BLOCKCHAIN-BASED FEDERATED LEARNING WITH SMPC MODEL VERIFICATION AGAINST POISONING ATTACK FOR HEALTHCARE SYSTEMS

[1]DR. B.SARADA, [2]PATLOLLA VARSHA, [3]NALLURI SINDHU, [4]MEDURI UMA

[1]Associate Professor, Department of School of Computer Science & Engineering, **MALLAREDDY ENGINEERING COLLEGE FOR WOMEN**, Maisammaguda, Dhulapally Kompally, Medchal Rd, M, Secunderabad, Telangana

[2,3,4]Student, Department of School of Computer Science & Engineering, **MALLAREDDY ENGINEERING COLLEGE FOR WOMEN**, Maisammaguda, Dhulapally Kompally, Medchal Rd, M, Secunderabad, Telangana..

ABSTRACT

Due to the rising awareness of privacy and security in machine learning applications, federated learning (FL) has received widespread attention and applied to several areas, e.g., intelligence healthcare systems, IoT-based industries, and smart cities. FL enables clients to train a global model collaboratively without accessing their local training data. However, the current FL schemes are vulnerable to adversarial attacks. Its architecture makes detecting and defending against malicious model updates difficult. In addition, most recent studies to detect FL from malicious updates while maintaining the model's privacy have not been sufficiently explored. This paper proposed blockchain based federated learning with SMPC model verification against poisoning attacks for healthcare systems. First, we check the machine learning model from the FL participants through an encrypted inference process and remove the compromised model. Once the participants' local models have been verified, the models are sent to the blockchain node to be securely aggregated. We conducted several experiments with different medical datasets to evaluate our proposed framework.

I.INTRODUCTION

The Internet of Things (IOT) has been applied in various services, including the healthcare domain. The integration of IOT in the healthcare system is also known as the Internet of Medical Things (IOMT). With the development of IOMT, many healthcare devices are interconnected, allowing devices to exchange information among medical experts and Artificial Intelligence (AI) based services. This interconnectivity helps healthcare industries like hospitals to improve the efficiency and quality of their services. In the medical diagnosis field, medical imaging devices facilitate the process of early diagnosis and treatment for medical staff.

Due to this interconnectivity, medical image retrieval is made easy, resulting in extensive data with wide variations. Consequently, medical image analysis has become a challenging task for medical experts and is prone to human error. In recent years, the success of Deep Learning (DL) in computer vision tasks has provided a significant breakthrough in medical image classification tasks. Several studies of DL in medical imaging fields have shown promising results by providing accurate and efficient diagnoses [1].

As shown in Figure 1, cloud computing is one paradigm that emerged to solve the availability of computing and storage resources. Therefore, the cloud is usually used to deploy the DL model for training and data inference. However, sending the raw data from the IOMT cluster to the cloud will be very expensive. This is where edge computing, like edge servers, will be advantageous to process the data before sending it to the cloud.

It is known that a high-performing Deep Learning (DL) model requires a large and diverse dataset for its training. This large-scale dataset is often obtained from multi-institutional or multi-national data accumulation and voluntary data sharing in the healthcare industry. While massive data collection is essential for the deep learning process, sharing patients' data raises privacy concerns and relative regulations such as the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA). Due to the rising concerns, healthcare institutions may be prevented from sharing their medical datasets. In some cases where sharing is possible, some restrictions are applied, resulting in inadequate data sharing.

In recent studies, [2] proposed a federated learning model that allows parties to collaboratively train a model by sharing local model updates with a parameter server. Intuitively, this method is safer than centralized training because machine learning models learn from healthcare IOMT data without relying on a third-party cloud to hold their data [3]. However, federated learning also presents some challenges that may limit its applications in real-world case scenarios. For example, federated learning remains vulnerable to various attacks that may result in leakage of private data [4] or poisoned learning model [5]. Also, the participants in the current FL setup cannot verify the authenticity of the machine learning model. To protect FL participants' privacy, the existing defense method mainly focuses on ensuring the confidentiality of the machine learning gradients. Differential Privacy (DP) [6], [7] is one of the commonly used methods to preserve the privacy of the learning model. Adding DP to a federated learning scenario can improve the privacy of the participants models. However, adding noise into machine learning gradients will reduce the

