International Journal on Applications in Information and Communication Engineering
Volume 9 : Issue 2 : April 2023, pp 26 – 31 www.aetsjournal.com ISSN (Online) : 2394 - 6237

ENABLING VERIFIABLE AND DYNAMIC RANKED SEARCH OVER OUTSOURCED DATA

S. SHANMUGAPRIYA , E. NANDHINI

***Abstract*—** The project **"ENABLING VERIFIABLE AND DYNAMIC RANKED SEARCH OVER OUTSOURCED DATA"Is** developed tooCurrently,Cloud computing as a promising computing paradigm is increasingly utilized as potential hosts for users' massive dataset. Since the cloud service provider (CSP) is outside the users' trusted domain, existing research suggests encrypting sensitive data before outsourcing and adopting Searchable Symmetric Encryption (SSE) to facilitate keyword-based searches over the ciphertexts. However, it remains a challenging task to design an effective SSE scheme that simultaneously supports sublinear search time, efficient update and verification, and on-demand information retrieval. To address this, we propose a Verifiable Dynamic Encryption with Ranked Search (VDERS) scheme that allows a user to perform top-K searches on a dynamic document collection and verify the correctness of the search results in a secure and efficient way. Specifically, we first provide a basic construction, VDERS0, where aranked inverted index and a verifiable matrix are constructed to enable verifiable document insertion in top-K searches. Then, an advanced construction, VDERS⋆, is devised to further support document deletion with a reduced communication cost. Extensive experiments on real datasets demonstrate the efficiency and effectiveness of our VDERS scheme.

I. INTRODUCTION

As a promising computing paradigm, cloud computing has drawn great attention from both research and industry communities. Because of the benefits of low costs, flexibility, and scalability, it has become a prevalent trend for users to outsource their massive datasets to clouds and delegate a cloud service provider (CSP) to manage data storage and offer query services. Due to security and privacy concerns, existing research suggests encrypting data before outsourcing. However, data encryption makes keyword-based searches over ciphertexts a challenging problem. This is even harder for efficient top-K searches in a dynamic and malicious cloud environment.

Let us consider the following scenario. Alice outsources archived emails to the cloud, where each email is indexed by the sender's name and ranked in descending order of the receipt date. For example, for a set of emails indexed by keyword Bob, the email received on April 2 has a higher rank than the email received on April 1. To keep keyword and document contents secret, Alice uploads them in the encrypted forms to the cloud. There could be hundreds of documents matching a specific keyword, and the consumed costs will be extensive if all the matched documents are returned to and decrypted by the user. Therefore, Alice may want to perform a top-K search to retrieve the most recent emails. Moreover, Alice may want to store only the emails received in the last three months for monetary saving. For example, when entering May, Alice will delete all emails received before February.

In the above application scenario, the adopted encryption scheme should meet the following requirements: (1) Ranked search. The user is allowed to perform a top-K search to retrieve the best-matched documents. (2) Dynamic. The user is able to update (add and delete) documents stored in the cloud. (3) Verifiability. The malicious CSP may delete encrypted documents not commonly used to save memory space, or it may forge the search results to deceive the user. Even if the CSP is honest, a virus or worm may tamper with encrypted documents. Therefore, the user should have the ability to verify the correctness of the search results. (4) Efficiency. The user can efficiently perform

S. Shanmugapriya , Assistant Professor, Department of Computer Applications,Erode Sengunthar Engineering College (Autonomous), Perundurai, Erode. (Email : riyashanmu@gmail.com)
E. Nandhini , PG Scholar , Department of Computer Applications,Erode Sengunthar Engineering College (Autonomous), Perundurai, Erode. (Email : nilaworkswith@gmail.com)

searches, updates, and verifications on a set of encrypted documents.

