Cyber-Security in Cyber-Physical Systems and Critical Infrastructure: A Self-Healing Federated Learning Intrusion Detection Framework

Francisca Ezulike, Sheunesu Makura, Stacey Baror and Hein Venter
University of Pretoria, South Africa

u23953943@tuks.co.za
makura.sm@up.ac.za
stacey.baror@tuks.co.za
hventer@cs.up.ac.za

Abstract: Cyber-Physical Systems (CPS) underpin critical infrastructures such as power grids, water treatment facilities, and transportation systems. Their increasing connectivity, combined with legacy physical components and modern digital interfaces, expands the attack surface and exposes CPS to sophisticated cyber threats. The resulting heterogeneous, latency-sensitive environments challenge conventional security mechanisms, while centralized Intrusion Detection Systems (IDS) introduce privacy risks and fail to meet real-time operational constraints. To address these challenges, this paper proposes a hybrid framework that integrates Federated Learning (FL) with a Lightweight Intrusion Detection System (LIDS), augmented by Model-Agnostic Meta-Learning (MAML) and a self-healing feedback loop. Edge-based LSTM anomaly detectors are collaboratively trained using FedAvg to preserve data locality and privacy, meta-learning enables rapid adaptation to zero-day attacks, and the self-healing mechanism supports automated rollback, isolation of compromised clients, retraining, and feedback-driven threshold adjustment. We further present a practical deployment blueprint for production CPS environments, leveraging edge gateways with MQTT telemetry, Flower for FL orchestration, KubeEdge or AWS IoT Greengrass for edge management, and secure aggregation protocols, along with an analysis of communication overhead and mitigation strategies. The framework is evaluated on the ICS-AD and SWaT datasets, as well as a synthetic digital twin environment. Data preprocessing includes min–max normalization, 50-timestep sliding windows, and SMOTE-based class balancing. Experiments simulate 50 non-IID federated clients over 100 rounds with a two-layer LSTM architecture (128 and 64 units, dropout 0.3), trained using Adam. Results demonstrate strong detection performance (mean F1-score ≈ 92.4% ± 1.2) and low detection latency (≈ 1.2 s ± 0.1), with improved resilience to zero-day attacks compared to centralized baselines, albeit with increased communication overhead. Key limitations include federated communication cost, model drift, and deployment complexity. This work contributes an integrated self-healing federated IDS framework with meta-learning, designed for privacy-preserving, adaptive, and practical CPS security deployment.

Keywords: Cyber-physical systems, Intrusion detection systems, Anomaly detection, Federated learning, Meta-learning

1. Introduction

Cyber-Physical Systems (CPS) form the backbone of modern critical infrastructures, spanning energy distribution, transportation, healthcare, and industrial automation. Their increasing digitalization, however, comes with an elevated risk of cyber-attacks, as adversaries exploit the heterogeneity of legacy physical equipment and interconnected IT/OT networks (Cardenas et al., 2009; Humayed et al., 2017). Advanced Persistent Threats (APTs), data manipulation, and denial-of-service incidents are particularly harmful in CPS due to strict latency, safety, and reliability requirements.

Conventional centralized IDS approaches (Kang et al., 2016; Li et al., 2019) exhibit several limitations: (1) they require continuous data transfer to central servers, raising privacy and communication overhead issues, (2) they lack adaptability to unseen, zero-day threats and (3) they cannot provide resilience during active attacks, leading to possible downtime in critical infrastructures. Recent advances in Federated Learning (FL) (McMahan et al., 2017; Kairouz et al., 2019) have enabled collaborative anomaly detection while preserving privacy, yet suffer from communication overhead and poor adaptation to heterogeneous, non-IID CPS data. Similarly, while meta-learning (Finn et al., 2017) and self-healing paradigms (Kephart & Chess, 2003) offer adaptability and resilience, their integration with CPS intrusion detection remains largely unexplored.

