ELSA: Edge Lightweight Searchable Attribute-based encryption Multi-keyword Scalability

Jawhara Aljabri
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
j.aljabri.1@research.gla.ac.uk

Anna Lito Michala
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
annalito.michala@glasgow.ac.uk

Jeremy Singer
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
jeremy.singer@glasgow.ac.uk

***Abstract*—The digitalisation of industrial manufacturing needs the support of systems technology to enhance the efficiency of manufacturing operations, product quality, and smart decisions. This digitalisation can be achieved by the industrial internet of things (IIoT). IIoT has played a powerful role in smart manufacturing by performing real-time analysis for a large volume of data. One possible approach to perform these operations in a secure and privacy-preserving manner is to utilise cryptographic solutions. In previous work, we proposed searchable encryption with an access control algorithm for IIoT based on an edge-cloud architecture, namely ELSA. This paper extends ELSA to illustrate the correlation between the number of keywords and ELSA performance. This extension supports annotating records with multiple keywords in trapdoor and record storage and allows the record to be returnable with single-keyword queries. In addition, the experiments demonstrate the scalability and efficiency of ELSA with an increasing number of keywords and complexity.**

***Index Terms*—Industrial Internet of Things, multi-keyword search, searchable encryption, search time.**

I. Introduction

Industry 4.0 is this century's revolution of the sector which started with the introduction of the Internet of Things (IoT), broadly refer to as *Industrial Internet of Things* (IIoT) [1], [2]. IIoT solutions are proposed for remote maintenance, quality control, product traceability, product life-cycle management and service optimization [3]. These aspects enforce a requirement for multi-actor access to the collected data such as insurers, customers, employees, and consultants. On the other hand, as with individuals, businesses also have privacy considerations often translating to competitive advantage or security. Thus, data could compromise the factory while controlled access to higher level information could be advantageous [4]. In this scenario data is often processed on the cloud. In this work we make the assumption that the cloud is not trusted [5].

Protecting data at rest has been a significant research domain in recent years utilising cryptographic primitives, access control (AC) policies [6], and searchable encryption (SE) [7]. The SE method however requires partial decryption on the cloud; it makes assumptions regarding the ordering of the data, and associates data with sets of keywords. To address these challenges, we recently published the Edge Lightweight Searchable Attribute-based encryption (ELSA) [8] (Fig. 1); a keyword-based searchable encryption multi-authority (MA) access control (AC) for IIoT devices assisted by a three-tier edge computing architecture.

978-1-6654-2141-6/22/$31.00 ©2022 IEEE

A. Contributions

In this paper we present an extension to ELSA to make the search queries more flexible allowing the retrieval of results that partially match the assigned keyword set provided by the user. Further, we present additional experiments investigating scalability. Our contributions are as follows:

- ELSA extension to support annotating records with multiple keywords in trapdoor and record storage and allowing the record to be returnable with single-keyword queries
- ELSA scalability experiments demonstrating search time remaining in the region of 10^2 ms for as many as 1000 keywords.

The paper is organised in four sections. Section II presents the state of the art in the domain of SE with AC for IoT applications. Section III describes the ELSA method briefly and discusses the methodology for the extension of the lookup tables compared to the state of the art. Section IV presents the scalability experiment methodology along with results. Finally, conclusions and future work are discussed Section V.

II. Related Work

SE allows the user to perform secure searches over encrypted data without compromising the data confidentiality. AC mechanisms are employed to dictate who has access to the data through access policies. A full review of both for the application domain of IIoT was previously presented in [9]. Since the early 2000's, SE [10] and asymmetrical SE [11] have been investigated as a method to allow retrieving only the required records of encrypted data. The method associates keywords to data records which can be retrieved through as single or multi-keyword search [12], [13].

