International Journal of Science Engineering and Advance Technology (IJSEAT)
ISSN 2321-6905, Vol. 9, Issue 9, December -2021

Efficient Authority Verification and Ranked Multi Keyword Search in Cloud

Kombathula Chanti Babu[1], Apvdl kumar[2]
#1 Student of M.Tech (CSE), Department of Computer Science & Engineering,
#2 Associate Professor, Depart of Computer Science & Engineering, BVCITS, Amalapuram AP, India.

Abstract–

Under the single owner concept, encrypted cloud data has sparked various research projects. In fact, however, most cloud servers do not serve a single owner; rather, they enable several owners to share the benefits of cloud computing. The problem of recovering encrypted data from the cloud is perplexing. For recovering scrambled data from the cloud, a variety of search procedures are used. This work revolves around a set of keyword search instruments applied to encrypted data, resulting in great data recovery proficiency. Because many believe that sensitive data must be scrambled before outsourcing to cloud servers in order to ensure client data security, searching through encrypted data is a method of remarkable interest for the cloud computing moment. Strategies from multiple spaces are used to provide a productive and secure search technique over scrambled data. Keyword search, it is assumed, is the finest way for searching encrypted data in the Cloud. It is more productive than searching for a single term.

Key words- Cloud computing, outsourced data, encrypted data, ranked keyword search

I. Introduction

Cloud computing has been viewed as another model of big business IT framework that can extract significant value from computing, storage, and applications, and engage clients to accept unrestricted, accommodating, and on-request organise access to a shared pool of configurable figuring assets with incomparable effectiveness and low financial overhead [1]. The two persons and the effort are drawn in by these interacting with highlights and decide to outsource their data to the cloud rather than obtaining programming and equipment to manage the info themselves. Regardless of the various reasons for interest in cloud administrations, outsourcing sensitive data (such as email, singular, association account data, government records, and so on.) to faraway servers raises security worries. Client-facing cloud specialist cooperatives (CSPs) may get access to sensitive data without permission. Encrypting data before outsourcing is a common way to handle secure data protection [2]. On the other side, there would be a significant penalty in terms of data convenience. For example, current keyword-based data recovery algorithms, which are widely utilised on plaintext data, cannot be clearly connected with scrambled data. It's clearly impossible to download all of the data from the cloud and decode it locally. In the simplest terms, cloud computing refers to storing and accessing data and projects via the Internet rather than your computer's hard disc. The cloud is merely an example of the Internet. Your hard drive isn't a part of cloud computing. The term "neighbourhood storage and computing" refers to the act of storing data on a hard drive or running programmes from it. All you need is a PC that is physically close to you, which means accessing your data is quick and easy for that one PC or others on the neighbourhood network. Working off your hard drive has been the way the PC industry has operated for a long time; some would argue that it is still superior to cloud computing, for reasons I'll explain shortly. To be termed "cloud figuring," you must be able to access your data or projects through the Internet, or have that data synchronised with other data via the Internet. In a large corporation, you may be aware of everything that goes on on the other side of the relationship; nevertheless, as a single client, you may never be aware of the massive data processing that goes on on the other end. With an online association, the result is the same: cloud computing should be possible everywhere, at any time.

II. Related work

Cloud computing transforms how information technology (IT) is consumed and managed,