Although Searchable Symmetric Encryption (SSE) allows a user to retrieve desired documents in a privacy-preserving way, existing SSE schemes only partially address the above requirements. To simultaneously satisfy all these properties, this paper proposes a Verifiable Dynamic Encryption with Ranked Search (VDERS) scheme that allows the user to perform updates and top-K searches on cipher texts in a verifiable and efficient way. Our main idea is to construct a verifiable matrix to record the ranking information and encode it with RSA accumulator [3]. Furthermore, a ranked inverted index is built from a collection of documents to facilitate efficient top-K searches and updates. Specifically, we first provide a basic construction, denoted by VDERS0, which enables verifiable document insertion operations. Then, we provide an advanced construction, denoted by VDERS⋆, which not only can support efficient deletion operations, but also can reduce communication costs without outsourcing the verifiable matrix. Our main contributions

are summarized as follows:

- We propose a VDERS scheme to achieve dynamic and ranked searches in a cloud environment in an efficient and verifiable way.
- Two constructions are provided to achieve efficient top-K searches with support for verifiable updates.
- We theoretically analyze the security and performance of our scheme and conduct extensive experiments on real datasets to validate its effectiveness.

II. PAPER ORGANIZATION.

We introduce related work in Section 2 and provide the preliminaries in Section 3. After the overview of this work in Section 4, we provide our basic and advanced VDERS constructions in Section 5 and Section 6, respectively. We evaluate the proposed scheme in Section 7. Finally, we conclude the paper in Section 8.

1) OBJECT DETECTION- AN OVERVIEW

2.2 Supporting Data Dynamics Various auditing schemes have been proposed to support data dynamics. However, to the best of our knowledge, none of these have been able to achieve data dynamics that are as rapid as in our proposed scheme. In this subsection, we discuss public auditing with data dynamics which selectively preserves privacy against the TPA and/or the cloud.

We believe that there are no technical difficulties associated with achieving data privacy in public auditing. It can be attained by encrypting each block and generating the auditing metadata corresponding to the ciphertext. Nevertheless, we classify the existing auditing schemes in terms of privacy for clarity. We start by describing public auditing that does not consider privacy. 2.2.1 Auditing Without Privacy Elway et al. [9] designed the Dynamic Provable Data Possession (DPDP) scheme, a dynamic version of PDP, by supporting the updating of stored data. It uses a rank-based authenticated skip list to authenticate the tag information of challenged or updated data blocks before the verification procedure. Wang et al. [10] proposed a public auditing scheme that combines an HLA with a Merkle Hash Tree (MHT) to support dynamic data operations. However, in this scheme, the MHT needs to be re-constructed once the data has been updated. Zhu et al. [11] proposed a dynamic auditing scheme for cloud data based on a fragment structure, random sampling, and index hash tables. Their scheme is similar to ours in the sense that it does not involve the coupling of the auditing metadata and the index. They view a file as a group of sectors, where a set of sectors form a block on which auditing metadata is generated. Unfortunately, if any data dynamics for a sector occurs, every subsequent sector is affected. This can lead to the replacement of previous auditing metadata with new metadata, thereby resulting in inefficient data dynamics. Another line of researches that aimed at reducing the cost relating to data dynamics are batch update [40], index switcher [19], and dynamic hash table [14], [20].

Specifically, a batch update of data dynamics is to reduce the amount of computation cost relating to

dynamic data updates by performing and verifying multiple update operations at once [40]. Although it may avoid repetitive computations relating to data dynamics, the update delay can be unacceptable in time-critical applications. The index switcher approach provides a way of switching the encoded indices between block indices and tag indices [19], while dynamic hash table-based approaches aim to decouple the indices from tags [14], [20]. However, all of these schemes incur a significant tradeoff: the TPA-side computation cost relating to verification increases linearly with the number of challenged blocks. Given c number of challenged blocks, the TPA computes c pairings [14] or c exponentiations in a cyclic group [19], [20]. By contrast, the proposed scheme only requires two pairings and two exponentiations regardless of the number of challenged blocks. 2.2.2 Auditing with Privacy Jules et al. showed how POR can provide public auditing while preserving data privacy against the TPA [7]. Specifically, a POR client can separate data encryption and verification privileges, and delegate the verification capability to the TPA for public verification. Wang et al. [12] proposed a privacy-preserving public auditing scheme using random masking to prevent data content from being disclosed to the TPA. Similarly, Liu et al. [13] proposed a secure and efficient public auditing scheme using a homomorphic hash function and random masking. Shah et al.'s auditing scheme uses standard encryption (e.g., AES), but their scheme requires decryption in the cloud before any further computations can be performed [15]. Ramaiah et al. [17] proposed a privacy-preserving public auditing scheme, which encrypts each data block using somewhat homomorphic encryption. Thus, both the cloud and the TPA are not able to learn the content of a user's data while enabling integrity check of cloud data. However, this scheme only offers limited data dynamics, such as block modification and appending.