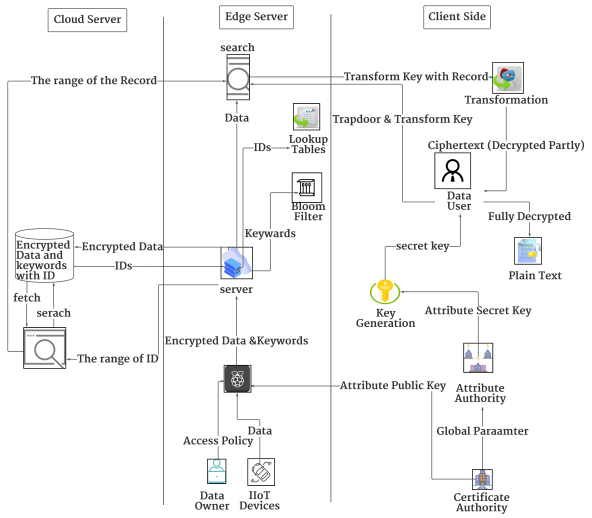

Fig. 1. Proposed ELSA architecture, data residence, and the domains of trust.

With the rising use of the cloud the Public Key Encryption with Keyword Search (PEKS) method has been combined with Attribute-Based Encryption (PEKS-ABE) [14]. However, protecting the privacy of user keys is an open challenge. The use of private cloud has also been investigated to improve data-privacy and performance in Ciphertext-Policy Attribute-Based Encryption (CP-ABE) [15]. The limitation of CP-ABE methods is the use of a central authority for attribute authorisation.

Extensions of keyword-based searchable encryption work [16], [17] have eliminated the central authority without considering the computational overhead. Both approaches depend on bilinear pairing operations which are expensive in terms of memory requirements, making them unsuitable for IIoT applications.

For IIoT the complexity of the challenge increases as real-time execution and processing capabilities impose new requirements. SE and AC algorithms were not originally designed for embedded processors. Thus, the community has focused on improving efficiency of SE or AC in isolation. Lightweight SE with security improvements has been proposed in [18]. Attribute-based multi-keyword search schemes with CP-ABE were investigated in [19] to improve the accuracy of returned records to support dynamic IoT applications.

Most recently a low computational complexity SE-AC scheme was proposed for applications in IoT [20]. Their fine-grained AC, multi-keyword search, lightweight decryption, and a multi authority environment schemes (LSABE-MA) can support single keyword and multi-keyword searching while maintaining lightweight decryption. It also improves privacy preventing leakage in transit. The method meets the low latency requirements of IoT and supports improved security against chosen-keyword and the chosen-plaintext attacks. However, LSABE-MA does not guarantee real-time interactions, nor does it investigate the impact on returned record accuracy or data bandwidth. Accuracy, bandwidth, and real-time guarantees are considered important criteria for industrial applications [8], [9], [21].

LSABE-MA is based on sequential search of all the encrypted data records. Depending on the location of the record the search time can introduce latency larger than the required real-time guarantees. Hence, ELSA [8] expanded LSABE-MA to address this deficiency in the searching method and associated time as well as the bandwidth utilised. ELSA improved performance by an order of magnitude. This was achieved through suggesting an improved organisation of the data and an edge-cloud architecture. An edge server was proposed to cluster data indices by keyword leading to better than linear search performance while maintaining accuracy over the results.

Specifically for accuracy in LSABE-MA, when a record is annotated with a set of keywords the record can only be returned if a user requests records with the exact matching set of keywords. For example, a data record stored in the database annotated with keyword(k) set: k_1, k_2, k_3, will not be retrieved by a search query that requests records matching k_1. In plain words, if a user needs to see all the encrypted images of 'cats' we could assume that they would want to see those that simultaneously have 'cats and dogs' as well. Unfortunately, LSABE-MA will not return any images with both 'cats and dogs'. Thus, the existing searching approach limits the extent of the returned records and demands users to know the exact set of keywords which might be unsustainable for scalability. This is a deficiency in terms of the LSABE-MA scheme. ELSA on the other hand uses lookup tables on the edge server that could be used to identify unique keywords within a set associated with a record. This extension is presented in the following section.

III. Methods

The cloud-edge architecture of the ELSA method is presented in Fig. 1. ELSA process queries over the encrypted data on the trusted edge to improve privacy. Partial decryption required for SE takes place on the trusted edge server. Also ELSA improves user key privacy protection by handling keys on the edge. The cloud server, not being trusted, is responsible for storage of the encrypted dataset (Fig. 1 left). The added benefit of the edge server is the potential reduction of load on cloud communication bandwidth, while reducing the latency for query results. Thus, ELSA reduces overall core traffic. The edge (Fig. 1 middle) is responsible for handling incoming requests from the user. Using a Bloom filter the edge can eliminate queries that would yield no result. Additionally, a query optimiser reduces the scope of the search for the cloud server. The edge is also responsible for establishing and enforcing the access policies directly linking to the IIoT devices and the data owner.