promising increased cost efficiency, accelerated development, shorter time-to-market, and the capacity to grow applications on demand (Leighton, 2009). [1]. According to Gartner, while the growth rose enormously between 2008 and afterwards, there is undeniably a vital development towards the cloud computing model, and the focal areas might be enormous (Gartner Hype-Cycle, 2012). Nonetheless, given the cloud's current status Although computing is growing and changing at a rapid pace both theoretically and practically, the genuine/legally binding, money-related, organisation quality, between operability, security, and assurance issues continue to be major concerns. We depict various administrations and relationship types of suitable figuring and recognise essential Difficulties in this section. 2.2 Cloud security issues for users in general Kui Ren, Cong Wang, and Qian Wang, Kui Ren, Cong Wang, and Qian Wang, Kui Ren, Cong Wang Cloud figuring is discussed in this study as the most energising computing development in data innovation. Despite this, security and protection are viewed as necessary roadblocks to its widespread adoption. The authors describe a few basic security concerns and encourage the evaluation of security solutions for a secure open cloud environment. The most up-to-date word for the long-ago envisioned concept of computers as a utility is cloud computing. The cloud provides convenient, on-demand access to a unified pool of programmable computing assets that can be sent rapidly and efficiently with no management overhead. With its low-priority features, cloud computing enables a significant shift in perspective in how to send and deliver computing administrations, namely, it enables computing outsourcing to the point where both individuals and organisations can avoid incurring significant capital costs when purchasing and overseeing programming and equipment, as well as managing the associated operational overhead. [1] 2.3 Cloud storage using cryptography Kamara, S., and Lauter, K. In this paper, we will look upon To investigate the issue of constructing a secure cloud storage benefit over an open cloud foundation where the customer does not completely trust the professional organisation. To present a few designs that combine later and non-standard cryptographic natives in an abnormal state with the goal of achieving our goal. To investigate the benefits that such an engineering would provide to both clients and specialist cooperatives, as well as to provide an overview of current breakthroughs in cryptography, particularly as they relate to cloud storage. [2] A cryptographic method that is totally homomorphic. C. The upper crust, To offer the primary totally homomorphism encryption system in this work, which addresses a key outstanding issue in cryptography. Given the encryptions E(m1);...; E() of m1;...; mt, one can competently process a reduced figure message that scrambles f(m1;... ; m) for any effectively calculable capacity f with such a scheme. Rivest et al. raised this issue in 1978. There are several uses for completely homomorphic encryption. For example, it allows a client to ask a confidential question to a search engine; the client submits a scrambled question, and the search engine figures out a compact encrypted response without ever looking at the question. It also enables searching on encrypted data, allowing a client to store scrambled documents on a remote record server and have the server recover only those documents that (when unscrambled) satisfy some Boolean constraint, despite the fact that the server cannot decode the records without the help of others. Completely homomorphic encryption improves the capability of secure gathering calculation even further. Our research begins with a "boot-strappable" encryption scheme that works when the capacity f is equal to the scheme's own unscrambling capacity. Then I'll explain how boot-strappable encryption becomes totally homomorphic encryption via recursive self-implanting. The development makes use of challenging challenges on perfectly cross-sectioned cross-sections. [3] 2.4 Software security and reproduction on rams that have been neglected R. Ostrovsky and O. Goldreich, [4] To demonstrate a potential treatment of programming security in this work. To distil and figure out the key issue of learning about a programme from its execution, and to reduce this issue to the question of on-line reproduction of a self-assured programme on a sloppy RAM. Then we'll show our main result: a successful reproduction of a subjective (RAM) programme on a probabilistic absence.

III. Cloud Computing Models

The cloud provides resources to users in a variety of ways. It is made available to users by service providers and hosted by cloud vendors. The numerous cloud services and layered architecture of

cloud services are explained in Figures 2.1 and 2.2. The following are the different cloud service models:

Software as a Service (SAAS) provides users with the necessary software, network, and operating system. Users are not required to install them on their computers. It is an application that can be accessed from anywhere in the globe as long as we have a computer connected to the internet.

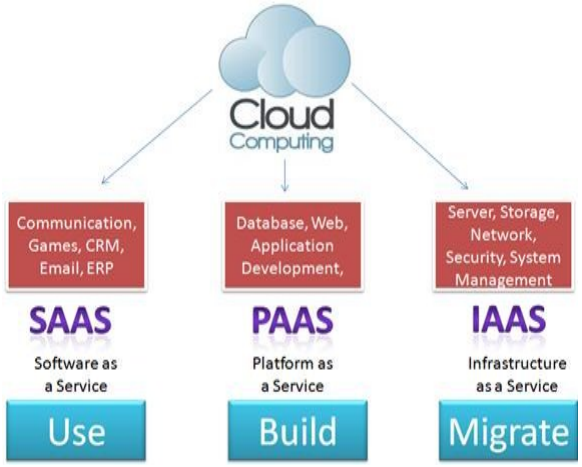

Figure 1: Different cloud services

- ✓ Platform as a Service (PAAS) provides users with a network and operating system. It's a platform that allows developers to make their own apps. Users do not need to manage their virtual machines or their operating systems at this tier.

- ✓ Infrastructure as a Service (IAAS) "Hardware as a Service" is another term for it. It refers to the cloud's physical environment, which includes storage, networking, and other resources. The user has access to the storage and network.

- ✓ The most widely used layer of cloud computing is the application layer, which is where users deploy their apps. The platform layer is necessary for the user to launch programmes. By obtaining appropriate resources and deploying large numbers of virtual machines (VMs) on hardware, the infrastructure layer enables user requests for computational resources. The server layer refers to the hardware layer. It is a representation of the physical hardware that supplies resources. Hardware resources are inexpensive.

