Research Article

Secured Multikeyword Search over Encrypted Cloud Data Based On Quality and Usability

B.Ananthi[1], S.V.Priyadharshini[2], and M.Ramesh[3]

[1]Assistant Professor, Department of Computer Science & Engineering, Vivekanandha College of Engineering for Women(Autonomous), Tiruchengode - 637205.Email:ananthi@vcew.ac.in
[2]PG Scholar, Department of Computer Science & Engineering, Vivekanandha College of Engineering for Women (Autonomous), Tiruchengode - 637205. Email:svpdharshini2015@gmail.com
[3]Assistant Professor, Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women(Autonomous) Tiruchengide- 637205. Email:ramesh@vcew.ac.in

Article History: Received: 5 April 2021; Accepted: 14 May 2021; Published online: 22 June 2021

Abstract

Searchable encryption permits to upload encrypted documents to a remote honest-but-suspicious server and query that data at the server without the papers having to be decrypted first. With the advent of cloud computing, data owners are encouraged to move their sophisticated data management systems from local sites to commercial public clouds for greater flexibility and cost savings. However, in order to safeguard data privacy, sensitive data must be encrypted before being outsourced, rendering traditional data utilization based on plaintext keyword search outdated. As a result, implementing an encrypted cloud data search service is critical. Given the huge number of data users and documents in the cloud, it is critical for the search service to support multi-keyword queries and result similarity ranking in order to meet the effective data retrieval requirement. In this paper propose the Secured Multikeyword Search over Encrypted Cloud Data, which is based On Quality and Usability of cloud data transmission and storage. Further we used triple DES (Data Encryption Standard) Algorithm of encryption and decryption key for secure authentication process. Here we used different key sizes are used in cryptographic process. Our analysis shows that the suggested approach is secure against adaptive chosen-keyword attacks. This solution is highly efficient and ready to be applied in real-world cloud storage systems, and it also achieves better encryption and decryption execution speed. .

Keyword: cloud, encryption, Machine Learning, Standard Hilbert curve and Secured Multikeyword search.

1 INTRODUCTION

Searchable Encryption enables querying of encrypted data in the cloud without decryption. Nonetheless, due of the fundamentally dissimilar relationship between the variables, most SE solutions are focused on SQL queries and cannot readily be applied to spatial data. In order to enable query services on encrypted spatial data, space filling curves have commonly been used to convert the original positions of POIs to one-dimensional index values. A space filling curve is one that crosses across every partition of a closed space without intersecting itself [1 and 2]. In this method, each point in multidimensional space will be mapped as a value to one-dimensional space. The standard Hilbert curve (SHC), a sort of space filling curve, is employed as a building block in many schemes for spatial data processing, which can protect the confidentiality of outsourced geographical data and enable successful spatial enquiries. Using the transformation key and the original geographic query, users can create query tokens to search across the encrypted spatial data [3]. As a result, fine-grained verification capability authorization is supported, which means that only users whose verification structure matches the permitted region can validate the query result [4]. Cloud storage is a computer data storage system that stores digital data in logical pools known as "the cloud." Physical storage is frequently distributed over multiple servers (potentially in different locations), and the physical environment is typically owned and maintained by a hosting company. Companies only need to pay for the storage they use, which is usually an average of consumption over the course of a month. This is not to say that cloud storage is less expensive; rather, it has ongoing costs rather than upfront expenditures [5].

A spatial database is one that is designed to store and query data that represents objects specified in a geometric space. The depiction of simple geometric objects such as points, lines, and polygons is supported by the majority of spatial databases. Some spatial databases are capable of handling more complicated structures such as 3D objects, topological coverage, linear networks, and TINs. While traditional databases have evolved to

support multiple numeric and character data types, such databases require additional capability to efficiently process geographical data types, and developers have frequently included geometry or feature data types. The Simple Features specification was created by the Open Geospatial Consortium (OGC), which establishes standards for adding geographical capabilities to database systems. The application of Factor Analysis in Geography, made so simple by modern computers, has been widespread but not always prudent. Because the retrieved vectors are specified by the data matrix, it is not possible to compare factors from different censuses. One method is to combine many census matrices into a single table that can then be analyzed [6]. This, however, presupposes that the variables' definitions have not changed over time and results in very huge, difficult-to-manage tables. In Los Angeles, however, it has played the previously unnoticed function of Downtown as an organizational core for the entire city for several decades. Many enterprises have employed multi storage cloud techniques in the cloud computing environment to mitigate disaster recovery and boost user flexibility. The data that the client has outsourced to a storage server may become corrupted. As a result, the integrity of the client's sensitive data is jeopardized.

