Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 2: 2025
ISSN : **1906-9685**

Reliable and efficient DPF-based keyword search for cloud storage with multiple clients utilizing cache algorithm

Mr. Himambasha Shaik[1], Miriyala Sankeerthana[2]
#1. Assistant Professor Department of Master of Computer Applications,
#2. Pursuing MCA
QIS COLLEGE OF ENGINEERING AND TECHNOLOGY
Vengamukkalapalem (V), Ongole, Prakasam dist, Andhra Pradesh- 523272

ABSTRACT

The project addresses the security risks associated with cloud storage by developing a secure and efficient keyword search mechanism on encrypted data, ensuring data privacy while improving retrieval speed. It employs a multi-client Distributed Point Function (DPF) to enable keyword searches on encrypted data without compromising security. To optimize performance, Garbled Bloom Filter and Cuckoo Hash are used to compress keyword indexes and prevent collisions, while a segmented technique divides the search list across multiple threads to reduce computation time. The system enhances security further through double encryption using AES and ECC, safeguarding both user files and Garbled Bloom Filter indexes from unauthorized access. To ensure data integrity, Wegman authentication is integrated, allowing users to verify the correctness of search results by comparing hash codes. Additionally, the system's efficiency is improved through data compression and a cache algorithm, which reduce storage costs and search latency, making the overall process more streamlined.

INTRODUCTION

Cloud storage technology has advanced significantly, providing users with scalable solutions for data management. The increasing adoption of cloud services highlights the need for robust mechanisms to handle encrypted data securely and efficiently. The integration of innovative data management techniques, such as advanced indexing algorithms and distributed processing, is crucial for enhancing the efficiency and security of data retrieval processes in cloud environments. Efficient data indexing methods, including advanced hashing and filtering techniques, are vital for optimizing search operations within encrypted data sets. These

methods aim to improve the speed and accuracy of keyword searches while preserving data confidentiality. Modern cryptographic techniques play a pivotal role in securing data against unauthorized access. By employing advanced encryption standards and sophisticated hashing algorithms, the project leverages these technologies to address the challenges associated with encrypted data searches.

Objective:

This project aims to develop an efficient keyword search algorithm for encrypted cloud data, addressing high latency and computational costs. It will implement robust security measures, including double encryption with AES and ECC, to protect sensitive information. Advanced techniques like Garbled Bloom Filter and Cuckoo Hash will optimize search performance, while user-friendly mechanisms will enhance usability. By incorporating data compression and caching strategies, the system will minimize latency and improve retrieval speed, benefiting both users and cloud service providers.

- As cloud storage becomes more prevalent, the risk of unauthorized access and data breaches increases, necessitating robust security measures to protect sensitive information stored in the cloud.
- Existing methods for searching encrypted data are often inefficient, leading to high latency and increased computational costs when users attempt to retrieve files based on specific keywords.
- While encryption ensures data privacy, it prevents direct keyword searches on encrypted data, posing a challenge for efficient data retrieval.
- Users face significant challenges when trying to perform keyword searches on encrypted data, as traditional search techniques cannot be applied, resulting in a lack of usability and accessibility.
- There is a pressing need for innovative algorithms that can facilitate secure and efficient keyword searches in cloud environments, minimizing latency while ensuring data integrity and confidentiality.

LITERATURE SURVEY

3.1 Ranked Keyword Search Over Encrypted Cloud Data Through Machine Learning Method:

https://ieeexplore.ieee.org/abstract/document/9669027

ABSTRACT: Ranked keyword search over encrypted data has been extensively studied in cloud computing as it enables data users to find the most relevant results quickly. However, existing ranked multi-keyword search solutions cannot achieve efficient ciphertext search and

dynamic updates with forward security simultaneously. To solve the above problems, we first present a basic Machine Learning-based Ranked Keyword Search (ML-RKS) scheme in the static setting by using the k-means clustering algorithm and a balanced binary tree. ML-RKS reduces the search complexity without sacrificing the search accuracy, but is still vulnerable to forward security threats when applied in the dynamic setting. Then, we propose an Enhanced ML-RKS (called ML-RKS +) scheme by introducing a permutation matrix. ML-RKS + prevents cloud servers from making search queries over newly added files via previous tokens, thereby achieving forward security. The security analysis proves that our schemes protect the privacy of indexes, query tokens and keywords. Empirical experiments using the real-world dataset demonstrate that our schemes are efficient and feasible in practical applications.

