Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)

Volume7-Issue2, January 2024
Paper: 4

Optimizing Privacy and Efficiency Using Multi threaded Approach for Association Rule Mining

BHAVADHARANI.R[1], KANSIYA.S.K[2], MENAKA.K[3], NARMATHA.A[4], MANIVANNAN.K.M.TECH (Ph.D)[5]

[1,2,3,4] Students, and [5] Faculty
Dept. of Information Technology,, V.S.B.
Engineering College, Karur-639111.
India.
manivannan.vsbec@gmail.com
bhavaravi16@gmail.comkansiyakansi19@gmail.c
ommenakakalisamy@gmail.comnarmathanarmath
a461@gmail.com

Abstract: One major difficulty in cloud computing is to efficiently mine frequent item sets and association rules on encrypted outsourced data while maintaining privacy. In this research, we present a complete solution in a two-cloud approach for privacy-preserving association rule mining and frequent item set mining on outsourced data. Through the use of additive secret sharing and secure computing protocols, our method allows clouds to mine frequent item sets and association rules in a safe manner without jeopardizing data security. To further enhance the mining and query operations' efficiency, we also present optimization techniques. Time-consuming operations like L2S and SSC are parallelized through the use of multi-thread programming, which significantly reduces execution time. According to experimental data, clouds can accomplish the mining operation in significantly less time when multi-threading is used, which makes the scheme extremely effective. We also expect hardware improvements in real-world deployment scenarios, where Cloud Service Providers and Evaluators use virtual machines to leverage more powerful hardware configurations, further improving system performance. All in all, our suggested approach guarantees mining on encrypted data while maintaining privacy and, by means of optimization techniques, attains notable gains in performance, which qualify it for real-world implementation in cloud contexts. Extensive examination of the privacy and efficiency guarantees of the suggested methods is provided, and real-world dataset experiments confirm that the suggested schemes do, in fact, incur little overhead in terms of computing and communication.

Keywords: Association rules, L2S and SSC, Multi-thread programming

INTRODUCTION:

The proliferation of mobile communication technology and the widespread adoption of Internet-of-Things (IoT) devices have ushered in an era of unprecedented data generation. With vast amounts of data being produced daily, there arises a critical need for effective analysis techniques to extract valuable insights. In response to this demand, the paradigm of data mining-as-a-service in cloud computing environments has gained prominence, enabling individuals and organizations to leverage cloud platforms for data outsourcing and mining tasks.

To address these challenges, researchers have explored two primary avenues: randomization-based and cryptography-based solutions. While randomization-based approaches offer efficiency, they often sacrifice result accuracy. In contrast, cryptography-based schemes provide stronger security guarantees and accurate mining results. One such approach involves encrypting data using fully homomorphic cryptosystems (FHE) before outsourcing. However, FHE incurs significant computational overhead, rendering it impractical for large-scale dataset mining.

In this context, this paper focuses on addressing privacy issues in the context of frequent itemset mining and association rule mining—two fundamental techniques with diverse applications ranging from market prediction to network traffic management. By proposing secure computation protocols based on additively homomorphic cryptosystems and additive secret sharing, we aim to enable secure mining by multiple clouds while preserving the confidentiality of datasets, query data, and mining results. Additionally, our scheme supports both cloud-mined results and user-defined thresholds, catering to the needs of diverse stakeholders in the cloud computing ecosystem. Through experimental validation, we demonstrate the efficiency and effectiveness of our approach compared to existing state-of-the-art solutions.

By advancing privacy protection for association rule mining and queries in cloud computing environments, we contribute towards fostering trust and confidence in cloud-based data mining services, thereby unlocking the full potential of data-driven insights in various domains.

PRIVACY AND EFFICIENCY PROCESS USING AUTOMATION:

Our paper, " Optimizing Privacy And Efficiency Using Multithreaded Approach For Association Rule Mining**"** the adoption of a multithreaded approach for association rule mining presents a promising avenue for optimizing both privacy and efficiency. Through thorough testing, including unit, integration, system, and acceptance testing, the viability and effectiveness of this approach can be validated. Unit and integration testing ensure the robustness and compatibility of individual components, while system testing verifies the system's overall performance, scalability, and adherence to privacy regulations. Acceptance testing, involving stakeholders, serves as the final validation step, ensuring that the system meets user expectations and regulatory standards.

By combining these testing methodologies, organizations can confidently deploy a multithreaded association rule mining system that not only maximizes efficiency but also prioritizes the protection of sensitive data, fostering trust and compliance in an increasingly data-driven landscape.