Provable Data Possession (PDP) is a technique that allows us to validate the integrity of outsourced data in an untrusted server without having to retrieve the complete file from the cloud server or keep a duplicate of the data. It is implemented using Identity Based Encryption (IBE) and other PDP techniques [7]. Dobson and Fisher's critical paper concerns the potential misuse of location information. Aside from the legislative framework, there are numerous technical techniques to protecting privacy through the use of privacy-enhancing technologies (PETs). PETs range in complexity from simple on/off switches to sophisticated PETs that include anonymization techniques or cryptographic algorithms [8]. Only a few LBSs have such PETs; for example, Google Latitude has an on/off button and lets users to lock their position to a freely definable area. Furthermore, it is unknown how users perceive and trust different PETs. It is the only study that examines user perception of cutting-edge PETs. Another set of techniques included in the PETs is location obfuscation, which slightly alters the users' location in order to mask their true location while still representing their position and receiving services from their LBS provider. LBS relies heavily on mobile texting. SMS messaging, in particular, has been utilized in conjunction with a variety of LBS applications, such as location-based mobile advertising. SMS is the primary technology for delivering mobile advertising and marketing campaigns to mobile phones [9].

2 LITERATURE REVIEW

M. Talha et.al., [10] has proposed in this paperDatabase redistributing is a typical distributed computing worldview that permits information proprietors to exploit its on-request capacity and computational assets. The primary test is keeping up information classification concerning untrusted parties i.e., cloud specialist organization, just as giving applicable question brings about ongoing to confirmed clients. Existing methodologies either bargain classification of the information or experience the ill effects of high correspondence cost between the worker and the client. To beat this issue, we propose a double change and encryption conspire for spatial information, where encoded inquiries are executed completely at the administration supplier on the encoded database and scrambled outcomes are come back to the client. The client issues scrambled spatial range questions to the specialist organization and afterward utilize the encryption key to unscramble the inquiry reaction returned.

In this study, Miao et al. [11] suggest a Fog computing, an outgrowth of cloud computing, outsources encrypted sensitive data to multiple fog nodes on the edge of the Internet of Things to reduce latency and network congestion (IoT). Existing cipher text retrieval systems, on the other hand, rarely focus on the fog computing environment, and the majority of them continue to impose significant computational and storage overhead on resource-constrained end users. We first provide a Lightweight Fine-Grained Cipher Text Search (LFGS) system in fog computing by combining Cipher text-Policy Attribute-Based Encryption (CP-ABE) and Searchable Encryption (SE) technologies, which can conduct fine-grained access control and keyword search simultaneously.

Xu et al., [12] built a basic query function, range query, which has been employed in a variety of applications such as SQL retrievals, location-based services, and computational geometry. Meanwhile, as data volume grows at an exponential rate, users are increasingly attracted to store data in the cloud in order to save on local storage and computational costs. However, a long-standing issue is that the user's data may be completely disclosed to the cloud server because it has full data access privileges. To overcome this issue, one frequent method is to encrypt raw data before outsourcing it, however this significantly limits data availability and operability. We

present an Efficient and Geometric Range Query (EGRQ) technique for searching and restricting data access across encrypted spatial data in this paper. We employ secure KNN computation, polynomial fitting, and order-preserving encryption to perform a safe, efficient, and accurate geometric range query over cloud data. In our EGRQ, we then provide a novel geographical data access control technique to refine user rights..

Yang Ji et.al., [13] has proposed in this paper with the expansion of distributed computing and dataset an administration (DaaS), an ever increasing number of associations and people redistribute their information to an outsider specialist co-op. While getting a charge out of the advantages of cloud-based information redistributing, the information proprietors are at the danger of losing control of information uprightness and get to the board. In this exhibition, we present a framework called vABS, which empowers evident Attribute-Based Search over shared cloud information. The vABS framework receives the normal DaaS engineering, in which the worker gives search administrations to clients for the benefit of information proprietors.

