IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203114

Privacy Preserving and Keyword Search for Multi Tenancy Cloud

N.Ranjith Reddy, N.Srivalli, P.Pranith Kumar, Dr.Khushbhu Doulani

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

Associate Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

ABSTRACT: cloud service models intrinsically cater to multiple tenants. In current multi-tenancy model, cloud service providers isolate data within a single tenant boundary with no or minimum cross-tenant interaction. With the booming of cloud applications, allowing a user to search across tenants is crucial to utilize stored data more effectively. However, conducting such a search operation is inherently risky, primarily due to privacy concerns. Moreover, existing schemes typically focus on a single tenant and are not well suited to extend support to a multi-tenancy cloud, where each tenant operates independently. In this article, to address the above issue, we provide a privacy preserving, verifiable, accountable, and parallelizable solution for "privacy-preserving keyword search problem" among multiple independent data owners. We consider a scenario in which each tenant is a data owner and a user's goal is to efficiently search for granted documents that contain the target keyword among all the data owners. We first propose a verifiable yet accountable keyword searchable encryption (VAKSE) scheme through symmetric bilinear mapping. For verifiability, a message authentication code (MAC) is computed for each associated piece of data. To maintain a consistent size of MAC, the computed MACs undergo an exclusive OR operation. For accountability, we propose a keyword-based accountable token mechanism where the client's identity is seamlessly embedded without compromising privacy. Furthermore, we introduce the parallel VAKSE scheme, in which the inverted index is partitioned into small segments and all of them can be processed synchronously. We also conduct formal security analysis and comprehensive experiments to demonstrate the data privacy preservation and efficiency of the proposed schemes, respectively.

I. INTRODUCTION

1. GENERAL

CLOUD computing has had a profound impact on data management. It offers massive storage and computing resources, payment-on-demand, and flexible scalability. Motivated by these advantages, thousands of clients are opting for cloud services. One typical application area is healthcare, and some applications are Healthvana [1] and CDPHP [2]; both the platforms are the tenants of Amazon [3]. Healthvana stores patient reports and CDPHP stores doctor information. It is desirable for a patient to search both the datasets to find the most suitable doctor by matching the patient data with the doctor information. For example, HIV patients store their reports in Healthvana and seek for suitable doctors from CDPHP. However, such a search across tenancies is challenging. Each tenant is an independent data owner and must abide the privacy laws, such as HIPAA [4], which are enforced to protect individuals' medical data privacy. In addition, for their own interests, companies treat patient data as an asset and tend to maintain complete control over it. Data encryption is the best practice for maintaining data privacy. Each data owner encrypts their data before outsourcing it to the cloud. This guarantees the confidentiality of the data but greatly reduces their utility. A user must download an entire dataset in order to retrieve one piece of data. Considering data utility and privacy, Song et al. [5] introduced the primitives of symmetric searchable encryption (SSE). SSE is a keyword search technique that allows search over the cipher text without decryption. Goh et al. [6] proposed a secure index to improve search efficiency. Subsequently, Curtmola et al. [7] formalized the security definition of SSE and proposed two constructions that corresponded to non adaptive semantic security and adaptive semantic security.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203114

In early research, most works on SSE focused on the honestbut-curious cloud service provider (CSP). In such a model, the search result is fully trusted and the CSP is assumed to honestly follow the protocol specification. Search results in practice may contain corrupted data due to underlying hardware/software failures. In addition, for self-interest, the CSP may deviate from the protocol specification. For example, to reduce computational costs, CSP may randomly choose data as a search result. To mitigate this problem, Chai and Gong [8] proposed verifiable SSE, where the search result includes not only retrieved documents but also proof of the correctness and completeness of the search. The correctness of the search means that the returned search result matches the query. completeness of the search means that the retrieved data has not been tampered with. In addition, Chen et al. [9] proposed an authenticated Merkle hash tree to verify the search result. Although significant progress has been made by the existing constructions [8], [9], the verifiable property comes at the high cost of extra storage and computation. There is still room for solutions that are more practical. Recently, with increasing demand for users (e.g., physicians), previous SSE constructions, providing a client either full access to the data or no access, expose their short term. It is desirable to design a fine-grained access control mechanism to enable data owners to selectively grant grant clients access to their data. To achieve this goal, Han et al. [10] proposed to apply attribute-based encryption [11] to solve this problem and provided a general solution in the context of public-key keyword search scenarios. With this design, only a keyword search request that matches the predefined access structure can retrieve the target document. The above searchable encryption schemes rely on public key encryption, which is inefficient compared with symmetric encryption. Moreover, none of them are suitable for use with dynamic dynamic access structures since the access structure is associated with either a key or cipher text. Any change in the access structure may result in all of the ciphertext or keys being renewed. Furthermore, all the mentioned works failed to allow a client to search the data from multiple data owners, where each data owner encrypts their data with a unique key. The existing SSE schemes only support a client to search over a single data owner [5], [12], [13].

