ESANN 2025 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 23-25 April 2025, i6doc.com publ., ISBN 9782875870933. Available from http://www.i6doc.com/en/.

SecureBFL: a Blockchain-enhanced federated learning architecture with MPC

Tanguy Vansnick[1], Leandro Collier[2] and Saïd Mahmoudi[1]

1- Faculty of Engineerging University of Mons - ILIA
Rue de Houdain 9, 7000 Mons - Belgium

2- Applied Research Center - CETIC
Avenue Jean Mermoz 28, 6041 Charleroi - Belgium

Abstract. The increasing demand for data in machine learning raises significant privacy concerns. Federated Learning (FL) enables multiple entities to train models collaboratively without sharing raw data. However, centralized FL (CFL) relies on a central server, making it vulnerable to poisoning attacks and single points of failure (SPOF). Decentralized FL (DFL) addresses these issues by removing the central server. This paper proposes a novel DFL architecture integrating blockchain for resisting attacks and Multi-Party Computation (MPC) for secure model parameter transfer. This architecture enhances security and confidentiality in collaborative learning without compromising result quality.

1 Introduction

Federated Learning (FL) [1] addresses privacy concerns by enabling multiple devices to train models locally, with models aggregated into a global model using algorithms like FedAvg [2]. This decentralized approach mitigates the privacy risks associated with centralized model training.

The most common FL architecture is centralized federated learning (CFL), where local model parameters (e.g., weights or gradients) are sent to a central server for aggregation. However, CFL relies on a single point of failure and is vulnerable to attacks. Decentralized Federated Learning (DFL) [3] eliminates the central server by sharing trained models among clients, who use consensus mechanisms for aggregation. Some DFL architectures incorporate blockchain technology to enhance traceability, transparency, and security during aggregation. Despite these systems' decentralized nature, it remains possible to infer sensitive data from local models through inference attacks, thereby threatening privacy.

We propose an architecture that integrates **Private Blockchain and Multi-party Computation (MPC)** to eliminate single points of failure, mitigate poisoning attacks, and further enhance data privacy. This approach achieves security and data privacy, providing a secure and private federated learning environment.

ESANN 2025 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 23-25 April 2025, i6doc.com publ., ISBN 9782875870933. Available from http://www.i6doc.com/en/.

2 Privacy and Security in Federated Learning

Integrating federated learning (FL) with privacy-preserving techniques faces challenges in data privacy, network security, and attack mitigation. This section explores FL approaches to enhance privacy, security, and address attack vectors.

2.1 Privacy-Preserving Techniques

Various solutions address data privacy. Differential Privacy (DP) [4] adds noise to training data or gradients to protect individual privacy, which can reduce model accuracy. Homomorphic Encryption (HE) [5] allows computations on encrypted data, though it introduces significant computational overhead, limiting its efficiency in large-scale systems. Additive Secret Sharing MPC [7] splits data into random shares that sum to the original value, distributing these shares among parties where no individual share reveals information about the input. This method enables efficient, secure aggregation of model updates while maintaining privacy, as reconstruction requires collecting all shares.

2.2 Security and reliability

FL systems often leverage distributed architectures like blockchain to enhance reliability and security. While public blockchains ensure tamper-proof records through consensus protocols, their transparency can expose sensitive details. Private blockchains, offer a more controlled environment, restricting access and visibility to authorized participants, which can mitigate some privacy concerns.

Existing Blockchain-Based Federated Learning (BFL) solutions address reliability through techniques like committee-based consensus [8], which withstand malicious participants. However, even private blockchains often store models or references on the chain, risking sensitive information exposure. Similarly, local collaboration approaches [9] bypass centralized models but fail to ensure privacy against adversarial attacks.

