Migration Letters

Volume: 21, No: S5 (2024), pp. 1913-1920

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

Enhancing Keyword Search Privacy In Encrypted Cloud Data Through Optimization

[1]Ms.A.Gomathy , [2]Aishwarya G, [3]Anudharshini K, [4]Anusha V, [5]Bharanika P

ABSTRACT:

As cloud computing gains prominence, organizations find themselves compelled to shift Moving their intricate data management systems from on-premise locations to private cloud providers, seeking enhanced flexibility and cost-effectiveness. Nonetheless, to safeguard sensitive data, it is crucial to encrypt it before externalizing. In order to address the demand for efficient data retrieval, these search services must support multi-keyword queries and offer similarity ratings for results, considering the substantial volume of data and documents stored in cloud storage. Efforts in searchable encryption often overlook the distinction between search results, typically concentrating on single-phrase or Boolean keyword searches. This research represents a notable progress as it introduces and tackles The complex problem of privacy-preserving multi-keyword ranking, ontology keyword

Mapping, and search over encrypted cloud data (EARM). Additionally, we establish a stringent set of privacy guidelines that must be followed for The deployment of a system for utilizing cloud data securely. Our preferred approach for measuring similarity between search queries and data documents involves the effective principle of "Enhanced Association Rule Mining coordinate matching," emphasizing the capture of as many matches as possible in the context of multiple-keyword semantics.

assessment, we employ the concept of "inner product similarity." Initially, we introduce a fundamental EARM technique using secure inner product computing, which is subsequently enhanced to satisfy various privacy criteria across two threat model levels

Keywords: *Encrypted Cloud Data, Cryptography, Boolean search.*

I INTRODUCTION:

In an era dominated by extensive cloud storage and data outsourcing, safeguarding the privacy and security of s[1]ensitive information has become paramount. A significant approach to tackle these challenges involves the implementation of authenticated Boolean search over encrypted cloud data. With an increasing number of individuals entrusting their data to cloud service providers, there is a rising need for robust and secure techniques to search for and retrieve this

[1]Assistant Professor, Department of Computer Science & Engineering, Sri Krishna college of Technology, Coimbatore, Tamil Nadu, India.

[2-5]Department of Computer Science & Engineering, Sri Krishna college of Technology, Coimbatore, Tamil Nadu, India.

data without compromising its confidentiality. To achieve a nuanced equilibrium between privacy protection and search functionality, this paper introduces a sophisticated methodology for efficiently validating Boolean searches over encrypted cloud data. By leveraging cryptographic techniques, the suggested methodology establishes a robust framework for conducting secure and privacy-preserving information retrieval in cloud environments. It simultaneously guarantees the morality of search results and the secretiveness of user data.In our data-driven society, cloud computing is anticipated to assume an increasingly significant role as a well-established computing paradigm [3]. Thorough evaluations have been undertaken to tackle potential security and privacyconcerns linked to the transfer of transferring user data and cerebral property to the cloud, coupled with the exploration of possible security remedies. One straightforward approach is the application of traditional encryption to safeguard data both at rest and in transit. Nevertheless, this method limits users from effectively sharing and searching encrypted data, thus negatively impacting their search experience. This constraint has prompted the emergence of techniques like Ciphertext-Policy Attribute-Based Keyword

Search (CPABKS) [3]. CP-ABKS methods provide nuanced access control for encrypted data, surpassing keyword-based search capabilities. Unlike conventional Searchable Encryption (SE) approaches, CP-ABKS addresses issues related to coarse-grained access control, cumbersome key management, and resource inefficiency stemming from multiple ciphertext copies.

Despite their benefits, the practical adoption of the utilization of CP-ABKS schemes is constrained by the computational and storage costs associated with current schemes.