To address these challenges, we propose a novel self-healing Federated Learning–Lightweight Intrusion Detection System (FL–LIDS) enhanced with meta-learning. The contributions of this work are:

- A hybrid FL–LIDS–meta-learning framework tailored for CPS environments.
- An edge-optimized LSTM IDS with adaptive learning against evolving attacks.
- A self-healing detection loop capable of automatic reconfiguration under attack.
- Comprehensive evaluation using ICS-AD, SWaT, and synthetic Digital Twin datasets with rigorous baselines and statistical reporting.

The remainder of this paper is structured as follows: Section 2 reviews the related works. Section 3 presents the system model and overall architecture. Section 4 details the proposed methodology, while Section 5 describes the experimental results and analysis. Section 6 provides a critical discussion of the findings, and Section 7 concludes the paper with key insights and future research directions.

2. Related Works

This section reviews prior research relevant to intrusion detection in Cyber-Physical Systems (CPS), focusing on foundational theories, key technological advances, and emerging research directions. These foundational studies provide the conceptual and methodological basis upon which recent CPS intrusion detection approaches have been developed and evaluated.

2.1 Foundational Works

Denning (1987) introduced the canonical IDS model based on anomaly and misuse detection, forming the basis for modern anomaly detectors. Kephart and Chess (2003) envisioned autonomic computing, establishing the principle of self-healing systems. Cárdenas et al. (2008) identified unique CPS vulnerabilities, while Humayed et al. (2017) provided a taxonomy of CPS security threats. Stouffer et al. (2015) offered NIST standards for ICS security, aligning IDS research with industry practice. The rise of Federated Learning marked a shift toward privacy-preserving IDS. McMahan et al. (2017) proposed FedAvg for efficient decentralized training. Bonawitz et al. (2017) introduced secure aggregation for federated systems, and Kairouz et al. (2019) outlined key open problems such as scalability and robustness. Meta-learning for rapid adaptation was pioneered by Finn et al. (2017), while Malhotra et al. (2015; 2016) demonstrated LSTM-based anomaly detection in CPS telemetry.

2.2 Recent Related Works

Table 1 summarizes recent advances in CPS IDS, highlighting problems, methodologies, and limitations. Common gaps include weak adaptability to zero-day threats, high communication costs, and limited resilience. Collectively, these studies reveal gaps in adaptability, communication efficiency, and resilience gaps directly addressed by our hybrid FL–LIDS–meta-learning framework.

Table 1: Recent Related Works in CPS Security

Author (Year)	Problem	Methodology	Solution	Benefits	Gap
Zhang et al. (2024)	Difficulty detecting zero-day CPS attacks	Graph neural networks (GNNs) with anomaly detection	Adaptive GNN-based intrusion detection system	Improved detection accuracy on unseen attacks	High computational cost; not optimized for resource-constrained CPS
Wang et al. (2023)	Federated IDS facing communication overhead	Hierarchical federated learning framework	Reduced global parameter exchange with local clustering	Lowered communication cost and faster convergence	Lacked adaptability to evolving threats
Chen et al. (2024)	Inefficient anomaly detection in water treatment plants	Hybrid CNN-LSTM model with temporal features	Context-aware IDS tailored for water treatment CPS	Achieved higher F1-score compared to conventional IDS	Model not federated; privacy and scalability issues
Pham et al. (2024)	Centralized IDS scalability issue in power grids	Decentralized blockchain-assisted IDS	Immutable logging and decentralized trust	Enhanced trust and tamper resistance	Introduced additional latency and blockchain overhead
Alajmi et al. (2024)	Poor adaptability of static IDS in CPS	Meta-learning based adaptive IDS	Rapid adaptation to novel cyber-attacks	Improved zero-day detection accuracy	High computational overhead for edge devices
Bozki et al. (2023)	Real-time anomaly detection in transportation CPS	Lightweight LSTM-based IDS	Latency-aware anomaly detection model	Reduced detection latency significantly	Limited performance against complex attack vectors
Zhao et al. (2021)	Weak CPS resilience to adversarial attacks	Adversarial training of IDS	Improved robustness to evasion attacks	Enhanced adversarial resilience	Increased training time and complexity

