IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

Privacy-Preserving Searchable Encryption with Access Control Using Secret Sharing for Secure Cloud Data Outsourcing

P. Dinesh[1], S. Dhineshkumar[2], B. Gowtham[3], T. Kalaiselvan[4]
Assistant Professor, Department of Computer Science and Engineering[1]
Students, Department of Computer Science and Engineering[2-4]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, Tamilnadu, India

Abstract: *Searchable encryption allows users to perform search operations on encrypted data without revealing its content. While many existing methods use public key and symmetric key encryption, public key approaches often involve high computational costs, making them less suitable for large databases in cloud environments. To overcome this issue, this paper proposes a privacy-preserving searchable encryption method based on secret sharing, known for its lower computational complexity. Secret sharing divides confidential data into multiple independent shares, improving both security and efficiency. Previous work introduced a searchable encryption method using secret sharing but did not support user access control. In this study, we present a searchable encryption framework with integrated user access control, using a (k, n) threshold secret sharing scheme. In the proposed system, each data item stored in the cloud has an owner, who can control access permissions for different users. A client-server model is used to perform secure computations between the data owner, authorized users, and $n \geq k$ cloud servers. We analyze the security of our system in terms of data distribution, query generation, and search processes, proving its resistance to honest-but-curious adversaries with knowledge of up to $k - 1$ servers. Additionally, we propose an improved method using an (n, n) additive secret sharing scheme for cases where $n = k$. The performance of the proposed methods is evaluated using Python, with comparisons made against existing secret sharing-based searchable encryption schemes. Results show that our approach offers better computational and communication efficiency while maintaining strong security guarantees.*

Keywords: Searchable Encryption, Secret Sharing, Access Control, Cloud Data Outsourcing, (k, n) Threshold Scheme, Additive Secret Sharing, Secure Computation, Honest-but-Curious Adversary, Privacy Preservation

I. INTRODUCTION

The rapid evolution of computer processing power and communication technologies has significantly enhanced the popularity of cloud services, including cloud storage. The recent shift to remote work and home-based employment has further solidified the importance of cloud storage for accessing and sharing data from virtually any location. Cloud storage is a cloud computing model that allows users to store data online via service providers, accessible through both public and private network connections [1], [2]. This model eliminates the need for individuals to purchase and manage their own storage systems, offering benefits such as scalability, flexibility, and reliability, while ensuring data accessibility at all times.

However, outsourcing sensitive data to remote cloud servers introduces significant challenges related to data privacy and security. Data breaches and unauthorized access are critical concerns, as they can lead to privacy violations and the exposure of sensitive information [3], [4], [5], [6]. While encryption methods, such as symmetric encryption, are commonly employed to protect data by transforming it into unreadable formats, they pose a challenge when it comes to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

searching through the encrypted content. Traditional keyword searches are not feasible on encrypted data, making it difficult to perform useful operations on encrypted storage.

To address this limitation, the concept of searchable encryption (or secure search) has emerged. This enables users to perform searches over encrypted data without revealing the content. The terms "searchable encryption," "secure search," and "searchable secret sharing" are used interchangeably in this paper to describe the ability to conduct searches over encrypted data, either in the form of cipher text (generated via key encryption methods) or shares (produced using threshold secret sharing schemes).

In addition, efficient access control systems are essential to managing who can access the data stored in the cloud. Conventional access control mechanisms rely on the attributes of the protected data and are typically governed by a trusted storage server. However, this assumption of trust becomes increasingly problematic in cloud environments, where the cloud provider itself may not be fully trustworthy. Therefore, there is a critical need for effective and secure access control mechanisms that can operate within the cloud, ensuring the integrity and confidentiality of sensitive data.

II. RELATED WORKS AND NOVELTY

Over recent years, several secure search methods have been introduced to enable secure access and search capabilities for outsourced data. These methods include secure data outsourcing [7]-[27], searchable encryption [28]-[49], and private information retrieval [50]-[52]. This study focuses on the use of searchable encryption, a secure computation approach for performing searches on encrypted databases. Searchable encryption allows users to confidentially outsource their data to public cloud servers while retaining the ability to perform selective searches [37].

