Privacy Preserving Cybersecurity Framework for Next Generation Digital Systems

Vinayak Musale[1], Basant Kumar[2]

[1] Lincoln University College, Malaysia; [2] Modern College of Business and Science, Oman

pdf.vinayak@lincoln.edu.my[1]; basant@mcbs.edu.om[2]

Abstract: The widespread adoption of IoT devices, edge computing and cloud services has created highly distributed digital environments that are increasingly vulnerable to sophisticated cyber threats. Although collaborative threat detection can improve security, sharing sensitive operational data among organizations raises concerns related to privacy, trust and regulatory compliance. To address these challenges, this paper proposes a unified privacy aware cybersecurity framework that supports collaborative intrusion detection without transferring raw security data. The proposed architecture combines federated learning for decentralized model development, differential privacy to protect shared model updates, permissioned blockchain to ensure secure and trustworthy collaboration and Explainable Artificial Intelligence to provide transparent and interpretable security decisions. Within this framework, participating organizations train intrusion detection models using locally available data, while only privacy protected model parameters are exchanged to construct a global detection model. Blockchain technology preserves the integrity of collaborative learning by maintaining verifiable records of model exchanges, whereas XAI assists security analysts in understanding the reasoning behind detected threats. The proposed framework offers a scalable and privacy conscious foundation for securing distributed digital infrastructures, including smart cities, healthcare, industrial automation and smart intelligent transportation systems. Future research will focus on prototype implementation and comprehensive performance evaluation using benchmark cybersecurity datasets.

Keywords: Blockchain; Distributed Cybersecurity; Explainable AI; Federated Learning; Security; Privacy.

1. Introduction

The rapid growth of IoT, edge computing and cloud technologies has transformed modern digital infrastructures by enabling intelligent applications across healthcare, industry, transportation and smart cities. However, the increasing connectivity of these systems also expands the cyberattack surface and makes securing distributed environments more challenging. Conventional cybersecurity solutions rely on centralized data collection, which raises privacy concerns, increases communication overhead and creates single points of failure. Although artificial intelligence has improved cyber threat detection, most of the learning models require centralized datasets, limiting collaboration in privacy sensitive environments where organizations cannot readily share security data. Federated learning addresses this limitation by allowing multiple participants to jointly train machine learning models by keeping raw data within local infrastructures. Nevertheless, exchanged model parameters may still reveal sensitive information, requiring additional privacy preserving mechanisms.

Differential privacy provides mathematical protection against information leakage by disturbing model updates before they are shared, while permissioned blockchain establishes trust, integrity and accountability among collaborating organizations. In addition, use of Explainable Artificial Intelligence improves the transparency of automated security decisions that helps analysts to understand the reasoning behind detected threats and supporting regulatory compliance.

Motivated by these challenges, this paper proposes a unified privacy preserving cybersecurity framework that integrates federated learning, differential privacy, blockchain based trust management and explainable AI. The proposed architecture provides a secure collaborative threat intelligence along with preserving data confidentiality, strengthening trust among distributed participants and improving the interpretability of AI-driven cybersecurity decisions. The framework provides a scalable foundation for securing next generation IoT, edge and cloud environments.

2. Related work

Recent advances in cybersecurity have focused on improving intrusion detection, privacy preservation and secure collaboration in distributed environments. Deep learning models have demonstrated high accuracy in detecting sophisticated cyberattacks by learning complex patterns from network traffics. For example, CNN-LSTM and CNN-GRU architectures have shown promising performance in IoT intrusion detection. Most existing approaches depend on centralized datasets and restrictive their use in privacy sensitive environments [1], [2].

Federated learning overcomes this challenge by enabling collaborative model training without sharing raw data [3], [4]. As per literature review, Kairouz *et al.* [3] highlighted its potential for distributed machine learning, whereas Sarker *et al.* [4] demonstrated its application in healthcare IoT systems. However, model updates may still reveal sensitive information, making differential privacy essential for protecting exchanged parameters [5].