III. LITERATURE SURVEY

1) ABSTRACT

The Internet of Things (IoT) forms a foundation for cyber-physical systems. We propose an efficient and secure authentication scheme for machine-to-machine (M2M) networks in IoT enabled cyberphysical systems. Smart objects and smart devices over CPS are capable of capturing a variety of multimedia contents; interact with each other and also with the physical world in a fully automatic manner without human interference. The proposed scheme allows any pair of entities in an M2M network to mutually authenticate each other and agree on a session key for communicating data in a secure and efficient way. The authentication process does not incorporate the M2M service provider, and hence eliminates the burden of managing the authentication of massive scale devices at the edge of the network. The burden of the authentication process is offloaded and distributed on the gateways under the authority of this M2M service provider.

The proposed scheme requires the mobile user to hold only one secret key provided by the M2M service provider, by which, he can roam randomly in the M2M network and authenticate to any of the gateways in the domain. Then, this authenticated gateway allows the mobile user to authenticate with any sensor node in the domain. In the proposed scheme, the authentication process does not rely on any public key cryptographic operations.

Authentication is achieved using very few hash invocations and symmetric key encryptions. Therefore, the scheme is suitable for environmental sensors which are limited in resources (computation, storage, and energy). We analyze the security of the proposed scheme using BAN logic, which is widely accepted as a framework for the assessment of authentication protocols and also using ProVerif. We assess the efficiency of the proposed scheme and compare with some recently proposed schemes.

2)SYSTEM ANALYSIS

Atenas et al. [5] proposed Provable Data Possession (PDP), in which a public verifier can check the correctness of a user's stored data in the cloud. PDP utilizes an RSA-based Homomorphic Linear Authenticator (HLA) for outsourced data. Because an HLA can be aggregated, it is possible to

compute an aggregated HLA that authenticates a linear combination of individual data blocks. Using sampling strategies, the public verifier is able to audit the integrity of the outsourced data without retrieving the entire data set. However, this scheme does not consider dynamic operations for outsourced data. To support dynamic operations, Atenas et al. [6] designed an improved PDP scheme using symmetric keys and a cryptographic hash function. However, this scheme only supports a limited number of verification challenge queries. In addition, it does not support block insertion, though append-type insertion is possible. Jules and Kali ski [7] defined another scheme called Proofs of Retrievability (POR). The POR scheme incorporates special blocks called sentinels, which are randomly embedded into the data for detection purposes. However, they restrict operations on the updated data.

Wang et al. [10] proposed a public auditing scheme that combines an HLA with a Merkle Hash Tree (MHT) to support dynamic data operations. However, in this scheme, the MHT needs to be re-constructed once the data has been updated. Zhu et al. [11] proposed a dynamic auditing scheme for cloud data based on a fragment

structure, random sampling, and index hash tables. Their scheme is similar to ours in the sense that it does not involve the coupling of the auditing metadata and the

index.

Similarly, Liu et al. [13] proposed a secure and efficient public auditing scheme using a homomorphic hash function and random masking. Shah et al.'s auditing scheme uses standard encryption (e.g., AES), but their scheme requires decryption in the cloud before any further computations can be performed [15]. Ramaiah et al. [17] proposed a privacy-preserving public auditing scheme, which encrypts each data block using somewhat homomorphic encryption. Thus, both the cloud and the TPA are not able to learn the content of a user's data while enabling integrity check of cloud data. However, this scheme only offers limited data dynamics, such as block modification and appending.

Elway et al. improved a dynamic PDP scheme in which data dynamics is supported, but its communication and computation complexities are O (log n), where n denotes the number of blocks in a file [39]. Moreover, the data owner should engage in verifying the integrity of the outsourced data.

DISADVANTAGES

- An existing methodology doesn't implement Homomorphic Hash Function method.
- The system not implemented Supporting Data Dynamics concept.