Symmetric encryption is often employed to support searchable encryption. However, a persistent challenge is that the searcher must be either the initial data owner or have previously received the symmetric key. Additionally, ensuring strong privacy guarantees in searchable encryption typically requires additional secure computations over the encrypted data. This can be accomplished using encryption methods that allow functional evaluation of encrypted data, such as homomorphic encryption (HE) or secret sharing-based secure multi-party computation. However, the real-world application of HE in cloud computing is hindered by its high computational complexity.

In contrast, (k, n) threshold secret sharing offers a more efficient computational approach, as demonstrated by the results in Table 1 (extracted from [53]). As shown, secure computation using secret sharing is significantly faster than fully homomorphic encryption (FHE) [54]. However, it requires multiple computing servers to fully leverage the benefits of the secret sharing scheme, as opposed to relying on a single server in FHE. The cloud computing landscape has undergone significant transformation, shifting from single-provider data centers to multi-provider infrastructures that utilize decentralized computing resources from various providers [2].

This study adopts the (k, n) threshold secret sharing scheme, which is known for its low computational overhead, to enable secure outsourced databases with search capabilities. In a (k, n) threshold secret sharing scheme, a secret sss is transformed into nnn distinct shares, which are assigned to nnn servers. The secret can be reconstructed by collecting any kkk shares, while fewer than kkk shares reveal no information about the secret [55], [56]. Previous studies, such as those by Hadavi et al. [12], [16], [18], Ito et al. [40], and Kamal et al. [57], [58], [59], have successfully implemented secure outsourced databases using this approach.

The method proposed by Ito et al. [40], based on Boneh et al.'s framework [29], features index-based searchable encryption, which excludes public-key encryption. This method uses an index to store document collections and facilitates efficient keyword searches. However, it has limitations, particularly in securing user search patterns. Kamal et al. [58], [59], [60] introduced improvements to address some of these challenges but did not include mechanisms for user access control.

In summary, secure search methods using key encryption, such as homomorphic encryption, benefit from needing only a single cloud server but involve high computational costs for both encryption and search operations. In contrast, secret sharing-based methods offer lower computational costs than HE; however, many do not address search pattern security or provide efficient user access control. Therefore, this study aims to develop an improved and efficient secure search method that mitigates information leakage, including access pattern leakage, through the implementation of a secret sharing framework.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

The novelty of this study is summarized as follows:

- We propose an improved and efficient secure search method based on secret sharing that mitigates information leakage, such as access pattern leakage, by implementing additional random numbers via a one-time pad approach.
- We introduce a novel user access control mechanism that enables data owners to decide who can access their private data using secure computation based on a secret sharing scheme. This approach ensures minimal computational overhead and enhanced security compared to traditional password- based systems.

III. PROPOSED METHODOLOGY

The proposed methodology revolves around ensuring the security, privacy, and efficiency of keyword-based search operations over encrypted data in a cloud environment. The key components of the methodology include secure computation, threshold secret sharing, user access control, and randomized query generation to prevent data leakage and unauthorized access. Below is a detailed breakdown of each step in the methodology

3.1 Data Encryption:

The data owner encrypts the data (e.g., text documents) using symmetric encryption (e.g., AES).
The encryption key is split into n shares using a (k, n) threshold secret sharing scheme. This means that at least k shares are needed to reconstruct the encryption key.

3.2 Index and Share Distribution:

The encrypted data and search index (mapping keywords to documents) are split into shares and stored across multiple cloud servers. Each server holds only a part of the data or index.

3.3 User Access Control:

Only authorized users, as defined by the data owner, can generate search queries. Access control policies are applied to restrict unauthorized search attempts.
Users must authenticate securely before accessing the data.

3.4 Query Generation and Randomization:

When users want to search, they generate a query based on keywords.
The query is encrypted and randomized to ensure privacy, making it difficult for any server to link the query with the user.

