WWW.IJRPR.COM

**International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Privacy Preserving Geo-Location Data Allocation : A Machine Learning Reach

***A.Charumathi*** [#1]***, G.Shanthi*** [#2]

[#1] M.Sc, Department of Computer Science, Kamban College of Arts and Science for Women, Tiruvannamalai-606603 .

[#2] Assistant professor, Department of Computer Science, Kamban College of Arts and Science for Women, Tiruvannamalai- 606603.

**ABSTRACT**

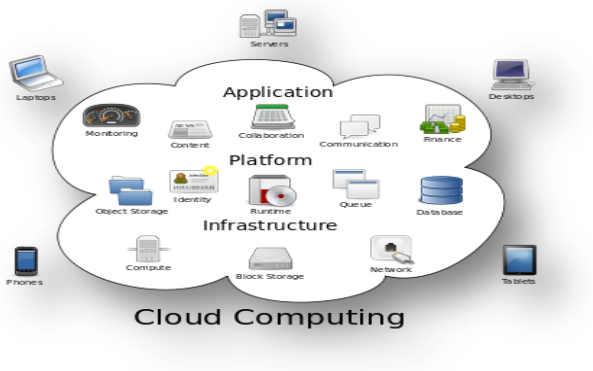
For the sake of privacy, cloud data owners prefer to outsource documents in encrypted form. As a result, developing efficient and trustworthy ciphertext search tools is critical. One difficulty is that the relationship between documents is generally hidden throughout the encryption process, resulting in significant search accuracy performance deterioration. In addition, the amount of data stored in data centers has increased dramatically. This will make developing ciphertext search methods capable of providing efficient and accurate online information retrieval on vast volumes of encrypted data even more difficult. A hierarchical clustering method is suggested in this research in order to accommodate more search semantics and to answer the demand for rapid ciphertext search in a huge data environment. The suggested hierarchical approach groups documents based on their lowest relevance thresholds, then divides the resulting clusters into sub-clusters until the maximum cluster size is attained. During the search phase, this strategy can achieve a linear computational complexity despite the fact that the size of the document collection grows exponentially. In this work, a structure known as the minimal hash sub-tree is devised to validate the legitimacy of search results. Experiments were carried on using a collection set compiled from IEEE Xplore. The results show that as the number of documents in the dataset grows, the suggested method's search time grows linearly, whereas the standard method's search time grows exponentially. Furthermore, the new method outperforms the previous method in terms of rank privacy and document relevancy.

KEYWORDS: Clustering, Cipher text

## 1.INTRODUCTION:

**Cloud computing** The utilization of computing resources (hardware and software) offered as a service through a network is known as cloud computing.(In most cases, the Internet). The name stems from the widespread use of a cloud-shaped symbol in system diagrams as a metaphor for the complicated architecture it encompasses.

Cloud computing entrusts a user's data, software, and processing to remote services. Cloud computing refers to the use of managed third-party services to make hardware and software resources available over the Internet. Typically, these services provide access to powerful software programs and high-end server computer networks.




## 2. LITERATURE SURVEY

[1] K. Ren, C. Wang, Q. Wang, and others, The most intriguing computing paradigm shift in information technology today is cloud computing. However, security and privacy are seen as major roadblocks to widespread adoption. The authors present a list of major security concerns and

encourage more research into security solutions for a secure public cloud environment.

[2] Gentry, C. We present the first fully homomorphic encryption system, which addresses a long-standing issue. Given the encryptions E(m1),..., E(mt) of m1,..., m t, a compact ciphertext that encrypts f(m1,..., m t) for every efficiently computable function f can be efficiently computed.

There are several uses for fully homomorphic encryption. It enables encrypted search engine inquiries, for example, where a search engine can provide you with a succinct encrypted answer to your (boolean) question without even knowing what your inquiry was. It also allows you to search encrypted data; you can store encrypted data on a remote server and have the server retrieve just the files that (once decrypted) fit your search criteria.

[3] G. Di Crescenzo, R. Ostrovsky, and G. Persiano, D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano,

We investigate the problem of searching encrypted data using a public key scheme. Consider the case of user Bob, who sends an encrypted email to user Alice using Alice's public key. An email gateway wants to see if the email has the keyword "urgent" in it so that it can route it appropriately. On the other hand, Alice does not want the gateway to be able to decrypt all of her messages. We develop and build a method that allows Alice to give the gateway a key that allows the gateway to check whether the word "urgent" is a keyword in the email without learning anything more about it. This approach is referred to as Public Key Encryption with Keyword Search. Consider a mail server that maintains multiple messages that have been publicly encrypted for Alice by others. Alice can transmit a key to the mail server using our approach, allowing the server to identify all mails containing a specified term.