These expenses, as demonstrated in the mentioned schemes, increase proportionally with the complicatedness of access regulations or the amount of users' attribute data.SE empowers a client to encrypt data and generate search tokens for querying a storage server. By employing a token, The server has the capability to execute searches on the encrypted data and provide the corresponding encrypted files. Informally, the security of an SSE scheme is affirmed when: (1) the ciphertext remains non-informative about the underlying data; (2) the combination of ciphertext with a search token reveals only the pertinent search result; and (3) the creation of search tokens is exclusive to the possession of the secret key. SSE finds its primary application in the creation of searchable crypto cloud storage, ensuring end-to-end security without compromising functionality.Additionally, SSE is employed in the advancement of graph encryption methods and controlled disclosure systems [5].The growing practice of entrusting encrypted data to the cloud has sparked renewed enthusiasm for Searchable Encryption (SE) systems, enabling users to explore encrypted data. Whether symmetric or asymmetric [16], current SE schemes, as exemplified in, are often tailored for a singular function, rendering them impractical [16].The primary goal of the research is to improve the privacy of keyword searches. This could involve protecting user search queries, ensuring that search results reveal minimal information, and overall enhancing the confidentiality of the search process. The focus is on data stored in the cloud, and specifically, this data is encrypted. This suggests that the research aims to improve the efficiency, effectiveness, or security of keyword searches through the application of optimization methods. Optimization could be related to algorithms, cryptographic techniques, or other strategies to enhance the overall search process.

II BACKGROUND STUDY

CP-ABKS schemes, as pioneered by Ming Zeng and collaborators [3], Provide a robust solution for users with constrained resources who aim to securely upload, share, and search data in the cloud. These schemes offer assistance for both precise access control and the retrieval of ciphertext based on keywords. In the realm of cloud searching systems, the conventional approach often neglects the practical consideration of multi-owner scenarios or fails to address the risk of malevolent cloud servers providing inaccurate search results. To address these constraints, our initial strategy introduces an enhanced Verifiable Fine grained keyword Search technique tailored for the static owners environment. method achieves a condensed ciphertext length, rapid ciphertext translation, expedited search processing, and dependable verification of search results.

As per the findings of Yinbin Miao and colleagues [6], Searchable Encryption (SE) plays a pivotal role in maintaining data security while facilitating cloud utilization. The CP-ABKS approach employs CP-ABE to accomplish Retrieval based on keywords and precise access control are both supported. However, existing CP-ABKS schemes encounter difficulties when dealing with a sole attribute authority, resulting in resource-intensive procedures like verification of user certificates and distribution of secret keys. [6]. Furthermore, this leads to a performance bottleneck concentrated at a singular point within decentralized cloud systems. .The goal is to address these margins and ease the computational and storage burdens on resource-constrained devices within cloud systems [6].Dawn Xiaodong Song and colleagues [7] delve into cryptographic techniques for encrypted data searching in this chapter, accompanied by security proofs for the resulting cryptosystems. Our techniques offer several notable advantages. They are inherently secure, providing verifiable confidentiality in encryption, ensuring that the untrusted server remains oblivious to the plaintext when presented only with the ciphertext. Additionally, our methods guarantee query isolation during searches, preventing the untrusted server from gaining any additional knowledge about the plaintext beyond the search result. In addition, our methods support regulated searching, preventing the untrusted server from executing searches for arbitrary words without explicit user authorization. Moreover, these methods allow for covert queries, granting users the ability to direct the untrusted server to search for a confidential phrase without disclosing it [7]. The approaches we introduce are uncomplicated, effective and involve minimal space and transmission overhead.

This makes them well-suited for practical applications in the current context.

Chengjun Cai and colleagues [10]. The blockchain technology has spearheaded the decentralization of applications, proving its utility beyond cryptocurrency. In decentralized storage solutions like Storj and Sia, blockchain serves as the foundation for an open platform in the sharing economy, delivering secure and loyal file-outsourcing
services. Despite these advancements, the widely utilized keyword search functionality over encrypted data has not been integrated. To address this gap, our approach involves implementing searchable encryption techniques within a decentralized framework. However, this initial strategy falls short in ensuring service integrity, as decentralized storage is prone to threats from both clients and service peers. Clients might intentionally defame their service peers to avoid making payments, and in turn, service peers might provide inaccurate or incomplete findings. To address these risks, we employ smart contracts to secure
encrypted search logs, also referred to as proof, on the blockchain. Additionally, we establish a fair system for managing claims and facilitating equitable payouts. Employing as an example, a dynamic and efficient searchable encryption scheme, we devise a specific strategy that maintains encrypted search functionality while promoting a healthy ecosystem. This strategy encourages service peers to exert genuine efforts, fostering collective responsibility for service reliability. Our implementation is developed in both Python and Solidity, with subsequent testing of search performance and transaction costs conducted on the Ethereum blockchain [10].