Jianbing Ni et al., [14] has proposed in this paperPortable crowd sensing empowers a horde of people to helpfully gather information for extraordinary intrigue clients utilizing their cell phones. The accomplishment of versatile crowd sensing to a great extent relies upon the taking an interest versatile clients. The more extensive cooperation, the all the more detecting information are gathered; by the by, the more reproduce information might be produced, subsequently bringing superfluous substantial correspondence overhead. Subsequently it is basic to dispose of copy information to improve correspondence proficiency, information duplication.

3 PROPOSED METHODOLOGY

We define and solve the difficult problem of privacy-preserving multi-keyword ranking ontology keyword mapping and search over encrypted cloud data (MROS), as well as develop a set of stringent privacy requirements for such a safe cloud data usage system to be realized. We select the efficient concept of coordinate matching from among several multi-keyword semantics. We present the Secured Multikeyword Search (SMS) problem over encrypted cloud data (ECD) and build a set of privacy standards for such a safe cloud data utilization system..

We first propose a basic Secured multi keyword ranking ontology keyword mapping and search technique based on safe inner product computation, which we subsequently develop to fulfill various privacy needs. The Ranked result displays the first k retrieval results. In addition, we propose an alert system that will generate alerts when an unauthorized user attempts to access data from the cloud; the alert will be generated in the form of a mail and a message..

3.1 Advantages of Proposed System:

1. MLKS-Tree multi-keyword ranking ontology keyword mapping and search over encrypted cloud data (Machine Learning keyword Search).

2. Coordinate matching based on inner product similarity.

3. Secured multi-keyword ranked ontology keyword mapping and search : Create search systems that allow multi-keyword queries and give result similarity rating for important data retrieval, rather than returning undifferentiated results.

4. Privacy: To prevent the cloud server from learning more information from the dataset and index, as well as to comply with privacy standards. Efficacy combined with good performance: The above functionality and privacy criteria should be met with minimal communication and processing overhead.

Secure Privacy Preserving Keyword Search (SPKS) authorizes a cloud service provider to decrypt data and provide a file containing keywords.

This method reduces computing and communication overhead while also providing users with query and data privacy. It devises six techniques for searching encrypted data efficiently. The flow of a Secure multi-keyword search (SMS) with coordinate matching is shown. To begin, a public/private key pair is generated using the TRIPLE DES Algorithm for Key Generation. Second, the encrypts all of the content in the file and the keywords, which are then saved on the server. Third, in the retrieving phase, where the user creates a trapdoor and passes it to CSP, compute is employed (cloud service provider). Fourth, TRIPLE DES with MROS determines whether the keyword is present in the encrypted data. Fifth, Decrypt primarily for CSP to partially decrypt the intermediate

result and provides the cipher text and partially decrypted material. Sixth, the user launches Recovery to decode the plain text. As a result, it provides semantic security against plain text attacks.

When searching in encrypted data with a multi-Ranking keyword search, the following actions must be taken.

- Multi-dimensional queries are translated to their sms secure multikeyword search coordinate matching counterparts.
- Attributes are defined hierarchically, i.e., attribute hierarchy
- The MROS Index and Classified Average Precision algorithms generate the indexes and capabilities, respectively.
- It is a multi-round protocol between a server and a single user based on a single keyword.
- It employs per index file, with each page including a keyword. The keyword index is encrypted using heuristic pseudorandom functions and pseudorandom bits.

During the setup phase the user selects a secret key to encrypt the file. The index and file content are then sent to the server by the user. During the retrieval phase, when the user wishes to search for or retrieve a file from the server, the user retrieves the index file and then computes the keyword using the secret key. The computed key is sent to the server, which compares it to the file and sends it to the user. The bloom filter strategy is outperformed by the per-index file scheme that uses pseudorandom functions. This approach fails when a large number of keywords are used.