2.3 Addressing Privacy and Security Attacks

FL systems face various vulnerabilities that can seriously affect privacy and security. One significant threat is poisoning attacks [10], where malicious participants submit false updates to the model, leading to incorrect predictions or system failures. Many traditional FL systems also rely on a centralized server for model aggregation. If this server is compromised, it poses a single point of failure [11], jeopardizing the entire system's integrity. Another concern is inference attacks [12], where attackers may try to deduce sensitive information from the aggregated model, posing a significant privacy risk without direct access to the training data. Combining techniques such as MPC with robust security protocols, including consensus algorithms or Blockchain-based models, is essential to address these vulnerabilities. These strategies can help establish a more resilient architecture for federated learning, enhancing its security and privacy assurances.

ESANN 2025 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 23-25 April 2025, i6doc.com publ., ISBN 9782875870933. Available from http://www.i6doc.com/en/.

2.4 Combined Privacy and Security Solutions

Several hybrid approaches have tried to address both privacy and security. One approach combines **MPC** with a two-server setup, where the servers securely aggregate model updates. However, this method has its drawbacks: if either of the servers fails, the entire process collapses, making it a more significant risk than a traditional **SPOF** due to the reliance on both servers for system operation [13]. Another approach also combining MPC (with Replicated Secret Sharing), blockchain and FL exists, but although it guarantees that the models sent and received are identical, it doesn't prevent the model before sending from being poisoned [6]. Without MPC, there are approaches integrates **DP** with **Proof of Federation (PoF)**, which not only strengthens data privacy but also uses **blockchain** to verify the integrity of participants within the federated network [14]. These hybrid models offer promising enhancements to FL systems, but they must balance the trade-off between model accuracy, efficiency, and security.

3 Proposed framework

The proposed architecture, shown in Figure 1, consists of three actors: **Devices**, **Clusters**, and **Nodes**. Each device trains a local model using its data and is linked to a single node. Multiple devices form a cluster, which communicates and shares information. Nodes aggregate models and maintain the blockchain. FL is divided into two stages and produces three types of models. Each device first trains a **local model**. These local models are aggregated within clusters to form a **cluster model**. Finally, the nodes use consensus mechanisms to aggregate cluster models into a **global model** shared with all devices for the next iteration.

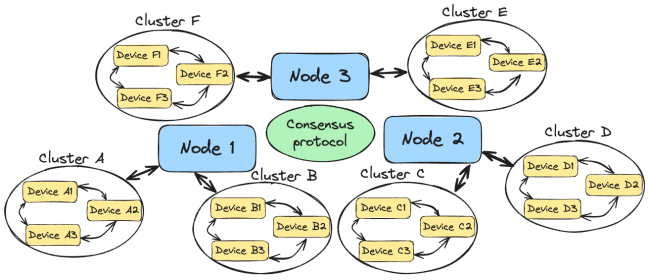

Fig. 1: Overview of the proposed BFL architecture

3.1 Key Processes

Cluster Generation: Before training, nodes form clusters by randomly grouping connected devices. Each cluster must include a minimum number of devices to ensure MPC. Clusters are regenerated before every new training process to enhance system robustness.

ESANN 2025 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 23-25 April 2025, i6doc.com publ., ISBN 9782875870933. Available from http://www.i6doc.com/en/.

Global Model Creation: Initially, nodes propose a first common global model with random weights. For next iterations, nodes aggregate multiple cluster models to create a new global model. The consensus mechanism validates these models before adding them to the blockchain. After validation, nodes distribute the global model to all devices.

Model Training and Aggregation: As illustrated in Figure 2, devices train their local models using the global model and their data. These local models are divided into fragments using additive secret sharing MPC, where each fragment represents a random share that sums to the original model parameters. These fragments are shared within the cluster, ensuring no single device can reconstruct the complete model updates. Once all fragments are shared, a cluster model is aggregated by summing the corresponding shares.