ELSA improves the search process by creating a lookup table in the edge to store the keywords with the unique number for each encrypted record before sending these records to the cloud. This lookup table process the search query (trapdoor) using a clustering algorithm to accelerate the search process.

The extended ELSA scheme uses this lookup table to support the multi-keywords scenario. It implements the multi keywords support so that system users can access the required data by issuing a query with any number of keywords. The key differences between multi-keyword architecture in the proposed system and single-keyword architecture are presented in Fig. 2. The single-keyword architecture allows data users to find data corresponding with the identical sequences of keywords encrypted with the required data record. ELSA supports multi-keyword association as separate entries in the lookup table for each keyword against the same unique identifier of the record. This association can be in trapdoor and record storage, with the record being returnable. Supporting multi-keyword search in the system will avoid returning a reduced subset of results to the system user. Specifically, the lookup table in ELSA stores each keyword with a unique record number. Therefore, the multi-keyword trapdoor will be processed efficiently. The output of this will be all the records containing any of the requested keywords.

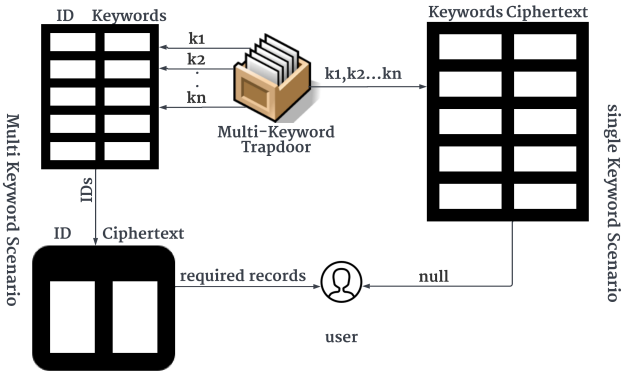

Fig. 2. Single VS Multi-keyword Scenario.

The following experiments consider to the extended ELSA scheme based on the multi-keyword scenario from two aspects:

- Multi-keywords in lookup table.
- Multi-keywords in trapdoor (query).

IV. Evaluation

This section outlines the experiments and results by discussing the influence of number of keywords in the lookup table and the trapdoor on the ELSA performance. Let N_{KL} be the number of keywords in the lookup table , N_{KT} be the number of keywords in the trapdoor.

We run the client application on an edge device with Intel 2.3 GHz Core i9 processor and 16GB RAM for evaluation setup. In addition, deployed the server code on a docker container hosted on a DigitalOcean cloud provider located in the UK. The plan for the cloud provider was CPU-Optimised, with one dedicated CPU, 2-32 vCPUs, 50 GB backing storage, 2GB RAM/CPU and 2TB Bandwidth.

For the experiment we used a synthetic dataset. The dataset consists of temperature, CO_2, and humidity values. For evaluation purposes, we considered the $\{CO_{2_{normal}}$, $CO_{2_{high}}$,$CO_{2_{low}}$, $humidity_{normal}$, $humidity_{high}$, $humidity_{low}$, $temperature_{normal}$, $temperature_{high}$, $temperature_{low}\}$ keywords. We used two different cases of data based on the following criteria:

- percentage of representation of one of keywords (in this experiment is $CO_{2_{high}}$),
- and number of keywords.

The two cases are (i) the sparse dataset case where the $CO_{2_{high}}$ represent 5% of keywords, and (ii) the dense case where the $CO_{2_{high}}$ represent 40%.

A. Multi-Keyword in Lookup Table

The specific trapdoor generated for this evaluation in each case was constant and contained one keyword, which is $CO_{2_{high}}$. The lookup table generated by frequency of N_{KL} starts from 100 to 1000 keywords.

