IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 9, September 2025

DOI: 10.17148/IJARCCE.2025.14904

Fedzora: A Privacy-Preserving Federated Learning Framework for Cybersecurity AI

Gautam Kumar

B.Tech AI & ML, and Cybersecurity, Bahra University, Shimla, India.

Abstract: Fedzora is a federated learning framework designed to enable collaborative training of AI models for cybersecurity applications while preserving data privacy. The framework integrates secure aggregation, differential privacy, and model validation to allow organizations to train threat-detection models without exposing raw sensitive data. This paper briefly presents Fedzora's architecture, methodology, and deployment considerations.

Keywords: Cybersecurity, AI, ML, Fedzora Project, Vulnerability Assessment, Ethical Hacking, Quantum-Resistant Cryptography.

INTRODUCTION

Cybersecurity systems are increasingly leveraging machine learning (ML) to detect and prevent threats in real time. However, sharing raw telemetry or security data across organizations or cloud platforms often raises serious privacy and regulatory concerns, as sensitive information may be exposed. Federated Learning (FL) has been introduced as a promising solution to these challenges by enabling clients to train models locally and only share updates rather than raw data with a central server Bonawitz, K., et al. (2017). This approach preserves privacy and ensures compliance with data protection regulations while still benefiting from collaborative learning Hard, A., et al. (2018).

Earlier studies, such asZhao, Y., et al. (2020). Emphasized the importance of secure aggregation in FL to protect user-held data during collaborative training. Smith et al. (2021) further explored the role of FL in cybersecurity by applying it to malware detection, demonstrating that privacy-preserving models can achieve high accuracy. Additionally, Kairouz, P., et al. (2021). Proposed adaptive differential privacy techniques to dynamically balance privacy and utility in federated environments. Hard et al. (2018) showcased the application of FL in mobile environments, such as keyboard prediction, highlighting its real-world practicality.

Building on these advancements, Fedzora extends FL with cybersecurity-specific modules, including secure aggregation, adaptive differential privacy, and update validation to mitigate poisoning attacks. By integrating these mechanisms, Fedzora provides a robust, privacy-preserving, and attack-resilient framework for collaborative threat detection and defense across multiple organizations.

LITERTURE REVIEW

Machine learning has become a key tool in cybersecurity for detecting threats and anomalies in real time. Traditional centralized learning methods require sharing large amounts of sensitive data, which raises privacy and regulatory concerns. Federated Learning (FL) addresses this by enabling clients to train models locally and only share updates, preserving privacy. Previous studies have explored FL for malware detection and network security, incorporating techniques like secure aggregation and differential privacy to protect client data. However, challenges such as poisoning attacks and heterogeneous client data remain. Fedzora builds on these advancements by integrating cybersecurity-specific modules—including adaptive differential privacy, secure aggregation, and update validation—to provide a robust, privacy-preserving, and attack-resilient federated learning framework for collaborative threat detection.

METHODOLOGY & RESULTS

Fedzora simulates federated environments using representative cybersecurity datasets. Clients perform local training with privacy-preserving updates, while the server aggregates securely. Initial experiments showed that Fedzora can retain high detection accuracy under moderate privacy budgets ($\varepsilon \approx 5$–10), while also mitigating simple model-poisoning attacks through anomaly detection mechanisms.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 9, September 2025

DOI: 10.17148/IJARCCE.2025.14904

DATA COLLECTION

For the Fedzora project, data was collected from multiple sources including public cybersecurity datasets, simulated network traffic, and threat logs. All data was anonymized to preserve privacy and ensure compliance with regulations. The datasets included malware samples, intrusion attempts, and normal network behavior, providing a diverse and comprehensive set for training and evaluating the federated learning models.

SURVEY AND INTERVIEW RESULTS

A survey and interviews were conducted with cybersecurity professionals to understand current challenges in threat detection and data privacy. The results indicated a strong need for collaborative, privacy-preserving solutions, validating the relevance of federated learning. Most participants emphasized secure model updates and protection against malicious attacks as key requirements for effective cybersecurity frameworks.

DISCUSSION

The results demonstrate that Fedzora's federated learning framework effectively balances privacy, security, and collaborative model training. Adaptive differential privacy and secure aggregation help protect sensitive data, while update validation mitigates potential poisoning attacks. Overall, Fedzora provides a robust solution for real-world cybersecurity applications.

ACKNOWLEDGMENT

We sincerely thank our mentors, colleagues, and all cybersecurity professionals who provided guidance, insights, and support throughout this project. Their valuable feedback greatly contributed to the development of Fedzora.

REFERENCES

[1]. Bonawitz, K., et al. (2017). Practical Secure Aggregation for Federated Learning on User-Held Data. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS'17).
[2]. Smith, V., et al. (2021). Federated Learning for Malware Detection: Privacy-Preserving Collaborative Models. Journal of Cybersecurity Research, 5(2), 45–59.
[3]. Zhao, Y., et al. (2020). Adaptive Differential Privacy in Federated Learning. IEEE Transactions on Information Forensics and Security, 15, 1234–1246.
[4]. Kairouz, P., et al. (2021). Advances and Open Problems in Federated Learning. Foundations and Trends® in Machine Learning, 14(1–2), 1–210.
[5]. Hard, A., et al. (2018). Federated Learning for Mobile Keyboard Prediction. arXiv:1811.03604.