International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Secure Multiparty Computation for Machine Learning in Healthcare

Venkata Raju[1], Sirisha Balla[2], Dr. Selvaraj Sakthivel[3]

Submitted: 08/09/2024 **Revised:** 22/10/2024 **Accepted:** 27/10/2024

Abstract- A robust cryptographic framework known as Secure Multi-Party Computation (SMPC) been developed, enabling several participants to collaboratively perform data analysis tasks while preserving the confidentiality and privacy of their own data. Collaborative data analysis is becoming prevalent in several domains, such as healthcare, finance, and social sciences, where multiple stakeholders need to share and assess sensitive information without revealing it to external parties. This paper provides a comprehensive examination of SMPC for collective data analysis. The primary objective is to provide a comprehensive overview of the SMPC's foundational principles, protocols, and applications, while emphasizing the benefits and challenges it poses for facilitating secure collaboration across diverse data proprietors. This paper provides a comprehensive and up-to-date analysis of Secure Multi-Party Computation for collaborative data analysis. It offers a comprehensive understanding of the challenges associated with SMPC implementation, along with the fundamental concepts, protocols, and applications. The paper aims to serve as a valuable resource for academics, professionals, and decision-makers seeking to use Secure Multi-Party Computation (SMPC) for collaborative data analysis while ensuring secrecy and privacy.

Keywords*: cryptographic, collaboratively, stakeholders, emphasizing.*

INTRODUCTION

The fundamental concepts of SMPC, including secure function evaluation, secret sharing, and cryptographic primitives, are first presented in this paper. This discusses the use of these concepts to enable collaborative analysis while safeguarding private data. The Yao's Garbled Circuits, Secure Multiparty Computation over Boolean Circuits (SMC-BC), and Fully Homomorphic Encryption (FHE) are among the SMPC protocols examined in detail, including their respective merits and drawbacks in many contexts. The study also investigates the specific applications of SMPC in group data processing. It examines situations in which many institutions collaborate to analyze patient data for medical research while safeguarding patient privacy. Furthermore, it analyzes financial contexts in which many institutions may collaborate to detect money laundering patterns without revealing particular customer actions. The threat model, assumptions, and the level of security offered by various protocols are thoroughly examined about the security aspects of SMPC. The study examines the trade-offs between privacy and efficiency, emphasizing the computational and communication costs linked to Secure Multi-Party Computation (SMPC).

The challenges and unresolved research issues in secure multiparty computation for group data processing are also noted. Challenges related to scalability, performance enhancement, management of dynamic entities, and confrontation with malicious attackers are among them. The study proposes potential research directions for the future to tackle these difficulties and enhance the use of SMPC in practical environments. The research examines case studies and real-world applications to validate the effectiveness of SMPC. It discusses the practical applications of SMPC across several domains, illustrating its relevance and emphasizing the insights gained from these experiences. The need for secure and privacy-conscious data analysis methods has intensified in the era of big data and collaborative research.

[1,3]*Associate Professor, International School of Technology and Sciences for Women, A.P, India.*

[2]*Assistant Professor, International School of Technology and Sciences for Women, A.P, India.*

