SJIF Impact Factor (2024): 8.675| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 9 | Issue: 12 | December 2024 - Peer Reviewed Journal

AN EFFICIENT DATA SEARCH WITH MATCH KEYWORD WITH ENCRYPTED DATA

G Manasa[1], G Sravani[2], Suruchi W. Kitey[3]
[1,2,3]*UG Scholar,*[3]*Assistant Professor*
[1,2,3]*Department of Computer Science and Engineering*
[1,2,3]*Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India*

Article DOI: https://doi.org/10.36713/epra19357
DOI No: 10.36713/epra19357

ABSTRACT

With the continuous improvement of the security of cloud storage, more users upload private data to the cloud. However, a large number of encrypted data using independent keywords to create indexes not only directly increase the storage overhead, but also lead to the decline of search efficiency. Therefore, this paper proposes an efficient search method using features to match joint keywords (FMJK) on encrypted cloud data. This method proposes that keywords are selected from the non-duplicated keywords, which are extracted from the documents of the data owner, to generate a joint keyword, and all joint keywords form a keyword dictionary. Each joint keyword is matched with the feature of the document and the query keyword respectively, and file index the query trapdoor. Finally, the algorithm is used to calculate the data of the file index and the trapdoor. Theoretical analysis and experimental results show that the proposed method is more feasible and more effective than the compared schemes

1 INTRODUCTION

With the rapid development of science and technology, enterprises or individual users increasingly rely on storing a large number of data documents on cloud servers in order to share data quickly and remotely [1]. However, with the increasing demand, the cost of cloud server storage increases, the efficiency of search decreases, and privacy protection has become a focus of research [2]. In most of the existing cipher-text sorting retrieval methods, KNN (K Nearest Neighbor) technology is used to create indexes supporting cipher-text retrieval [3]–[5]. In the process of massive data encryption search, most of the search encryption schemes have high time complexity and large storage space, which are closely related to the encrypted key, the document index and query request dimensions [6], [7]. Reducing high dimensional data encryption is a solution to improve search efficiency [8]. Some researchers try to study how to enrich the flexibility of retrieval [9], [10], however they still cannot meet the retrieval requirements of a large number of data, and they cannot sort and filter useful data for authorized users. Therefore, in the face of different user needs, it is urgent to find a scheme that can not only guarantee privacy, but also improve retrieval efficiency and ensure query accuracy. In this paper, we propose an efficient search method using features to match joint keywords (FMJK) on encrypted cloud data based on the MRSE (Privacy-Preserving Multi Keyword Ranked Search over Encrypted Cloud Data) scheme [3]. First, it is necessary to extract the features of each document to accurately express its theme. Then each d keywords are randomly selected from the non-duplicated keywords, which are extracted from the documents of the data owner, to generate a joint keyword, and all joint keywords form a keyword dictionary and then the features of each document are matched with the joint keywords to create an index. The authorized user enters the query keywords to match the joint keywords to create a trapdoor, and finally calculates the safe inner product of the trapdoor and the index to return the top k results. The contributions of this article are as follows: (1) Each d randomly selected keywords form a joint keyword, which matches with the feature of the document to be mapped to one dimension of the index, which reduces the dimension size of the key, the index and the trapdoor, simplifies the matrix operation during encryption, and improves search efficiency. (2) Through the improved BM25 algorithm [11] to calculate the inner product of the document index and the trapdoor, which not only sorts quickly but also ensures query accuracy. (3) The randomness of joint keywords and the encryption process of expanding and splitting ensure privacy protection

1.1 OBJECTIVE

The authorized user enters the query keywords to match the joint keywords to create a trapdoor, and finally calculates the safe inner product of the trapdoor and the index to return the top k results.

SJIF Impact Factor (2024): 8.675| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 9 | Issue: 12 | December 2024 - Peer Reviewed Journal

PROBLEM STATEMENT

With the increasing adoption of cloud storage for private data, the use of encrypted data has become essential for ensuring data security. However, traditional methods of indexing encrypted data using independent keywords result in significant challenges like Increased Storage Overhead, Decreased Search Efficiency

To address these challenges, an efficient search method that optimizes storage and search efficiency for encrypted cloud data is necessary.

1.2 EXISTING SYSTEM

- Reducing high dimensional data encryption is a solution to improve search efficiency.
- The existing cipher-text sorting retrieval methods, KNN (K Nearest Neighbor) technology is used to create indexes supporting cipher-text retrieval.
- In the process of massive data encryption search, most of the search encryption schemes have high time complexity and large storage space, which are closely related to the encrypted key, the document index and query request dimensions.