3.5 Search Execution on Cloud Servers:

The encrypted query is sent to the cloud servers. Each server performs secure computations on its share of the encrypted data and returns partial results.
The computation ensures that servers cannot decrypt the query or documents.

3.6 Result Combination and Decryption:

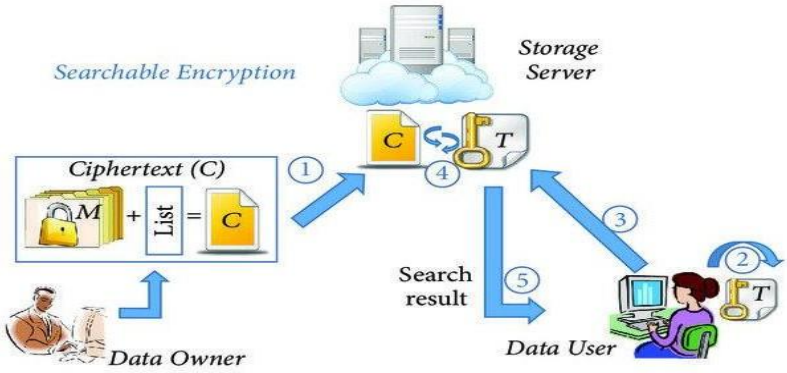
The user combines the partial results from the servers to generate the complete list of matching documents.
To decrypt the results, the user combines at least k shares of the encryption key to reconstruct the key and access the plaintext documents.

3.7 Security Features:

The system ensures that no server can access the entire data or key, preserving confidentiality. Query randomization prevents attackers from analyzing search patterns.
Access control ensures that only authorized users can perform searches.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

IV. SYSTEM ARCHITECTURE



1. Primary Entities

Data Owner: Responsible for uploading encrypted data and defining access policies.

Authorized Users: Users granted permission to perform search operations on the encrypted data.

Storage Servers: Hosts the encrypted data and facilitates search operations without revealing sensitive information.

2. Data Preparation and Keyword Extraction

The Enron email dataset [80] is utilized for evaluation. Keywords are extracted from each email by identifying the top three words with the highest Term Frequency-Inverse Document Frequency (TF-IDF)

3. Secret Sharing Scheme

A (k, n) threshold secret sharing scheme is employed to distribute the secret key among
n participants, such that any k participants can reconstruct the secret. This approach enhances security by ensuring that no single participant has access to the complete key.

4. Searchable Encryption with Secure Computation

The searchable encryption scheme allows authorized users to perform search operations on the encrypted data without revealing the search query or the data content. Secure computation techniques are applied to ensure that the search process does not leak any sensitive information.

5. User Access Control Mechanism

Access control policies are defined by the data owner to specify which users are authorized to perform search operations. These policies are enforced through secure computation protocols, ensuring that only authorized users can access the encrypted data.

V. PERFORMANCE EVALUATION

To validate the effectiveness and practicality of the proposed methods, we conducted a comprehensive performance evaluation using the Enron email dataset. This dataset is widely used in research involving information retrieval and email communication analysis due to its real-world structure and moderate size.

A. Evaluation Metrics

- We measured the performance of our proposed searchable encryption methods using the following key metrics: Computation Time: The time required to execute each phase of the protocol, including keyword extraction, query generation, secret sharing, and search operations.
- Communication Overhead: The total size of data exchanged between the data user and the cloud servers during query and result transmission.
- Scalability: The system's ability to handle increasing numbers of documents, users, and concurrent queries without significant degradation in performance.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

B. Experimental Setup

To simulate a secure multi-server environment, all operations were implemented in Python 3.10.12 and run on a single workstation configured to represent different system components:

CPU: Intel Core i7-14700K @ 3.40GHz
Memory: 128 GB RAM
Operating System: Ubuntu Desktop 22.04.5 LTS (64-bit)
Storage: 500 GB SSD

Although the test environment used a single physical machine, logical separation was maintained to simulate multiple users and servers. This allows for accurate comparative measurement of computation and communication costs between the proposed methods.

