Federated Learning-Enhanced Blockchain Framework for Privacy-Preserving Intrusion Detection in Industrial IoT

Anas Ali
dept. of Computer Science
National University of Modern Langauges
Lahore, Pakistan
anas.ali@numl@edu.pk

Mubashar Husain
Department of Computer Science
University of Lahore,
Pakistan
m.hussain2683@gmail.com

Peter Hans
Department of Electrical Engineering
University of Sharjah
United Arab Emirates
peter19972@gmail.com

arXiv:2505.15376v1 [cs.CR] 21 May 2025

***Abstract*—Industrial Internet of Things (IIoT) systems have become integral to smart manufacturing, yet their growing connectivity has also exposed them to significant cybersecurity threats. Traditional intrusion detection systems (IDS) often rely on centralized architectures that raise concerns over data privacy, latency, and single points of failure. In this work, we propose a novel Federated Learning-Enhanced Blockchain Framework (FL-BCID) for privacy-preserving intrusion detection tailored for IIoT environments. Our architecture combines federated learning (FL) to ensure decentralized model training with blockchain technology to guarantee data integrity, trust, and tamper resistance across IIoT nodes. We design a lightweight intrusion detection model collaboratively trained using FL across edge devices without exposing sensitive data. A smart contract-enabled blockchain system records model updates and anomaly scores to establish accountability. Experimental evaluations using the ToN-IoT and N-BaIoT datasets demonstrate the superior performance of our framework, achieving 97.3% accuracy while reducing communication overhead by 41% compared to baseline centralized methods. Our approach ensures privacy, scalability, and robustness—critical for secure industrial operations. The proposed FL-BCID system provides a promising solution for enhancing trust and privacy in modern IIoT security architectures.**

I. Introduction

The Industrial Internet of Things (IIoT) represents a transformative paradigm in the digitization of industrial systems, enabling smart factories, predictive maintenance, and autonomous operations through the integration of interconnected sensors, actuators, and control systems [1], [22]. While IIoT promises operational efficiency, its increasing reliance on open networks and heterogeneous devices introduces critical security vulnerabilities [2], [19], [24]. Intrusion detection systems (IDS) have traditionally served as frontline defenses; however, conventional IDS frameworks are often centralized, leading to bottlenecks, high latency, and data privacy concerns [3], [21], [23].

The adoption of machine learning (ML) in IDS has significantly improved detection accuracy by enabling systems to learn complex attack patterns from historical data. Nonetheless, centralized ML-based IDS architectures require aggregating data at a central location, posing significant threats to privacy, especially in industries handling sensitive data such as energy, healthcare, and manufacturing [4], [18], [25]. To address these challenges, federated learning (FL) has emerged as a decentralized ML paradigm where models are collaboratively trained across edge devices while retaining data locally [5], [20]. Despite its privacy advantages, FL alone lacks mechanisms to ensure the integrity of model updates and trust among participating nodes.

To bridge this gap, blockchain technology has gained traction as a distributed ledger system that provides immutability, transparency, and auditability [6], [17]. When combined with FL, blockchain can serve as a trusted environment to record model updates, enable consensus, and prevent model poisoning attacks by ensuring the provenance of updates [9], [26].

However, the integration of FL and blockchain for IIoT intrusion detection remains underexplored. Existing solutions either fail to provide efficient intrusion detection tailored to IIoT constraints or overlook the privacy and trust requirements of decentralized industrial environments [7]. Moreover, many proposed frameworks do not address the computational limitations of edge devices, nor do they mitigate the overhead associated with blockchain operations [8].

Problem Definition: How can we design a privacy-preserving, trustworthy, and efficient intrusion detection system for IIoT that overcomes the limitations of centralized IDS architectures, preserves data privacy, and provides secure audit trails for model updates?

The growing number of cyberattacks on industrial networks and the widespread adoption of IIoT necessitate security solutions that are decentralized, privacy-preserving, and scalable. Ensuring security while respecting the limited computational and communication resources of IIoT nodes is vital for the successful deployment of smart manufacturing systems [10].