2 PROPOSED SYSTEM

- This method proposes keywords are selected from the keywords, which are extracted from the documents of the data owner, to generate a joint keyword, and all joint keywords form a keyword dictionary.
- Each joint keyword is matched with the feature of the document and the query keyword respectively, and the result obtained by the former is regarded, while the result obtained by the latter is regarded as a dimension of the query trapdoor.
- The proposed a flexible multi-keyword query scheme which takes keyword and user access history into when generating the query result

2.1 SYSTEM ARCHITECTURE

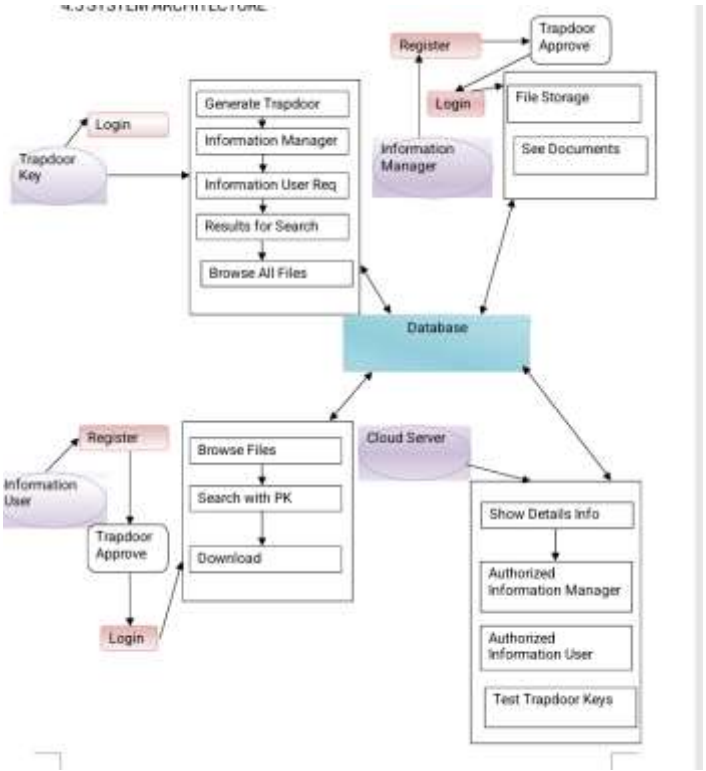

Fig2.1.1: Secure Information Management with Trapdoor Key Authorization

SJIF Impact Factor (2024): 8.675| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 9 | Issue: 12 | December 2024 - Peer Reviewed Journal

In Our Project it was Information manager has a register with all details. Then login it has a takes a permission with a trapdoor key authorizer has a accept a request. After accept the request it was a login. After login IM has an upload a data. and see the uploaded document. Then the information user has a register and then login takes a permission of trapdoor key. After approve login and see a document it will have a private key to protect a data from the attacker. Send a key to the verify. The cloud server has a login with id and password. Cloud server has a test the keys and then verify and send response in trapdoor. The Trapdoor has a send response to the user. Then the user has a download. Cloud server has information of an IM and IU authorized details.

2.2TASK OFFLOADING STATEGY

Stage:1- Preprocessing on the Client Side
Keyword Extraction:
Extract non-duplicated keywords from the documents before uploading. This minimizes redundant data processing in the cloud.
Joint Keyword Generation: Combine non-duplicated keywords to form joint keywords on the client side, reducing computational overhead on the cloud.
Feature Matching Preparation:
Generate initial document features locally, preparing them for efficient matching after upload.
Stage:2- Index Creation on the Cloud Side
Keyword Dictionary Management: Offload the storage and management of the joint keyword dictionary to the cloud. The cloud creates a centralized repository for joint keywords to support efficient query processing.
File Index Generation: Use the cloud's computational resources to generate file indices by associating joint keywords with document features. This reduces the client's resource burden.
Stage:3- Query Processing and Matching
Trapdoor Generation on the Client: Clients generate trapdoors (encrypted query representations) for their search requests locally, ensuring query privacy
Feature Matching in the Cloud: Offload the computationally intensive task of matching query trapdoors with file indices to the cloud. The cloud uses the precomputed keyword dictionary and document features for rapid search.
Stage:4- Result Retrieval and Post-Processing
Search Results Compilation on the Cloud: The cloud compiles a list of matching files and returns encrypted results to the client.
Decryption and Final Analysis on the Client: The client decrypts the results and performs any additional local filtering or analysis, ensuring sensitive data remains secure.