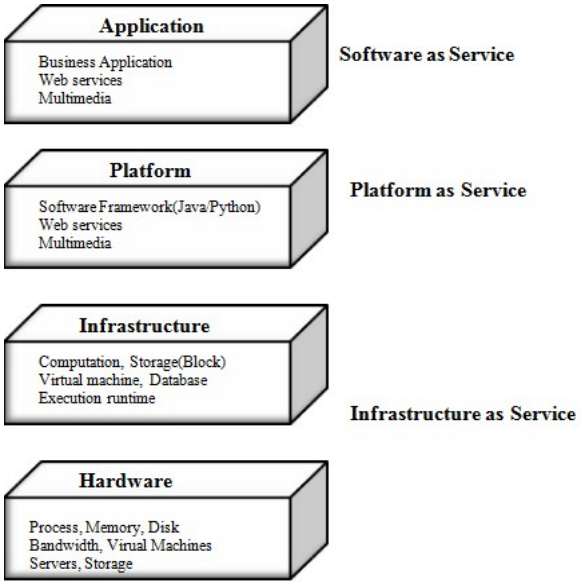

Figure 2: Layered Architecture of cloud services

IV. Trapdoor Generation:

Users register their identity tokens in order to obtain secrets that can be used to rewrite the information they have access to. Users only register the identity tokens linked with the Owner's sub ACPs, and the remaining identity tokens are registered with the cloud in a highly privacy-conscious manner. Throughout this section, it should be remembered that the cloud does not learn the identification traits of Users. When users register with the Owner, the Owner assigns them two sets of secrets for the attribute conditions in command that are shared throughout the sub ACPs in the ACPB cloud. One set is kept by the owner, while the other is given to the cloud. To prevent the cloud from decrypting the Owner's encrypted data, two distinct sets of keys are used.

V. Ranked Keyword Searching

Data owners share their outsourced data as cloud computing has become a vital aspect of the IT business. Due to the massive volumes of information available on the World Wide Web, a great number of people attempt to retrieve specific data files. Keyword-based search is one of the most popular techniques to do so. To use cloud data for a specific inquiry, keyword searches are performed. Such keyword search approaches have been frequently used in plain text search contexts to allow users to selectively retrieve files of interest (C.wang). A lot of effort has gone into making keyword search easier for users. However, there are few studies that consider the precise user query and give a sorted URL list based on it. The majority of keyword searchers are built in such a way that users can query a collection using clouds (7). Ranked keyword search is used to reduce unnecessary network bandwidth by avoiding returning useless data. In the "pay-as-you-go" cloud paradigm, this strategy is particularly appealing. Such a ranking procedure should not reveal any keyword-related information for privacy reasons. It is vital for such a ranking system to offer -keyword search in order to increase the search result accuracy as well as the user searching experience, as single keyword searches typically produce much too coarse results (5). To compute the relevance scores of a given request, information is retrieved from the matching files. If the ranking system allows for multiple keyword searches, the accuracy of search results can be improved, as well as the user's search experience. Users submit a list of keywords rather than just one to show that they are interested in a specific topic in all web search engines. The search procedure is narrowed down by each keyword in the user query.

VI. System Proposed

In these operations, there are three key actors: the cloud server, the data owner, and the data consumer. Data owners have their own collections of documents, making it difficult to keep track of them locally. Maintaining and storing documents locally is costly in terms of storage and incurs computational overhead. As a result, data owners are encouraged to outsource their document sets to the cloud in order to gain more flexibility.

However, prior to the transfer process, the data privacy issue confronted the owner, so she utilised encryption methods and outsourced the data in encrypted form, expecting the cloud server to give keyword retrieval service to the data owner or other approved users. Data privacy would be jeopardised as a result of information leaking, which is undesirable to the data owner. The data user has been given permission to do keyword retrieval on the outsourced data. The data user encrypts the query and sends it to the cloud server, which responds with the relevant files. After that, the data user can decrypt the files and use them.

A. Model in Vector Space

It's used to determine the accuracy of a rating. To find reliable ranking and similarity measurements, the TF-IDF rule is utilised. Where TF stands for the number of times a word appears in a document, and IDF is calculated by dividing the total number of documents in the collection by the number of documents that contain the term. The top-k retrieval result is returned. IDF = total number of documents in the collection divided by the number of documents that contain the phrase.