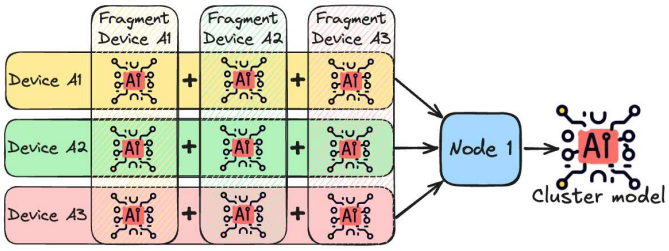

Fig. 2: Overview of the proposed BFL architecture

Validation: Once a cluster model is generated, it is validated against the global model using a threshold parameter T and a ceiling C. The condition is:

$$\min\left(L_{\text{global}} \times T, L_{\text{global}} + C\right) \quad \geq \quad L_{\text{cluster}}$$

Here, L_{cluster} represents the loss of the cluster model, and L_{global} is the loss of the global model. The blockchain maintains a sequential record of both models. When sufficient cluster models are collected, a new global model is aggregated and validated through consensus.

3.2 Model Storage

Due to the models' size, the blockchain does not store the models directly. Instead, a reference to the storage location and a hash of the model weights are stored. This ensures integrity and reduces blockchain storage requirements.

4 Experimentation

We conducted two experiments using an NVIDIA RTX 4090 GPU on three architectures: a **classic CFL**, a **verified CFL** (CFL-V), and a **BFL** architecture using consensus. For our tests, we set the threshold $T = 1.05$ and ceiling $C = 0.08$ for CFL-V and BFL, simulating random and targeted poisoning.

ESANN 2025 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 23-25 April 2025, i6doc.com publ., ISBN 9782875870933. Available from http://www.i6doc.com/en/.

In the first experiment with CIFAR-10 using MobileNet-v2 and 18 clients in clusters of 3 over 50 training rounds, the CFL architecture's accuracy dropped to 17.12% due to 6 poisoned clients. In contrast, CFL-V and BFL maintained accuracies of **81.83%** and **81.32%**, respectively. The second experiment with CIFAR-100 and ShuffleNet involved 45 clients. Here, CFL-V underperformed against BFL, as verification on local models did not generalize well. Targeted attacks had minimal effects on performance due to the dataset's 100 classes.

As detailed in Table 1, each architecture exhibited different trade-offs. While efficient in communication and energy consumption, CFL was highly vulnerable to poisoning. CFL-V improved robustness by filtering suspicious updates but incurred higher energy costs. BFL offered the best resilience, maintaining higher accuracy even under attack, but at the expense of significantly more significant communication overhead and energy consumption.

Architecture	**Metric**	**CIFAR-10, MobileNet-v2**			**CIFAR-100, ShuffleNet**		
		No Pois.	Rand (6)	Targ. (6)	No Pois.	Rand (15)	Targ. (15)
CFL	Loss	0.55	2.28	0.87	1.99	2.18	1.98
	Acc (%)	**84.49**	17.12	75.51	**50.37**	44.47	**50.43**
	Comm. (Mo)	7805.85	7805.85	7805.85	11803.88	11803.88	11803.88
	Energy (MJ)	0.33	0.37	0.30	0.29	0.27	0.32
CFL-V	Loss	0.64	0.68	0.81	2.54	2.57	2.55
	Acc (%)	83.76	**81.83**	76.59	34.83	33.98	34.54
	Comm. (Mo)	7805.85	7805.85	7805.85	11803.88	11803.88	11803.88
	Energy (MJ)	0.43	0.41	0.39	0.53	0.52	0.51
BFL	Loss	0.61	0.65	0.57	1.93	2.06	1.98
	Acc (%)	83.96	81.32	**84.28**	49.5	**45.86**	47.98
	Comm. (Mo)	48133.65	47561.16	48133.65	71315.45	23581.02	23581.02
	Energy (MJ)	0.94	0.98	0.95	1.5	1.84	1.82

Table 1: Performance comparison of CFL, CFL-V, and BFL architectures on CIFAR-10 (MobileNet-v2) and CIFAR-100 (ShuffleNet) under different poisoning scenarios.