3. PEFORMANCE EVALUATION

The Performance Evaluation ensures that the output be produced within the time limits, and the time taken by the system for compiling, giving response to the users and request being send to the system for to retrieve the results

3.1 RESULT AND IMPLEMENTATION

Fig 3.1.1- Registration of Manager in the Database

SJIF Impact Factor (2024): 8.675| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 9 | Issue: 12 | December 2024 - Peer Reviewed Journal

It is a registration form for the manager. Manager enter the details like name, email id, age and gender.

Fig 3.1.2: Registration of User in the Database.

It is a registration form for the user. User enter the details for the user login

Fig 3.1.3:Trapdoor key login in the application

Trapdoor key login is a secure mechanism .Here trapdoor function is used to allow authentication or access to registered in a controlled manner.

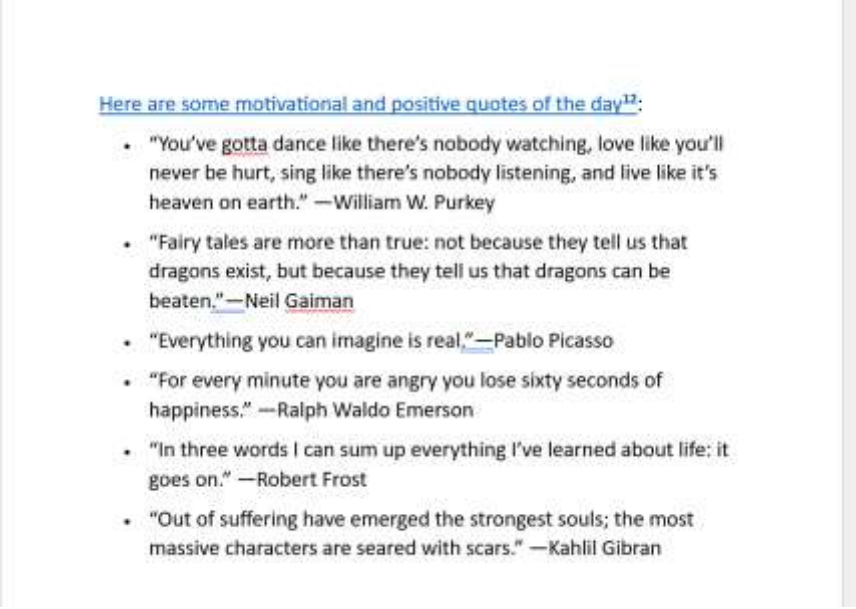

Fig 3.1.4 - Decrypted Data (What the User see with Private Key in Decryption Form)

SJIF Impact Factor (2024): 8.675| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 9 | Issue: 12 | December 2024 - Peer Reviewed Journal

This is the plain text inserted by the manger.

Fig 3.1.5-Encrypted Data (What the User see without Private Key in Encryption Form)

Here, the data was transformed using the cryptographic algorithms to make it unreadable without the correct decryption key.

4 CONCLUSIONS

In this paper, we propose an efficient search method using features to match joint keywords (FMJK) on encrypted cloud data. This method proposes that each keyword is randomly selected from the non-duplicated keywords, which are extracted from the documents of the data owner, to generate a joint keyword, and all joint keywords form a keyword dictionary, which greatly reduces the dimension of the keywords dictionary. Because the dimension of creating indexes and trapdoors is related to the dimension of keyword dictionary, it also reduces the dimension of the key, the indexes and the trapdoors, which improves the search efficiency.