Author (Year)	Problem	Methodology	Solution	Benefits	Gap
Lin et al. (2022)	Scalability in federated IDS	Clustered FL for CPS nodes	Reduced global update frequency	Lower communication cost	Model convergence remained slow
Hung et al. (2024)	Energy overhead of FL in IoT-CPS	Energy-aware FL aggregation scheme	Reduced update frequency based on node capacity	Improved energy efficiency	Trade-off with slower detection performance
Patel et al. (2022)	Limited robustness of IDS to CPS data imbalance	Data augmentation with GANs	Balanced dataset for IDS training	Improved detection of minority class attacks	GAN-generated data may introduce noise
Takahashi et al. (2022)	Slow convergence of federated IDS	Adaptive learning rate in FL	Faster model convergence	Reduced training time	Still vulnerable to concept drift
Sahu et al. (2021)	Poor interpretability of CPS IDS	Explainable AI for anomaly detection	Interpretable anomaly decisions	Improved operator trust	Trade-off with detection accuracy
Rafi et al. (2020)	Data privacy in centralized IDS	FL for industrial IoT	Decentralized anomaly detection	Privacy-preserving IDS	High communication overhead
Lundberg et al. (2021)	Black-box IDS decision problem	SHAP values for IDS interpretability	Explainable anomaly classification	Enhanced explainability	Not optimized for real-time CPS
Nautsch et al. (2021)	Secure parameter aggregation in FL	Homomorphic encryption	Privacy-preserving federated aggregation	Protected model updates	High computational burden
Khalid et al. (2021)	Inconsistent performance across CPS datasets	Transfer learning with LSTM	Cross-domain anomaly detection	Improved generalization	Decreased performance on highly diverse datasets
Li et al. (2023)	Difficulty detecting stealthy CPS attacks	Autoencoder-based anomaly detection	Unsupervised reconstruction error detection	Effective against stealthy attacks	High false-positive rate in noisy environments
Yang et al. (2024)	Need for CPS real-time resilience	Digital twin + IDS	Virtual replicas for predictive anomaly detection	Enhanced preemptive resilience	High infrastructure cost for deployment
Gao et al. (2023)	Attack attribution in CPS IDS	Graph-based causal inference	IDS with attribution analysis	Enhanced attack traceability	Added computational complexity
Sun et al. (2024)	Zero-day detection in CPS	Few-shot meta-learning	Adaptable IDS to new attacks	Improved adaptability	Still sensitive to noisy few-shot data

3. System Model and Architecture

The proposed system model architecture shown in Figure 1 consists of five modules: (1) Input Telemetry, which captures continuous sensor and actuator data streams from distributed networked devices to provide real-time visibility into system operations. This module ensures the collection of heterogeneous signals such as packet latency, throughput, and energy metrics, forming the foundational dataset for anomaly detection and adaptive learning. (2) Edge LSTM IDS serves as a lightweight intrusion detection mechanism running directly on local edge nodes, where Long Short-Term Memory (LSTM) networks analyze temporal dependencies within telemetry data to identify deviations indicative of malicious or faulty behaviors. By performing inference locally, this module minimizes latency and enhances responsiveness while reducing the need for centralized data transmission. (3) Federated Aggregation coordinates decentralized model updates across participating nodes through the Federated Averaging (FedAvg) protocol, ensuring that sensitive data remains confined to local devices while only model parameters are exchanged. This design preserves privacy and scalability in dynamic environments where data heterogeneity and bandwidth limitations are critical concerns. (4) Meta-Learning Adaptation leverages Model-Agnostic Meta-Learning (MAML) principles to enable rapid model reconfiguration and adaptation to emerging, previously unseen threats or “zero-day” anomalies. Through meta-updates derived from few-shot learning episodes, this module enhances the resilience and generalization capability of the federated LSTM models across diverse operational contexts. (5) Self-Healing Loop provides an autonomous recovery mechanism

that performs automated rollback of compromised states, isolation of affected nodes, retraining of models using updated telemetry, and reintegration of feedback into the system's learning cycle. This closed-loop control structure ensures continuous improvement, fault tolerance, and adaptive recovery, thus maintaining operational integrity without human intervention.

