENTATION RESULT

We compare the results of a single key word search ranked search and a multi keyword search ranked search over encrypted data in the cloud in this section, as demonstrated in the following figures. Exch Ranked Search and Multi Keyword Ranked Search Over Encrypted Data On Cloud in this Result. As a result, the existing system is a single keyword search system, while the proposed system is MRES.

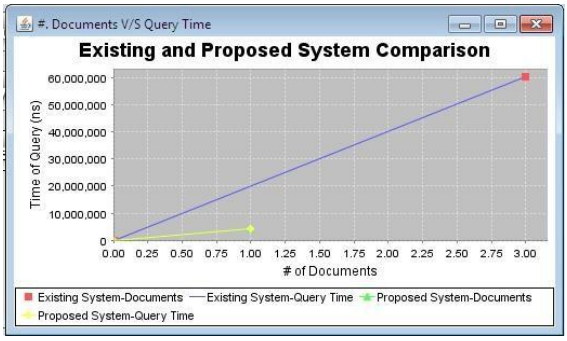

Figure 2: Graph of Number of Documents vs. Query Time

A comparison graph of the existing system and our system is shown in Figure 2. In the graph, the number of documents returned by the respective system's search result is plotted against the time it took to return the documents in the corresponding system. As seen in the graph, our system takes less time, approximately 5 nanoseconds, with the most specific result of one document, which is fewer than the three documents returned by the present system, which takes around 6 nanoseconds.

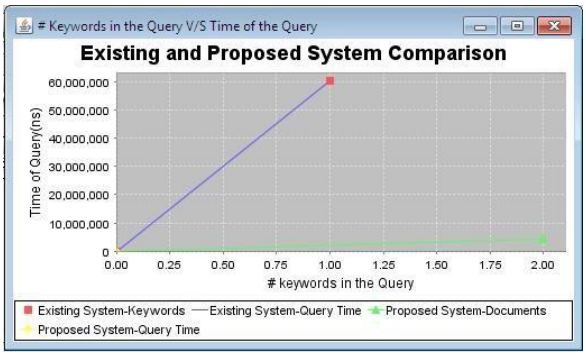

Figure 3: Graph of the Number of Keywords vs. Query Time

Figure 3 shows a comparison graph of the existing system and the system we implemented. The graph is plotted against the number of keywords used in the respective system's search and the amount of time it takes to complete the search. As seen in the graph, our system takes less time with numerous Keyword Query, taking less than 5 ns, whereas the present system takes roughly 6 ns even when only a single Keyword query is launched. As a result, our system outperforms the competition in every way.

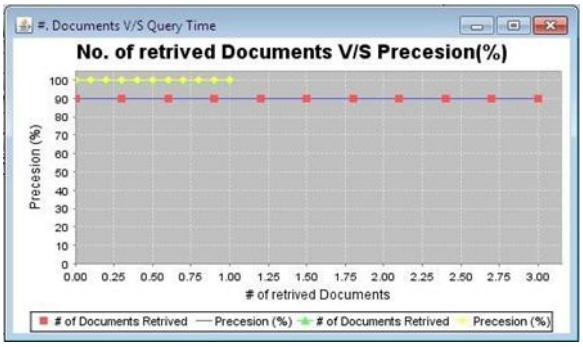

Figure 3: Number of Documents Retrieved vs. Precision

The comparison graph of the existing system and our implemented system plotted against the number of documents returned by each system V/S percentage precision is shown in Figure 4. (Perfectness). As demonstrated in the graph, our approach provides far better precession than the previous system.

The following are the main advantages of this method: 1) data security 2) privacy shield 3) auditing information to the data owner 4) audit aptitude aware data scheduling At this point, we'll assess our proposed scheme's performance in terms of the computing overhead introduced by each operation. The computing parameters are request and resources. When the number of requests increases at the same moment, it's time to see if they're all served in a reasonable amount of time. Each request's waiting time is recorded.

V. CONCLUSION AND FUTURE WORK:

In this paper, we define and solve the problem of multi-keyword ranked search over encrypted cloud data for the first time, as well as establish a set of privacy requirements. We choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and we use "internal product similarity" to quantitatively evaluate such similarity measure in the midst of various multi-keyword semantics. We present a basic idea of MRSE employing safe inner product computing to tackle the difficulty of enabling multi-keyword semantic without privacy breaches. Then, in two separate threat models, we present two revised MRSE techniques to meet various severe privacy requirements. We also look towards improving our ranked search method by adding support for other search semantics, such as TF IDF, and dynamic data operations. A thorough investigation of the suggested schemes' privacy and efficiency guarantees is provided, and experiments on a real-world dataset indicate that our proposed methods have low overhead on both computation and communication.

We'll have to work more in the future to improve data storage security on cloud storage services. This is a non-negotiable topic when it comes to cloud computing. We enhance the layers of authentication in order to implement that process. We'll be testing the integrity of the rank order in the search result in the future, presuming the cloud server is untrustworthy.

ACKNOWLEDGMENT

I'd like to thank my guide, Prof. Patil B.M., for his unwavering support and encouragement, as well as for his role as a driving factor behind the completion of this study. He has been quite helpful with all of his suggestions and hints. I'd like to express my gratitude to everyone who helped make it possible.

REFERENCES

[1]. Privacy preserving public auditing for Secure Cloud Storage", Cong Wang, Sherman S.-M. Chow, Qian Wang, KuiRen.
[2]. Li, S. Yu, N. Cao, and W. Lou. Authorized private keyword search over encrypted data in cloud
[3]. Wang, K. Ren, S. Yu, K. Mahendra, and R. Urs, "Achieving Usable and Privacy-Assured Similarity Search over Outsourced Cloud Data," Proc. IEEE INFOCOM, 2012.
[4]. International Journal of Advance Research, IJOAR.org Volume 3, Issue 2, February 2015, Online: ISSN 2320-9194
[5]. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc.IEEE INFOCOM, pp. 829-837, Apr, 2011.
[6]. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill. Order-reserving symmetric encryption. In Proceedings of Eurocrypt'09, volume 5479 of LNCS. Springer, 2009.
[7]. Kuchi Ravi Kishore, et al International Journal of Computer and Electronics Research [Volume 4, Issue 2, April 2015]
[8]. Secure Ranked Keyword Search over Encrypted Cloud Data , IEEE PAER, 2010.
[9]. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill. Order-preserving symmetric encryption. In Proceedings of Eurocrypt'09, volume 5479 of LNCS. Springer,
[10]. ShibaSampat Kale et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014, 7093-7096
[11]. International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 2, February 2014).