C. Dataset Preprocessing

- From the Enron email corpus:
- Each email was parsed and tokenized.
- The top three keywords were extracted per email using TF-IDF scoring.
- Each keyword was indexed and encrypted using our secure computation approach.
- This enabled realistic and consistent testing of keyword-based query operations.

VI. CONCLUSION

In this study, we introduced two secure and efficient searchable encryption methods that incorporate secure computation and a (k, n) threshold secret sharing scheme, along with user access control mechanisms. Our approach ensures that only authorized data users, as designated by the data owner, can generate valid search queries and execute search operations. Additionally, by leveraging secure computation, we proposed a query generation algorithm capable of randomizing search queries, thereby enhancing the privacy of user search patterns.

We demonstrated that our proposed methods are executable within a reasonable timeframe when implemented in Python, and our performance evaluations indicate improvements over conventional secret sharing-based approaches. This makes our solution suitable for deployment in multi-server environments, including cloud storage as a service.

VII. FUTURE WORK

In future work, we plan to implement a logical search mechanism for multi-keyword queries and incorporate verification functions to guard against malicious adversaries. We also aim to conduct a comprehensive survey of recent adversarial attacks, particularly in the field of machine learning [88], and propose a more robust security framework capable of defending against both honest-but-curious and malicious adversaries.

Moreover, we will explore the integration of the zero trust security models, which is becoming increasingly vital in modern, mobile, and cloud-based computing environments [89]. Our future research will focus on efficient fine-grained access control through secure computation, with an emphasis on strict authentication and authorization mechanisms. These enhancements aim to support complex layered access management scenarios, such as IoT data access systems. Finally, we intend to conduct a thorough real/ideal security analysis based on the framework proposed by Curtmola et al., ensuring a comprehensive assessment of our scheme's security guarantees.

REFERENCES

[1]A.WebServices.(May2024).WhatisCloudStorage?[Online].Available:https://aws.amazon.com/what-is/cloud- storage
[2]B. Varghese and R.Buyya, ''Next generation cloud computing: New trends and research directions,''FutureGener.Comput.Syst. vol.79, pp. 849–861, Feb.2018.
[3]M.K.Morol,''Data security and privacy in cloud computing platforms: A comprehensivereview,''Int.J.CurrentSci.Res.Rev.,vol.5, no.5,pp.1–9,May2022.
[4]K.Ren,C. Wang, and Q.Wang,''Security challenges for the public cloud, ''IEEE Internet Comput.,vol.16,no.1,pp.69–73,Jan.2012.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

[5]M. Ali, S. U. Khan, and A. V. Vasilakos, ‘‘Security in cloud computing: Opportunities and challenges, ’’Inf.Sci., vol.305,pp.357–383,Jun.2015.
[6]H.Tabrizchi and M.Kuchaki Rafsanjani, ‘‘A survey on security challenges in cloud computing: Issues, threats, and solutions, ’’J.Super comput.,vol. 76, no. 12, pp. 9493–9532, Dec. 2020.
[7]D. Agrawal, A. E. Abbadi, F. Emekci, and A. Metwally, ‘‘Database management as a service: Challenges and opportunities,’’ in Proc. IEEE 25th Int. Conf. Data Eng., Mar. 2009, pp. 1709–1716.
[8]S. Wang, D. Agrawal, and A. E. Abbadi, ‘‘A comprehensive framework for secure query processing on relational data in the cloud,’’ in Workshop Secure Data Manage., vol. 6933, W. Jonker and M. Petkovi, Eds., 2011,pp.52–69.
[9]S.DeCapitanidi Vimercati,S.Foresti, S.Paraboschi, G.Pelosi, and P. Samarati, ‘‘Efficient and private access to outsourced data,’’ in Proc.31st Int. Conf. Distrib. Comput. Syst., Jun. 2011, pp. 710–719.
[10]X. Tian, C. Sha, X. Wang, and A. Zhou, ‘‘Privacy preserving query processing on secret share based data storage,’’ in Proc. 16th Int. Conf. Database Syst. Adv. Appl., Jan. 2011, pp. 108–122.