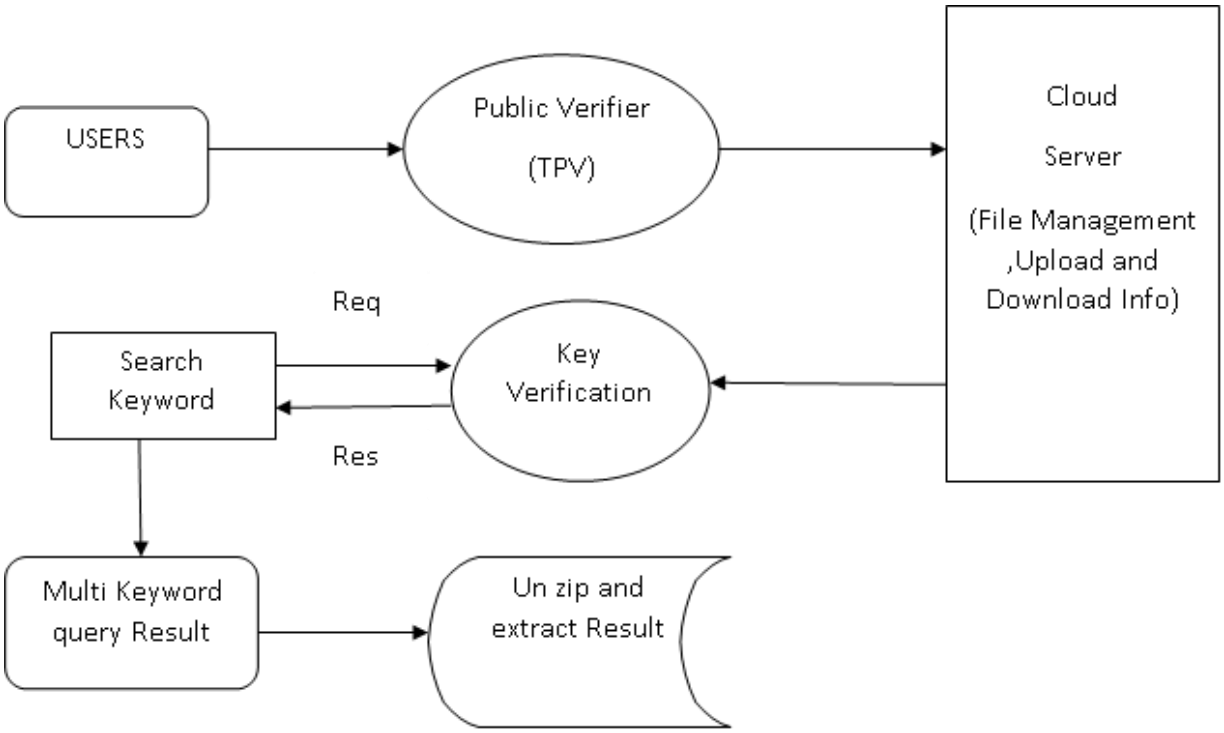


Figure 1: System Flow Diagram:

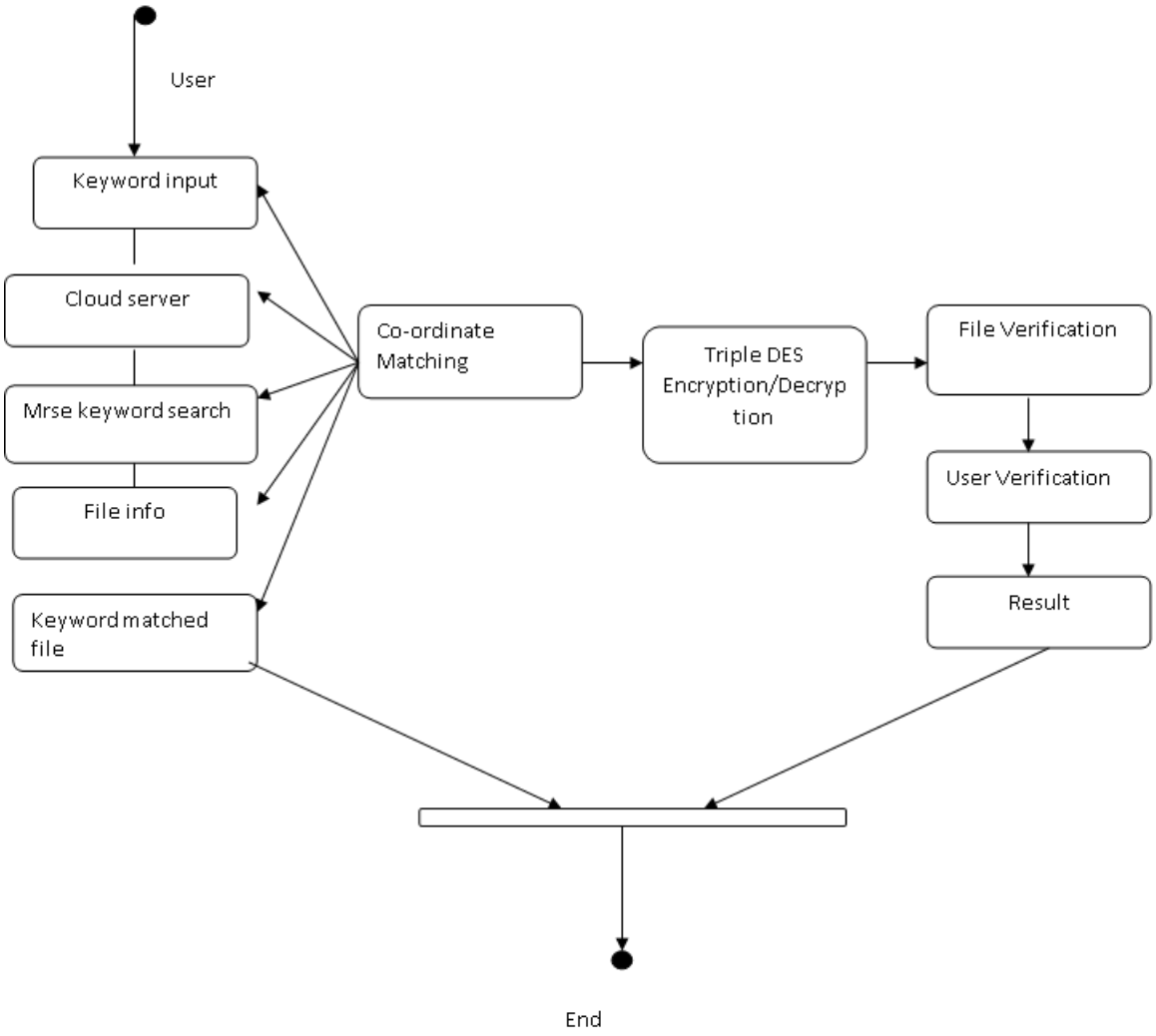


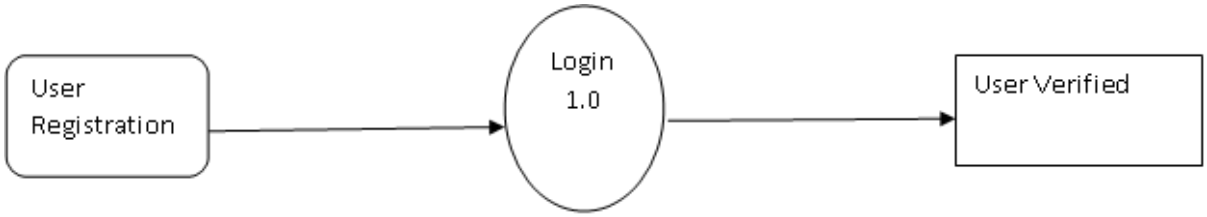
Figure 2: Process Flow Diagram

We offer a basic SMS strategy based on secure inner product computation, which is adopted from a secure technique, and then incrementally modify it to meet various privacy criteria in two threat models.

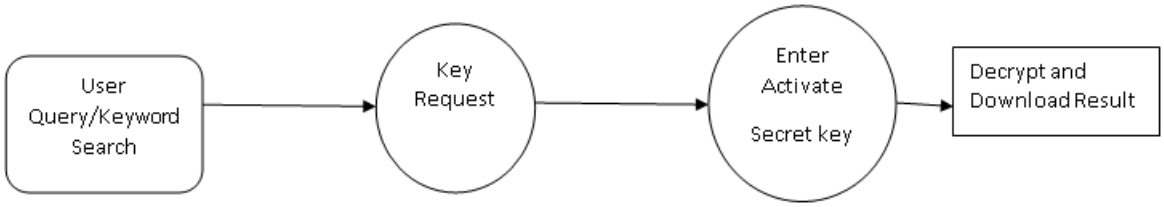
1) Demonstrating the issue of SMS over Encrypted Cloud Data

2) Propose two techniques based on the coordinate matching and inner product similarity principles.