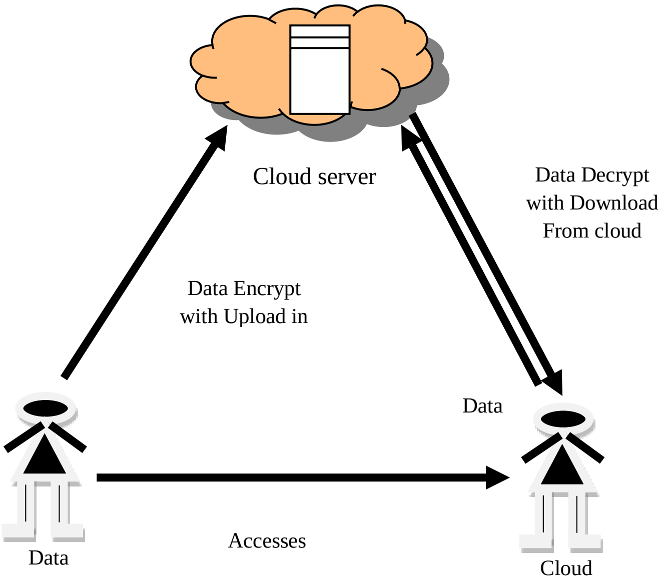
III PROBLEM DEFINITION

On a continuous basis, confidential data is transmitted and stored across the internet, necessitating ongoing assurances regarding privacy and security. However, such assurances are not consistently upheld. The unwarranted collection, sale, or exposure of private information frequently infringes upon the legal right to privacy held by data owners. This article delves into various privacy risks, concepts, legal frameworks, and types of personal information. It introduces Privacy-Enhancing Technologies (PETs), offering a survey that examines the features and capabilities of different models, metrics, anonymization techniques, and privacy tools. The paper also addresses current research challenges aimed at achieving heightened privacy levels in cloud computing and evaluates the relevance of the explored privacy mechanisms in today's cloud services.

IV PROPOSED MODEL

Creating a secure system for leveraging cloud data is crucial, particularly in addressing the intricate challenge of EARM. This entails defining and solving the complexities of the problem while enforcing a set of stringent privacy constraints. Among The different multi-keyword semantics,we choose to utilize the better concept of "coordinate matching." We elucidate a series of privacy principles for a secure system of utilizing cloud data, with a specific focus on the issue of secured multi keyword search (MKS).The selection of the coordinate matching rule is based on its effectiveness, aiming to maximize matches among the numerous multi-keyword semantics and determine the degree pertaining to the similarity between the search query and the data. To enhance the matching process, we introduce inner data correspondence to quantitatively formalize the measurement of this similarity

Fig.1.System Architecture



V RESULTS AND DISCUSSION

Encryption

In cryptography, an encryption scheme is employed wherein a document is encrypted together along a corresponding label. The secret code employed for encrypting documents can be selectively perforation, referring to designated set of label. The selectively punctured secret key retains the ability to decrypt ciphertexts, excluding those associated with tags that have been perforation on the secret code. The SE utilized in the SE is based on a perforation scheme. SE =(SE. Enc, SE. Dec) and a hash function H:KxN$^+$ K. SE is comprised of three algorithms (KeyGen, Enc, Dec), along with a message space M and a label space T.

$$(\mathrm{PKr}, \mathrm{SKr}) \leftarrow rkeygen(\lambda)$$

$$\mathrm{b} \in \mathcal{R}\{0,1\}, \mathrm{C}_{\mathrm{Wb}}=\mathrm{Encrypt}(\mathrm{W}_{\mathrm{b},}\mathrm{PK}_{\mathrm{r}})$$

Data Security

This section conducts a comprehensive security analysis of VBS. The information disclosed in our scheme is categorized into two primary aspects: the alter profile and the search profile. Concerning updates, the cloud server can solely access addition pairs and deletion pairs, lacking knowledge of Their particular content, except for the update operations.. Consequently, the update in VBS can be characterized as:

$$\int_{F}^{Udt} (op, \omega, ind) = (op)$$

VBS faces more intricate leakages in boolean searches compared to single keywords. Instead of examining the entire query, the approach identifies the smallest frequent each conjunctive query contains a keyword, as the boolean query is partitioned into multiple conjunctive queries.