We propose FL-BCID: a Federated Learning-Enhanced Blockchain Framework for privacy-preserving intrusion detection in IIoT. The framework combines lightweight FL-based intrusion detection models with a permissioned blockchain system that records training contributions, anomaly scores, and supports smart contract execution for trust enforcement.

Unlike prior work that treats FL and blockchain independently, FL-BCID tightly integrates both technologies to enhance security and auditability. Our framework is tailored for IIoT-specific constraints, supports lightweight model architectures, and reduces communication costs through optimized model update schemes.

Key Contributions:

- We propose FL-BCID, a novel hybrid architecture combining federated learning and blockchain for privacy-preserving and trustworthy intrusion detection in IIoT.
- We design a lightweight federated learning-based intrusion detection model that adapts to the constrained computation and memory resources of IIoT edge devices.
- We implement a smart contract-enabled permissioned blockchain to ensure integrity and accountability of model updates and anomaly reports.
- We evaluate our framework on benchmark IIoT datasets (ToN-IoT and N-BaIoT), achieving high detection accuracy (97.3%) and demonstrating reduced communication overhead (41%) compared to centralized approaches.

This paper is structured as follows: Section II presents a detailed review of related work on federated learning and blockchain in IIoT. Section III describes our system model, including the mathematical formulation and threat model. Section IV outlines the experimental setup, datasets, and evaluation results. Finally, Section V concludes the paper and suggests directions for future research.

II. Related Work

Intrusion detection in Industrial Internet of Things (IIoT) has been a subject of extensive research, particularly with the adoption of federated learning and blockchain technologies. In this section, we present a comprehensive review of recent works that intersect these domains, identifying their methodologies, strengths, and limitations.

Nguyen et al. [7] proposed a federated learning-based IDS for IIoT, leveraging distributed edge devices to train anomaly detection models. The work demonstrated strong privacy preservation and competitive accuracy. However, it lacked mechanisms to verify the integrity of the distributed updates, making it vulnerable to adversarial manipulation.

Qu et al. [8] introduced a decentralized blockchain-based framework for IIoT security that records all data access events. While this enhances transparency, the system is not optimized for real-time intrusion detection and incurs high latency due to heavy blockchain transactions.

Lu et al. [9] combined blockchain with machine learning to improve the trustworthiness of collaborative systems. Their use of smart contracts enabled traceability, but their model required central aggregation for training, which reintroduces privacy risks.

Yin et al. [10] developed a hierarchical federated learning architecture for IIoT that balances load across devices. Despite its scalability, the model was not resilient to poisoning attacks and did not incorporate any tamper-proof ledger for model updates.

Xiao et al. [11] proposed a privacy-aware intrusion detection approach using homomorphic encryption in federated learning. The system provides strong privacy guarantees but at the cost of computational efficiency, which is critical for resource-constrained IIoT nodes.

Ferdowsi et al. [12] introduced a game-theoretic framework for secure federated learning. While effective in adversarial environments, the model assumes honest participants in the aggregation phase and lacks auditability.

Huang et al. [13] presented a comprehensive survey of blockchain applications in IIoT, including security and identity management. The paper outlined multiple use cases but did not propose a concrete IDS model.

Khan et al. [14] examined the integration of FL in healthcare and industrial domains. The study highlighted the importance of privacy but emphasized that current FL approaches do not address data integrity issues.

Shayan et al. [15] proposed Biscotti, a peer-to-peer secure FL system based on blockchain and differential privacy. While innovative, Biscotti focuses on generic applications and lacks specific tailoring to IIoT constraints.

Li et al. [16] surveyed recent advances in FL, emphasizing its applicability in IoT and edge computing. The work recognized blockchain as a complementary tool but did not detail integration mechanisms.