Level 1:



Level 0



3.2 Admin Module:

This module is intended to assist the server in seeing details and uploading files while maintaining security. The log key is used by the administrator at login time. Change the log key before the admin logout. After logging in, the administrator can reset the password and check the user's downloading data as well as the number of file requests counted on the flowchart. After converting the Zip file format, the administrator can upload the file.

3.3 File upload Module:

This module is intended to assist the server in seeing details and uploading files while maintaining security. The log key is used by the administrator at login time. Change the log key before the admin logout. After logging in, the administrator can reset the password and check the user's downloading data as well as the number of file requests counted on the flowchart. After converting the Zip file format, the administrator can upload the file.

3.4 Ranking Result:

When a user requests data, the k-nearest neighbor algorithm is used to rank the requested data. The notion of co-ordinate matching is employed for ranking. After ranking, the user receives the expected query results.

4 EXPERIMENTAL SETUP AND RESULT

For the data owners and data consumers, several users are formed in a centralized location. We can see that once logged in, either of the users can access the system.

The flow of communication between data owners and data users takes place solely through the Data frames system, allowing the system to be secure. Because the contents are encrypted and stored in the cloud, public access to these files is not feasible.

The files or contents can only be read with the authorization of the data owners and after obtaining the secret key.

4.1 System Implementation

The most important stage in developing a successful system and providing users trust that the new system is workable and effective is implementation. Implementation of a customized application that will be used to replace an old one. As long as there are no big changes in the system, this form of convetion Triple DES is rather simple to handle. .

Table 1: Hardware Requirements:

System	:	Pentium Core 2 Duo
Hard Disk	:	80 GB
RAM	:	1 GB
Key Board	:	LG
Mouse	:	Logitech
Monitor	:	15 inch TFT Color monitor

4.2 Cloud Setup

In this module, we set up the data owner and the cloud server. As a result, the data owner will upload the data to the cloud server. When users outsource their private data to the cloud, the cloud service providers have control over and monitoring of the data, and the communication between users and the cloud is safe. Cloud Storage for Cryptography While the data put into this module may contain sensitive information, the cloud servers cannot be totally trusted to protect the data. As a result, outsourced files must be encrypted. Any type of information leaking that jeopardizes data privacy is considered inappropriate. In this model, we used a variety of searchable symmetric encryption techniques to enable cipher text search. The former ranks files only based on the amount of retrieved keywords, which reduces search accuracy. As a first stage, the executable version of the application must be developed and loaded in a common server machine that is available to all users, and the server must be network-connected. The final stage is to document the complete system, which includes the system's components and operational processes.

Each program is tested individually during development using data and has verified that this program linked together in the method stated in the program specification, the computer system and its surroundings are tested to the user's satisfaction. The designed system has been acknowledged and proven to be satisfactory for the user. As a result, the system will be installed very soon. A short operating procedure is supplied so that the user can quickly and easily grasp the many functions.

Encryption and decryption of data Result When we apply the Triple DES algorithm to the data, we receive encrypted data. This encrypted data is then stored on the cloud. After downloading and decrypting the file, the user can access the data. Keys for encryption and decryption are provided.

Ranking Result When a user requests data, ranking is performed on the provided data in order to employ the co-ordinate matching principle. After ranking, the user receives the expected query results.