Differential privacy has therefore become an important enhancement for collaborative learning systems. By introducing carefully calibrated perturbations into model updates, differential privacy minimizes the possibility of reconstructing confidential training data while maintaining acceptable prediction performance. Nguyen et al. [5] reported that differential privacy can provide formal privacy guarantees for cybersecurity applications without substantially affecting model utility. Blockchain further strengthens collaboration by ensuring data integrity, secure participation and trust among distributed entities [6], [7] while Explainable Artificial Intelligence (XAI) improves the transparency and interpretability of AI-based threat detection [8] [9]. Despite these advances, limited research has explored their integration within a unified cybersecurity framework. This work addresses that gap by combining federated learning, differential privacy, blockchain, and XAI into a comprehensive privacy-preserving cybersecurity architecture. The proposed system features and supports privacy preserving collaborative threat intelligence across IoT, edge and cloud environments.

3. Proposed Framework and Methodology

3.1 Framework Overview

The proposed framework enables collaborative cyber threat detection along with preserving the privacy of participating organizations. IoT devices, edge nodes, and cloud platforms independently train local intrusion detection models using their own datasets instead of sharing raw security data. , As illustrated in Fig. 1, the architecture consists of five interconnected layers namely Federated Learning, Differential Privacy, Blockchain based Trust Management, Explainable Artificial Intelligence and Security Operations. This layered design facilitates secure collaboration, trusted information sharing and transparent threat analysis across distributed digital environments.

3.2 Federated Learning

Locally, each participating organization trains an intrusion detection model using network traffic, authentication logs and system events. Only model parameters are transmitted to the aggregation server, where the Federated Averaging (FedAvg) algorithm combines local updates to construct a global model.

$$W^{t+1} = \sum_{k=1}^{N} \frac{n_k}{n} W_k^t$$

where W_k^trepresents the local model parameters, $n_k$denotes the local training samples, and $n$is the total number of samples across all participants. The updated global model is redistributed for subsequent learning rounds, allows collaborative threat intelligence without exposing the sensitive datasets.

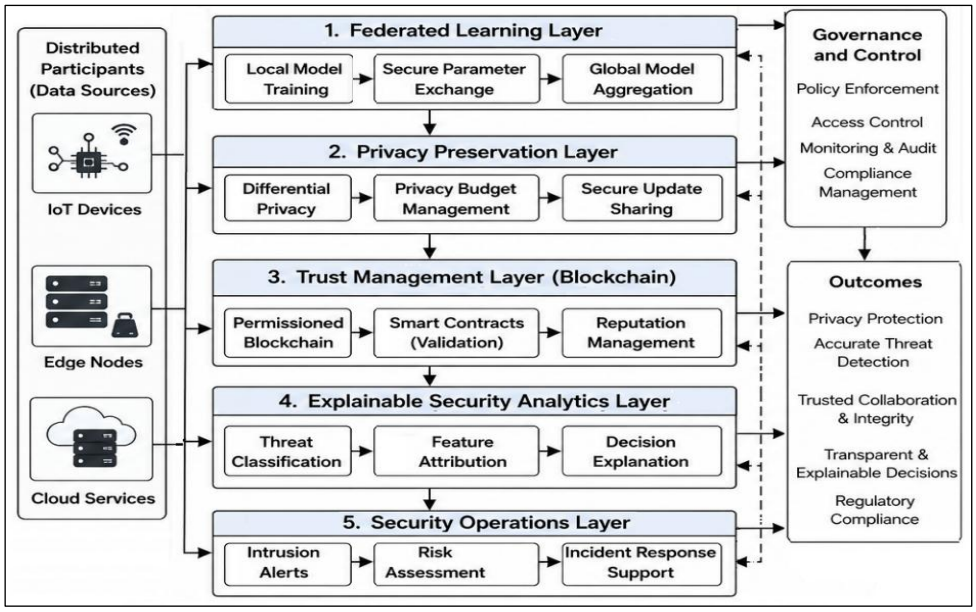

Figure 1. Proposed Privacy Preserving and AI-driven Cybersecurity Framework

3.3 Differential Privacy

Although federated learning protects raw data, exchanged model parameters may still leak confidential information. To reduce this risk, differential privacy is applied by injecting Gaussian noise into model updates before transmission.

$$\widetilde{W} = W + \mathcal{N}(0, \sigma^2)$$

where W is the original parameter vector and $\mathcal{N}(0, \sigma^2)$ represents the added noise. This mechanism strengthens privacy while maintaining acceptable detection performance.