In Fig. 3, we present the effect of the number of keywords in the lookup table N_{KL} on ELSA performance. The search time increases linearly as the N_{KL} increases. The extended ELSA scheme makes better use of the multi-keywords in the sparse case (5%) than dense case (40%). However, the computational cost of searching on the lookup table with 1000 keywords does not exceed 132 ms in the worst case, while the computational cost of the sparse case requires as little as 20 ms. As expected, the keyword representation ratio will directly affect the computational cost in this phase, but there is only a time gap of 132 to 20 ms between the above two cases, which could be acceptable but should be evaluated on a real-world dataset for validation.

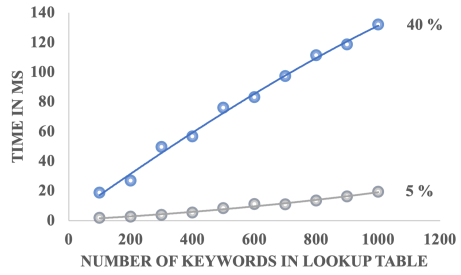

Fig. 3. Search Time linearly increasing with no. keywords.

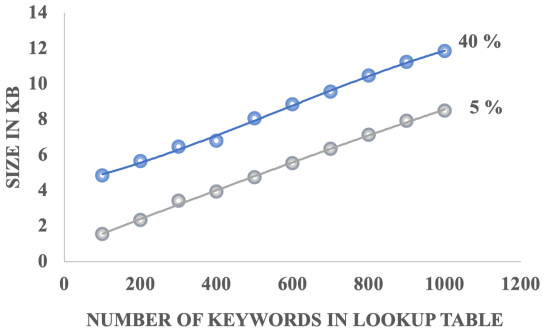

Fig. 4. Lookup Table Size linearly increasing with no. keywords.

Further to investigate scalability in terms of memory used, we measured the lookup table size for both the sparse case

(5%) and dense case (40%) with different values of N_{KL}. As presented in Fig. 4, both cases remain below 14KB even for the very unrealistic scenario of 1000 keywords. Obviously, the lookup table size has an approximately linear relationship with N_{KT} values. However, it does not grow to an unsupported size for the edge server.

B. Multi-Keyword in Trapdoor

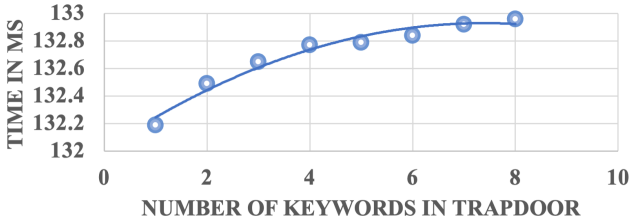

Fig. 5. Search Time reaching a steady state below 3s regardless of no. keywords.

In this experiment, N_{KT} take a value from one keyword to eight keywords, and the lookup table contains 1000 keywords (worst case in previous experiment). Fig. 5 presents the experimental result of search time under different values of N_{KT}. When the value of N_{KT} is 1 and 8, the computational cost of searching is 132.19 ms and 132.92 ms respectively. However, it is not linear and saturates below 133 ms, which is a promising result for scalability, and remains below 3 seconds which is marginally noticeable in terms of user experience.

V. Conclusion

This paper presents an extension to ELSA and the development of a scalable and efficiently searchable encryption scheme with access control for IIoT utilising an edge-cloud architecture. First, we extend the ELSA scheme, which uses lookup tables on the edge server to identify unique keywords within sets of keywords. This extension can provide better accuracy of returned records with partial query searches. The experimental results for scalability with varying numbers of keywords demonstrate acceptable trade-off for query performance. The results show that ELSA's performance remains within acceptable limits from the perspective of user experience and memory utilisation when increasing the number of keywords. However, the lookup table size increases linearly. Therefore, in future work we will aim to minimise the lookup table size and summarise the data records by integrating ELSA with ML methods. This integration will eliminate records of unnecessary data that do not add value to further processing. The result would minimize all of the lookup table size, the cloud storage and the bandwidth utilisation taking full advantage of the edge architecture benefits.

Acknowledgment

This work was funded under scholarship TBU331 provided by University of Tabuk in Saudi Arabia.

References

[1] M. Hermann, T. . Pentek, and B. Otto, "Design Principles for Industrie 4.0 Scenarios: A Literature Review," Tech. Rep. [Online]. Available: www.snom.mb.tu-dortmund.de

[2] P. Mathur, *IoT Machine Learning Applications in Telecom, Energy, and Agriculture*, 2020.