learning model accuracy [7]. DP is also ineffective in mitigating poisoning attacks while maintaining model performance resulting in a faulty global model. To tackle the poisoning attack, existing research on anomaly detection [8],[9] has been explored. However, the existing methods cannot eliminate all the poisoned models and cause the accuracy of the global model to be reduced. Also, they perform the anomaly detection method in a plaintext model. This will lead to another issue where the attacker can perform a parameter stealing attack [10] and a membership inference attack [11]. Thus, a verifiable and secure anomaly detection method for federated learning scenarios is needed.

This paper proposes a privacy-preserving verification method to eliminate poisoned local models in a federated learning scenario. The proposed method eliminates the compromised local model while guaranteeing the privacy of the local model's parameters using an SMPC-based encrypted inference process. Once the local model is verified, the verified share of the local model is sent to the blockchain for the aggregation process. SMPC-based aggregation is used to perform the secure aggregation between the blockchain and the hospital. After the aggregation process, the global model is stored in tampered-proof storage. Later, each hospital receives the global model from the blockchain and verifies the authenticity of the global model. The contributions of our work are summarized as follows:

_ Propose a new block chain-based federated learning architecture for healthcare systems to ensure the security of the global model used for classifying disease.

_ Design a privacy-preserving method for local model anomaly detection in a Federated learning scenario with SMPC as the underlying technology. Our encrypted model verification method eliminates the poisoned model while protecting the local model privacy from membership inference attacks and parameter stealing.

_ Propose an SMPC-based secure aggregation in the block chain as a platform to decentralize the aggregation process.

_ We present a verifiable machine learning model for federated learning participants using block chain in the IOMT scenario.

The rest of this paper is organized as follows. Section II defines the problem and design goals. Section III discusses the related work. Then, we present the system architecture and introduce the proposed frameworks in Section IV. Next, we describe the experimental setup and evaluation results of the proposed work in Section V. Finally, a conclusion is drawn in Section VII.

II. PROBLEM SCENARIO AND DESIGN GOALS

A. Problem Scenario

To discuss and highlight the current issues with current federated learning, we use an IOMT-enabled hospital scenario (see Fig. 2). Assume that several smart hospitals are placed in different regions with varying patient demographics and diseases. Each smart hospital is equipped with a cluster of IOMT devices. The IOMT devices will be used to scan the patient to detect a severe disease. In The current IOMT scenario, IOMT devices will act as data sources since the IOMT devices are resource-constrained and cannot perform any machine learning algorithm. Hence, each hospital has an edge server with computing resources to execute the machine learning tasks using the local datasets. Nevertheless, due to dataset limitations, the machine learning model accuracy generated from the local datasets is relatively low. Therefore the edge server from each hospital participates in the federated learning platform. In the federated learning platform, locally trained models from the hospital's edge server are collected and aggregated to produce a highly accurate machine learning model without sending private datasets to the cloud provider. Later, the aggregated or global model is sent back to the edge server for another round of federated learning processes. Once the global model reaches the desired accuracy, it will be used to recognize the disease more accurately.

Although the aforementioned federated learning scenario improves the overall machine learning accuracy, it suffers from the following security risks:

_ Risks of local model security: In the current setup of federated learning, every party that sends their local model is sent to the cloud for the aggregation process without checking the model's validity. This traditional FL method introduces the risk of a local model being poisoned. For example, an attacker can perform a poisoning attack and train the model using poisoned data, leading to a faulty local model. Since healthcare data are critical, sending plaintext local models to the cloud can pose privacy risks. Therefore, validating and securing the local model is required to prevent it from various security aspects.

_ Risks of generating a biased aggregated model: The model aggregation process of the local model is performed on the cloud services that can be tampered with and produce a biased global model. For example, an attacker can include a poisoned local model during the aggregation process that may lead the global model to have a false classification. Hence, a secure aggregation method is required to encounter the current security problem.