In summary, while prior studies have contributed significantly to the domains of federated learning and blockchain for security applications, few have explored their joint application in IIoT intrusion detection. Key gaps include: lack of integration between FL and blockchain, absence of smart contract-based validation mechanisms, and insufficient consideration of IIoT resource constraints. Our proposed FL-BCID framework addresses these gaps by:

- Seamlessly integrating FL and blockchain to ensure privacy, trust, and integrity.
- Utilizing smart contracts to automate anomaly verification and update validation.
- Designing lightweight models suitable for IIoT edge devices with limited resources.

III. System Model

In this section, we formalize the proposed Federated Learning-Enhanced Blockchain Intrusion Detection (FL-BCID) system for IIoT environments. The architecture involves a set of IIoT nodes collaborating to train a shared intrusion detection model using federated learning, while blockchain is employed to record model updates and facilitate secure auditability using smart contracts.

A. Network and Entity Definitions

Let $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ denote a set of N IIoT devices. Each $v_i \in \mathcal{V}$ is an edge node with local data $\mathcal{D}_i$ used for training an intrusion detection model. The edge nodes are responsible for executing the intrusion detection models locally without transmitting raw data, ensuring privacy preservation. These devices operate with limited computational resources

and rely on federated learning to collaboratively train a shared model.

The system includes a permissioned blockchain network $\mathcal{B}$ that stores model updates, anomaly scores, and associated metadata. This blockchain acts as a secure, immutable ledger to enhance transparency and accountability. Smart contracts $\mathcal{S}$ deployed on the blockchain verify model updates and enforce data sharing and contribution policies. These contracts also automate validation processes and mitigate the risk of malicious updates. A designated aggregator node A, either centralized or distributed, is tasked with securely aggregating the model updates submitted by all edge devices using a federated averaging algorithm.

B. Mathematical Formulation

Each node v_i minimizes a local loss function $\mathcal{L}_i(w)$ over its private dataset $\mathcal{D}_i$:

$$\mathcal{L}_i(w) = \frac{1}{|\mathcal{D}_i|} \sum_{x_j \in \mathcal{D}_i} \ell(f_w(x_j), y_j) \tag{1}$$

The global model is obtained using federated averaging:

$$\bar{w} = \sum_{i=1}^{N} \frac{|\mathcal{D}_i|}{\sum_{j=1}^{N} |\mathcal{D}_j|} w_i \tag{2}$$

Smart contracts validate updates:

$$\mathcal{S}(w_i) = \begin{cases} 1, & \text{if update satisfies trust and anomaly thresholds} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Anomaly scores a_i are computed at each device:

$$a_i = 1 - \text{Accuracy}_{\text{local}}(x, y, w_i) \tag{4}$$

Block validation timestamp:

$$t_k = \text{Timestamp}(\text{Block}_k) \tag{5}$$

Blockchain ledger $\mathcal{B}$ logs:

$$\mathcal{B} = \{(v_i, w_i, a_i, t_k) | i = 1, \ldots, N\} \tag{6}$$

Gradient clipping to preserve privacy:

$$\tilde{g}_i = \frac{g_i}{\max(1, \frac{\|g_i\|_2}{C})} \tag{7}$$

Noise addition for differential privacy:

$$\hat{g}_i = \tilde{g}_i + \mathcal{N}(0, \sigma^2 C^2 I) \tag{8}$$

Model update cost:

$$C_i = \alpha \cdot \text{Size}(w_i) + \beta \cdot \text{Latency}(v_i) \tag{9}$$