II. SCOPE OF THE PROJECT

The scope of the project **"Privacy-Preserving and Keyword Search for Multi-Tenancy Cloud"** includes the development of a secure and efficient system that allows multiple tenants to store and search their encrypted data on a shared cloud infrastructure. The primary focus is to enable keyword-based search functionality over encrypted data without revealing the actual content of the data or the search queries to the cloud service provider or other tenants. The system ensures data confidentiality, query privacy, and strict access control by implementing advanced cryptographic techniques such as searchable encryption. It supports multi-tenancy by logically isolating data and access rights for each tenant, ensuring that users can only access their own data. Additionally, the system is designed to be scalable and efficient, providing fast and accurate search results with minimal overhead, even as the number of tenants and data volume increases.

III. OBJECTIVE

The main objective of this project is to design and implement a secure system that allows multiple cloud tenants to perform keyword-based searches on their encrypted data stored in the cloud, while ensuring data privacy, query confidentiality, and access control. The system aims to prevent the cloud service provider and unauthorized users from accessing the content of the data or the search queries. It also ensures proper data isolation between different tenants in a multi-tenant cloud environment. Additionally, the project focuses on maintaining efficient performance, enabling scalable and fast search operations without compromising security or usability.

IV. PROBLEM STATEMENT

In a multi-tenancy cloud environment, multiple users or organizations share the same infrastructure and storage services. While cloud computing offers scalability and cost-efficiency, it raises serious concerns about **data privacy and security**, especially when sensitive information is stored. Traditional cloud services do not allow **secure** keyword-based searches on encrypted data, forcing users to either decrypt their data (compromising security) or rely on the cloud provider (compromising privacy). Additionally, in a multi-tenant setting, there's a risk of **data leakage across tenants**, unauthorized access, and **lack of query confidentiality**. Therefore, there is a critical need for a system that enables privacy-preserving keyword search over encrypted data in a way that supports multiple tenants, ensures **strong isolation**, and protects both data and query privacy from other users and the cloud provider—without compromising performance or usability.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203114

EXISTING SYSTEM

o Moreover, existing schemes typically focus on a single tenant and are not well suited to extend support to a multi-tenancy cloud, where each tenant operates independently.

o In this to address the above issue, we provide a privacy preserving, verifiable, accountable, and parallelizable solution for "privacy-preserving keyword search problem" among multiple independent data owners.

o We consider a scenario in which each tenant is a data owner and a user's goal is to efficiently search for granted documents that contain the target keyword among all the data owners.

EXISTING SYSTEM DISADVANTAGES

- No Security authentication.
- It cannot be security to a data.

V. LITERATURE SURVEY

Recent research in the field of privacy-preserving keyword search for multi-tenancy cloud environments has seen significant advancements. Traditional techniques such as Searchable Symmetric Encryption (SSE), introduced by Curtmola et al., have been enhanced to support dynamic data operations with forward and backward privacy, ensuring minimal data leakage during updates. A study by Sun et al. (2023) proposed a ranked search mechanism using secure index structures and order-preserving encryption, allowing users to retrieve the most relevant documents without revealing actual data or queries. Zhang et al. (2023) addressed multi-tenancy specifically by developing a framework that uses identity-based encryption to isolate tenants and protect against cross-tenant data inference. Meanwhile, Chen et al. (2024) incorporated blockchain technology to make search operations auditable, enhancing trust and transparency in shared environments. Another contribution by Li et al. (2024) focused on lightweight searchable encryption for IoT-cloud systems, which is also applicable to multi-tenant scenarios where computational efficiency is critical. Their model supports fuzzy keyword search, improving the system's usability. These recent works collectively highlight the growing emphasis on privacy, scalability, and efficiency in secure cloud search solutions.

VI. PROPOSED SYSTEM

➢ We first propose a verifiable yet accountable keyword searchable encryption (VAKSE) scheme through symmetric bilinear mapping.

➢ For verifiability, a message authentication code (MAC) is computed for each associated piece of data. To maintain a consistent size of MAC, the computed Macs undergo an exclusive OR operation.

➢ For accountability, we propose a keyword-based accountable token mechanism where the client's identity is seamlessly embedded without compromising privacy. Furthermore, we introduce the parallel VAKSE scheme, in which the inverted index is partitioned into small segments and all of them can be processed synchronously.

➢ We also conduct formal security analysis and comprehensive experiments to demonstrate the data privacy preservation and efficiency of the proposed schemes, respectively.

VII. PROPOSED SYSTEM ADVANTAGES

Enhanced Data Privacy: All tenant data is encrypted before storage, ensuring that neither the cloud provider nor other tenants can access the plaintext information.

Secure Keyword Search: Enables efficient keyword search directly over encrypted data using **searchable encryption**, without compromising data or query confidentiality.