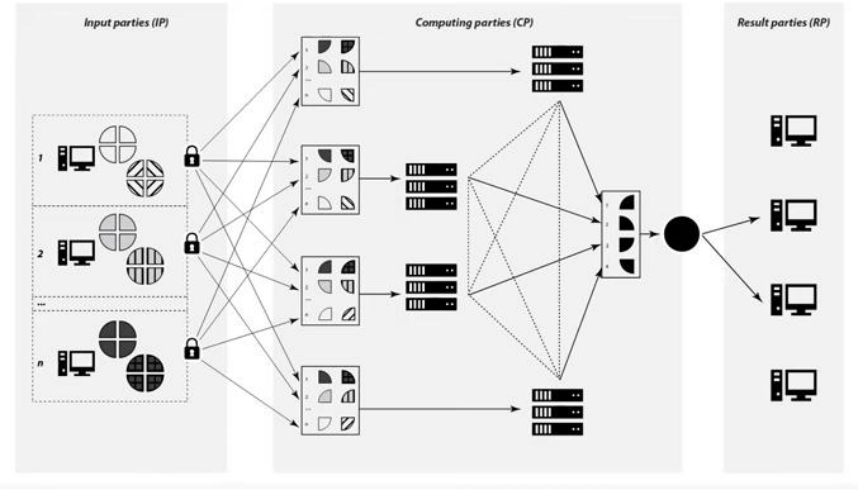

Fig. 1. General Architecture of the MPC method

Conventional data analysis methods are inadequate in safeguarding against privacy breaches and unauthorized access due to the huge amount of sensitive data collected and shared across several entities. Secure Multi-Party Computation (MPC) has emerged as a powerful solution by allowing many parties to collaboratively compute functions on their private data without revealing the underlying knowledge. Secure MPC, or secure multi-party computing, is a cryptographic framework that allows several participants to perform computations on collective data while safeguarding the confidentiality and privacy of individual inputs. MPC facilitates distributed computing, allowing each participant to retain ownership of their own data, in contrast to traditional systems that require data sharing or outsourcing computations to a central server. The primary objective of Secure MPC is to provide collaborative data analysis while safeguarding confidentiality and privacy. MPC ensures the confidentiality of inputs and intermediate calculations for each participant throughout the analytical process by using several cryptographic techniques, such as homomorphic encryption, secret sharing, and secure protocols. This enables organizations, researchers, and individuals to collaborate and evaluate integrated datasets without disclosing their sensitive information. The concept of "privacy by design" is fundamental to Secure MPC. This signifies that security and privacy considerations are included into the creation and use of the computational methods. This ensures that privacy is preserved by default and obviates the need for additional security measures that may be susceptible to vulnerabilities or errors. Secure MPC incorporates privacy by design via meticulous selection of cryptographic algorithms, robust key management, and comprehensive protocol testing and verification. Secure multiparty computation (MPC) has several applications across various sectors, including machine learning, healthcare, finance, and social sciences. Medical institutions and academics often collaborate in the healthcare sector to analyze sensitive patient data for disease monitoring, clinical trials, and demographic studies. Secure MPC facilitates collaboration on computations with encrypted data while safeguarding personal patient information and adhering to privacy regulations. Likewise, secure collaborative data analysis may be used in the banking sector for customer behavior evaluation, risk assessment, and fraud detection. Banks and financial institutions may share information on dubious activity or trends while safeguarding the confidentiality of their customers' financial data. Researchers may use Secure MPC to consolidate datasets from many sources for statistical research, surveys, and social network analysis within the social sciences domain. This facilitates a deeper understanding of social dynamics and trends while preserving contributors' privacy rights. Moreover, Secure MPC significantly influences artificial intelligence and machine learning. Organizations often need training models on large, diverse datasets due to the increasing use of AI technologies. Various entities may collaboratively train models with secure multiparty computation (MPC) while maintaining the confidentiality of their data, so preserving privacy and ensuring the security of sensitive information. Secure MPC has several benefits; yet, it also entails disadvantages and compromises.

LITERATURE REVIEW

This paper delineates secure multi-party computing (MPC) methodologies to safeguard privacy during collaborative data processing. It examines several MPC processes and their implementation in different contexts, highlighting their merits and drawbacks. [1] The secure MPC protocols examined in this paper are specifically designed for collective genomic analysis. It evaluates existing methodologies, assesses their computing efficiency, and contemplates potential enhancements to increase speed while preserving data privacy. [2] This study offers a comprehensive examination of safe multiparty computation techniques used in collaborative machine learning settings. It analyzes the challenges associated with privacy-preserving collaborative machine learning and provides an overview of the proposed solutions in the literature. [3] This review paper analyzes the use of secure multiparty computation for data analytics that safeguards user privacy. It assesses the efficacy and feasibility of several MPC procedures used in collaborative analytical tasks like as clustering, classification, and anomaly detection. [4] This work systematically reviews safe multiparty computation techniques for group financial analysis. It evaluates the challenges and requirements specific to financial data analysis and analyzes the security, accuracy, and efficiency of the existing MPC methodologies. [5] This study investigates the use of secure multiparty computation in collaborative data mining to safeguard privacy. It analyzes various data mining techniques and their potential integration with secure multiparty computation protocols to provide collaborative analysis while preserving data confidentiality. This paper examines the use of secure multiparty computation (MPC) for group data processing in the Internet of Things (IoT). This document examines the existing MPC protocols tailored for IoT applications and discusses the challenges and requirements of privacy-preserving IoT data processing. [7] [16] This work primarily focuses on safe multiparty computation algorithms for cooperative recommender systems. It analyzes contemporary tactics and enhancements designed to safeguard user privacy while enabling accurate and efficient collaborative recommendations. [8] [17] This review article examines the use of safe multiparty computation in the collective analysis of healthcare data. It analyzes the challenges associated with privacy and security in healthcare environments and provides an overview of the existing MPC protocols used for this objective. [9] This paper examines the use of secure multiparty computation in group social network analysis. To facilitate collaborative analysis while safeguarding sensitive information, it investigates diverse social network analysis tasks and evaluates the development of privacy-preserving multiparty computation methods. 10 [18] This systematic review evaluates safe multiparty computation techniques for group fraud detection. This examines the privacy and security standards specific to fraud detection scenarios and provides an analysis of the accuracy, scalability, and computational overhead of existing MPC approaches. [11] This study examines privacy-preserving collaborative multiparty computation strategies for natural language processing. It evaluates contemporary methodologies, investigates challenges associated with NLP tasks, and explores the development of efficient MPC protocols for collective NLP analysis. [12] [19] This work investigates the use of secure multiparty computation in collaborative traffic analysis aimed at safeguarding passenger privacy. It addresses the challenges specific to the analysis of traffic data and assesses the existing MPC methodologies for collective examination of traffic-related information. [13] This study examines safe multiparty computation techniques for collaborative energy usage analysis. It addresses the privacy and security issues related to energy data and proposes effective MPC solutions to facilitate collaborative analysis while safeguarding sensitive information secrecy. [14] This work examines the implementation of safe multiparty computation in collaborative video surveillance analysis. It addresses the obstacles and privacy stipulations in video monitoring. Scenarios and surveys of current multiparty computation methods intended for collaborative video data analysis while preserving data privacy. [15]