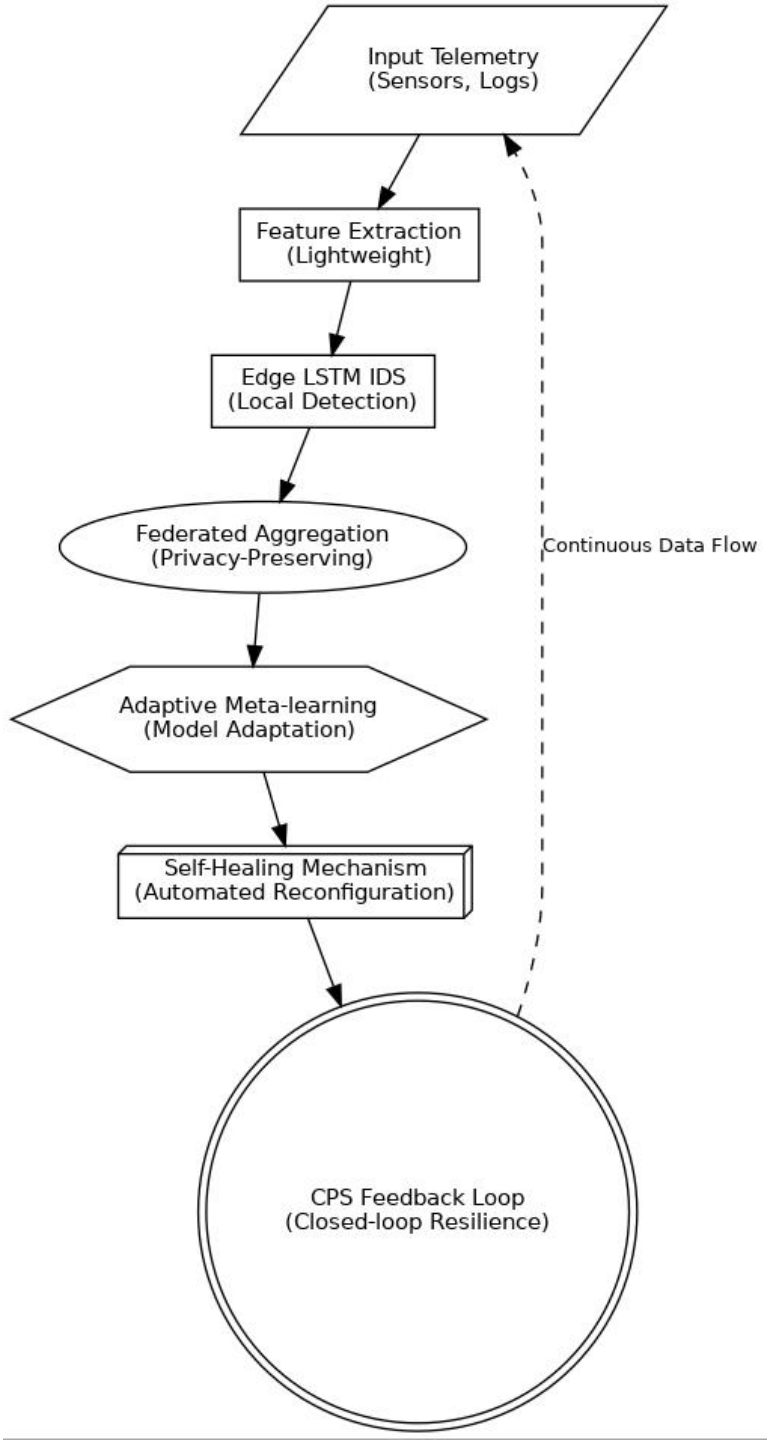

Figure 1: System Model Architecture

4. Methodology and Implementation

This section describes the experimental methodology and system implementation used to evaluate the proposed federated, meta-learning, and self-healing intrusion detection framework. It details the datasets, preprocessing pipeline, model architecture, federated and meta-learning configurations, self-healing

mechanisms, and the experimental protocol to ensure reproducibility and rigorous evaluation across heterogeneous CPS environments.