[4] A. Perrig, D. X. Song, D. Wagner, and D. Wagner, To avoid security and privacy issues, it is preferable to keep data on data storage servers such as mail servers and file servers in encrypted form. However, this frequently entails sacrificing utility for security. For example, if a client wants to obtain only documents that contain specific terms, it wasn't clear how to let the data storage server execute the search and respond to the query without jeopardizing data confidentiality. For the challenge of searching, we discuss our cryptographic techniques.

[5] M. Mitzenmacher and Y.-C. Chang, We'll take a look at the following issue: A user U wants to save his data on a remote file server S in an encrypted format. Later, user U wants to quickly retrieve some of the encrypted files containing (or indexed by) specified keywords while keeping the keywords hidden and the security of the remotely stored files intact. For example, a user may desire to keep encrypted old e-mail messages on a server controlled by Yahoo or another large provider, and then retrieve certain messages while on the go with a mobile device.

## 3. PROPOSED SYSTEM:

A vector space model is employed in this research, and each document is represented by a vector, allowing each document to be viewed as a point in a high-dimensional space. All of the documents can be split into many groups due to the relationships between them. First, the cloud server will scan the categories for the smallest desired sub-category. The cloud server will then choose the required k documents from the smallest sub-category. The user chooses the value of k beforehand and sends it to the cloud server. If the current sub-category is unable to satisfy the k documents, the cloud server will go back to its parent and select the documents from its sibling categories

## 4. METHODOLOGY

### 4.1 Data Owner Module

This module assists the owner with registering such details as well as providing login information. This module assists the owner in encrypting and uploading his file using the RSA technique. This ensures that the files are safe from unauthorized access. The data owner has a collection of documents *F ={f1; f2; :::; fn}* that he wishes to send to a cloud server in encrypted form while still being able to search them for efficient use. The data owner creates a safe searchable tree index I from document collection F before generating an encrypted document collection C for F in our approach. The data owner then outsources the encrypted collection C and secure index I to a cloud server, and securely distributes the key information for trapdoor generation and document decryption to approved data users.

### 4.2Data User Module

This module contains the login information for users who have registered. This module assists the client in searching the file using numerous key phrases and receiving an accurate result list based on the user query. Before entering the activation code, the user will select the appropriate file, register their details, and receive an activation code through email. After that, the user can download and extract the Zip file. Data users have permission to view the documents of the data owner. The authorized user can create a trapdoor TD according to search control techniques to get k encrypted documents from the cloud server using t query keywords. The data user can then use the shared secret key to decrypt the documents.

.

### 4.3 Cloud Server

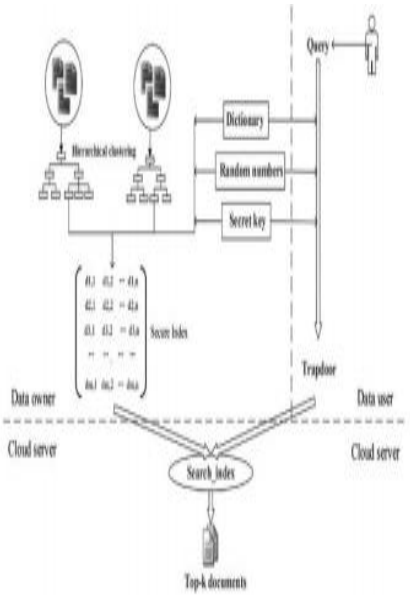
This module assists the server in encrypting the document using the AES algorithm, converting the encrypted document to a Zip file with an activation code, and then sending the activation code to the user for download. For the data owner, the cloud server stores the encrypted document collection C and the encrypted searchable tree index I. The cloud server searches the index tree I after getting the trapdoor TD from the data user, and then returns the appropriate collection of top-k ranked encrypted documents. Furthermore, after the server receives the update information from the data owner, it must update the index I and document collection C to reflect the new information.

### 4.4 Hierarchical clustering Search Module

These modules ensure that the user can utilize hierarchicalclusteringsearch to find files that are often searched.