IV. PROPOSED SYSTEM

Our novel auditing challenge-response protocol reduces the computation cost to the TPA significantly. Specifically, the TPA-side computation cost with respect to verification is a constant number of pairings and exponentiations in a cyclic group, while prior works require those operations linearly with the number of challenged blocks. The proposed scheme facilitates pre-computation capabilities such that, after sending an auditing request to the cloud, the TPA can pre-compute all exponentiation operations needed for the subsequent phase. Note that these computations can be performed in parallel with the cloud's efforts to compute the auditing response. According to our experiment, the TPA can notify users of the auditing results within 4 milliseconds, while previous schemes require 1.0 to 1.6 seconds [12], [13], [40].

Lastly, the proposed scheme is compatible with any symmetric-key encryption algorithm such that the blocks are encrypted any encryption algorithm of the data owner's choice. Data confidentiality is preserved against the cloud due to the CPA-secure property of the underlying encryption algorithm. We prove that the cloud cannot learn the outsourced data during the auditing process, and the cloud cannot forge a valid proof in response to the auditing request from the TPA.

ADVANTAGES

We propose a novel public auditing scheme for encrypted data that supports extremely fast data dynamics. Asymptotic analysis demonstrates that the proposed scheme has O (1) complexity, while

that of previous schemes ranges from O (log n) to O(n).

Our novel auditing challenge-response protocol reduces the computation cost of the TPA significantly, thus increasing the verification speed for the auditing results. We analytically detail the performance of our constructions over prior work and experimentally confirm that the proposed data dynamics method is orders of magnitude faster than that of previous schemes.

We formally define the security model under which we rigorously prove security of the proposed scheme to show that data integrity and privacy is preserved in the presence of an untrusted cloud.

V. SYSTEM IMPLEMENTATION

1) DATA OWNERS

In this module, the data owner performs operations such as Upload Blocks, Verify Block (Data Auditing), Update Block, Delete File, View Uploaded Blocks.

2) USER

In this module, he logs in by using his/her user's name and password. After Login receiver will perform operations like View All Data Owner Files, Request File, View File Response, and Download File.

3) THIRD PARTY AUDITOR

In this module, the sector can do following operations View Hash Table, View Attackers, View File Updated or Deleted, View Results.

4) CLOUD SERVICE PROVIDER

The Service Provider manages a server to provide data storage service and can also do the following operations such as View Data Owners, View End Users, View Hash Table, View File Request, View Transactions, View Attackers, View Results, View File Time Delay Results, View File Throughput Results.

VI. CONCLUSION

In this paper, we propose a public auditing scheme for encrypted data that supports extremely fast data dynamics. The proposed scheme supports data dynamics at a constant cost irrespective of the number of blocks. Our auditing challenge-response protocol requires a constant number of pairings and exponentiations, which significantly increases the verification speed for the auditing results. The proposed scheme ensures data confidentiality and integrity against the cloud server. During the auditing process, the TPA can verify the correctness of the proof without decrypting it and without key exposure, due to the homomorphic hash function. Security and performance analysis shows that the proposed scheme requires minimal extra computation while guaranteeing data privacy and integrity.

VII. REFERENCES

[1] Armbruster, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Kaminski , A., ... &Zaharie, M. (2010). A view of cloud computing . Communications of the ACM, 53(4), 50-58.

[2] Ren, K., Wang, C., & Wang, Q. (2012). Security challenges for the public cloud. IEEE Internet Computing, 16(1), 69-73.

[3] Song, D., Shi, E., Fischer, I., & Shankar, U. (2012). Cloud data protection for the masses. Computer, 45(1), 39-45.

[4] Wei, L., Zhu, H., Cao, Z., Dong, X., Jia, W., Chen, Y., &Vasilios, V. (2014). Security and privacy for storage and computation in cloud computing. Information sciences, 258, 371-386.

[5] Atenas, G., Burns, R., Carmela, R., Herring, J., Kisser, L., Peterson , Z., & Song, D. (2007, October). Provable data possession at untrusted stores. In Proceedings of the 14th ACM conference on Computer and communications security (pp. 598-609).

[6] Atenas, G., Di Pietro, R., Mancini, L. V., &Tzadik, G. (2008, September). Scalable and efficient provable data possession. In Proceedings of the 4th international conference on Security and privacy in communication networks (pp. 1-10).