5 Conclusion

Our experiments demonstrate that the BFL architecture is effective and reliable. In contrast, CFL is prone to a single point of failure and suffers performance degradation from poisoning attacks. The decentralized structure and consensus mechanism of our framework ensure stability and accuracy, even with compromised devices.

Acknowledgement

This research was partially supported by BelSPO AIDE AICall2023 and ARIAC SPW 2010235 Projects, as well as by the InforTech research Institute of UMONS. We would like to express our gratitude to Dr. Volker Strobel (ULB) for his valuable expertise and guidance, and to Maxime Gloesener for his help and key insights.

ESANN 2025 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 23-25 April 2025, i6doc.com publ., ISBN 9782875870933. Available from http://www.i6doc.com/en/.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas, editors. *Communication-efficient learning of deep networks from decentralized data*, in Artificial Intelligence and Statistics, PMLR, pp. 1273-1282, 2017.

[2] P. Qi, D. Chiaro, A. Guzzo, M. Ianni, G. Fortino, and F. Piccialli. *Model aggregation techniques in federated learning: A comprehensive survey*, in Future Generation Computer Systems, Elsevier, vol. 150, pp. 272-293, 2024.

[3] Q. Li, Z. Wen, Z. Wu, S. Hu, N. Wang, Y. Li, X. Liu, and B. He. *A survey on federated learning systems: Vision, hype and reality for data privacy and protection*, in IEEE Transactions on Knowledge and Data Engineering, IEEE, vol. 35, no. 4, pp. 3347-3366, 2021.

[4] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. S. Quek, and H. V. Poor, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3454–3469, 2020.

[5] X. Lessage, L. Collier, C.-H. B. Van Ouytsel, A. Legay, S. Mahmoudi, and P. Massonet, "Secure federated learning applied to medical imaging with fully homomorphic encryption," in *Proc. 2024 IEEE 3rd International Conference on AI in Cybersecurity (ICAIC)*, IEEE, 2024, pp. 1–12.

[6] M. Fan, Z. Zhang, Z. Li, G. Sun, H. Yu, J. Kang, and M. Guizani, "SecureVFL: privacy-preserving multi-party vertical federated learning based on blockchain and RSS," in *Digital Communications and Networks*, Elsevier, 2024.

[7] S. Truex, N. Baracaldo, A. Anwar, T. Steinke, H. Ludwig, R. Zhang, and Y. Zhou, "A hybrid approach to privacy-preserving federated learning," in *Proc. 12th ACM Workshop on Artificial Intelligence and Security*, 2019, pp. 1–11.

[8] Y. Li, C. Chen, N. Liu, H. Huang, Z. Zheng, and Q. Yan, "A blockchain-based decentralized federated learning framework with committee consensus," *IEEE Network*, vol. 35, no. 1, pp. 234–241, 2020.

[9] N. Onoszko, G. Karlsson, O. Mogren, and E. L. Zec, "Decentralized federated learning of deep neural networks on non-iid data," *arXiv preprint arXiv:2107.08517*, 2021.

[10] V. Tolpegin, S. Truex, M. E. Gursoy, and L. Liu, "Data poisoning attacks against federated learning systems," in *Proc. 25th European Symposium on Research in Computer Security (ESORICS 2020), Part I*, Guildford, UK, Sep. 2020, pp. 480–501.

[11] A. Gholami, N. Torkzaban, and J. S. Baras, "Trusted decentralized federated learning," in *Proc. 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, 2022, pp. 1–6.

[12] Y. Chen, Y. Gui, H. Lin, W. Gan, and Y. Wu, "Federated learning attacks and defenses: A survey," in *Proc. 2022 IEEE International Conference on Big Data (Big Data)*, 2022, pp. 4256–4265.

[13] Y. Khazbak, T. Tan, and G. Cao, "MLGuard: Mitigating poisoning attacks in privacy preserving distributed collaborative learning," in *Proc. 2020 29th International Conference on Computer Communications and Networks (ICCCN)*, 2020, pp. 1–9.

[14] 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. 1513–1525, 2020.