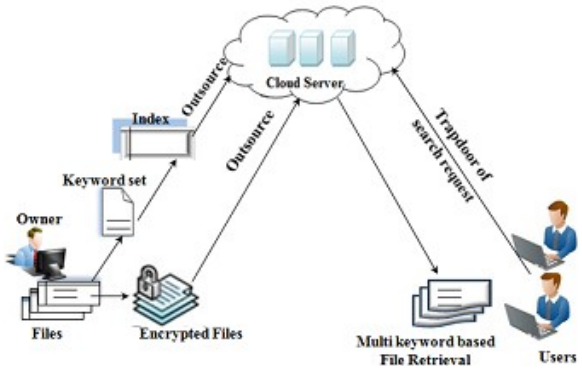

Fig. 3. Architecture of proposed system

B. Improve the Secure Indexing Scheme We use the cosine measure to evaluate similarity scores in order to produce reliable -keyword ranked search. As demonstrated in the suggested system, we divide the original long document index vector into many sub index vectors, with each sub index representing a subset of keywords and becoming a component of the ith level of the index tree. The query vector is split in the same way that the document index vector is split.

By adding the scores from each level, the final similarity score for document 'd' can be calculated. The cloud server determines the relevance document d to query Q based on these similarity scores and sends the top most relevant document to the user. The document index vector and query index vector are both well protected utilising the level wise secure inner product method.

MD Algorithm (C)

The MD method is used to identify the k-best match in a database with an MDB-tree structure. The attribute domain is represented by the MDB tree, and each attribute in that domain has an attribute value.

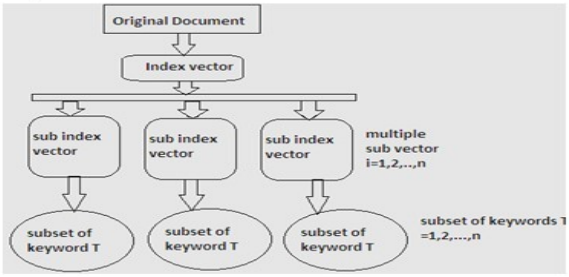

Fig. 4. Mechanism of document index formation

D. Check File Status

Proposed system announces a third party auditor to audit user file request for checking integrity of corresponding file. Audit result from third party would be helpful for cloud service provider to enhance cloud based service platform.

Proposed Algorithms

A. *Algorithm for Top Result selection:*

1) Input

Take variable ‘k’ like a number and list source of selected item

2) Initialization:

Set pointers tk & tid as a null

3) Iteration phase

a. For all i ϵ source do

Insert(tk,(i, index))

b. End for

c. For all tuple ϵ tk do

tid.append(tuple[1])

d. End for

4) Output:

tid

B. *Algorithm for Insertion:*

1) Input

Take list tk to stored the top scoring items

Tuple(i,index)

2) Iteration

a. If length(tk) < k then

Insert(i, index) into tk in ascending order of items

b. Else

For all element ϵ tk do

If i< element[0] then Continue

Else Discard tk[0], insert(i, index) into tk in ascending order of item

EndIf

EndFor

EndIf

VII. Experimental Result

Some outcomes are resulting from this scheme:

A. Response Time

Fig.3 shows a graph in which time require to get search result after adding number of documents in database. If database size increases then time require to get result increases.

Results must require less time for MD search as compare to MRSE technique.

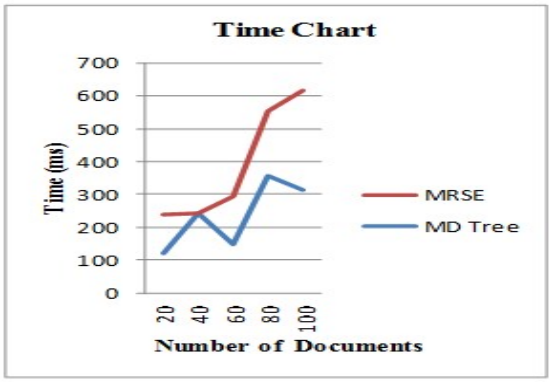

Fig. 5. Response Time

B. Encryption time

Fig.4 shows a graph in which graph shows the expected comparative analysis for time requires to encrypt keywords using both techniques.