_ Risk of receiving faulty global model: In the existing federated learning method, the global model generated from the cloud will be sent back to each edge server in the hospitals. However, the hospital cannot verify the global model they received. The attacker can intercept

and alter the global model. As a result, the hospital received a faulty global model. From this problem, a global model verification method is required to ensure the integrity of the global model.

B. Design Goals With the risks and threats mentioned above, our goals for preserving privacy in Federated Learning can be decomposed into three aspects as follows:

_ Robustness: The proposed work should have the ability to prevent the adversary from poisoning federated learning. This allows the federated learning participant to learn from a benign global model to improve their model accuracy. Also, a robust aggregation method needs to be developed to secure the aggregation process from an attacker.

_ Privacy: The prior work [12] has shown that an attacker can perform a poisoning attack to decrease the global model accuracy by miss-classifying the machine learning model.To protect the federated learning participants, checking the participant's local learning model while maintaining the local model privacy itself is essential.

_ Verifiability: The designed method should have the ability to verify the machine learning model, specifically the global model. Since the adversary may alter or poison the global model. In the current federated learning scenario, the participant received the global model from the cloud without knowing the model's authenticity.

III.EXISTING SYSTEM

In FL, data privacy is achieved by sending the model to the client and performing local training. Later, the locally trained model will be collected by the central server and aggregated into a global model. With this method, the participants only shared the local model and did not send any datasets. However, FL itself is not sufficient to provide a privacy guarantee. Some research has been performed to secure the FL architecture. The author in [6] and [7] enhance the data privacy in FL with differential privacy (DP) by adding noise in the local datasets. In [7], also anonymize the end-user by adding a proxy server. However, the experiment result show there is a significant accuracy reduction. This privacy-preserving method is unsuitable for FL in healthcare systems since accuracy is essential for the inference process.

Zhang et al. [13] use fully homomorphic encryption (FHE) to perform aggregation and training processes by performing a batch encryption method. However, all the homomorphic encryption methods are unusable for healthcare scenarios since the training process takes significant time. Authors in [14], [15], and [16] have successfully performed an adversarial attack on FL architecture. The authors have demonstrated a poisoning attack on the local client's datasets. The poisoned model will be generated and impact the global

model. Based on the existing attack, DP and FHE method is insufficient against the poisoning attack. In [17], the author proposed a privacy-enhanced FL against poisoning adversaries. To secure the machine learning model, they encrypt the model using linear homomorphic encryption. Since they encrypt the model from the first round of FL, the training process will take longer than regular machine learning. After the participants finish the encrypted training process, The local model will send to the server for encrypted aggregation. Based on the results of their experiments, their aggregation method reduces the accuracy of the machine learning model.

Blockchain is known for its immutability and is used for tampered-proof storage. The use of blockchain can track the local or global model for audibility purposes. Combining blockchain with FL can ensure the machine learning model's integrity. Author in [18] proposed verifiable aggregation for FL. Their method follows the concept of blockchain, where they use the hash to compute the digest for verification. Nonetheless, the aggregation and hashing process is performed on a single server. The correct utilization of blockchain technology can overcome the problem. In tackling the issue, [19] proposed decentralized privacy using blockchain-enabled FL. They use blockchain to store and verify the model using cross-validation, but the participant is connected to the same blockchain. In their framework, the participant can use other's local models, which leads to privacy issues.

The work on [20] uses a smart contract to verify the global model. The use of smart contracts can audit the authenticity of the global model. However, they did not perform any checks on the local or global model. Also, the local model is not sent to the blockchain, and not possible to perform any audit process. From the proposed work, they can not handle any poisoning attack.

Disadvantages

- The system didn't implement a verifiable Federated Learning (FL) scenario that leverages SMPC to perform an encrypted local model verification process and secure aggregation on the blockchain node.

- The federated learning scenario not allows each participant to collaboratively train the machine learning model locally with their local datasets. Later the machine learning model will send to the cloud for the model aggregation process.