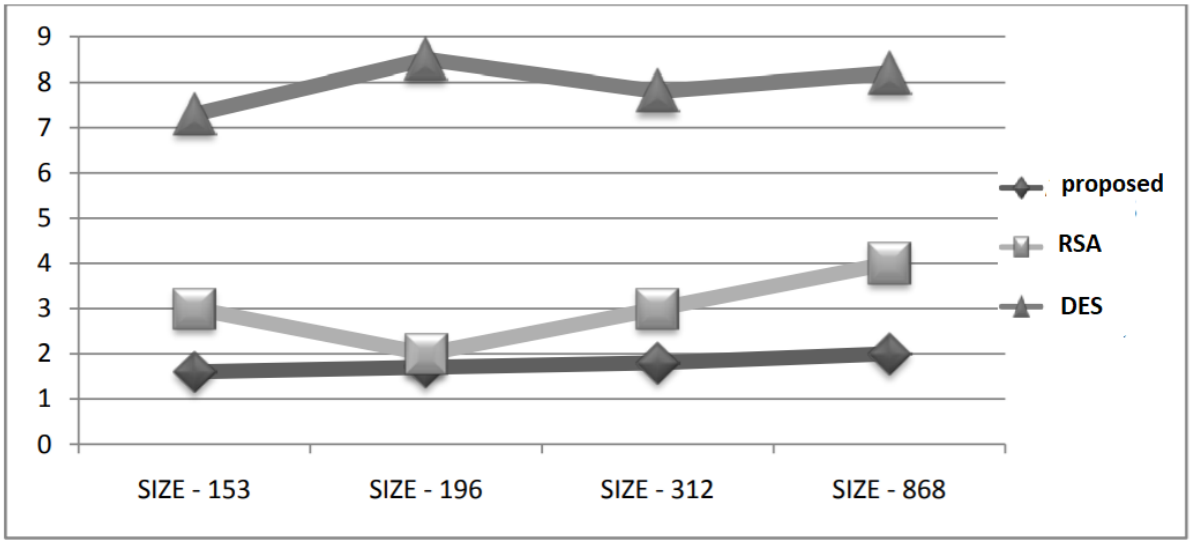


Figure 3: Comparison of Encryption Time among proposed method with DES and RSA

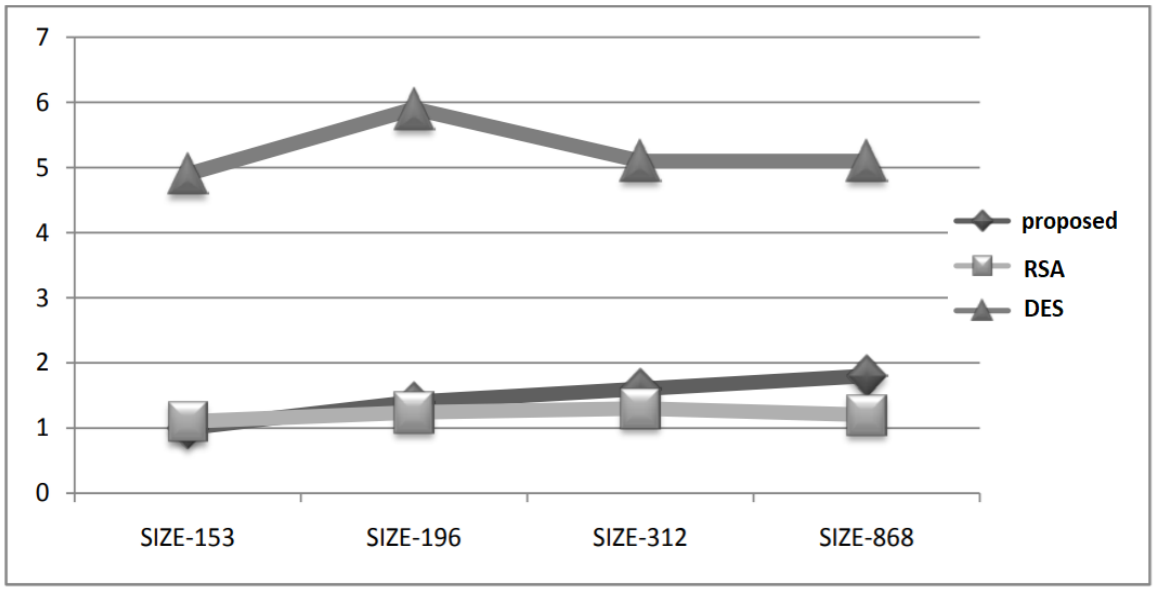


Figure 4: Comparison of Decryption Time among proposed method with DES and RSA

By examining Figs. 4 and 5, which indicate the time required for encryption and decryption on various file sizes using three techniques. The RSA algorithm takes substantially longer than the DES, RSA, and Triple DES algorithms. The Triple DES algorithm has a very slight difference in the time it takes to encrypt and decrypt data.

