Efficient Privacy-Preserving Skyline Query Processing on Encrypted Cloud Data

Jasvika S[1], Dr.Sreejith Vignesh B P[2]
[1]*Junior Researcher, Department of Information Technology*
[2]*Associate Professor & Head, Department of Information Technology*
[1,2]*Sri Krishna Adithya college of arts and science*

***Abstract*—Cloud computing has turned out to be a popular choice for data storage. However, the storage of data in the cloud has raised serious privacy issues. In order to mitigate this problem, data owners usually use encryption to secure their data before it is uploaded to the cloud. Even though encryption is used to secure data, it becomes difficult to perform complex database operations. One of the complex database operations is the skyline query, which is used to retrieve the best data according to multiple criteria without the need to specify preference weights. Performing skyline queries over encrypted data is challenging, as the cloud server will not be able to access the data to compare it. This paper aims to describe a method of processing skyline queries over encrypted data in the cloud, ensuring data privacy. In this method, the cloud server will be able to perform the skyline query without knowing the data content. This will be achieved by using encryption with efficient query.**

***Index Terms*—Skyline Query, Encrypted Data, Privacy-Preserving Query Processing, Cloud Computing, Secure Data Outsourcing, Multi-Criteria Decision Making, Data Security, Encrypted Database Processing.**

I. INTRODUCTION

Cloud computing is one of the most popular techniques used for data storage and management. For data privacy, data owners often use data encryption techniques to encrypt their data before storing them in the cloud. But data encryption makes it difficult to process complex database queries. Skyline queries can be used to identify the best data. But processing skyline queries in encrypted data is difficult because the server cannot access the data. In this work, we have used skyline queries to process encrypted data in the cloud.

II. PROBLEM MOTIVATION WITH REAL - WORLD STATISTICS

Large volumes of data are frequently stored on cloud computing, and research indicates that over 80% of businesses use cloud services for data management. However, there are privacy and security issues with storing sensitive data on cloud servers. Before uploading their data to the cloud, data owners typically encrypt it to safeguard sensitive information. Although encryption guarantees security, it also makes complex queries, like skyline queries, challenging to execute. Users must choose the best options based on a variety of factors, including price, quality, and distance, in many real-world applications. Thus, effective techniques that enable the processing of skyline queries on encrypted data while preserving data privacy are required.

III. LITERATURE REVIEW & REVIEW OF RECENT RELATED STUDIES

Methods for handling skyline queries on encrypted data kept in cloud environments have been investigated in a number of studies.[1] Researchers have developed privacy-preserving methods that enable cloud servers to carry out skyline query operations without disclosing private data.[2] To guarantee data confidentiality during query processing, these approaches primarily employ encryption techniques.[3]

Enhancing the effectiveness of skyline query processing on sizable encrypted datasets is another area of recent research [4]. To cut down on computation time without sacrificing privacy, strategies like secure comparison protocols and

optimized query algorithms have been developed. [5]. However, scalability and high computational cost continue to be issues for many current methods.[6] In order to create more effective and safe techniques for processing skyline queries on encrypted cloud data, more research is required.[7]

IV. DATASET DESCRIPTION

Multi-dimensional records with multiple attributes that are necessary for processing skyline queries are included in the dataset used in this study. Values like price, rating, distance, and other pertinent factors that aid in determining the best outcomes are included in each record.[8] To guarantee data security and privacy, all attribute values are encrypted prior to the dataset being stored on the cloud server. Skyline query operations are then carried out using the encrypted dataset without disclosing the original data values.[9] This dataset aids in assessing the suggested skyline query processing method's effectiveness and capacity to protect privacy.

EXISTING SYSTEM

- Skyline queries are executed on plaintext data in conventional database systems, allowing the server to directly access and compare attribute values like price, rating, distance, or quality.
- Without giving each attribute a weight, these techniques enable the effective identification of ideal records based on a variety of criteria, assisting users in making better decisions.
- Data is kept on cloud servers in many applications to lower storage costs and increase accessibility. However, there are security and privacy issues when sensitive data is stored on the cloud.
- Before sending their datasets to cloud servers, companies frequently encrypt them to safeguard sensitive information. The original data cannot be read by unauthorized users thanks to encryption.

PROPOSED SYSTEM

In cloud environments, the suggested system offers a safe framework for handling skyline queries over encrypted data. The suggested method allows skyline query computation directly on encrypted datasets while maintaining data confidentiality, in contrast to conventional methods that require access to plaintext data. Sensitive data is kept safe throughout storage and query processing thanks to the system. The framework enables cloud servers to carry out skyline operations without disclosing the true attribute values by fusing encryption techniques with effective query processing methods.[10]

A. Layer of Data Encryption

In this layer, the data owner encrypts the dataset before outsourcing it to the cloud server. Secure encryption algorithms are used to transform each attribute value in the dataset into an encrypted format.[11] This procedure guarantees that unauthorized parties cannot access or alter sensitive data. After that, the encrypted dataset is kept in a cloud environment so that skyline query processing can safely utilize it.[12]

B. Query Processing Layer

The query processing layer is tasked with the processing of skyline queries received from users. When a user submits a skyline query based on various parameters such as price, distance, and quality, the cloud server processes the query based on secure comparison methods. The secure comparison methods enable the server to establish dominance relationships between records without decrypting the actual values of the records.

C. Result Decryption Layer

After the skyline computation is completed, the encrypted results are sent back to the user. The user then decrypts the results using the appropriate decryption key to obtain the final skyline records. This layer ensures that only authorized users can access the query results while the cloud server remains unaware of the actual data values.