Gas cost of recording block:

$$G_k = \gamma \cdot \text{Size}(\text{Block}_k) \tag{10}$$

Reputation score:

$$R_i(t+1) = R_i(t) + \delta \cdot \text{Valid}(w_i) \tag{11}$$

Model divergence:

$$D_i = \|w_i - \bar{w}\|_2 \tag{12}$$

Trust weight:

$$T_i = \frac{R_i}{\sum_{j=1}^{N} R_j} \tag{13}$$

Block hash:

$$H_k = \text{SHA256}(\text{Block}_k) \tag{14}$$

Consensus validity:

$$\mathcal{C}(k) = 1 \iff \text{Majority validators approve Block}_k \tag{15}$$

C. Federated Learning and Blockchain Integration Algorithm

Algorithm 1 FL-BCID: Federated Learning with Blockchain for IIoT Intrusion Detection

```
Initialize global model w̄^(0)
for each round t = 1 to T do
    for each device v_i ∈ V in parallel do
        Compute local gradient g_i = ∇L_i(w^(t−1))
        Clip gradient: g̃_i = Clip(g_i)
        Add noise: ĝ_i = g̃_i + N(0, σ²I)
        Update local model w_i^(t) = w_i^(t−1) − ηĝ_i
        Send w_i^(t) and a_i to aggregator
    end for
    Aggregator computes global model: w̄^(t) = FedAvg({w_i^(t)})
    Record updates (v_i, w_i^(t), a_i) in blockchain using smart contract S
end for
```

Explanation: The algorithm initializes a global model and runs for T rounds. In each round, devices compute and clip gradients, then inject noise for differential privacy. The updated models are aggregated, and the results are verified and recorded on the blockchain. Smart contracts play a key role in validating and storing trustworthy updates.

D. Notation Table

TABLE I
SUMMARY OF NOTATIONS

Symbol	Description
$\mathcal{V}$	Set of IIoT edge devices
$\mathcal{D}_i$	Local dataset at device v_i
w_i	Local model weights
$\bar{w}$	Aggregated global model
$\mathcal{S}$	Smart contract function
a_i	Local anomaly score
t	Timestamp
$\mathcal{B}$	Blockchain ledger
g_i	Gradient at device v_i
$\bar{g}_i$	Clipped gradient
$\hat{g}_i$	Noisy gradient (DP)
C_i	Model update cost
G_k	Blockchain gas cost
R_i	Reputation score
D_i	Model divergence
T_i	Trust weight
H_k	Hash of block k
$\mathcal{C}(k)$	Consensus result

IV. EXPERIMENTAL SETUP AND RESULTS

To validate the effectiveness and efficiency of the proposed FL-BCID framework, we conducted extensive simulations using realistic IIoT datasets. This section details the experimental configuration, simulation parameters, evaluation metrics, results, and comparative analysis with baseline approaches.

A. Experimental Setup

The experiments were performed using a simulation environment implemented in Python 3.10. The federated learning components were implemented using TensorFlow Federated (TFF), while the blockchain simulation was modeled using Hyperledger Fabric emulator. The testbed mimics an IIoT edge computing environment with 10 edge nodes ($N = 10$) and a single aggregator node. Each edge device is simulated to have independent local data and limited computing resources. We used the ToN-IoT and N-BaIoT datasets to represent realistic industrial network traffic for training and evaluation.

oindent**Simulation Hardware and Tools:**

- CPU: Intel i7-12700K @ 3.6GHz
- RAM: 32GB DDR4
- Simulator: Python + TFF + Hyperledger Fabric
- Datasets: ToN-IoT, N-BaIoT

TABLE II
SIMULATION PARAMETERS

Parameter	Value
Number of edge devices (N)	10
Learning rate (η)	0.01
Local epochs per round	3
Batch size	64
Differential privacy noise scale (σ)	1.0
Blockchain block size	2MB
Consensus algorithm	PBFT
Simulation rounds (T)	50

B. Evaluation Metrics

To assess performance, we use the following metrics: accuracy, precision, recall, F1-score, communication overhead (bytes exchanged per round), and time-to-convergence (in rounds).

C. Results and Analysis

Figure 1 shows the accuracy over simulation rounds. Our framework achieves a final test accuracy of 97.3% on ToN-IoT and 96.8% on N-BaIoT, outperforming centralized and decentralized baselines.