IV.PROPOSED SYSTEM

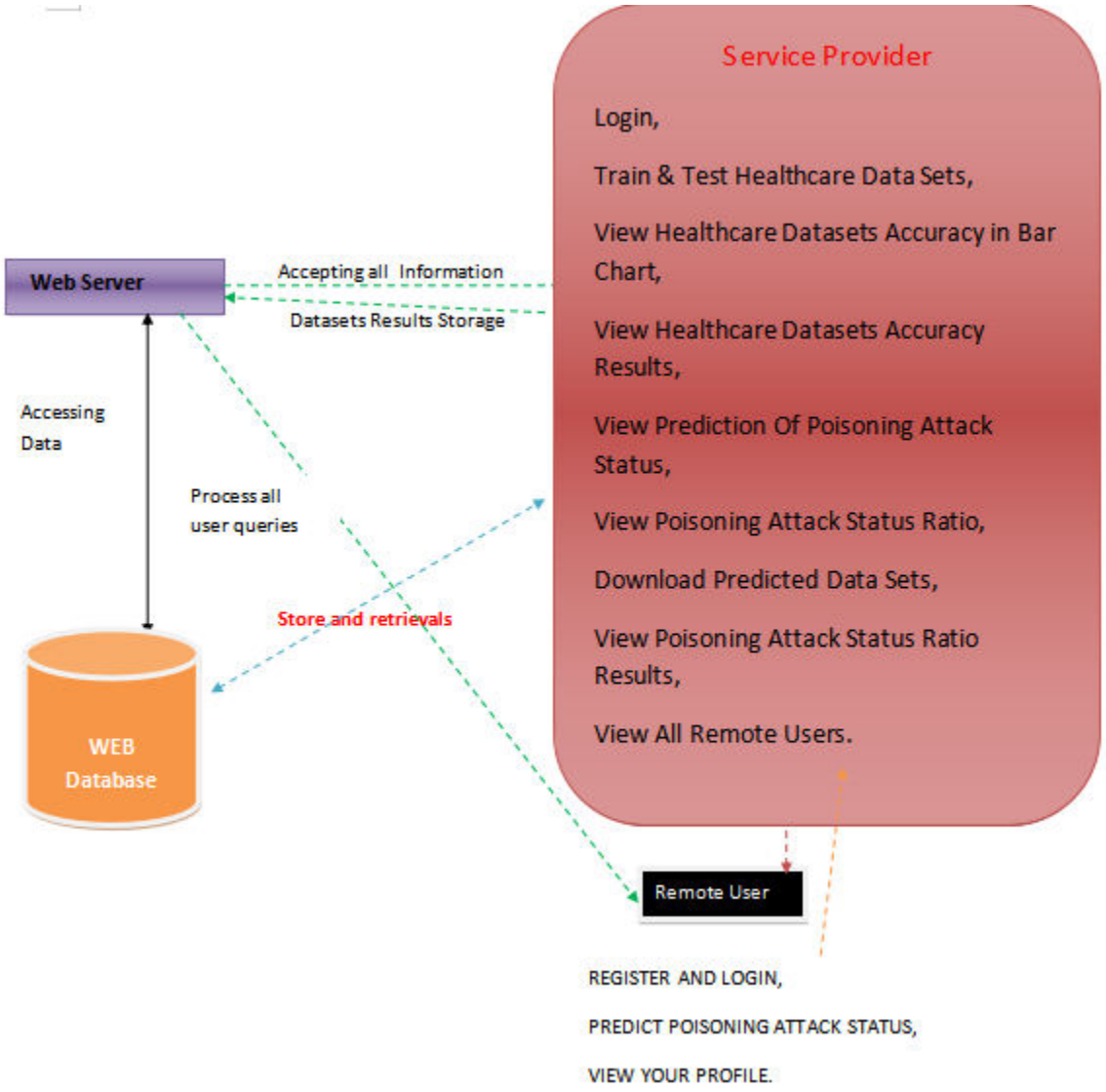
The system proposes a privacy-preserving verification method to eliminate poisoned local models in a federated learning scenario. The proposed method eliminates the compromised local model while

guaranteeing the privacy of the local model's parameters using an SMPC-based encrypted inference process. Once the local model is verified, the verified share of the local model is sent to the blockchain for the aggregation process. SMPC-based aggregation is used to perform the secure aggregation between the blockchain and the hospital. After the aggregation process, the global model is stored in tampered-proof storage. Later, each hospital receives the global model from the blockchain and verifies the authenticity of the global model

Advantages

- Propose a new blockchain-based federated learning architecture for healthcare systems to ensure the security of the global model used for classifying disease.

