ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2025, Volume 6 Issue 2
www.ijisea.org

Maintenance of Privacy and Security in Decentralized Federated Learning Using Blockchain Technology

M.Sreedevi[1] , CH.Mounika[2],CH.Pavan Manikanta Chari[3],G.Mahalakshmi[4],G.Sandeep[5]
Professor[1],UG Student[2,3,4,5]
Computer Science and Engineering
Amrita Sai Institute of Science & Technology
Paritala, Andhra Pradesh , India

ABSTRACT

Federated learning has been rapidly evolving and gaining popularity in recent years due to its privacy-preserving features, among other advantages. Nevertheless, the exchange of model updates and gradients in this architecture provides new attack surfaces for malicious users of the network which may jeopardize the model performance and user and data privacy. For this reason, one of the main motivations for decentralized federated learning is to eliminate server-related threats by removing the server from the network and compensating for it through technologies such as blockchain. However, this advantage comes - at the cost of challenging the system with new privacy threats.

Key words: Block chain , Federated learning , security ,decentralized federated Learning, Reviews

Abbreviations: FL

I.INTRODUCTION

With the widespread use of machine learning over the recent years, new concerns have been raised regarding user and data privacy. The data-driven nature of these intelligent models necessitates gathering users' data to constantly improve and maintain the operating statistical model. This issue becomes more problematic in large-scale distributed systems with millions of users such as mobile networks.

Aside from privacy issues, in large-scale systems, communicating user data may pose an overhead to the network. This is while information technology and intelligent devices are evolving at a rapid pace, and in the wake of it, there is an explosion of data at the edge of the network. Given the potential benefits of this data collection process for improving their model, organizations tend to make the best use of it for knowledge extraction with minimal waste of data. Thus, modern intelligent systems struggle to find the optimal trade-off between user privacy and service quality.

Inspired by recent breakthroughs in distributed optimization, Federated Learning (FL) has proposed as a potential solution to resolve the aforementioned challenges. In contrast to distributed machine learning, FL proposes training local models at the edges of the network and then sharing the model parameters to a central server that aggregates the received information and then updates all the client models.

Despite the breakthrough made by FL, the proposed architecture was not flawless and demanded further research endeavours on the topic. To begin with, although the user data is not being shared in the network, communicating local parameters is still vulnerable to sniffing attacks which will lead to stealing model parameters for launching an inference attack to extract sensitive information from the local training data at the edges of the networks.

The process of local update preparation is similar to that of FL; however, exchanging model parameters and model aggregation is mostly undertaken through P2P communication or blockchain technology.

Despite the security and efficiency features of DFL, this architecture is not flawless. For instance, incorporating blockchain into FL may come at the cost of making the system defenseless against blockchain-related security adversaries. The trustability of the DFL participants is another issue of concern that requires further research.

II.RELATED WORK

Here is a grammatically correct, well-structured, and professionally formatted version of your content:

Decentralized Learning: A Literature Survey

Federated Learning (FL) was proposed to reduce the risks associated with data ownership in collaborative training of deep learning models. Before the introduction of FL, collaborative training required massive data exchanges between participants in a distributed machine learning framework, where clients would frequently send local training data to a central server. The distributed model was trained on the accumulated data gathered from across the network.

However, this approach raised concerns about data privacy and excessive communication overhead, which motivated the invention of FL. In simple terms, FL allows each client to locally train a model of the same architecture using its own data, and only the resulting model parameters (e.g., neural network weights) are sent to the central server. The server then aggregates these parameters and updates the global model, which is subsequently distributed back to all clients. In each training round, clients initialize their local models with the updated global parameters before continuing with local training and submitting updates to the server.

Decentralized Learning in the Healthcare Domain

Decentralized learning is transforming the healthcare industry by addressing critical issues related to privacy, security, and data sharing. Key applications include:

1. Privacy-Preserving Collaboration:

Decentralized learning enables hospitals, research institutions, and healthcare providers to collaboratively train machine learning models without sharing sensitive patient data. This approach aligns with compliance requirements under regulations such as HIPAA and GDPR.

2. Medical Imaging:

Techniques like Federated Learning (FL) and Split Learning (SL) are applied to enhance diagnostic accuracy in areas such as cancer detection, X-ray interpretation, and MRI analysis.

3. Wearable Devices:

Wearable health devices can contribute to decentralized learning by performing local training. This enables health monitoring and predictive analytics while maintaining user data privacy, particularly for managing chronic conditions and early detection.