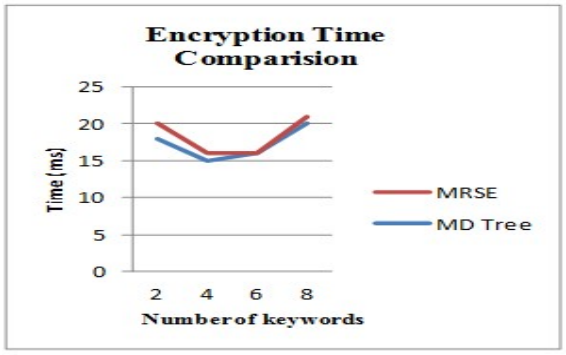

Fig. 6. Encryption time comparison

VIII. Concluding Remarks & Prospects

The prior work primarily focused on giving privacy to cloud data through the use of -keyword ranked search over encrypted cloud data using an efficient similarity measure of co-ordinate matching. A basic idea of MRSE using safe inner product computation was previously given in previous work. There was a need for more true privacy, which is what this paper addresses. The cloud user is given a unique ID in this system, which ensures strict privacy. To secure the user's data on the cloud from the cloud service provider and third-party user, this user ID is kept secret from both the CSP and the third-party user. As a result, the confidentiality of the user's data is preserved by concealing the user's identity. The challenge of searching encrypted cloud data using ranked -keyword (MRSE) is defined and solved in this work. The appropriate similarity measurement of "coordinates matching" and "inner product similarity," i.e., possibilities of many matches for capturing the documents from query search perceptible assessments for similarity measures, comes out of separate -keyword semantics. Using safe inner product computation and archive privacy requirements in two remote thread models as the foundation for the MRSE. Experiments based on real-world data reveal that low-overhead processing and communication are now possible. The cloud server will be treated as a trusted state in the future, with the integrity of the rank order in search results being checked.

References

[1] Hui Cui, Zhiguo Wan, Robert H. Deng, Guilin Wang, and Yingjiu Li “efficient and expressive keyword search over encrypted data in cloud”, IEEE JOURNAL OF , VOL. , NO. , 2016.

[2] Hongwei Li, Yi Yang,Tom H. Luan, Xiaohui Liang, Liang Zhou, Xuemin (Sherman) Shen, “Enabling FineGrained Multi-Keyword Search Supporting Classified Sub-Dictionaries over Encrypted Cloud Data”, IEEE Transactions On Dependable And Secure Computing, Vol. 13, No. 3, May/June 2016.

[3] Wei Zhang, Yaping Lin, Sheng Xiao, JieWu, Siwang Zhou, “Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing”, IEEE Transactions On Computers, Vol. 65, No. 5, May 2016.

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

[5] Chi Chen, Xiaojie Zhu, Peisong Shen, Jiankun Hu, Song Guo, Zahir Tari, Albert Y. Zomaya, “An Efficient Privacy-Preserving Ranked Keyword Search Method”, IEEE Transactions On Parallel And Distributed Systems, Vol. 27, No. 4, April 2016.

[6] Zhangjie Fu, Kui Ren, Jiangang Shu, Xingming Sun, Fengxiao Huang“, Enabling Personalized Search

over Encrypted Outsourced Data with Efficiency Improvement," IEEE Transactions On Parallel And Distributed Systems, Vol. 27, No. 9, September 2016.

[7] Zhihua Xia, Xinhui Wang, Xingming Sun,Qian Wang, Member, "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.

[8] Jingbo Yan, Yuqing Zhang, Xuefeng Liu, "Secure Multikeyword Search Supporting Dynamic Update and Ranked Retrieval", Services and applications, China Communications, 2016.

[9] Chia-Mu Yu, Chi-Yuan Chen, and Han-Chieh Chao, "Privacy-Preserving Multi-keyword Similarity Search Over Outsourced Cloud Data", 1932-8184 © 2015 IEEE Systems Journal.

[10]Wenhai Sun, Bing Wang, Ning Cao, Ming Li, Wenjing Lou, Y. Thomas Hou, Hui Li, "Verifiable PrivacyPreserving Multi-Keyword Text Search in the Cloud Supporting Similarity-Based Ranking", IEEE Transactions On Parallel And Distributed Systems, Vol. 25, No. 11, November 2014.

Authors

Kombathula Chanti Babu is pursing M.TECH (CSE) in the Department of Computer Science and Engineering from BVC Institute of Technology & Science, Batlapalem, Amalapuram, AP, India.

Apvdl kumar is working as Associate Professor in Department of Computer Science & Engineering, BVC Institute of Technology & Science, Batlapalem, Amalapuram, AP, India.