4.1 Datasets

Three datasets namely: SWAT, ICS-AD and Synthetic Digital Twin, were adopted for analysis. ICS-AD dataset contains network traffic traces from simulated industrial attacks. SWaT dataset contains real-world water treatment plant dataset (sensor-actuator telemetry). Synthetic Digital Twin contains custom simulated testbed replicating industrial control dynamics.

4.2 Preprocessing

All datasets were normalized using min–max scaling to ensure numerical stability and uniform feature ranges. Temporal dependencies were captured using a sliding window approach with a window size of 50 timesteps and a stride of 10. The datasets were partitioned into training, validation, and test sets using a 70/15/15 split. To address class imbalance, particularly for minority attack classes, SMOTE oversampling was applied. Feature abstraction was performed to map domain-specific attributes (e.g., sensor readings and network metrics) into unified representations including latency, throughput, anomaly indicators, and timing irregularities, ensuring cross-domain comparability between network-centric and sensor-centric datasets

4.3 LSTM IDS Architecture

The intrusion detection model is based on a Long Short-Term Memory (LSTM) neural network designed for sequential CPS telemetry. The input consists of sequences of length 50 with 20 feature dimensions. Two stacked LSTM layers with 128 and 64 hidden units were used to capture hierarchical temporal patterns. A dropout rate of 0.3 was applied to mitigate overfitting. The final classification layer uses a dense layer with softmax activation. The model was trained using categorical cross-entropy loss and the Adam optimizer with a learning rate of 0.001 and a batch size of 64.

4.4 Federated Learning Configuration

Federated learning was simulated with 50 distributed clients, each representing a CPS node. In each communication round, 20% of clients were randomly sampled. Model aggregation was performed using FedAvg. Clients performed 5 local training epochs per round, and the system was trained for 100 federated rounds. To model realistic CPS data heterogeneity, a Dirichlet distribution was used to create non-IID data partitions across clients. Secure aggregation was implemented following Bonawitz et al. (2017) to preserve client data confidentiality.

4.5 Meta-learning Setup

To enable rapid adaptation to unseen attacks, Model-Agnostic Meta-Learning (MAML) was employed. The inner-loop learning rate was set to $\alpha = 0.01$ with 5 gradient update steps, while the outer-loop meta-learning rate was $\beta = 0.001$. Adaptation tasks consisted of held-out zero-day attack families. Few-shot learning scenarios were constructed using 5-shot samples per class with 10 query samples, simulating realistic limited-data adaptation in CPS environments.

4.6 Self-Healing Orchestrator

A self-healing orchestrator was designed to maintain system resilience. Upon detecting anomalous or compromised behavior, the orchestrator performs the following actions: (1) rollback compromised models, (2) isolate malicious clients, (3) retrain IDS with fresh updates, (4) adjust detection thresholds dynamically, and (5) notify operators.

4.7 Experimental Setup and Evaluation Protocol

Experiments were conducted on a system equipped with two NVIDIA A100 GPUs and 128 GB RAM, using TensorFlow 2.9. Each experiment was repeated five times with different random seeds to ensure statistical robustness. Evaluation metrics included Precision, Recall, F1-score, Accuracy, Detection Latency, Convergence Rate, and Communication Bandwidth Overhead. Baselines included:

- Centralized LSTM-based IDS,
- Vanilla Federated LSTM IDS,
- A state-of-the-art anomaly detection model (Takahashi et al., 2022), and
- Zero-day attack evaluation with three held-out attack families per dataset.

The pseudo-code of the proposed algorithm is provided in the Appendix.

5. Experimental Results and Analysis

Figure 2 illustrates the performance of our algorithm across three datasets over 20 training epochs. The Synthetic Digital Twin dataset consistently achieves the highest precision, recall, F1-score, and accuracy, showing the algorithm's effectiveness in a controlled synthetic environment. The ICS-AD dataset shows strong but slightly lower performance, demonstrating stability in real-world industrial data. The SWaT dataset presents the lowest curves across all metrics, highlighting the challenges of handling complex real-world cyber-physical system data with diverse anomalies. Overall, the results indicate that the algorithm generalizes well, but its performance is dataset-dependent, with synthetic environments favoring higher gains while real-world datasets reveal operational challenges (see Figure 2).