  Design a privacy-preserving method for local model anomaly detection in a Federated learning scenario with SMPC as the underlying technology. Our encrypted model verification method eliminates the poisoned model while protecting the local model privacy from membership inference attacks and parameter stealing.
- Propose an SMPC-based secure aggregation in the blockchain as a platform to decentralize the aggregation process.
- We present a verifiable machine learning model for federated learning participants using blockchain in the IoMT scenario.

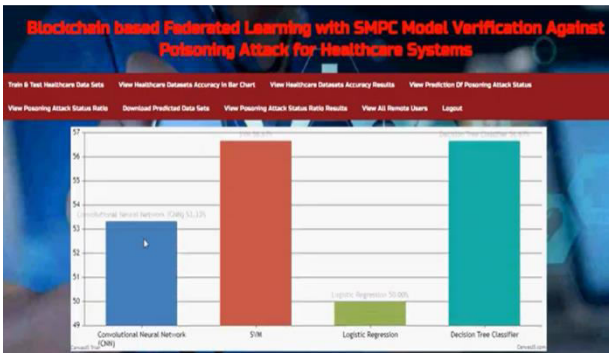


V. MODULES

Service Provider

In this module, the Service Provider has to login by using valid user name and password.

After login successful he can do some operations such as Train & Test Healthcare Data Sets, View Healthcare Datasets Accuracy in Bar Chart,



View Healthcare Datasets Accuracy Results,

View Prediction Of Poisoning Attack Status,View Poisoning Attack Status Ratio, Download Predicted Data Sets, View Poisoning Attack Status Ratio Results, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like register and login, predict poisoning attack status,

view your profile.

VI.CONCLUSION

This paper proposes block chain-based federated learning with a secure model verification for securing healthcare systems. The main objective is to ensure the local model is poisoned-free while maintaining privacy and providing verifiability for the federated learning participants.

In this framework, we perform a privacy-preserving verification process on the local model before the aggregation process. To preserve privacy on the local model, the verification is performed through an encrypted inference supported by SMPC protocol. This method allows the verifier to check the model with encrypted models and images. Once the local model is verified, the verified share of the local model is sent to the block chain node. Block chain and the hospital will perform SMPC-based secure aggregation. Once the majority of nodes have the same result, the global model is stored in the block chain. Later, the tamper-proof storage will distribute the updated global model to every hospital that joins the federated learning round.

In the experiment, we use Convolutional Neural Network (CNN) based algorithms with several medical datasets to generate local models and aggregate them under FL settings. Our experiment results show that the model encrypted verification process can eliminate all the participants' poisoned models while maintaining the privacy of the local model. In addition, we can recover up to 25% for the global model accuracy. It is essential to mention that our secure inference processing time is almost similar to the original inference process.

In the future, we plan to develop an efficient consensus mechanism for block chain-based aggregation. In this paper, we assume that all hospitals use the homogeneous model and use the same setup to generate their respective local models. However, we plan to broaden our work in the future to support a heterogeneous model in block chain-based federated learning.

VII.REFERENCES

[1] L. Sun, X. Jiang, H. Ren, and Y. Guo, "Edge-cloud computing and artificial intelligence in internet of medical things: Architecture, technology and application," IEEE Access, vol. 8, pp. 101 079–101 092, 2020. [2] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, "On the convergence of fedavg on non-iid data," arXiv preprint arXiv:1907.02189, 2019.

[3] Z. Yu, S. U. Amin, M. Alhussein, and Z. Lv, "Research on disease prediction based on improved deepfm and iomt," IEEE Access, vol. 9, pp. 39 043–39 054, 2021.