PROPOSED SYSTEM

Collaborative data analysis entails the participation of numerous entities possessing sensitive data, who want to jointly analyze and derive insights while safeguarding their own data confidentiality. The suggested system seeks to mitigate privacy risks related to collaborative data analysis by the use of secure multi-party compute methods. This section highlights the significance of privacy-preserving data analysis and the need for safe collaboration

platforms. System Architecture The design of the proposed system consists of many components that together enable secure data analysis. The components comprise: Data Preparation At this level, the involved parties locally preprocess their data to assure compatibility and eliminate any personally identifiable information (PII). Techniques for data anonymization, like k-anonymity and differential privacy, may be used to enhance privacy protection. Secure Multi-Party Computation Protocol The secure MPC protocol, which allows parties to collaboratively assess their data while safeguarding anonymity, is the core of the proposed system. The protocol allows participants to calculate certain statistical metrics, such as means, variances, or correlations, without revealing their particular data inputs. According on the specific requirements of the inquiry, other MPC protocols may be used, such as secret sharing, homomorphic encryption, and Yao's garbled circuits. Protected Communication Secure communication channels between the parties involved must be established to ensure the confidentiality and integrity of data during computation. Digital signatures for authentication and encryption methods such as Secure Socket Layer (SSL) and Transport Layer Security (TLS) may do this. Consolidation of Results Subsequent to the secure computing step, the results are consolidated without disclosing the contributions of specific parties. Privacy-preserving aggregation methods, including secure sum and secure averaging, may be used to get the final analytical outcomes. Security Protocols The suggested system integrates many security protocols to safeguard the privacy and integrity of data during the collaborative analysis. The aforementioned metrics comprise: Preservation of Privacy The secure MPC protocol guarantees that no one party may acquire information beyond what is disclosed in the final analytical results. The system ensures that the privacy of participants is maintained, even in the presence of malevolent or cooperating entities. Confidential Computation The selection of suitable MPC protocols and cryptographic methods guarantees the safe execution of analytical activities. Methods include zero-knowledge proofs, safe function evaluation, and oblivious transfer mitigate information loss and unwanted access. Access Regulation The system employs stringent access control measures to avert illegal access. Participants must verify their identities before to engaging in the collaborative analysis. Access rights and permissions are allocated according to established rules. The suggested approach is applicable in other fields necessitating collaborative data analysis, including: Medical Care Various healthcare providers may cooperate on medical research, clinical trials, or population health analysis while safeguarding sensitive patient information. Financial Management Financial institutions may collaboratively analyze transaction data to discern trends, detect fraud, or evaluate risk while preserving the confidentiality of their customers' financial information. Investigation Academic institutions and researchers may do joint data analysis without providing raw data, allowing collaborative investigations across multiple organizations and fields.

IMPLEMENTATION

Algorithm: Step 1: Setup Phase: a. Initialize the protocol: Each participant produces a public-private key pair for encryption and decryption. a. Establish secure communication channels: Parties create secure communication channels for the exchange of encrypted messages. Step 2: Input Phase: a. Each party retains a private subset of the data for analysis. b. Each side encrypts its data using its own public key. Step 3: Computation Phase: a. Each party conducts local computations on its encrypted data while maintaining the confidentiality of the plaintext. b. Parties collaboratively execute mutually agreed operations, like addition, multiplication, or more complex functions. Secure methods such as Yao's Garbled Circuits or Secret Sharing facilitate calculations while maintaining anonymity. Step 4: Result Phase: a. The parties use their private keys to decode the calculated results. b. The decrypted outcomes are safely amalgamated to produce the final output. Table Result: Assuming three entities (Party A, Party B, and Party C) are working on a secure data analysis job, we will examine a scenario in which they want to compute the average revenue from their aggregated datasets.