As a result, an overloaded notation is introduced to articulate the leakage function for the boolean query q. Regarding the explore pattern, it can be described as:

$$EP(q) = \int_{i-1}^{k} (\omega i$$

Performance of Experiments

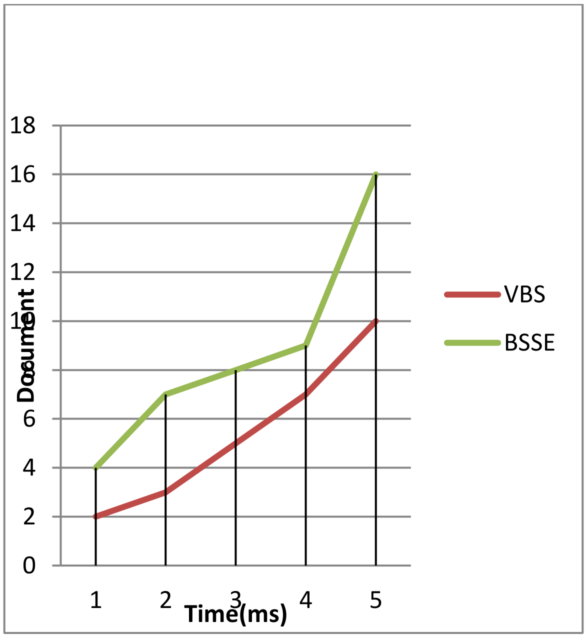


Fig.2. Data Search Graph

As depicted in Figures 2, the index setup time demonstrates for both schemes fluctuations based on the quantity of documents and keywords. The time measurement excludes the transformation of a keyword. It is observed that with an increase in the number of documents or keywords, the setup time for these methods also proportionally rises. Notably, for Verifiable Boolean Search over encrypted data (VBS), which only requires generating a label for each document, the setup phase consumes the least time. In contrast, Boolean Symmetric Searchable Encryption (BSSE) necessitates additional operations to support result verifiability.

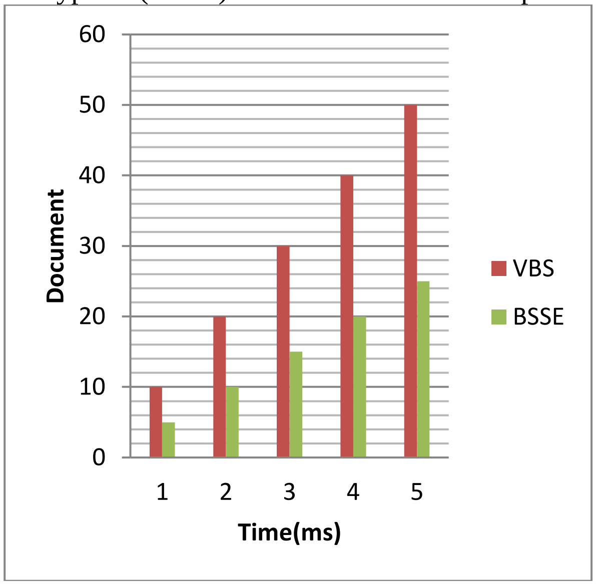


Fig.3. Data Storage

In Figure 3, The computational and storage burdens of both BSSE scheme and the basic VBS increase in correlation with the number of search results. Notably, the BSSE method exhibits lower efficiency in the verification process compared to our basic VBS. This is attributed to the fact that the BSSE scheme necessitates decrypting these search results before initiating the verification of search results.

VI CONCLUSION AND FUTURE WORK

This paper is predominantly centered on developing a practical, verifiable boolean search with the dual objectives of enhancing search efficiency and supporting result verification. The initial focus on VBS was extended to incorporate both forward and backward privacy. Formal security proofs were established for both schemes, and a comprehensive performance evaluation was conducted. However, an inherent drawback of the extended scheme are evident. is the increased computational costs during index construction and updates, stemming from the use of SE for encrypting document identity. In dealing with efficiency trade-offs, we have not specifically addressed the information leak of VBS in both search and access pattern. As a result, future efforts are aimed at improving the efficiency of SE or devising a scheme that is both new and more efficient capable of achieving complete backward privacy.

REFERENCE

[1] Y. Guo, M. Wang, C. Wang, X. Yuan, and X. Jia, “Privacy-preserving packet header checking over in-the-cloud middleboxes,” IEEE Internet of Things J., vol. 7, no. 6, pp. 5359–5370, Jun. 2020.

[2] Y. Guo, H. Xie, Y. Miao, C. Wang, and X. Jia, “FedCrowd: A federated and privacy-preserving crowdsourcing platform on blockchain,” IEEE Trans. Services Comput., to be published, doi: 10.1109/TSC.2020.3031061.