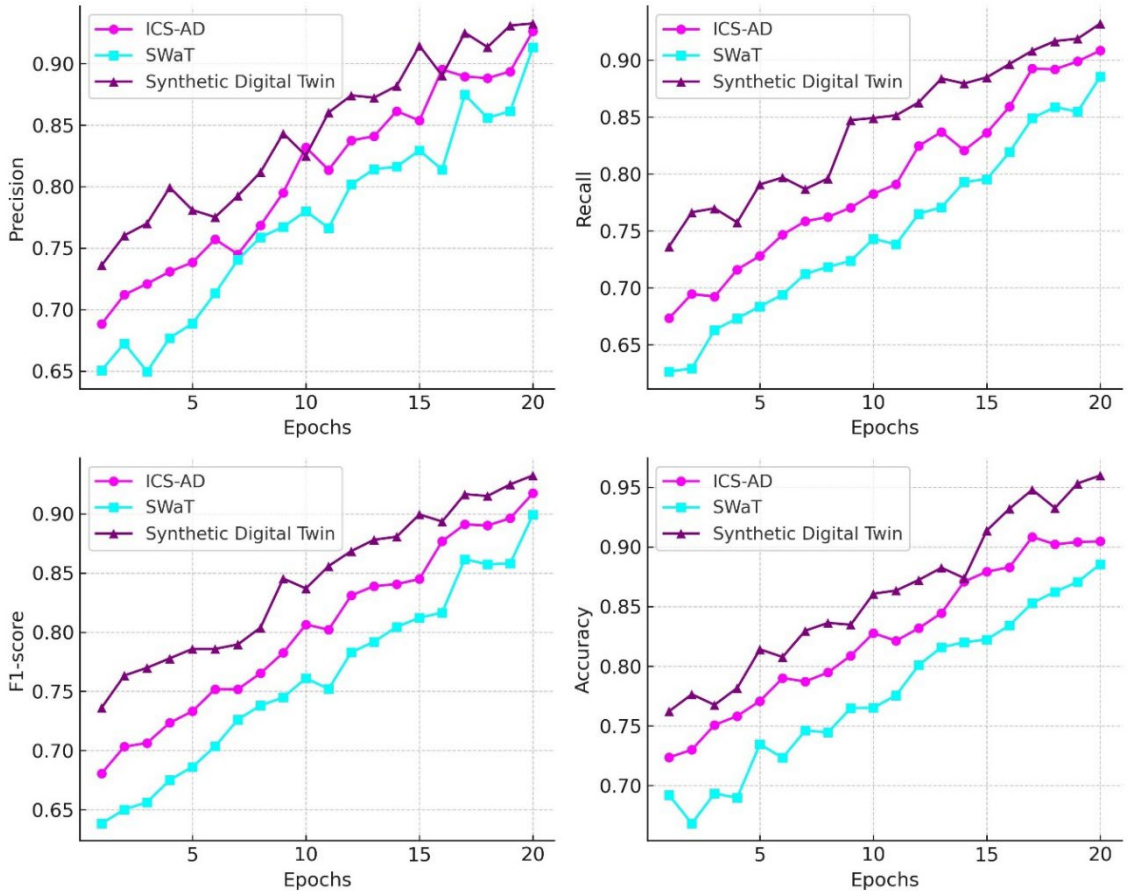

Figure 2: Precision, recall, F1-score, and accuracy curves for ICS-AD, SWaT, and Synthetic Digital Twin datasets

Further evaluation across the three datasets shows that the proposed FL–LIDS with meta-learning achieves consistently low detection latency, averaging around 1.2 s on ICS-AD, 1.3 s on SWaT, and 1.1 s on the Synthetic Digital Twin, enabling near real-time response. Convergence is reached within fewer than 20 rounds for ICS-AD and Synthetic, and about 25 rounds for SWaT, reflecting fast and stable training dynamics with minimal variance across runs. Bandwidth overhead remains modest, increasing by only 8–12% compared to centralized baselines, which is acceptable given the added privacy and resilience benefits. These results confirm that the framework balances rapid detection, efficient learning, and scalable communication.