[3] Y. Yu, R. Chen, H. Li, Y. Li, and A. Tian, "Toward Data Security in Edge Intelligent IIoT," *IEEE Network*, vol. 33, no. 5, pp. 20–26, 2019.

[4] Web, Searching T H E and Devices, F O R At-risk, "IoT FOR BUSINESS Take Manufacturing ' s Shift Your Manufacturing Shift to Lightspeed to Lightspeed," 2020.

[5] J. Blömer and N. Löken, "Cloud architectures for searchable encryption," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*, 2018, pp. 1–10.

[6] K. Pothong, I. Brass, M. Carr, L. Tanczer, S. Security, R. Nicolescu, B. Craggs, E. Lupu, A. Rashid, C. Maple, S. Wakenshaw, M. Taddeo, J. Lindley, S. Cannizzaro, R. Procter, and P. Coulton, "Editors of the Cybersecurity of the Internet of Things: PETRAS Stream Report 03 Privacy and Trust 05 Adoption and Acceptability," 2019.

[7] G. S. Poh, J.-J. Chin, W.-C. Yau, K.-K. R. Choo, and M. S. Mohamad, "Searchable symmetric encryption: designs and challenges," *ACM Computing Surveys (CSUR)*, vol. 50, no. 3, pp. 1–37, 2017.

[8] J. Aljabri, A. L. Michala, and J. Singer, "ELSA: a keyword-based searchable encryption for cloud-edge assisted industrial internet of things," in *22nd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid 2022),Taormina (Messina), Italy, 16-19 May 2022*. IEEE/ACM, 2022.

[9] J. Bader and A. L. Michala, "Searchable encryption with access control in industrial internet of things (iiot)," *Wireless Communications and Mobile Computing*, vol. 2021, 2021.

[10] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proceeding 2000 IEEE symposium on security and privacy. S&P 2000*. IEEE, 2000, pp. 44–55.

[11] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *International conference on the theory and applications of cryptographic techniques*. Springer, 2004, pp. 506–522.

[12] R. Chen, Y. Mu, G. Yang, F. Guo, X. Huang, X. Wang, and Y. Wang, "Server-aided public key encryption with keyword search," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 12, pp. 2833–2842, 2016.

[13] H. Li, Y. Yang, T. H. Luan, X. Liang, L. Zhou, and X. S. Shen, "Enabling fine-grained multi-keyword search supporting classified sub-dictionaries over encrypted cloud data," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 3, pp. 312–325, 2015.

[14] Y. Zhou, N. Li, Y. Tian, D. An, and L. Wang, "Public key encryption with keyword search in cloud: A survey," *Entropy*, vol. 22, no. 4, pp. 1–24, 2020.

[15] S. Qi, Y. Lu, W. Wei, and X. Chen, "Efficient Data Access Control with Fine-Grained Data Protection in Cloud-Assisted IIoT," *IEEE Internet of Things Journal*, vol. 4662, no. c, pp. 1–1, 2020.

[16] 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 Transactions on Dependable and Secure Computing*, vol. 18, no. 4, pp. 1667–1680, 2019.

[17] Q. Huang, G. Yan, and Y. Yang, "Privacy-preserving traceable attribute-based keyword search in multi-authority medical cloud," *IEEE Transactions on Cloud Computing*, 2021.

[18] B. Chen, L. Wu, N. Kumar, K.-K. R. Choo, and D. He, "Lightweight Searchable Public-key Encryption with Forward Privacy over IIoT Outsourced Data," *IEEE Transactions on Emerging Topics in Computing*, vol. PP, no. c, pp. 1–1, 2019.

[19] Y. Miao, X. Liu, R. H. Deng, H. Wu, H. Li, J. Li, and D. Wu, "Hybrid keyword-field search with efficient key management for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3206–3217, 2019.

[20] K. Zhang, J. Long, X. Wang, H.-N. Dai, K. Liang, and M. Imran, "Lightweight Searchable Encryption Protocol for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 3203, no. c, pp. 1–1, 2020.

[21] A. L. Michala, I. Vourganas, and A. Coraddu, "Vibration edge computing in maritime iot," *ACM Transactions on Internet of Things*, vol. 3, no. 1, pp. 1–18, 2021.