Decentralized Learning for Security and Privacy

Decentralized learning enhances privacy and security in collaborative machine learning settings through the following mechanisms:

1. Privacy Preservation:

Sensitive data remains on local devices, reducing the risk of exposure. Techniques such as Secure Aggregation and Differential Privacy are used to protect individual data during model training.

2. Cryptographic Mechanisms:

Technologies like Homomorphic Encryption and Attribute-Based Encryption are integrated to ensure secure computation and fine-grained access control within decentralized learning frameworks.

3. Robustness Against Attacks:

Decentralized learning eliminates single points of failure, making systems more resilient to adversarial attacks and server breaches. This is especially important when central trust is limited or absent.

4. Regulatory Compliance:

By avoiding centralized data storage, decentralized approaches help organizations comply with strict data privacy laws such as GDPR and HIPAA.

Implementing Blockchain Technology in Federated Learning

Integrating blockchain with federated learning introduces a new layer of privacy, transparency, and trust in decentralized environments. Key benefits include:

1. Decentralized Identity (DID):

Blockchain supports self-sovereign identities, allowing users to control their personal data and selectively share it, eliminating reliance on centralized identity providers.

2. Data Integrity and Transparency:

The immutable nature of blockchain ensures that data and model updates are tamper-proof, facilitating transparent and secure management of sensitive information.

3. Privacy-Preserving Techniques:

Zero-Knowledge Proofs (ZKPs) allow users to validate claims without disclosing the underlying data—ideal for scenarios like identity verification.

4. Differential Privacy:

Combining blockchain with differential privacy adds controlled noise to datasets, preserving user anonymity while maintaining the utility of aggregated insights.

Here is a corrected, polished, and professionally formatted version of your

III. METHODOLOGY

Algorithms

Decentralized Federated Learning (DFL) integrated with blockchain utilizes a range of algorithms to ensure **privacy**, **security**, and **efficient model training**. The key algorithms involved include:

1. Consensus Algorithms:

- **Proof of Authority (PoA):** Used in permissioned blockchain networks to ensure efficient and secure consensus.
- **Proof of Stake (PoS)** and **Delegated Proof of Stake (DPoS):** Employed in blockchain systems to validate transactions and maintain decentralization.

2. Model Aggregation Techniques:

- **Weighted Average:** Aggregates local models based on their performance and contribution to the global model.
- **Ensemble Methods:** Combines predictions from multiple models to improve robustness and generalization.

3. Privacy-Preserving Mechanisms:

- **Homomorphic Encryption:** Enables computation on encrypted data without the need for decryption.
- **Zero-Knowledge Proofs (ZKP):** Allows verification of data validity without exposing the actual data.

4. Optimization Algorithms:

- **Stochastic Gradient Descent (SGD):** A widely used optimization method for training machine learning models in federated settings.
- **Adaptive Weight Calculation:** Dynamically adjusts weights based on the quality of local data to ensure fair and effective training.

System Design Document Overview

The system design outlines critical aspects of the DFL framework and includes:

- **System Requirements**
- **Operating Environment**
- **System and Subsystem Architecture**
- **File and Database Design**
- **Input/Output Formats**
- **Human-Machine Interfaces**
- **Processing Logic and External Interfaces**

Data Collection

The **NSL-KDD dataset** is used for intrusion detection. It contains symbolic features such as protocol type, service, and flag. These features are processed as follows:

- **One-Hot Encoding:** Symbolic data is converted into binary vectors to be recognized by machine learning algorithms.
  - Example: The protocol_type feature with values like tcp, udp, and icmp is transformed as:
    - tcp → [1, 0, 0]
    - udp → [0, 1, 0]
    - icmp → [0, 0, 1]

Data Preprocessing

Before training, the dataset undergoes thorough cleaning to remove:

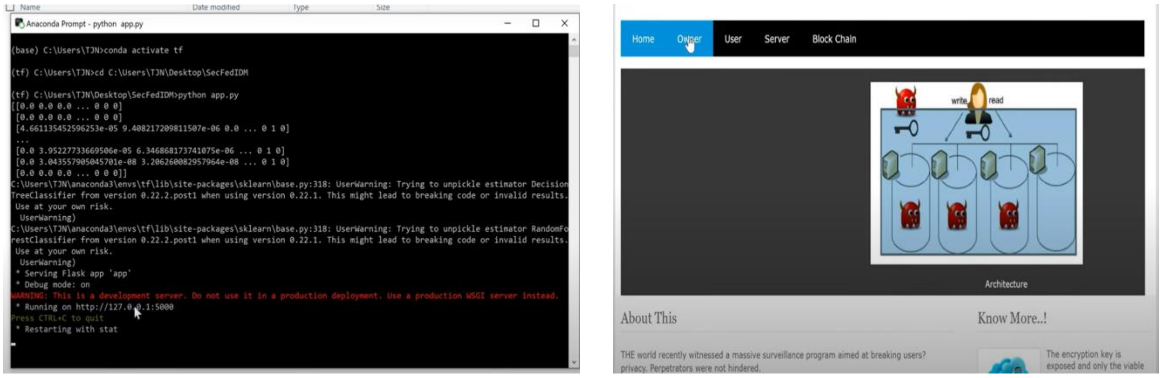
- **Noisy data**
- **Duplicate entries**
- **Missing or infinite values**

This step ensures data quality and consistency for better model performance.

Train-Test Split and Model Fitting

- The dataset is divided into **training** and **testing** subsets.
- This split allows the evaluation of the model’s generalization ability on unseen data.
- After splitting, **model fitting** is performed, where the training data is used to optimize the model’s parameters and minimize error.

Let me know if you'd like this reformatted for a journal manuscript or thesis template.



```
(base) C:\Users\TJN>conda activate tf

(tf) C:\Users\TJN>cd C:\Users\TJN\Desktop\SecFedIDM

(tf) C:\Users\TJN\Desktop\SecFedIDM>python app.py
[[0.0 0.0 0.0 ... 0 0 0]
 [0.0 0.0 0.0 ... 0 0 0]
 [4.661135452596253e-05 9.408217209811507e-06 0.0 ... 0 1 0]
 ...
 [0.0 3.95227733669506e-05 6.346868173741075e-06 ... 0 1 0]
 [0.0 3.043557905045701e-08 3.206260082957964e-08 ... 0 1 0]
 [0.0 0.0 0.0 ... 0 0 0]]
C:\Users\TJN\anaconda3\envs\tf\lib\site-packages\sklearn\base.py:318: UserWarning: Trying to unpickle estimator Decision
TreeClassifier from version 0.22.2.post1 when using version 0.22.1. This might lead to breaking code or invalid results.
 Use at your own risk.
  UserWarning)
C:\Users\TJN\anaconda3\envs\tf\lib\site-packages\sklearn\base.py:318: UserWarning: Trying to unpickle estimator RandomFo
restClassifier from version 0.22.2.post1 when using version 0.22.1. This might lead to breaking code or invalid results.
 Use at your own risk.
  UserWarning)
 * Serving Flask app 'app'
 * Debug mode: on
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
 * Restarting with stat
```

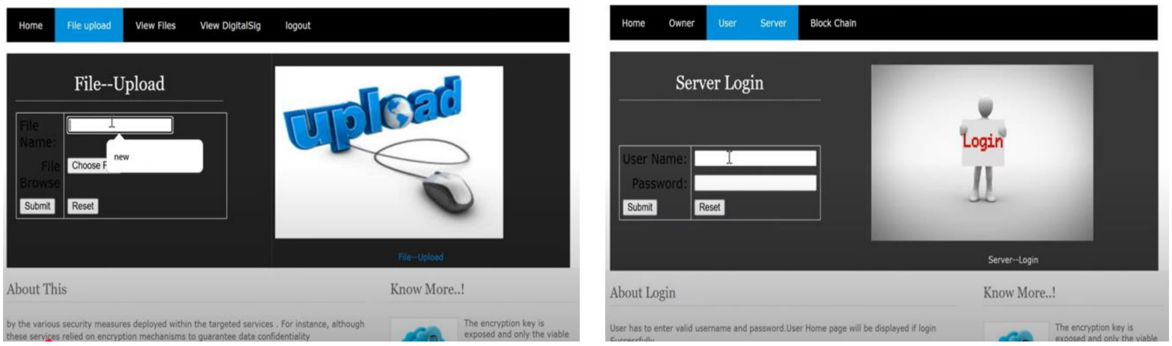
About This

THE world recently witnessed a massive surveillance program aimed at breaking users? privacy. Perpetrators were not hindered.

Know More..!

The encryption key is exposed and only the viable

Figure 1: Shows main test and owner login



File--Upload

About This

by the various security measures deployed within the targeted services . For instance, although these services relied on encryption mechanisms to guarantee data confidentiality

Know More..!