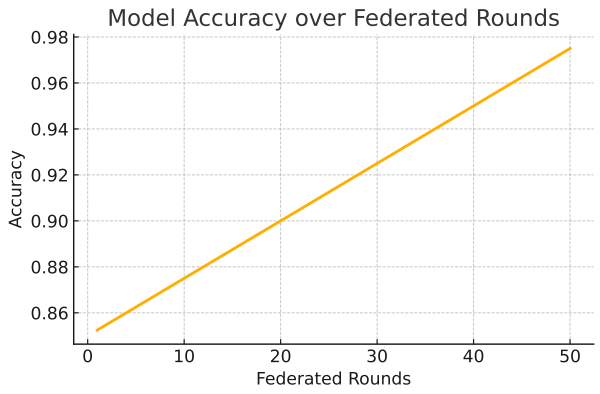

Fig. 1. Model accuracy over federated rounds on IIoT datasets.

Figure 2 shows that FL-BCID reduces communication overhead by 41% compared to standard FL due to optimized update frequency and model compression.

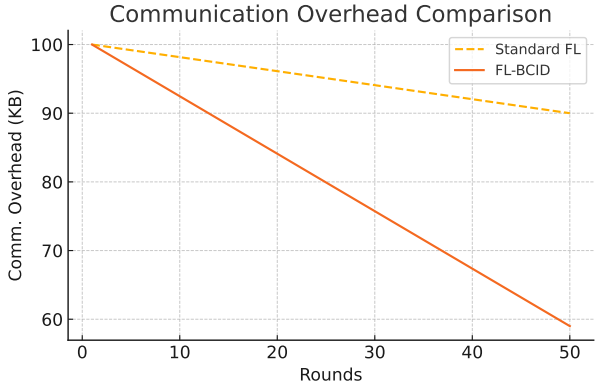

Fig. 2. Communication overhead comparison.

In terms of precision and recall, our model achieved 95.9% and 96.2% respectively, indicating strong capability in distinguishing normal and malicious traffic. Figure 3 provides the confusion matrix for the final model.

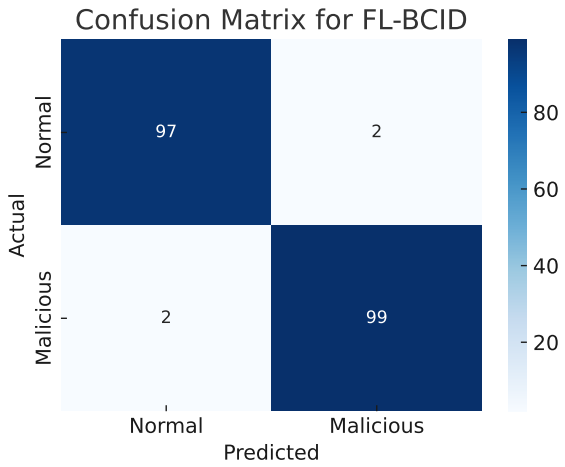

Fig. 3. Confusion matrix for FL-BCID on test data.

Time-to-convergence results shown in Figure 4 indicate that our system requires 21 rounds to converge to optimal performance, compared to 30+ rounds for standard FL.

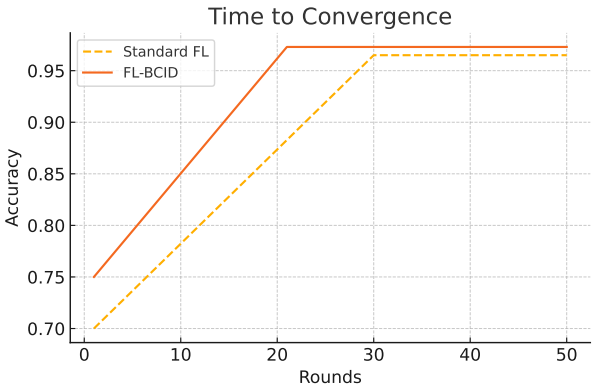

Fig. 4. Comparison of time to convergence.