Tenant Isolation: The system ensures **strict separation** between different tenants' data and search operations, preventing any possibility of data leakage or cross-tenant access.

VIII. FUTURE ENHANCEMENT

Future enhancements for the proposed system on privacy-preserving keyword search in a multi-tenancy cloud environment aim to further improve its security, efficiency, and usability. One major advancement could be the integration of artificial intelligence or machine learning to optimize search accuracy and provide more relevant results based on user behavior or context. The system can also be enhanced to support **semantic search**, allowing users to search by meaning rather than exact keywords, thereby improving usability. To cater to a broader range of users, future versions could be optimized for **mobile and edge devices**, enabling secure cloud access even on low-power platforms.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203114

Incorporating **blockchain technology** for secure and tamper-proof logging of user activities would add an auditable layer of trust. Additionally, implementing **zero-knowledge proofs (ZKP)** could allow the cloud server to prove correct execution of search queries without revealing any sensitive information. Expanding the system to be compatible with **multi-cloud or hybrid cloud architectures** would increase flexibility and resilience. Furthermore, providing **real-time indexing**, **API support**, and **multi-level user access control** would make the system more adaptable to enterprise environments. These enhancements collectively aim to make the system more intelligent, secure, and applicable to a wider range of real-world use cases.

IX. CONCLUSION

In conclusion, the proposed system for privacy-preserving keyword search in a multi-tenancy cloud environment addresses critical challenges related to data confidentiality, query privacy, and tenant isolation. By leveraging advanced cryptographic techniques such as searchable encryption and implementing strict access controls, the system ensures that sensitive data remains secure while still allowing efficient and accurate search capabilities. This approach not only protects users' information from unauthorized access by the cloud provider or other tenants but also maintains performance and scalability suitable for real-world applications. As cloud adoption continues to grow across industries that handle sensitive data, such as healthcare and finance, this system provides a robust solution for balancing the benefits of cloud computing with the essential need for privacy and security.

REFERENCES

[1] (2022). Healthvana. [Online]. Available: https://healthvana.com
[2] (2022). CDPHP. [Online]. Available: https://www.cdphp.com
[3] (2022). Customer Success Stories. [Online]. Available: https://aws. amazon.com/solutions/case-studies/
[4] (2022). HiPAA. [Online]. Available: http://www.cms.hhs.gov/ HIPAAGenInfo
[5] D. Xiaoding Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Secur. Privacy. (S&P), May 2000, pp. 44–55.
[6] E.-J. Goh, "Secure indexes," Cryptol. ePrint Arch., Oct. 2003.
[7] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," J. Comput. Secur., vol. 19, no. 5, pp. 895–934, 2011.
[8] Q. Chai and G. Gong, "Verifiable symmetric searchable encryption for semi-honest- but-curious cloud servers," in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2012, pp. 917–922.
[9] C. Chen et al., "An efficient privacy-preserving ranked keyword search method," IEEE Trans. Parallel Distrib. Syst., vol. 27, no. 4, pp. 951–963, Apr. 2016.
[10] F. Han, J. Qin, H. Zhao, and J. Hu, "A general transformation from KP ABE to searchable encryption," Future Gener. Comput. Syst., vol. 30, pp. 107–115, Jan. 2014.
[11] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryp tion for fine-grained access control of encrypted data," in Proc. 13th ACM Conf. Comput. Commun. Secur., Oct. 2006, pp. 89–98.
[12] R. Brinkman, L. Feng, J. Doumen, P. H. Hartel, and W. Jonker, "Efficient tree search in encrypted data," Inf. Syst. Secur., vol. 13, no. 3, pp. 14–21, May 2004.
[13] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: Protecting confidentiality with encrypted query process ing," in Proc. 23rd ACM Symp. Oper. Syst. Principles (SOSP), 2011, pp. 85–100.
[14] J. Wang and S. S. Chow, "Omnes pro uno: Practical multi-writer encrypted database," in Proc. 31st USENIX Secur. Symp. (USENIX Security), 2022, pp. 2371–2388.
[15] J. Li, K. Ren, B. Zhu, and Z. Wan, "Privacy-aware attribute-based encryption with user accountability," in Proc. Int. Conf. Inf. Secur. Berlin, Germany: Springer, 2009, pp. 347–362. [16] J. Li, X. Chen, S. S. M. Chow, Q. Huang, D. S. Wong, and Z. Liu, "Multi-authority fine-grained access control with accountability and its application in cloud," J. Netw. Comput. Appl., vol. 112, pp. 89–96, Jun. 2018.
[17] A. Soleimanian and S. Khazaei, "Publicly verifiable searchable sym metric encryption based on efficient cryptographic components," Des., Codes Cryptogr., vol. 87, no. 1, pp. 123–147, 2019.
[18] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy preserving multi-keyword ranked search over encrypted cloud data," IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 1, pp. 222–233, Jan. 2013.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com