The encryption key is exposed and only the viable

Server Login

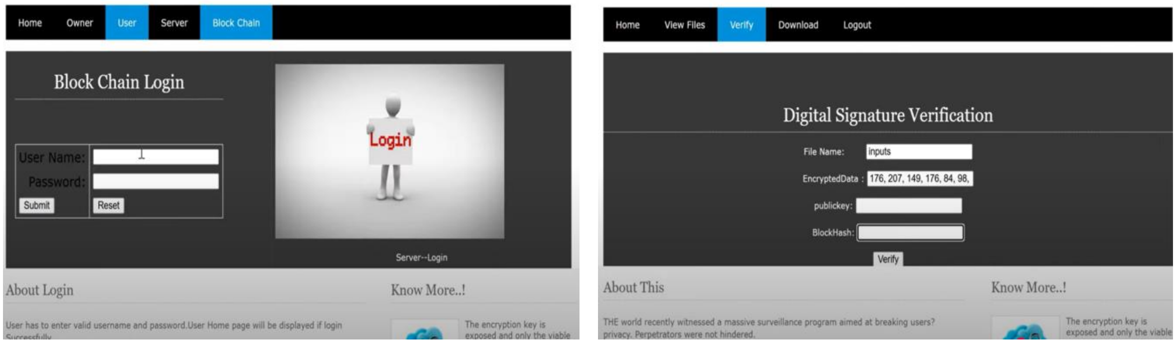
About Login

User has to enter valid username and password.User Home page will be displayed if login Successfully

Know More..!

The encryption key is exposed and only the viable

Figure 2: Shows uploading a file and server login



Block Chain Login

About Login

User has to enter valid username and password.User Home page will be displayed if login Successfully

Know More..!

The encryption key is exposed and only the viable

Digital Signature Verification

File Name: inputs

EncryptedData : 176, 207, 149, 176, 84, 98,

publickey:

BlockHash:

About This

THE world recently witnessed a massive surveillance program aimed at breaking users? privacy. Perpetrators were not hindered.

Know More..!

The encryption key is exposed and only the viable

Figure 3: Shows to create a key login to blockchain and Digital signature

V.CONCLUSIONS

The integration of FL and blockchain technologies has alleiated the need for a server in the network. This is followed by a number of advantages such as efficient communication and the elimination of a single point of failure in the federation. While there are a limited number of surveys on the security analysis of FL, since the previous studies were all based on centralized architectures, further analysis is required to study the new paradigm of DFL from a security perspective. This work first reviewed common trends and preliminaries of FL and blockchain. It then performed a security analysis on DFL by identifying possible threats and defense mechanisms in such systems.

VI.DISCUSSIONS

DFL relies on efficient communication protocols to exchange model updates directly between clients. This reduces communication overhead compared to centralized federated learning.

Network Topologies: Various network structures, such as peer-to-peer or hierarchical models, are explored to optimize the flow of information among clients.

ACKNOWLEDGEMENT(S):

It gives me immense pleasure to thank my guide Dr. P. SREEDEVI for having been a source of constant inspiration, precious guidance and generous assistance during the project work.Without her close monitoring and valuable suggestions, this work would not have taken this shape. I feel that his help is substitutable and unforgettable.

I take the privilege to record my thanks to H.O.D of CSE Dr. M. CHIRANJEEVI. for his encouragement, cooperation and valuable suggestions to complete project in time.

I take the privilege to record my thanks to Directories od academics of college Dr. P CHIRANJEEVI for his encouragement, cooperation and valuable suggestions to complete project in time.

I am very thankful to honorable Principal Dr M. SASIDHAR and I express my sincere gratitude and deep sense of respect for making me available.

REFERENCES:

[1] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, “Federated optimization: Distributed machine learning for on-device intelligence,” *arXiv preprint*, arXiv:1610.02527, 2016.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, vol. 54, 2017, pp. 1273–1282.

[3] C. Ma, J. Li, M. Ding, H. H. Yang, F. Shu, T. Q. S. Quek, and H. V. Poor, “On safeguarding privacy and security in the framework of federated learning,” *IEEE Network*, vol. 34, no. 4, pp. 242–248, 2020.

[4] 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.

[5] M. Shayan, C. Fung, C. J. M. Yoon, and I. Beschastnikh, “Biscotti: A blockchain system for private and secure federated learning,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 7, pp. [Incomplete Page Numbers], 2021.