Detecting Threats in Edge IoT Networks Using Federated Learning and Digital Twin

Kanita Jerin Tanha, Md Mahinur Alam, Md Raihan Subhan, and Taesoo Jun
Pervasive Intelligent Computing Laboratory, Department of IT Convergence Engineering,
Kumoh National Institute of Technology, Gumi, South Korea
(kanitajerin17, mahinuralam213, raihan, and taesoo.jun)@kumoh.ac.kr

***Abstract*—This paper introduces an innovative approach to enhance intrusion detection in Edge IoT networks by integrating Federated Learning (FL) with Digital Twin technology. As the Internet of Things (IoT) continues to expand, the complexity of connected systems grows, making them susceptible to a wide range of security threats, including cybersecurity risks and other vulnerabilities. Traditional Intrusion Detection Systems (IDS) often fall short in addressing the scalability challenges and resource limitations of IoT environments, especially in decentralized, edge-based networks. By leveraging Digital Twin for simulating real-world attack scenarios and FL for decentralized model training, this framework ensures privacy preservation and real-time adaptability. The proposed system effectively detects a variety of threats, demonstrating high detection 98.51% accuracy with minimal latency while overcoming resource constraints in edge IoT networks.**

***Index Terms*—Digital Twin, edge IoT, federated learning (FL), intrusion detection.**

I. Introduction

In the era of IoT has become integral to modern systems, connecting devices across industries like healthcare, transportation, and smart cities. As the number of IoT devices grows, so does the complexity of securing these networks. Traditional IDS often struggle to address the dynamic nature of IoT environments, especially with the limited computational resources of edge devices.

FL offers a solution by enabling edge devices to collaboratively train machine learning models without sharing sensitive data, improving security while preserving privacy. This decentralized approach allows continuous updates to IDS models, ensuring they can adapt to new threats [1] [2] and digital twin technology enhances IDS by creating virtual replicas of IoT systems, enabling real-time monitoring and simulation of attack scenarios. Digital twin can also generate synthetic data to improve model training, increasing detection accuracy [3].

Together, FL and Digital Twin offer a powerful, scalable framework for modernizing intrusion detection in IoT networks, ensuring real-time, adaptive, and privacy-preserving security [4]. The key contribution of this papers are: (1) FL for distributed and privacy-preserving model training, enabling efficient threat detection across multiple edge devices. (2) Digital twin technology enables a virtual representation of the physical system to simulate attack scenarios and provide rich, synthetic data for model training.

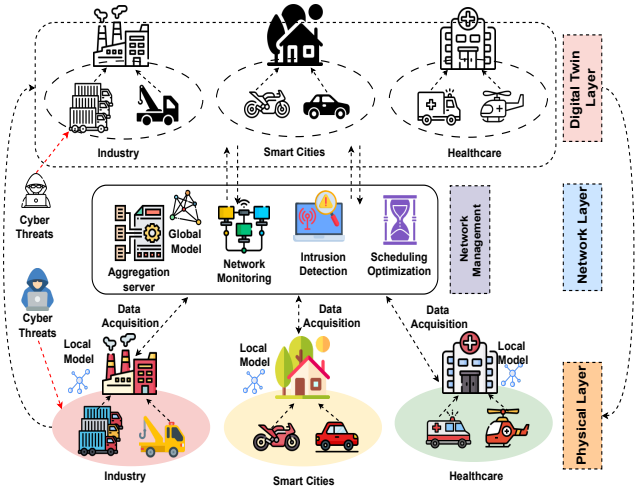

Fig. 1: Proposed system for intrusion detection in edge IoT networks integrating digital twin and FL.

II. Proposed System

The proposed system architecture is composed of Edge IoT devices, a central aggregation server, and a Digital Twin model. The edge devices generate real-time data from their sensors, which is used to detect intrusions within the network. Instead of relying on a centralized model, FL allows local models to be trained on the edge devices using local data, while ensuring that sensitive information never leaves the devices. The aggregated model is then updated and distributed back to the devices to improve the overall system's ability to detect new attack patterns. Digital twin technology is employed to create a virtual replica of the IoT network, simulating various attack scenarios. The simulated data generated by the Digital Twin is used to enhance model training, providing a broader set of attack scenarios for improving the detection capability of the system. Fig. 1 depicts a multi-layered system where local models on physical edge devices detect cyber threats, with updates aggregated via Digital Twin and Network Management layers to enable real-time intrusion detection, monitoring, and collaborative defense across industry, smart cities, and healthcare IoT environments. This combination of FL and digital twin offers the flexibility and scalability

TABLE I: Comparison of the proposed method with existing approaches.