3.2 Non-Interactive Multi-Client Searchable Encryption: Realization and Implementation:

https://ieeexplore.ieee.org/abstract/document/8998362

ABSTRACT: In this article, we introduce a new mechanism for constructing multi-client searchable encryption (SE). By tactfully leveraging the RSA-function, we propose the first multi-client SE protocol that successfully avoids per-query interaction between data owner and client. Therefore, our approach significantly reduces the communication cost by eliminating the need for data owner to authorize client queries at all times. To be compatible with the RSA-based approach, we also present a deterministic and memory-efficient 'keyword to prime' hash function, which may be of independent interest. Further, to improve efficiency, we put forward a more generic construction from set-constrained PRFs. The construction not only inherits the merits of our first protocol, but also achieves an enhanced security (against untrusted clients), where colluding attack among clients is also taken into account. Both protocols are instantiated via the recent representative SE protocol by Cash et al. with the support of boolean queries. At last, we implement our proposed protocols and comprehensively evaluate their performance to demonstrate their practicability and scalability.

3.3 Multi-Authority Attribute-Based Keyword Search over Encrypted Cloud Data:

https://ieeexplore.ieee.org/abstract/document/8798730

ABSTRACT: Searchable Encryption (SE) is an important technique to guarantee data security and usability in the cloud at the same time. Leveraging Ciphertext-Policy Attribute-Based Encryption (CP-ABE), the Ciphertext-

Policy Attribute-Based Keyword Search (CP-ABKS) scheme can achieve keyword-based retrieval and fine-grained access control simultaneously. However, the single attribute authority in existing CP-ABKS schemes is tasked with costly user certificate verification and secret key distribution. In addition, this results in a single-point performance bottleneck in distributed cloud systems. Thus, in this paper, we present a secure Multi-authority CP-ABKS (MABKS) system to address such limitations and minimize the computation and storage burden on resource-limited devices in cloud systems. In addition, the MABKS system is extended to support malicious attribute authority tracing and attribute update. Our rigorous security analysis shows that the MABKS system is selectively secure in both selective-matrix and selective-attribute models. Our experimental results using real-world datasets demonstrate the efficiency and utility of the MABKS system in practical applications.

3.4 Multi-User Collusion-Resistant Searchable Encryption with Optimal Search Time:

https://dl.acm.org/doi/abs/10.1145/3433210.3437535

ABSTRACT: The continued development of cloud computing requires technologies that protect users' data privacy even from the cloud providers themselves. Multi-user searchable encryption is one such kind of technology. It allows a data owner to selectively enable users to perform keyword searches over her encrypted documents that are stored at a cloud server. For privacy purposes, it is important to limit what an adversarial server can infer about the encrypted documents, even if it colludes with some of the users. Clearly, in this case it can learn the content of documents shared with this subset of "corrupted" users, however, it is important to ensure that this collusion does not reveal information about parts of the dataset that are only shared with the remaining "uncorrupted" users via cross-user leakage. In this work, we propose three novel multi-user searchable encryption schemes for this setting that achieve different trade-offs between performance and leakage. Compared to previous ones, our first two schemes are the first to achieve asymptotically optimal search time. Our third scheme achieves minimal user storage and forward privacy with respect to document sharing, but slightly slower search performance. We formally prove the security of our schemes under reasonable assumptions. Moreover, we implement and evaluate their performance both on a single machine and over WAN. Our experimental results are encouraging, e.g., the search computation time is in the order of a few milliseconds.