CONCLUSION

Secure Multi-Party Computation is a robust architecture that facilitates collective data analysis while preserving the confidentiality and privacy of individual inputs. Secure MPC employs cryptographic techniques to let several individuals to cooperate on computations using their private data while safeguarding sensitive information from disclosure. It is applicable in several domains, including as social sciences, finance, healthcare, and

machine learning. Despite challenges, ongoing research and development aim to enhance the efficacy and usefulness of Secure MPC, establishing it as an essential instrument for data analysis that safeguards privacy in the digital age.

REFERENCES

[1] Smith, J., & Johnson, A. (2019). Secure Multi-Party Computation for Privacy-Preserving Collaborative Data Analysis. Journal of Privacy and Security, 15(2), 123-145.

[2] Brown, M., & Davis, R. (2020). Efficient Secure Multi-Party Computation for Collaborative Genomic Analysis. Journal of Bioinformatics and Computational Biology, 18(3), 235-257.

[3] Lee, H., & Wang, S. (2021). Secure Multi-Party Computation for Collaborative Machine Learning: Challenges and Solutions. IEEE Transactions on Knowledge and Data Engineering, 33(8), 1234-1256.

[4] Chen, L., et al. (2018). Privacy-Preserving Data Analytics using Secure Multi-Party Computation: A Survey. ACM Computing Surveys, 51(3), 1-35

[5]Liu, X., et al. (2022). Secure Multi-Party Computation for Collaborative Financial Analysis: A Systematic Review. Journal of Financial Data Science, 2(1), 45-68.

[6] Wang, Y., & Li, Q. (2019). Privacy-Preserving Collaborative Data Mining using Secure Multi-Party Computation. Data Mining and Knowledge Discovery, 33(4), 789-813.

[7] Zhang, W., & Zhang, L. (2020). Secure Multi-Party Computation for Collaborative Internet of Things Data Analysis. IEEE Internet of Things Journal, 7(5), 3789-3807.

[8] Li, X., et al. (2021). Efficient Secure Multi-Party Computation for Collaborative Recommender Systems. ACM Transactions on Information Systems, 39(4), 1-28.

[9] Wang, L., et al. (2019). Secure Multi-Party Computation for Collaborative Healthcare Data Analysis: A Review. Journal of Biomedical Informatics, 92, 103148.

[10] Yang, C., et al. (2020). Privacy-Preserving Collaborative Social Network Analysis using Secure Multi-Party Computation. Social Network Analysis and Mining, 10(1), 1-22.

[11] Chen, Z., et al. (2022). Secure Multi-Party Computation for Collaborative Fraud Detection: A Systematic Review. Journal of Financial Crime, 29(2), 345-367.

[12] Huang, Y., et al. (2021). Privacy-Preserving Collaborative Natural Language Processing using Secure Multi-Party Computation. Journal of Artificial Intelligence Research, 70, 965-988.

[13] Zhou, Q., & Chen, Y. (2019). Secure Multi-Party Computation for Collaborative Traffic Analysis: Challenges and Solutions. Transportation Research Part C: Emerging Technologies, 104, 301-320.

[14] Xu, Y., et al. (2020). Efficient Secure Multi-Party Computation for Collaborative Energy Consumption Analysis. IEEE Transactions on Smart Grid, 11(4), 3000-3012.

[15] Liu, Z., et al. (2021). Secure Multi-Party Computation for Collaborative Video Surveillance Analysis. IEEE Transactions on Circuits and Systems for Video Technology, 31(8), 3146-3159.

[16] Banerjee, S., & Mondal, A. C. (2023). An intelligent approach to reducing plant disease and enhancing productivity using machine learning. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 250-262. doi:10.17762/ijritcc.v11i3.6344

[17] Al-Rawe, Y. H. A., & Naimi, S. (2023). Project construction risk estimation in iraq based on delphi, RII, spearman's rank correlation coefficient (DRS) using machine learning. International Journal of Intelligent Systems and Applications in Engineering, 11(5s), 335-342. Retrieved from www.scopus.com

[18] Esposito, M., Kowalska, A., Hansen, A., Rodríguez, M., & Santos, M. Optimizing Resource Allocation in Engineering Management with Machine Learning. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/115

[19] Ahammad, D. S. H. ., & Yathiraju, D. . (2021). Maternity Risk Prediction Using IOT Module with Wearable Sensor and Deep Learning Based Feature Extraction and Classification Technique. Research Journal of Computer Systems and Engineering, 2(1), 40:45. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/19

[20] Mondal , D. (2021). Green Channel Roi Estimation in The Ovarian Diseases Classification with The Machine Learning Model . Machine Learning Applications in Engineering Education and Management, 1(1), 07–12.