Model	Accuracy (%)	Pecision (%)	Recall (%)	F1-Score (%)
CNN-MLP	89.78	90.01	89.78	89.28
DCFL-Chain	92.04	91.87	91.04	92.92
FCL-SBLS	93.69	93.91	93.71	93.81
Proposed	**98.51**	**97.24**	**98.23**	**97.81**

required for modern Edge IoT environments. Here, we explain the process of continuous model updates in the proposed FL system for intrusion detection. The global model θ_t is updated by aggregating the local models θ_i^t from each participating edge device. The update rule for the global model is defined as:

$$\theta_t = \frac{1}{N} \sum_{i=1}^{N} \theta_i^t$$

Where θ_t represents the global model parameters at time t, θ_i^t denotes the local model parameters of the i-th edge device at time t, and N is the total number of devices participating in the FL process.

This equation ensures that the model remains up-to-date by aggregating the knowledge of all participating devices while keeping the data privacy intact.

III. Performance Analysis

The performance of the proposed FL system integrated with digital twin technology is evaluated using CIC-IDS 2017 dataset to assess its effectiveness in detecting intrusions within Edge IoT networks. The system is tested over multiple rounds of federated training involving several edge devices to simulate a realistic distributed IoT environment. As the system continuously trains, the model parameters are updated iteratively to minimize the loss. The model update process is governed by the following gradient descent rule:

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L(\theta_t)$$

Where η is the learning rate, and $\nabla_\theta L(\theta_t)$ denotes the gradient of the loss function $L(\theta_t)$ with respect to the model parameters at time t.

Table I summarizes the comparative evaluation between the proposed system and existing FL-based intrusion detection approaches. The proposed system achieves an accuracy of 98.51%, which is significantly higher than the baseline models such as CNN-MLP [2], DCFL-Chain [1], and FCL-SBLS [4]. Similarly, the proposed method demonstrates improved precision 97.51%, recall 97.24%, and F1-score metrics 97.81%, highlighting its superior ability to correctly identify intrusion events while minimizing false alarms.

Fig. 2 illustrates the robustness of the proposed IDS model, depicting a consistent increase in detection accuracy as the number of edge IoT devices increases. This scalability indicates that the system can effectively maintain high detection performance in expanding IoT environments.

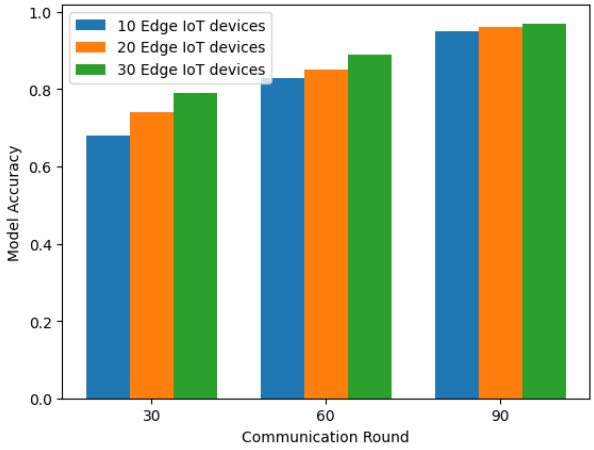

Fig. 2: Comparison of the accuracy of proposed IDS model with different number of IoT device.

IV. Conclusion

This paper presented a novel intrusion detection framework for Edge IoT networks based on the integration of FL and digital twin technology. The proposed system enables real-time, adaptive, and privacy-preserving detection of cyber threats, leveraging both local model training and simulated attack scenarios. The experimental results demonstrate that the system achieves a prediction accuracy of 98.51% while maintaining low latency and ensuring privacy. Future work will explore the use of more advanced machine learning models and optimization techniques to further enhance the system's performance and scalability.

Acknowledgment

This research was funded by the Innovative Human Resource Development for Local Intellectualization Program (IITP-2025-RS-2020-II201612, 34%) through IITP under MSIT, the Basic Science Research Program (2018R1A6A1A03024003, 33%) through NRF, and ITRC Program (IITP-2025-RS-2024-00438430, 33%) funded by MSIT through IITP.

References

[1] M. M. Alam, M. Golam, E. A. Tuli, M. R. Subhan, D.-S. Kim, and T. Jun, "Dcfl-chain: Digital-twin-based collaborative fl-integrated energy consumption prediction for smart factory," , pp. 310–311, 2024.

[2] A. Zainudin, R. Akter, D.-S. Kim, and J.-M. Lee, "Federated learning inspired low-complexity intrusion detection and classification technique for sdn-based industrial cps," *IEEE Transactions on Network and Service Management*, vol. 20, no. 3, pp. 2442–2459, 2023.

[3] M. M. Alam, G. Mohtasin, M. R. Subhan, D.-S. Kim, and T. Jun, "Federated semi-supervised digital twin for enhanced human-machine interaction in industry 5.0," in *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2024, pp. 1270–1275.

[4] X. He, Q. Chen, L. Tang, W. Wang, T. Liu, L. Li, Q. Liu, and J. Luo, "Federated continuous learning based on stacked broad learning system assisted by digital twin networks: An incremental learning approach for intrusion detection in uav networks," *IEEE Internet of Things Journal*, vol. 10, no. 22, pp. 19 825–19 838, 2023.