REFERENCES

1. *Z. Wan and R. H. Deng, ''VP Search: Achieving verifiability for privacy preserving multi-keyword search over encrypted cloud data,'' IEEE Trans. Depend. Secure Compute, vol. 15, no. 6, pp. 1083–1095, Nov./Dec. 2016.*
2. *Y. Yang, H. Lin, X. Liu, W. Guo, X. Zheng, and Z. Liu, ''Blockchain based verifiable multi-keyword ranked search on encrypted cloud with fair payment,'' IEEE Access, vol. 7, pp. 140818–140832, 2019.*
3. *N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, ''Privacy-preserving multikey word ranked search over encrypted cloud data,'' in Proc. IEEE INFOCOM, Apr. 2011, pp. 829–837.*
4. *W. K. Wong, D. W.-L. Cheung, B. Kao, and N. Maroulis, ''Secure k NN computation on encrypted databases,'' in Proc. ACM SIGMOD Int. Conf. Manage. Data, Jun. 2009, pp. 139–152.*
5. *Z. Fu, X. Sun, Q. Liu, L. Zhou, and J. Shu, ''Achieving efficient cloud search services: Multi-keyword ranked search over encrypted cloud data supporting parallel computing,'' IEICE Trans. Commun., vol. E98.B, no. 1, pp. 190–200, 2015.*
6. *Z. Xia, X. Wang, X. Sun, and Q. Wang, ''A secure and dynamic multikey word ranked search scheme over encrypted cloud data,'' IEEE Trans. Parallel Disturb. Syst., vol. 27, no. 2, pp. 340–352, Jan. 2016.*
7. *W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, ''Verifiable privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking,'' IEEE Trans. Parallel Disturb. Syst., vol. 25, no. 11, pp. 3025–3035, Nov. 2014.*
8. *L. Liu and Z. Liu, ''A novel fast dimension-reducing ranked query method with high security for encrypted cloud data,'' Chin. J. Electron., vol. 29, no. 2, pp. 344–350, Mar. 2020.*
9. *W. Zhang, Y. Lin, and G. Qi, ''Catch you if you misbehave: Ranked keyword search results verification in cloud computing,'' IEEE Trans. Cloud Compute., vol. 6, no. 1, pp. 74–86, Mar. 2015.*
10. *Z. Guan, X. Liu, L. Wu, J. Wu, R. Xu, J. Zhang, and Y. Li, ''Cross-lingual multi-keyword rank search with semantic extension over encrypted data,'' Inf. Sci., vol. 514, pp. 523–540, Apr. 2020.*
11. *M. Murata, H. Nagano, R. Mukai, K. Kashino, and S. Satoh, ''BM25 with exponential IDF for instance search,'' IEEE Trans. Multimedia, vol. 16, no. 6, pp. 1690–1699, Oct. 2014.*
12. *D. Xiaodong Song, D. Wagner, and A. Perrigo, ''Practical techniques for searches on encrypted data,'' in Proc. IEEE Symp. Secure. Privacy (S&P), May 2000, pp. 44–55.*
13. *R. Cartmel, J. Garay, S. Kamara, and R. Ostrovsky, ''Searchable symmetric encryption: Improved definitions and efficient constructions,'' J. Compute. Secure., vol. 19, no. 5, pp. 895–934, Jan. 2011, Doi: 10.3233/JCS-2011- 0426.*
14. *R. Li, Z. Xu, W. Kang, K. C. Yow, and C.-Z. Xu, ''Efficient multi-keyword ranked query over encrypted data in cloud computing,'' Future Gener. Compute. Syst., vol. 30, no. 1, pp. 179–190, Jan. 2014, Doi: 10.1016/j. future.2013.06.029.*

SJIF Impact Factor (2024): 8.675| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 9 | Issue: 12 | December 2024 - Peer Reviewed Journal

15. *J. Wang, H. Ma, T. Qiang, L. Jin, H. Zhu, S. Ma, and X. Chen, ''A new efficient verifiable fuzzy keyword search scheme,'' J. Wireless Mobile Newt., Ubiquitous Compute., Dependable Appl., vol. 3, no. 4, pp. 61–71, Dec. 2012.*
16. *Z. Fu, X. Wu, C. Guan, X. Sun, and K. Ren, ''Toward efficient multikey word fuzzy search over encrypted outsourced data with accuracy improvement,'' IEEE Trans. Inf. Forensics Security, vol. 11, no. 12, pp. 2706–2716, Jul. 2016, Doi: 10.1109/TIFS.2016.2596138.*
17. *Z. Fu, F. Huang, K. Ren, J. Weng, and C. Wang, ''Privacy-preserving smart semantic search based on conceptual graphs over encrypted outsourced data,'' IEEE Trans. Inf. Forensics Security, vol. 12, no. 8, pp. 1874–1884, Aug. 2017, Doi: 10.1109/TIFS.2017.2692728.*
18. *R. Zhao, H. Li, Y. Yang, and Y. Liang, ''Privacy-preserving personalized search over encrypted cloud data supporting multi-keyword ranking,'' in Proc. 6th Int. Conf. Wireless Commun. Signal Process. (WCSP), Oct. 2014, pp. 1–6.*
19. *C. Chen, X. Zhu, P. Shen, J. Hu, S. Guo, Z. Tari, and A. Y. Zomaya, ''An efficient privacy-preserving ranked keyword search method,'' IEEE Trans. Parallel Distribute. Syst., vol. 27, no. 4, pp. 951–963, Apr. 2016.*
20. *RFC Index. Accessed: May 25, 2020. [Online]. Available: https://www.rfc-editor.org/rfc-index-100a.html*