V. RESEARCH DESIGN METHODOLOGY

The proposed research work adopts a design science research methodology that aims to design and evaluate a secure framework for processing skyline queries on encrypted cloud data. The research methodology is divided into the following steps:

Problem Identification:
Identify the problems associated with the existing skyline query processing systems, especially the problem of processing skyline queries on encrypted data in an efficient manner.

Data Encryption and Modelling:
Model the data using various attributes such as price, rating, distance, and quality, which are necessary for skyline processing. The data is then encrypted before being sent to the cloud server for processing to ensure data privacy.

Secure Skyline Query Processing:
Design a secure skyline query processing system that enables the cloud server to compare the encrypted values of the attributes and determine the non-dominated records without compromising data privacy.

Algorithm Development:
Develop an efficient skyline query algorithm that securely checks the dominance of the encrypted data and produces the skyline result while ensuring data privacy.

VI. MODEL COMPARISON

Evaluation of the Proposed Framework
The proposed framework for processing skyline queries on encrypted data is compared with two existing methods: traditional plaintext skyline query systems and basic encrypted database query systems.

Traditional Skyline Query Systems
Traditional skyline query systems are designed to work directly on plaintext data stored in traditional databases. These systems enable efficient comparison of multiple attributes to determine the best records based on dominance criteria. Although these systems enable fast query processing and the generation of correct query results, they do not guarantee data privacy when the data is outsourced to cloud platforms. In this case, the server has access to the original values of the data, which can lead to the disclosure of sensitive information to unauthorized parties or third-party service providers.

Basic Encrypted Database Methods
Basic encrypted database methods are designed to mainly guarantee data privacy by encrypting data before it is stored in the cloud. Although these methods guarantee data confidentiality, most of them do not enable complex query operations such as skyline queries. In most cases, the data has to be decrypted before the query processing operation can be performed.

VII. INTEROPERABILITY AND DATA INTEGRATION

The proposed framework provides support for data integration and interoperability within cloud environments for secure skyline query processing on encrypted data. The framework is intended to be compatible with various cloud platforms and database management systems, ensuring that encrypted data can be stored, retrieved, and processed without any compatibility problems. The framework enables seamless interaction between users, data owners, and cloud servers through standardized data formats and secure communication channels.
For cloud-based environments, the framework is integrated with cloud storage services where encrypted data is stored securely. The skyline queries entered by users are processed by the cloud server through secure comparison methods without accessing the actual data values. This integration provides efficient query processing with preserved data privacy. Moreover, the framework comprises a data integration layer that aggregates and manages multi-attribute data necessary for skyline query processing. Data from various sources can be integrated and encrypted before being uploaded to the cloud. This integrated data enables the processing of skyline queries on various attributes such as cost, rating, distance, or quality.

VIII. CONCLUSION

This work has offered a secure method for skyline query processing on encrypted data in cloud storage systems. The proposed system enables decision-making queries with multiple criteria while preserving the confidentiality of critical data. The system ensures the confidentiality of the data by encrypting the data before storing it in the cloud and using secure comparison algorithms during query processing.

The proposed system enhances data security while allowing efficient computation of skyline queries. The system enables privacy-preserving data analysis and helps decision-makers achieve optimal solutions without violating data confidentiality. In general, the system has shown that skyline query processing can be efficiently done in cloud-based data storage systems.

REFERENCE

[1] S. Börzsönyi, D. Kossmann, and K. Stocker, "The Skyline Operator," Proceedings of the IEEE International Conference on Data Engineering (ICDE), 2001.

[2] W. Balke, U. Güntzer, and J. X. Zheng, "Efficient Distributed Skylining for Web Information Systems," Proceedings of the International Conference on Extending Database Technology (EDBT), 2004.

[3] K. Wang, B. C. M. Fung, and P. S. Yu, "Handing Encrypted Data in Cloud Databases for Secure Skyline Queries," IEEE Transactions on Knowledge and Data Engineering, 2014.

[4] H. Hacıgümüş, B. Iyer, and S. Mehrotra, "Providing Database as a Service with Encrypted Data," Proceedings of the IEEE International Conference on Data Engineering (ICDE), 2002.

[5] B. Hore, S. Mehrotra, and G. Tsudik, "Privacy-Preserving Skyline Queries over Encrypted Data," Proceedings of the International Conference on Data Engineering (ICDE), 2009.

[6] J. Vaidya and C. Clifton, "Privacy-Preserving Data Mining: Why, How, and When," IEEE Security & Privacy Magazine, 2004.

[7] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," Proceedings of the ACM Symposium on Theory of Computing (STOC), 2009.

[8] X. Liu, R. Lu, J. Ma, L. Chen, and B. Qin, "Privacy-Preserving Skyline Query Processing over Encrypted Cloud Data," IEEE Transactions on Cloud Computing, 2016.

[9] Sreejith, Vignesh. and Babu, B.P.M.R "Classifying the Malware Application in the Android based smart phones using Ensemble ANFIS Algorithm", International Journal of Networking and Virtual Organization, Vol 19 N2/3/42018.

[10] Sreejith, Vignesh. And RajeshBabu "Experimental research identifications on Malware detection by embedding C4.5 algorithm and SVM in smart Phones" Perspectivasci journal on Information sciences" Vol 22 Special issue (2017).

[11] Sreejith Vignesh et al, "Machine Learning algorithms to control the security issues in android applications" Sambodhi UGC Care Journal ISSN: 2249-6661, Vol-43, No-4, (VI) October December (2020)

[12] B P Sreejith Vignesh, "Application of IPF to achieve CSR Routing in Adhoc Networks" Asian journal of Computer Science and Technology, ISSN 2249-0701, Volume 9 No.2 July-December 2020 pp 18-23