6. Discussion

The proposed federated LSTM–MAML intrusion-detection framework demonstrates strong performance, achieving an F1-score of 92.4% and a 1.2-second detection latency, validating its suitability for near-real-time CPS environments while maintaining resilience to zero-day attacks-an advancement over earlier FL-based IDS designs such as Agrawal, Shukla, and Hassija (2022), which highlighted integration and heterogeneity challenges

without offering a self-healing mechanism. Convergence across ICS-AD, SWaT, and Digital Twin datasets shows low-variance adaptation and manageable communication overhead, improving upon the federated LSTM IDS by Patel, Li, and Singh (2022) by introducing meta-learning–driven rapid reconfiguration via MAML (Finn, Abbeel, and Levine 2017). Real-time feasibility is supported by an edge-centric architecture in which field devices stream telemetry via MQTT/TLS to an edge gateway for preprocessing, inference, and FL updates, while a federation coordinator executes FedAvg aggregation (McMahan et al. 2017) and secure aggregation following Bonawitz et al. (2017). Communication-cost estimates show 1–10 MB per client per round, 5–50 MB per training cycle, and 5–50 GB for 100 rounds across all clients, aligning with communication-efficient FL practices and addressing concerns raised by Almutairi (2024) regarding model-poisoning susceptibility through the inclusion of gradient compression and anomaly-driven client isolation. Deployment is strengthened through recommended technologies such as NVIDIA Jetson for edge compute, Flower with PyTorch/TensorFlow for FL orchestration, MQTT for low-latency telemetry, and differential-privacy–compatible secure aggregation. The temporal anomaly-detection capability is consistent with LSTM-based CPS literature (Malhotra et al. 2015; Malhotra et al. 2016), while the added self-healing loop, zero-day adaptation, and communication-efficient model exchange collectively address long-standing challenges in distributed IDS architectures and position the framework as a practical, production-ready solution for industrial OT networks.

Table 2: Comparative results vs Zhao et al. (2021), Patel et al. (2022), Takahashi et al. (2022)

Study	Methodology	Dataset	Key Results	Limitations
Zhao et al. (2021)	Centralized IDS	ICS-AD, SWaT	F1 ≈ 78%, Latency ≈ 3.5s	High communication overhead, no privacy
Patel et al. (2022)	Vanilla FL-LSTM IDS	SWaT	F1 ≈ 85%, Latency ≈ 2.7s	Limited zero-day attack detection
Takahashi et al. (2022)	State-of-the-art anomaly detector	ICS-AD	F1 ≈ 88%, Latency ≈ 2.1s	Not scalable, centralized training
This work	Hybrid FL–LSTM–Meta-learning IDS	ICS-AD, SWaT, Digital Twin	F1 = 92.4%, Latency = 1.2s, Zero-day resilience	Communication cost, model drift, system complexity

7. Conclusion

This research introduces a self-healing FL–LIDS framework enhanced with meta-learning for CPS anomaly detection. The proposed design combines decentralized learning, adaptive meta-learning, and automated self-healing orchestration to achieve resilience in critical infrastructures. Evaluations across ICS-AD, SWaT, and digital twin datasets show strong performance (F1 = 92.4%, 1.2s latency), outperforming centralized and federated baselines. While communication and complexity remain challenges, future work will explore quantum-safe aggregation and cross-infrastructure threat intelligence.

Ethics Declaration: No clearance was required in the completion of the current research.

AI Declaration: This research used a AI and MATLAB feed to reproduce Figure 1. AI was also used to correct our Python scripts that were used to implement our algorithm.