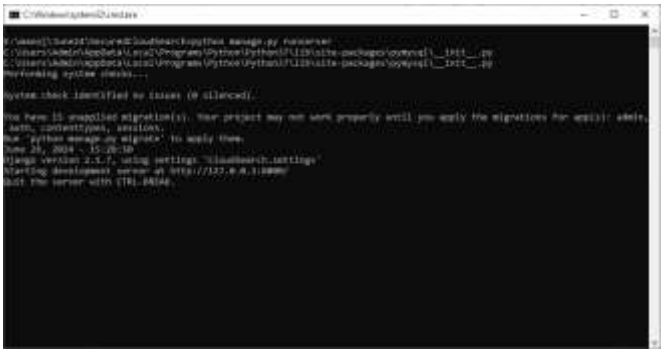
3.5 Harnessing Encrypted Data in Cloud for Secure and Efficient Mobile Image Sharing:

https://ieeexplore.ieee.org/abstract/document/7524781

ABSTRACT: Nowadays, large volumes of multimedia data are outsourced to the cloud to better serve mobile applications. Along with this trend, highly correlated datasets can occur commonly, where the rich information buried in correlated data is useful for many cloud data generation/dissemination services. In light of this, we propose to enable a secure and efficient cloud-assisted image sharing architecture for mobile devices, by leveraging outsourced encrypted image datasets with privacy assurance. Different from traditional image sharing, we aim to provide a mobile-friendly design that saves the transmission cost for mobile clients, by directly utilizing outsourced correlated images to reproduce the image of interest inside the cloud for immediate dissemination. First, we propose a secure and efficient index design that allows the mobile client to securely find from encrypted image datasets the candidate selection pertaining to the image of interest for sharing. We then design two specialized encryption mechanisms that support secure image reproduction from encrypted candidate selection. We formally analyze the security strength of the design. Our experiments explicitly show that both the bandwidth and energy consumptions at the mobile client can be saved, while achieving all service requirements and security guarantees.

SYSTEM ARCHITECTURE:

To run project double click on 'runServer.bat' file to get below screen



In above screen cloud server started and now open browser and enter URL as http://127.0.0.1:8000/index.html and press enter key to get below page

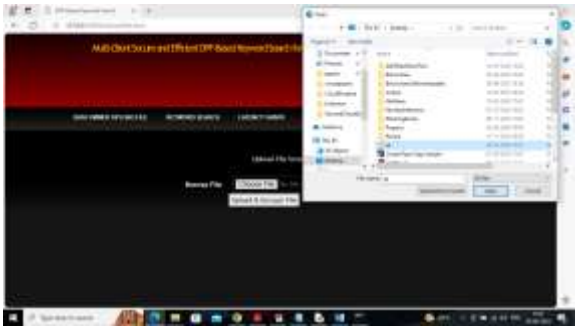
In above screen click on 'Register Here' link to get below page

In above screen user is entering sign up details and then press button to register user with cloud and get below page

In above screen user sign up completed and similarly you can register other user and now click on 'User' link to get below login screen

In above screen user is login and after login will get below page

In above screen 'Data Owner' can click on 'Data Owner Upload File' link to get below page



In above screen selecting and uploading '.txt' file and then click on 'Open and Upload' button to upload file to cloud and get below page

In above screen file uploaded to cloud and can see 'Wegman' generated authentication for loaded file and now click on 'Keyword Search' link to get below search screen

In above screen giving some queries and then press button to get below output

In above screen can see list of searched files for given query and similarly you can search for any query

In above screen giving another query and below is the search result

In above screen can see list of searched file and now can click on 'Click Here to Download' link to download file

In above screen can see file is downloaded and now click on 'Latency Graph' link to get below graph

In above graph x-axis represents 'number of query searched' and y-axis represents latency or time and can see propose is faster than existing as propose algorithm will segment query and search via multiple threads and now click on 'Storage Cost Graph' link to get below page

In above graph x-axis represents algorithm names and y-axis represents storage space and in both propose and extension storage we can see extension got less storage space. Now click on 'Access Cache' link to view and download all previous searched keywords and file like below screen

In above screen user can see keywords and the appropriate file searched for those keywords and user can view and download desired file without performing any keywords searched computations.

CONCLUSION

The proposed project enhances cloud storage security and efficiency by implementing advanced encryption techniques and authentication methods, ensuring robust protection and confidentiality for user data. By leveraging innovative algorithms such as Garbled Bloom Filters and Cuckoo hashing, the system significantly improves search efficiency, reducing latency and computational overhead for keyword searches on encrypted data. The user-centric design of the project offers an intuitive interface for file management and searches, simplifying the process for users and enhancing their experience. Additionally, the incorporation of data compression algorithms and caching mechanisms reduces storage costs and boosts overall system performance, making cloud storage more cost-effective. This project also establishes a strong foundation for future enhancements, such as integrating machine learning, supporting cloud environments, and advancing user authentication methods, which will further enhance the system's capabilities and adaptability to emerging technologies and user needs.