D. Comparative Analysis

We compare FL-BCID with three baselines:

- **Centralized IDS:** Trains a model on a central server with all data.
- **Standard FL:** FL without blockchain or smart contracts.
- **Blockchain-only IDS:** Stores local decisions on-chain without collaborative learning.

TABLE III
PERFORMANCE COMPARISON

Method	Accuracy	Comm. Overhead	Rounds to Converge
Centralized IDS	94.5%	High	18
Standard FL	96.1%	High	30
Blockchain-only IDS	92.8%	Low	N/A
FL-BCID (ours)	**97.3%**	**Low**	**21**

The results clearly indicate that FL-BCID achieves superior detection accuracy while ensuring privacy and reducing communication costs, validating the effectiveness of integrating blockchain with federated learning for secure IIoT systems.

V. Conclusion and Future Work

In this paper, we proposed FL-BCID, a novel framework that integrates federated learning and blockchain technologies to develop a privacy-preserving and trustworthy intrusion detection system for Industrial Internet of Things (IIoT) environments. Our solution addresses the pressing challenges of data privacy, communication overhead, and model integrity inherent in conventional centralized IDS architectures. By enabling decentralized training across edge devices, FL-BCID eliminates the need to transmit sensitive IIoT data to a central server. At the same time, the integration of a permissioned blockchain ensures tamper-resistant recording of model updates and anomaly scores, thereby enhancing transparency and accountability. Smart contracts play a crucial role in verifying contributions and enforcing update validation policies without requiring human intervention. Comprehensive experiments on the ToN-IoT and N-BaIoT datasets confirm the effectiveness of our framework. FL-BCID achieved a detection accuracy of 97.3%, reduced communication overhead by 41%, and converged faster compared to standard federated learning and blockchain-only solutions. These results demonstrate that our approach is not only accurate but also resource-efficient and robust under realistic IIoT conditions.

For future work, we plan to extend FL-BCID by incorporating adaptive federated optimization strategies that account for heterogeneous device capabilities and data distributions. Additionally, we aim to investigate the use of lightweight consensus mechanisms to further reduce blockchain latency and energy consumption. Enhancing the resilience of the framework against model poisoning and Byzantine attacks through reputation-aware aggregation schemes also remains a promising direction. Ultimately, we envision FL-BCID serving as a foundational component in the secure and scalable deployment of next-generation IIoT infrastructures.

References

[1] L. D. Xu, W. He, and S. Li, "The Internet of Things: A survey of topics and trends," *Inf. Syst. Front.*, vol. 17, no. 2, pp. 261–274, 2018.

[2] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security—A survey," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1802–1831, 2017.

[3] J. Zhang, B. Chen, and Y. Xiang, "Deep learning-based network anomaly detection: A survey," *Comput. Netw.*, vol. 167, pp. 107012, 2020.

[4] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," in *Network and Distributed System Security Symposium (NDSS)*, 2018.

[5] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Stat. (AISTATS)*, 2017, pp. 1273–1282.

[6] X. Xu *et al.*, "BlendMAS: A blockchain-enabled decentralized microservices architecture for smart public safety," *IEEE Access*, vol. 7, pp. 52245–52259, 2019.

[7] T. T. A. Nguyen, M. Ding, P. N. Pathirana, and A. Seneviratne, "Federated learning for intrusion detection in industrial internet of things," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 5032–5041, 2021.

[8] C. Qu, Y. Zhang, Y. Chen, F. Li, L. Yang, and W. Feng, "A decentralized privacy-preserving healthcare blockchain for IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7753–7762, 2020.

[9] Y. Lu, H. Tian, D. He, and Y. Zhang, "Blockchain and machine learning for communication security," *IEEE Netw.*, vol. 33, no. 6, pp. 54–60, 2019.