This module allows the user to download a file and then decrypt it using his secret key. This module allows the Owner to see the files that have been uploaded and downloaded. Not only does the suggested approach allow for multi-keyword queries and accurate result grouping, but it also allows for dynamic updates to document collections. The method is intended to keep the cloud server from knowing further details about the document collection, index tree, or query.

## 5. SYSTEM ARCHITECTURE:




## 6. RESULTS

We looked into ciphertext search in the context of cloud storage in this article. We investigate the difficulty of retaining semantic relationships between various plain documents and related encrypted documents, as well as a design strategy for improving semantic search performance. The MRSE-HCI design is also proposed to adapt to the needs of data explosion, online information retrieval, and semantic search. At the same time, a verifiable mechanism is proposed to ensure that search results are correct and complete. We also look at the search efficiency and security in the context of two prevalent threat models. The search efficiency, accuracy, and rank security are all evaluated using an experimental platform.

## 7. OUTPUT


**Fig 1: cloud data processing file**

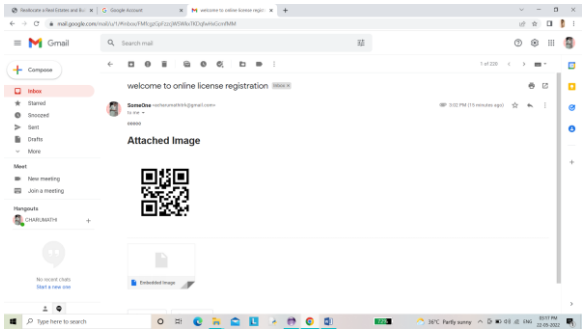



**Fig 2: Cloud security attachement**

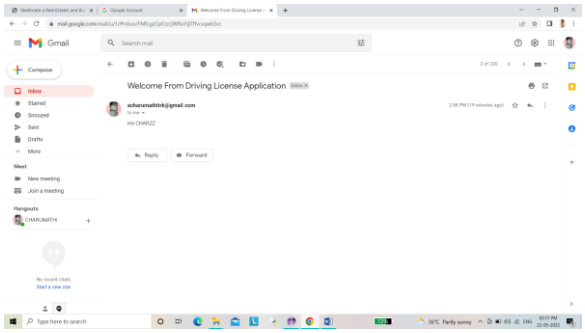



**Fig 3: Verification progression**

## 8.CONCULSION:

: A high-quality output is one that satisfies the end-needs user's and conveys information clearly. Any system's processing results are conveyed to users and other systems via outputs. It is decided how the information will be displaced for immediate use, as well as the hard copy output, in output design. It is the user's most essential and direct source of information. The system's relationship with the user is improved via efficient and intelligent output design.

## REFERENCES

[1] K. Ren, C.Wang, Q.Wang*et al.*, "Security challenges for the public cloud," *IEEE Internet Computing*, vol. 16, no. 1, pp. 69–73, 2012.

[2] S. Kamara and K. Lauter, "Cryptographic cloud storage," in *Financial Cryptography and Data Security*. Springer, 2010, pp. 136–149.

[3] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009.

[4] O. Goldreich and R. Ostrovsky, "Software protection and simulation on oblivious rams," *Journal of the ACM (JACM)*, vol. 43, no. 3, pp. 431–473, 1996.

[5] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Advances in Cryptology-Eurocrypt 2004*. Springer, 2004, pp. 506–522.

[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, "Public key encryption that allows pir queries," in *Advances in Cryptology-CRYPTO 2007*. Springer, 2007, pp. 50–67.

[7] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on*. IEEE, 2000, pp. 44–55.

[8] E.-J. Goh*et al.*, "Secure indexes." *IACR Cryptology ePrint Archive*, vol. 2003, p. 216, 2003.

[9] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Proceedings of the Third international conference on Applied Cryptography and Network Security*. Springer-Verlag, 2005, pp. 442–455.

[10] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in *Proceedings of the 13th ACM conference on Computer and communications security*. ACM, 2006, pp. 79–88.

[11] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–5.

[12] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in *Data Engineering (ICDE), 2012 IEEE 28th International Conference on*. IEEE, 2012, pp. 1156–1167.

[13] C. Wang, K. Ren, S. Yu, and K. M. R. Urs, "Achieving usable and privacy-assured similarity search over outsourced cloud data," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 451–459.

[14] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in *IEEE INFOCOM*, 2014.

[15] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in *Applied Cryptography and Network Security*. Springer, 2004, pp. 31–45.