ARCHITECTURE:

The existing system described in the provided content involves, employing a multi-threaded approach within an existing system can yield substantial benefits. By utilizing multiple threads, tasks can be divided and executed concurrently, reducing processing time and enhancing overall system efficiency. Moreover, with careful design considerations, such as implementing secure data partitioning and encryption techniques, privacy concerns can be addressed effectively. In proposed system, we aim to address the privacy and efficiency concerns of association rule mining by integrating a novel approach that combines differential privacy techniques with a distributed computing framework. Leveraging the power of differential privacy, we ensure that sensitive data remains protected while still allowing meaningful insights to be extracted. By incorporating a distributed computing framework such as Apache Spark, we can harness the parallel processing capabilities of modern computing clusters to improve efficiency and scalability. Through this approach, we seek to offer a robust solution that optimizes both privacy and efficiency for association rule mining tasks across diverse datasets and use cases.

1.Data Input: The process begins with the input of the dataset containing sensitive information.

2.Data Preprocessing: This step involves cleaning the data, handling missing values, and transforming the dataset into a suitable format for association rule mining.

3.Privacy Protection: In this stage, privacy-enhancing techniques are applied to the data to ensure that sensitive information is protected. This may involve techniques such as anonymization, differential privacy, or encryption.

4.Multi-threaded Association Rule Mining: The dataset is divided into multiple subsets, and each subset is

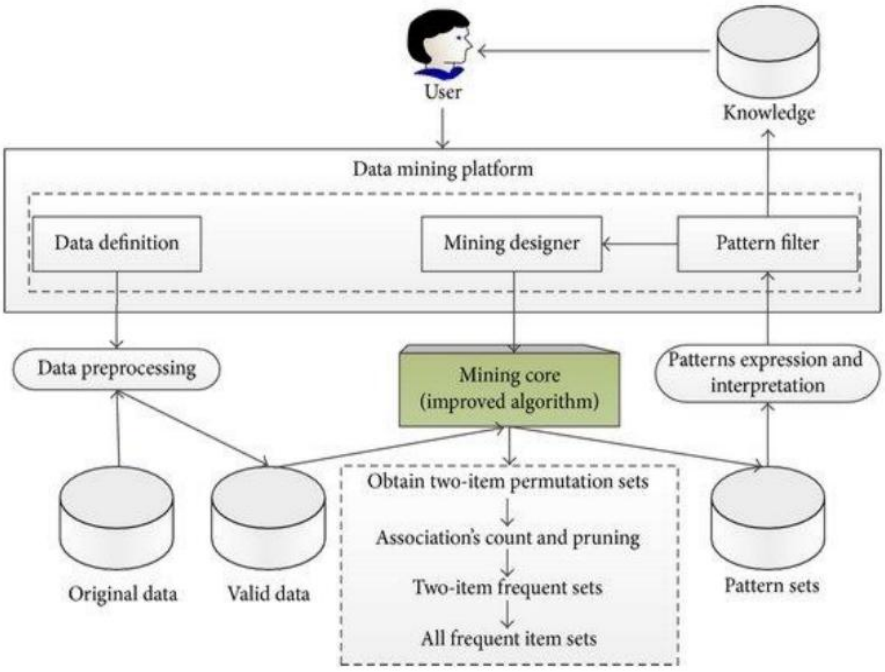
Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 4

processed simultaneously using multi-threading. Association rule mining algorithms, such as Apriori or FP-Growth, are applied to each subset to discover frequent item sets and association rules.

5.Rule Filtering: Once the association rules are generated, they are filtered based on predefined criteria. This may include minimum support and confidence thresholds, as well as privacy constraints to ensure that only relevant and privacy-preserving rules are retained.

6.Performance Evaluation: The efficiency and effectiveness of the mining process are evaluated using performance metrics such as execution time, memory usage, and the quality of the discovered rules.