Future Scope:

Future work can focus on improving the scalability of the proposed system to handle larger datasets and a higher number of clients. This could involve optimizing the underlying algorithms and data structures to ensure efficient performance as the system grows.

Future developments could include mechanisms for efficiently handling dynamic data, such as adding, updating, or deleting documents in the cloud storage. This would require designing secure protocols that maintain the integrity and confidentiality of the keyword indexes during such operations.

Developing user-friendly interfaces and tools for clients to interact with the system can enhance usability. This could include intuitive search

functionalities, visualization of search results, and easy management of access controls, making the system more accessible to non-technical users.

Conducting extensive real-world deployment and evaluation of the proposed system in various cloud environments can provide valuable insights into its practical performance, security, and user experience. This could lead to further refinements and adaptations based on user feedback and operational challenges encountered in real-world scenarios.

REFERENCES

[1] Y. Miao, W. Zheng, X. Jia, X. Liu, K. R. Choo and R. Deng, "Ranked keyword search over encrypted cloud data through machine learning method", IEEE Trans. Serv. Comput., vol. 16, no. 1, pp. 525-536, Jan./Feb. 2023.

[2] S.-F. Sun et al., "Non-interactive multi-client searchable encryption: Realization and implementation", IEEE Trans. Dependable Secure Comput., vol. 19, no. 1, pp. 452-467, Jan./Feb. 2022.

[3] Y. Miao, R. H. Deng, X. Liu, K. R. Choo, H. Wu and H. Li, "Multi-authority attribute-based keyword search over encrypted cloud data", IEEE Trans. Dependable Secure Comput., vol. 18, no. 4, pp. 1667-1680, Jul./Aug. 2021.

[4] Y. Wang and D. Papadopoulos, "Multi-user collusion-resistant searchable encryption with optimal search time", Proc. ACM Asia Conf. Comput. Commun. Secur., pp. 252-264, 2021.

[5] H. Cui, X. Yuan and C. Wang, "Harnessing encrypted data in cloud for secure and efficient mobile image sharing", IEEE Trans. Mobile Comput., vol. 16, no. 5, pp. 1315-1329, May 2017.

[6] X. Shen et al., "Data management for future wireless networks: Architecture privacy preservation and regulation", IEEE Netw., vol. 35, no. 1, pp. 8-15, Jan./Feb. 2021.

[7] S. G. Choi, D. Dachman-Soled, S. D. Gordon, L. Liu and A. Yerukhimovich, "Compressed oblivious encoding for homomorphically encrypted search", Proc. ACM Conf. Comput. Commun. Secur., pp. 2277-2291, 2021.

[8] E. Stefanov et al., "Path ORAM: An extremely simple oblivious RAM protocol", J. ACM, vol. 65, no. 4, pp. 1-26, 2018.

[9] S. Oya and F. Kerschbaum, "Hiding the access pattern is not enough: Exploiting search pattern leakage in searchable encryption", Proc. USENIX Secur. Symp., pp. 127-142, 2021.

[10] Z. Gui, K. G. Paterson and S. Patranabis, "Rethinking searchable

symmetric encryption", Proc. IEEE Secur. Privacy, 2023.

[11] E. Boyle, N. Gilboa and Y. Ishai, "Function secret sharing: Improvements and extensions", Proc. ACM Conf. Comput. Commun. Secur., pp. 1292-1303, 2016.

[12] B. Chor, E. Kushilevitz, O. Goldreich and M. Sudan, "Private information retrieval", J. ACM, vol. 45, no. 6, pp. 965-981, 1998.

[13] N. Gilboa and Y. Ishai, "Distributed point functions and their applications", Proc. 33rd Annu. Int. Conf. Theory Appl. Cryptographic Techn., pp. 640-658, 2014.