[4] W. Wei, L. Liu, M. Loper, K.-H. Chow, M. E. Gursoy, S. Truex, and Y. Wu, "A framework for evaluating client privacy leakages in federated learning," in European Symposium on Research in Computer Security. Springer, 2020, pp. 545–566.

[5] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," Future Generation Computer Systems, vol. 115, pp. 619–640, 2021.

[6] Y. Zhao, J. Zhao, M. Yang, T. Wang, N. Wang, L. Lyu, D. Niyato,and K.-Y. Lam, "Local differential privacy-based federated learning for internet of things," IEEE Internet of Things Journal, vol. 8, no. 11, pp. 8836–8853, 2021.

[7] B. Zhao, K. Fan, K. Yang, Z. Wang, H. Li, and Y. Yang, "Anonymous and privacy-preserving federated learning with industrial big data," IEEE Transactions on Industrial Informatics, vol. 17, no. 9, pp. 6314–6323, 2021.

[8] X. Wang, S. Garg, H. Lin, J. Hu, G. Kaddoum, M. Jalil Piran, and M. S. Hossain, "Toward accurate anomaly detection in industrial internet of things using hierarchical

federated learning," IEEE Internet of Things Journal, vol. 9, no. 10, pp. 7110–7119, 2022.

[9] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-learning-based anomaly detection for iot security attacks," IEEE Internet of Things Journal, vol. 9, no. 4, pp. 2545–2554, 2022.

[10] B. Wang and N. Z. Gong, "Stealing Hyperparameters in Machine Learning," in 2018 IEEE Symposium on Security and Privacy (SP), 2018, pp. 36–52.

[11] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning," in 2019 IEEE Symposium on Security and Privacy (SP), 2019, pp. 739–753.

[12] V. Tolpegin, S. Truex, M. E. Gursoy, and L. Liu, "Data poisoning attacks against federated learning systems," in European Symposium on Research in Computer Security. Springer, 2020, pp. 480–501.

[13] C. Zhang, S. Li, J. Xia, W. Wang, F. Yan, and Y. Liu, "Batchcrypt: Efficient homomorphic encryption for cross-silo federated learning," in 2020 fUSENIXg Annual Technical Conference (fUSENIXgfATCg 20), 2020, pp. 493–506.

[14] J. Zhang, B. Chen, X. Cheng, H. T. T. Binh, and S. Yu, "Poisongan: Generative poisoning attacks against federated learning in edge computing systems," IEEE Internet of Things Journal, vol. 8, no. 5, pp. 3310–3322, 2021.

[15] M. Fang, X. Cao, J. Jia, and N. Gong, "Local model poisoning attacks to byzantine-robust federated learning," in 29th fUSENIXg Security Symposium (fUSENIXg Security 20), 2020, pp. 1605–1622.

[16] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to backdoor federated learning," in International Conference on Artificial Intelligence and Statistics. PMLR, 2020, pp. 2938–2948.

[17] X. Liu, H. Li, G. Xu, Z. Chen, X. Huang, and R. Lu, "Privacy-enhanced federated learning against poisoning adversaries," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 4574–4588, 2021.

[18] X. Guo, Z. Liu, J. Li, J. Gao, B. Hou, C. Dong, and T. Baker, "Verifl: Communication-efficient and fast verifiable aggregation for federated learning," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1736–1751, 2021.

[19] Y. Qu, L. Gao, T. H. Luan, Y. Xiang, S. Yu, B. Li, and G. Zheng, "Decentralized privacy using blockchain-enabled federated learning in fog computing," IEEE Internet of Things Journal, vol. 7, no. 6, pp. 5171–5183, 2020.

[20] Z. Peng, J. Xu, X. Chu, S. Gao, Y. Yao, R. Gu, and Y. Tang, "Vfchain: Enabling verifiable and auditable federated learning via blockchain systems," IEEE Transactions on Network

Science and Engineering, vol. 9, no. 1, pp. 173–186, 2022.