[3] Y. Miao, R. H. Deng, K.-K. R. Choo, X. Liu, J. Ning, and H. Li, “Optimized verifiable fine-grained keyword search in dynamic multi-owner settings,” IEEE Trans. Dependable Secure Comput., vol. 18, no. 4, pp. 1804–1820, Jul./Aug. 2021.

[4] Q. Tang and X. Chen, “Towards asymmetric searchable encryption with message recovery and flexible search authorization,” in Proc. Symp. Inf. Comput. Commun. Secur., 2013, pp. 253–264.

[5] S. Kamara, C. Papamanthou, and T. Roeder, “Dynamic searchable symmetric encryption,” in Proc. Conf. Comput. Commun. Secur., 2012, pp. 965–976.

[6] Y. Miao, R. H. Deng, X. Liu, K.-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.

[7] D. X. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Proc. IEEE Symp. Secur. Privacy, 2000, pp. 44–55.

[8] P. Gupta and N. McKeown, “Packet classification on multiple fields,” ACM SIGCOMM CCR, vol. 29, no. 4, pp. 147–160, 1999.

[9] M. Naveed, S. Kamara, and C. V. Wright, “Inference attacks on property-preserving encrypted databases,” in Proc. of ACM CCS, 2015.

[10] C. Cai, J. Weng, X. Yuan, C. Wang, “Enabling reliable keywordsearch in encrypted decentralized storage with fairness,” IEEE TDSC, 2018.

[11] Y. Miao, J. Ma, X. Liu, X. Li, Z. Liu, and H. Li, “Practical attributebased multi-keyword search scheme in mobile crowdsourcing,” IEEE Internet of Things Journal, vol. 5, no. 4, pp. 3008–3018, 2018.

[12] W. Sun, S. Yu, W. Lou, Y. T. Hou, and H. Li, “Protecting your right: verifiable attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud,” IEEE Transactions on Parallel and Distributed Systems, vol. 27, no. 4, pp. 1187– 1198, 2016.

[13] Y. Miao, J. Ma, X. Liu, X. Li, Q. Jiang, and J. Zhang, “Attributebased keyword search over hierarchical data in cloud computing,” IEEE Transactions on Services Computing, vol. PP, pp. 1–14, 2017.

[14] Y. Miao, J. Ma, X. Liu, J. Weng, H. Li, and H. Li, “Lightweight fine-grained search over encrypted data in fog computing,” IEEE Transactions on Services Computing, vol. PP, pp. 1–14, 2018.

[15] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Proc. International conference on the theory and applications of cryptographic techniques (EUROCRYPT’04). Springer, 2004, pp. 506–522.

[16] Li, F., Ma, J., Miao, Y., Zhiquan, L., Choo, K. K. R., Liu, X., & Deng, R. (2021). Towards efficient verifiable boolean search over encrypted cloud data. IEEE Transactions on Cloud Computing.

[17] Y. Kaneita et al., "Insomnia Among Japanese Adolescents: A Nationwide Representative Survey", Sleep, vol. 29, no. 12, pp. 1543-1550, 2006.

[18] P. Singh, "Insomnia: A sleep disorder: Its causes, symptoms and treatments", International Journal of Medical and Health Research, vol. 2, no. 10, pp. 37-41, 2016.

[19] Sarah Graham, Colin Depp, Ellen E Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim, and Dilip V Jeste. Artificial intelligence for mental health and mental illnesses: an overview. Current psychiatry reports, 21(11):1–18, 2019.

[20] Taboada M, Brooke J, Tofiloski M, Voll K, Stede M. Lexicon-based methods for sentiment analysis. Computational
linguistics, 2011; 37(2): 267–307.

[21] Strapparava C, Valitutti A. Wordnet-affect: an affective extension of wordnet. In Proceedings of the 4th International
Conference on Language Resources and Evaluation, 2004; 1083–1086.

[22] Patra BG, Mazumdar S, Das D, Rosso P, Bandyopadhyay S. A Multilevel Approach to Sentiment Analysis ofFigurative Language in Twitter. In Proccdings of the Conference on Intelligent Text Processing and ComputationalLinguistics, 2016 (pp. 281–291).

[23] Rosenthal S, McKeown K. Age prediction in blogs: A study of style, content, and online behavior in
pre-and post-social media generations. In: Proceedings of the 49th Annual Meeting of the Association

for Computational Linguistics: Human Language Technologies-Volume 1. Association for Computational Linguistics; 2011. p. 763–772.