Fig. 1. Architecture diagram



FLOWCHART:

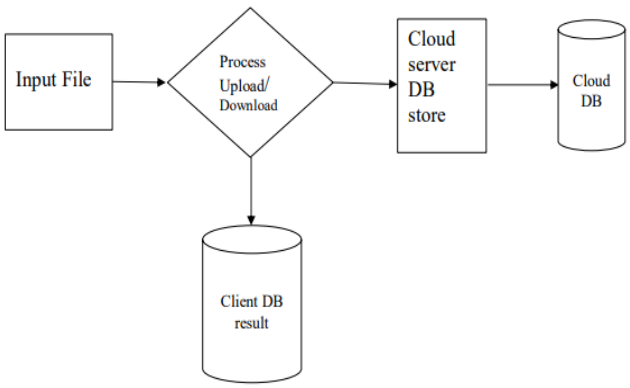
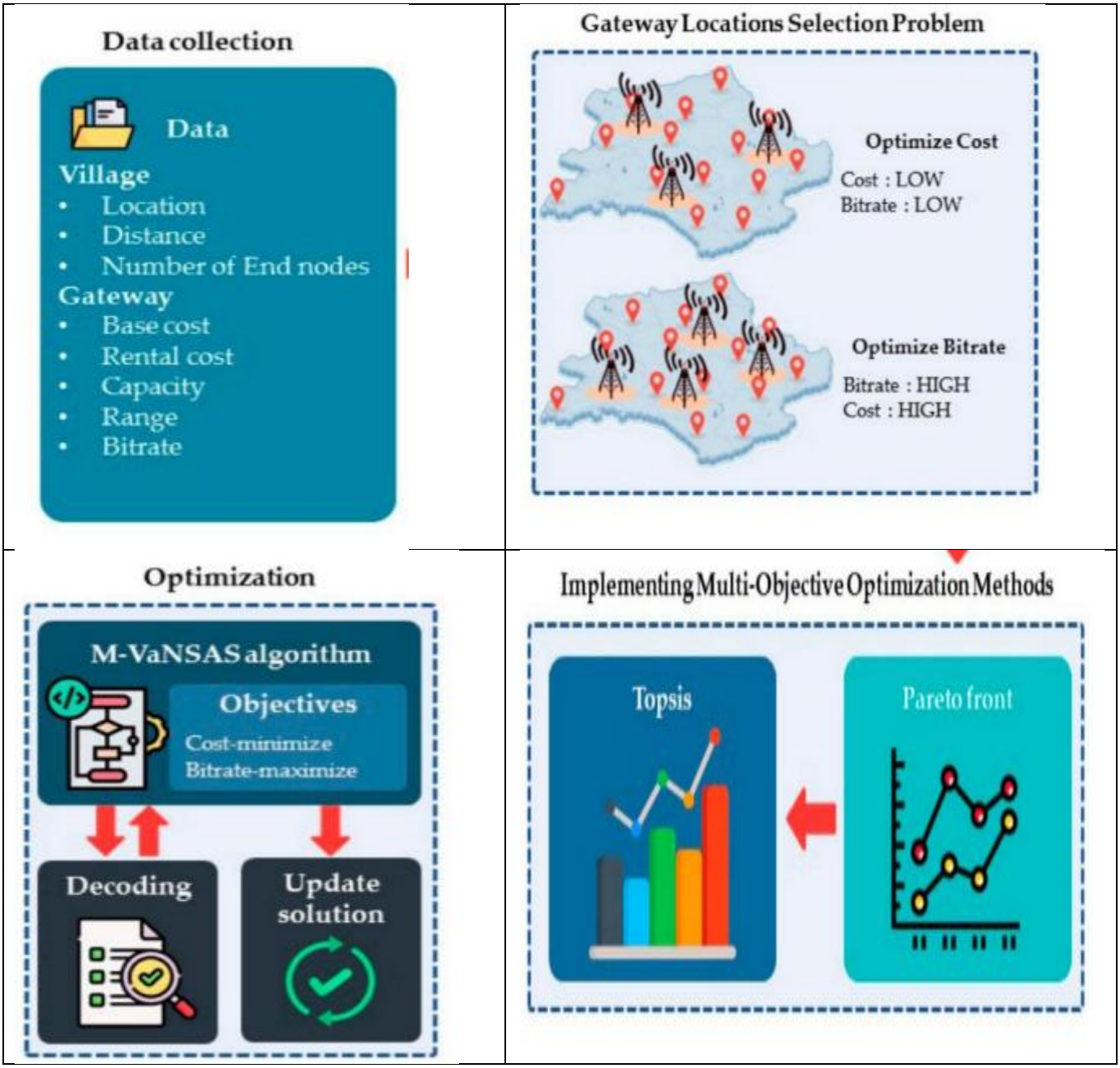


Fig. 2. Flowchart diagram

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 4

PICTORIAL REPRESENTATION:



Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 4

APPLICATIONS:

6.1 When it comes to expressing sets, BKCM have a significant spatial advantage over other data structures.

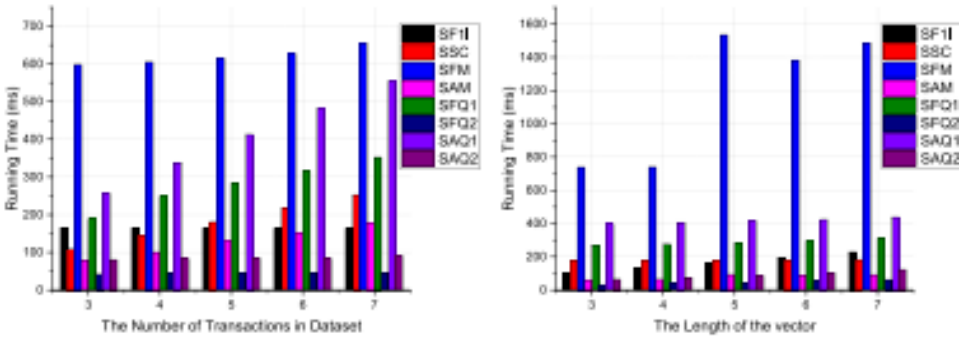
6.2 Before making a recommendation, BKCM filtering looks for user similarities between users based on their ratings

6.3 For the ideal suggestion, these techniques merely need to examine the products and user profile

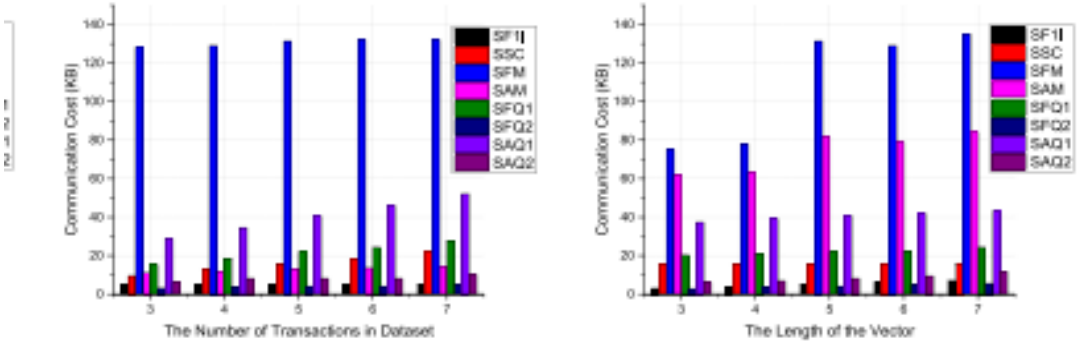
6.4 Improved Transparency: The content-based approach may inform you of the goods it recommends based on what attributes, whereas the collaborative technique provides you the suggestion because some anonymous individuals share your tastes

6.5 No cold start: in contrast to collaborative filtering, new things may be proposed before receiving a sizable user rating.

SAMPLE OUTPUT:



(a) Runtime ($n = 5$) (b) Runtime ($m = 5$)



(c) Communication Cost ($n = 5$) (d) Communication Cost($m = 5$)

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 4

CONCLUSION:

The adoption of a multithreaded approach for association rule mining presents a promising avenue for optimizing both privacy and efficiency.Through thorough testing, including unit, integration, system, and acceptance testing, the viability and effectiveness of this approach can be validated.By combining these testing methodologies, organizations can confidently deploy a multithreaded association rule mining system that not only maximizes efficiency but also prioritizes the protection of sensitive data, fostering trust and compliance in an increasingly data-driven landscape.

References:

[1] W. Wu, J. Liu, H. Wang, J. Hao, and M. Xian, "Secure and efficient outsourced k-means clustering using fully homomorphic encryption with ciphertext packing technique," IEEE Trans. Knowl. Data Eng., vol. 33, no. 10, pp. 3424– 3437, Oct. 2021.

[2] Wang, M. Li, and L. Xiong, "Fastgeo: Efficient geometric range queries on encrypted spatial data," IEEE Transactions on Dependableand Secure Computing, vol. 16, no. 2, pp. 245–258, 2019.

[3] Y. Chen, Q. Zhao, P. Duan, B. Zhang, Z. Hong, and B. Wang, "Veri f iable privacypreserving association rule mining using distributed de cryption mechanism on the cloud," Expert Syst. Appl., vol. 201, 2022, Art. no. 117086.

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)

Volume7-Issue2, January 2024
Paper: 4