3.4 Blockchain-Based Trust Management

A permissioned blockchain establishes trust among collaborating organizations by verifying participant identities and recording model updates in an immutable ledger. Smart contracts validate every transaction before aggregation. It also ensures data integrity, traceability and accountability throughout the collaborative learning process. The decentralized trust model eliminates reliance on a single authority and strengthens secure cooperation among distributed participants.

3.5 Explainable Security Analytics

To improve transparency, the Explainable Artificial Intelligence feature is incorporated after global threat prediction. Feature attribution techniques such as SHAP and LIME identify the factors influencing each prediction. By providing clear explanations for security predictions, the framework also helps analysts to identify genuine attacks, reduce false alarms and respond with greater confidence. The improved transparency also supports forensic investigations and compliance requirements.

3.6 Operational Workflow

The framework initiates with collecting security events from IoT devices, edge servers and cloud platforms. Each participant trains an intrusion detection model using locally available data, while only privacy protected model updates are shared through federated learning. Blockchain ensures the integrity of collaborative interactions, and XAI provides clear explanations for threat predictions before security alerts are issued. Based on these validated insights, appropriate response actions, such as administrator notification, traffic filtering, endpoint isolation or access control, are initiated. This integrated workflow enables timely and privacy preserving cyber defence across distributed environments.

4. Discussion

The proposed framework demonstrates how privacy-preserving collaborative cybersecurity can be achieved without centralized data sharing. By combining federated learning with differential privacy, organizations can jointly enhance intrusion detection while keeping sensitive security data within their local environments. This approach strengthens data confidentiality and supports compliance with modern privacy regulations.

The integration of blockchain further improves the framework by providing secure participation, data integrity, and trusted collaboration among distributed entities. In addition, Explainable Artificial Intelligence usage makes threat predictions easier to interpret, allows security analysts to verify alerts and respond more confidently to potential attacks.

Although the proposed architecture offers a comprehensive solution, practical deployment requires careful consideration of communication overhead, blockchain latency and the balance between privacy protection and detection performance. Nevertheless, its modular design enables flexible implementation across IoT, edge and cloud environments, making it a promising approach for real-world distributed cybersecurity applications.

5. Conclusion and Future Work

This paper proposed a privacy preserving cybersecurity framework that combines federated learning, differential privacy, blockchain, and Explainable Artificial Intelligence (XAI) to enable secure collaborative threat detection in distributed digital environments. The framework allows organizations to jointly build intelligent intrusion detection models by keeping sensitive security data within their local environments. By integrating privacy protection, trusted collaboration and transparent decision making, it provides a practical approach for securing next-generation cyber-physical systems such as smart cities, healthcare, industrial automation and intelligent transportation.

Future work will focus on developing a prototype and validating the framework using benchmark intrusion detection datasets to evaluate detection accuracy, scalability, communication efficiency, blockchain performance and privacy preservation under realistic deployment conditions.

References

[1] Z. Zhang, M. Li, and A. Alazab, "An intelligent CNN–LSTM based intrusion detection framework for autonomous IoT networks," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12105–12118, 2022.

[2] A. Alazab, N. Islam, and R. J. Alazab, "Deep learning fusion model for IoT threat detection using CNN–GRU," *Future Generation Computer Systems*, vol. 149, pp. 237–250, 2023.

[3] P. Kairouz *et al.*, "Advances and open problems in federated learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[4] I. H. Sarker, M. A. Hossain, and S. Khatun, "Privacy-preserving healthcare IoT framework using federated learning and blockchain," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 720–732, 2023.

[5] T. D. Nguyen, A. Berrada, and D. S. Kim, "Differential privacy-enabled machine learning for cybersecurity applications," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 456–470, 2022.

[6] A. Kumar and R. Singh, "Blockchain-driven AI security architecture for intelligent cyber-physical systems," *IEEE Access*, vol. 12, pp. 45332–45347, 2024.

[7] R. Bansal and M. Patel, "Blockchain-based trust management framework for autonomous digital systems," *Sensors*, vol. 23, no. 8, Art. no. 3921, 2023.

[8] X. Wang, H. Zhao, and K. Yu, "Hybrid privacy-preserving cybersecurity framework for Industry 5.0 systems," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 3058–3071, 2024.

[9] J. Chen, P. Liu, and D. Shen, "Deep reinforcement learning for automated cyber threat response," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 5, pp. 2568–2581, 2020.