[7] Jules, A., &Kali ski Jr, B. S. (2007, October). PORs: Proofs of retrievability for large files. In Proceedings of the 14th ACM conference on Computer and communications security (pp. 584-597).

[8] Shechem, H., & Waters, B. (2008, December). Compact proofs of retrievability. In International conference on the theory and application of cryptology and information security (pp. 90-107). Springer, Berlin, Heidelberg.

[9] Elway, C. C., Kupec¸ ¨ u, A., Hapaxanthous, C., &Tamasin, R. (2015). Dynamic provable data possession. ACM Transactions on Information and System Security (TISSEC), 17(4), 1-29.

[10] Wang, Q., Wang, C., Ren, K., Lou, W., & Li, J. (2010). Enabling public auditability and data dynamics for storage security in cloud computing. IEEE transactions on parallel and distributed systems, 22(5), 847-859.

[11] Zhu, Y., Wang, H., Hu, Z., Ahn, G. J., Hu, H., &Yua, S. S. (2011, March). Dynamic audit services for integrity verification of outsourced storages in clouds. In Proceedings of the 2011 ACM Symposium on Applied Computing (pp. 1550-1557).

[12] Wang, C., Chow, S. S., Wang, Q., Ren, K., & Lou, W. (2011). Privacy-preserving public auditing for secure cloud storage. IEEE transactions on computers, 62(2), 362-375.

[13] Liu, H., Zhang, P., & Liu, J. (2013). Public data integrity verification for secure cloud storage. Journal of networks, 8(2), 373.

[14] Shen, J., Shen, J., Chen, X., Huang, X., & Susilo, W. (2017). An efficient public auditing protocol with novel dynamic structure for cloud data. IEEE Transactions on Information Forensics and Security, 12(10), 2402-2415.

[15] Shah, M. A., Swaminathan, R., & Baker, M. (2008). Privacy-Preserving Audit and Extraction of Digital Contents. IACR Cryptal. print Arch., 2008, 186.

[16] Liu, J., Huang, K., Rong, H., Wang, H., & Xian, M. (2015). Privacy preserving public auditing for regenerating-code-based cloud storage. IEEE transactions on information forensics and security, 10(7), 1513-1528.

[17] Ramaiah, Y. G., & Kumari, G. V. (2013, July). Complete privacy preserving auditing for data integrity in cloud computing. In 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (pp. 1559-1566). IEEE.

[18] Wang, J., Chen, X., Huang, X., You, I., & Xiang, Y. (2015). Verifiable auditing for outsourced database in cloud computing. IEEE transactions on computers, 64(11), 3293-3303.

[19] Jinn, H., Jiang, H., & Zhou, K. (2016). Dynamic and public auditing with fair arbitration for cloud data. IEEE Transactions on cloud computing, 6(3), 680-693.

[20] Tian, H., Chen, Y., Chang, C. C., Jiang, H., Huang, Y., Chen, Y., & Liu, J. (2015). Dynamic-hash-table based public auditing for secure cloud storage. IEEE Transactions on Services Computing, 10(5), 701- 714.

[21] Naune, E. (2010). What Twitter learns from all those tweets. Technology Review, 28.

[22] Crohn, M. N., Freedman, M. J., &Manières, D. (2004, May). On the-fly verification of rate less erasure codes for efficient content distribution. In IEEE Symposium on Security and Privacy, 2004. Proceedings. 2004 (pp. 226-240). IEEE.

[23] Gennaro, R., Katz, J., Krawczyk, H., & Rabin, T. (2010, May). Secure network coding over the integers. In International Workshop on Public Key Cryptography (pp. 142-160). Springer, Berlin, Heidelberg.

[24] Bone, D., Lynn, B., &Shechem, H. (2001, December). Short signatures from the Weil pairing. In International conference on the theory and application of cryptology and information security (pp. 514-532). Springer, Berlin, Heidelberg.

[25] Xia, H., Lu, T., Shao, B., Ding, X., & Gu, N. (2014, May). Hermes: On collaboration across heterogeneous collaborative editing services in the cloud. In Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD) (pp. 655–660). IEEE.