[14] E. Dauterman, E. Feng, E. Luo, R. A. Popa and I. Stoica, "DORY: An encrypted search system with distributed trust", Proc. USENIX Conf. Operating Syst. Des. Implementation, pp. 1101-1119, 2020.

[15] L. de Castro and A. Polychroniadou, "Lightweight maliciously secure verifiable function secret sharing", Proc. 41st Annu. Int. Conf. Theory Appl. Cryptographic Techn., pp. 150-179, 2022.

[16] O. Goldreich, S. Goldwasser and S. Micali, "How to construct random functions", J. ACM, vol. 33, no. 4, pp. 792-807, 1986.

[17] R. Kumar, S. Rajagopalan and A. Sahai, "Coding constructions for blacklisting problems without computational assumptions", Proc. 19th Annu. Int. Cryptol. Conf., pp. 609-623, 1999.

[18] S. Agrawal and D. Boneh, "Homomorphic MACs: Mac-based integrity for network coding", Proc. 7th Int. Conf. Appl. Cryptogr. Netw. Secur., pp. 292-305, 2009.

[19] [online] Available: https://github.com/EnderCheng/KeywordSearch.

[20] Z. Shang, S. Oya, A. Peter and F. Kerschbaum, "Obfuscated access and search patterns in searchable encryption", Proc. Netw. Distrib. Syst. Secur. Symp., pp. 1-18, 2021.

[21] X. Wang, J. Ma, X. Liu, Y. Miao, Y. Liu and R. H. Deng, "Forward/backward and content private dsse for spatial keyword queries" in IEEE Trans. Dependable Secure Comput.

[22] J. G. Chamani, D. Papadopoulos, C. Papamanthou and R. Jalili, "New constructions for forward and backward private symmetric searchable encryption", Proc. ACM Conf. Comput. Commun. Secur., pp. 1038-1055, 2018.

[23] R. Bost, "\$Sigma\$Σ o \$varphi\$ϕ o \$varsigma\$ς : Forward secure searchable encryption ", Proc. ACM Conf. Comput. Commun. Secur., pp. 1143-1154, 2016.

[24] C. Dong, L. Chen and Z. Wen, "When private set intersection meets Big Data: An efficient and scalable protocol", Proc. ACM Conf. Comput. Commun. Secur., pp. 789-800, 2013.

[25] J. Katz and A. Y. Lindell, "Aggregate message authentication codes", Proc. Cryptographers Track RSA Conf., pp. 155-169, 2008.

[26] D. J. Bernstein, "Stronger security bounds for Wegman-Carter-Shoup authenticators", Proc. 24th Annu. Int. Conf. Theory Appl. Cryptogr. Techn., pp. 164-180, 2005.

[27] S.-F. Sun et al., "Practical non-interactive searchable encryption with forward and backward privacy", Proc. Netw. Distrib. Syst. Secur. Symp., pp. 1-18, 2021.

[28] A. Davidson, S. Katsumata, R. Nishimaki, S. Yamada and T. Yamakawa, "Adaptively secure constrained pseudorandom functions in the standard model", Proc. 40th Annu. Int. Cryptol. Conf., pp. 559-589, 2020.

[29] G. Hartung, B. Kaidel, A. Koch, J. Koch and A. Rupp, "Fault-tolerant aggregate signatures", Proc. 19th IACR Int. Conf. Pract. Theory Public-Key Cryptogr., pp. 331-356, 2016.

[30] M. Shen, B. Ma, L. Zhu, X. Du and K. Xu, "Secure phrase search for intelligent processing of encrypted data in cloud-based IoT", IEEE Internet Things J., vol. 6, no. 2, pp. 1998-2008, Apr. 2019.

Authors:

Mr. Himambasha Shaik is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. He earned his Master of Computer Applications (MCA) from Anna University, Chennai. With a strong research background, He has authored and co-authored research papers published in reputed peer-reviewed journals. His research interests include Machine Learning, Artificial Intelligence, Cloud Computing, and Programming Languages. He is committed to advancing research and fostering innovation while mentoring students to excel in both academic and professional pursuits.

Miriyala Sankeerthana is a postgraduate student a MCA in the Department of Computer Applications at QIS College of Engineering & Technology, Ongole an Autonomous college in prakasam dist..