5 CONCLUSION

Searchable encryption is a mechanism which enables secure searching on remote servers of encrypted data. In this work, we discover and solve the issue of searching encrypted cloud data with one term for the first time, and we offer a number of privacy restrictions. In order to successfully capture a resemblance between query keywords and outsourced documents, we chose an efficient principle of "coordinate co-ordination" i.e. as many matches as feasible between different multi-keyword sémantics. To respond to the difficulty of multi-semantic words without losing confidentiality. We first offer a basic Secured Multikeyword Search strategy based on safe inner product computing, which we significantly improve to meet privacy requirements in two threat models. By comparing DES, RSA, and the planned Triple DES. The proposed approach demonstrates a significant difference in the time required for the encryption and decryption processes. A thorough investigation into the privacy and efficiency guarantees of various methods is provided, as are real-world trials.

Reference

[1]. C. Guo, R. Zhuang, Y. Jie, K. Choo, and X. Tang, “Secure rangesearch over encrypted uncertain iot outsourced data,” IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1520–1529, 2018.

[2]. Y. Wang, Q. Wu, B. Qin, W. Shi, R. Deng, and J. Hu, “Identitybased data outsourcing with comprehensive auditing in clouds,” IEEE Transactions on Information Forensics and Security, vol. 12, no. 4, pp. 940–952, 2017.

[3]. X. Yao, R. Zhang, Y. Zhang, and Y. Lin, “Verifiable social data outsourcing,” in Proc. of IEEE Conference on Computer Communications, Atlanta, USA, May 2017, pp. 1-9

[4]. Y. Zhang, C. Xu, X. Liang, H. Li, Y. Mu, and X. Zhang, “Efficient public verification of data integrity for cloud storage systems from in distinguishability obfuscation,” IEEE Transactions on InformationForensics and Security, vol. 12, no. 3, pp. 676–688, 2017.

[5]. K. Yang, Q. Han, H. Li, K. Zheng, Z. Su, and X. Shen, “An efficient and fine-grained big data access control scheme with privacy preserving policy,” IEEE Internet of Things Journal, vol. 4, no. 2,pp. 563–571, 2017.

[6]. J. Li, R. Ma, and H. Guan, "Tees: An efficient search scheme over encrypted data on mobile cloud," IEEE Transactions on CloudComputing, vol. 5, no. 1, pp. 126–139, 2017

[7]. H. Tian, Y. Chen, C.-C. Chang, H. Jiang, Y. Huang, Y. Chen, and J. Liu, "Dynamic-hash-table based public auditing for secure cloud storage," IEEE Transactions on Services Computing, vol. 10, no. 5, pp. 701–714, 2017.

[8]. 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.

[9]. F. Tian, X. Gui, J. An, P. Yang, and X. Zhang, "A density-based space filling curve for location privacy-preserving," in Proc. of the 11th IEEE International Conference on Services Computing, Alaska,USA, Jun. 2014, pp. 131–138

[10]. M. Talha, I. Kamel, and Z. A. Aghbari, Facilitating secure and efficient spatial query processing on the cloud, IEEE Transactions on Cloud Computing, vol. 7, no. 4, 2019.

[11]. 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. 12, no. 5, pp. 772–785, 2019.

[12]. G. Xu, H. Li, Y. Dai, K. Yang, and X. Lin, Enabling efficient and geometric range query with access control over encrypted spatial data, IEEE Transactions on Information Forensics and Security,vol. 14, no. 4, pp. 870–885, 2019.

[13]. Y. Ji, C. Xu, J. Xu, and H. Hu, vabs: Towards verifiable attribute based search over shared cloud data, in Proc. of the 35th International Conference on Data Engineering, Macao, China, 2019, pp.2028–2031.

[14]. J. Ni, K. Zhang, Y. Yu, X. Lin, and X. Shen, Providing task allocation and secure deduplication for mobile crowd sensing via fog computing, IEEE Transactions on Dependable and Secure Computing, vol. 17, no. 3, pp. 581-594, 2018.