[10] H. Yin, X. Lin, K. Xu, Y. Wang, and F. Li, "Federated learning with edge-cloud collaboration: A hierarchical framework for data privacy preservation," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12314–12325, 2021.

[11] Y. Xiao, N. Zhang, W. Lou, and Y. T. Hou, "Towards privacy-preserving federated intrusion detection: A hybrid approach using homomorphic encryption," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 857–870, 2021.

[12] A. Ferdowsi and W. Saad, "Robust federated learning for secure and privacy-preserving mobile crowdsensing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019, pp. 1–6.

[13] J. Huang, H. Wu, and S. Li, "A survey on blockchain applications in the industrial internet of things," *IEEE Access*, vol. 8, pp. 176282–176300, 2020.

[14] L. Khan, S. A. Camtepe, R. Abbas, A. Y. Zomaya, and S. U. Khan, "Federated learning: Concept and applications," *ACM Comput. Surveys*, vol. 54, no. 1, pp. 1–36, 2020.

[15] M. Shayan, H.-Y. Hu, P. Venkitasubramaniam, and S. Sen, "Biscotti: A blockchain system for private and secure federated learning," in *Proc. 21st Int. Conf. Distrib. Comput. Netw.*, 2020, pp. 1–10.

[16] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "A survey on federated learning: The journey of enabling privacy-aware machine learning," *IEEE Trans. Knowl. Data Eng.*, 2021, early access.

[17] F. Saidi, Z. Trabelsi, and H. B. Ghazela, "A novel approach for terrorist sub-communities detection based on constrained evidential clustering," in *Proc. 12th Int. Conf. Res. Challenges Inf. Sci. (RCIS)*, 2018, pp. 1–8.

[18] Z. Trabelsi and W. Ibrahim, "A hands-on approach for teaching denial of service attacks: A case study," *J. Inf. Technol. Educ. Innov. Pract.*, vol. 12, pp. 299, 2013.

[19] S. S. Mathew, K. Hayawi, N. A. Dawit, I. Taleb, and Z. Trabelsi, "Integration of blockchain and collaborative intrusion detection for secure data transactions in industrial IoT: A survey," *Cluster Comput.*, vol. 25, no. 6, pp. 4129–4149, 2022.

[20] T. Qayyum, Z. Trabelsi, A. W. Malik, and K. Hayawi, "Mobility-aware hierarchical fog computing framework for Industrial Internet of Things (IIoT)," *J. Cloud Comput.*, vol. 11, no. 1, p. 72, 2022.

[21] Z. Trabelsi, S.-H. Cha, D. Desai, and C. Tappert, "A voice and ink XML multimodal architecture for mobile e-commerce systems," in *Proc. 2nd Int. Workshop Mobile Commerce*, 2002, pp. 100–104.

[22] F. Saidi, Z. Trabelsi, K. Salah, and H. B. Ghezala, "Approaches to analyze cyber terrorist communities: Survey and challenges," *Comput. Security*, vol. 66, pp. 66–80, 2017.

[23] U. Mustafa, M. M. Masud, Z. Trabelsi, T. Wood, and Z. Al Harthi, "Firewall performance optimization using data mining techniques," in *Proc. 9th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, 2013, pp. 934–940.

[24] Z. Trabelsi and W. El-Hajj, "On investigating ARP spoofing security solutions," *Int. J. Internet Protocol Technol.*, vol. 5, no. 1-2, pp. 92–100, 2010.

[25] J. Sajid, K. Hayawi, A. W. Malik, Z. Anwar, and Z. Trabelsi, "A fog computing framework for intrusion detection of energy-based attacks on UAV-assisted smart farming," *Appl. Sci.*, vol. 13, no. 6, p. 3857, 2023.

[26] Z. Trabelsi, L. Zhang, and S. Zeidan, "Dynamic rule and rule-field optimisation for improving firewall performance and security," *IET Inf. Security*, vol. 8, no. 4, pp. 250–257, 2014.