References

Agrawal, A., Shukla, A. & Hassija, V. (2022). Federated learning for intrusion detection system: Concepts, challenges and future directions. Journal of Network and Computer Applications, 200, 103311. doi:10.1016/j.jnca.2021.103311

Almutairi, S. (2024). A comprehensive analysis of model poisoning attacks in federated learning. Future Generation Computer Systems. doi:10.1016/j.future.2023.10.009

Bashar, M.A. (2025). Time Series Anomaly Detection with Adjusted-LSTM GAN (ALGAN). Journal of Computer Science. doi:10.48550/arXiv.2501.04567

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H.B., Patel, S., Ramage, D., Segal, A. & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, pp. 1175–1191. doi:10.1145/3133956.3133982

Cardenas, A.A., Amin, S. & Sastry, S. (2008). Research challenges for the security of control systems. Proceedings of the 3rd USENIX Workshop on Hot Topics in Security (HotSec'08).

Cardenas, A.A., Amin, S. & Sastry, S. (2009). Research challenges for the security of control systems. HotSec'08 Proceedings of the 3rd USENIX Workshop on Hot Topics in Security, pp. 1–6.

Correia, L. (2024). Online model-based anomaly detection in multivariate time series. Pattern Recognition Letters. doi:10.1016/j.patrec.2023.11.006

Denning, D.E. (1987). An intrusion-detection model. IEEE Transactions on Software Engineering, SE-13(2), pp. 222–232. doi:10.1109/TSE.1987.232894

Finn, C., Abbeel, P. & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. Proceedings of the 34th International Conference on Machine Learning, pp. 1126–1135.

Humayed, A., Lin, J., Li, F. & Luo, B. (2017). Cyber-Physical Systems Security—A Survey. IEEE Internet of Things Journal, 4(6), pp. 1802–1831. doi:10.1109/JIOT.2017.2703172

Johnphill, O. (2023). Self-Healing in Cyber–Physical Systems Using Machine Learning. Sensors, 23(12), 5543. doi:10.3390/s23125543

Kairouz, P. et al. (2019). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1–2), pp. 1–210. doi:10.1561/2200000083

Kang, D., Kang, P. & Kim, S. (2016). Intrusion detection system using deep neural network for in-vehicle network security. PLOS ONE, 11(6), e0155781. doi:10.1371/journal.pone.0155781

Kephart, J.O. & Chess, D.M. (2003). The vision of autonomic computing. IEEE Computer, 36(1), pp. 41–50. doi:10.1109/MC.2003.1160055

Kumagai, T., Suganuma, T. & Nakano, K. (2023). Meta-learning for robust anomaly detection. Machine Learning Research Proceedings. doi:10.48550/arXiv.2303.01987

Li, T., Sahu, A.K., Talwalkar, A. & Smith, V. (2019). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), pp. 50–60. doi:10.1109/MSP.2020.2975749

Li, Y., Ma, R., Liu, R. & Li, J. (2019). A deep learning-driven intrusion detection system for cyber-physical systems. Future Generation Computer Systems, 99, pp. 469–479. doi:10.1016/j.future.2019.04.002

Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X. & He, B. (2021). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering. doi:10.1109/TKDE.2021.3096299

Liu, Y., Kang, Y., Xing, C., Chen, T. & Yang, Q. (2020). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 34(9), pp. 4237–4256. doi:10.1109/TKDE.2020.2981314

Liu, Z. et al. (2022). SASH: Efficient secure aggregation based on SHPRG for federated learning. IEEE Transactions on Dependable and Secure Computing. doi:10.1109/TDSC.2022.3172320

Malhotra, P., Vig, L., Shroff, G. & Agarwal, P. (2015). Long short term memory networks for anomaly detection in time series. Proceedings of the 23rd European Symposium on Artificial Neural Networks, pp. 89–94.

Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P. & Shroff, G. (2016). LSTM encoder-decoder for multi-sensor anomaly detection. arXiv preprint arXiv:1607.00148. doi:10.48550/arXiv.1607.00148

Mathur, A.P. & Tippenhauer, N.O. (2016). SWaT: A water treatment testbed for research and training on ICS security. Proceedings of CySWater@CPSWeek, pp. 31–36. doi:10.1109/CySWater.2016.7469060

Ma, W. (2025). A defense method against multi-label poisoning attacks. Scientific Reports, 15, 6654. doi:10.1038/s41598-025-6654-3

McMahan, H.B., Moore, E., Ramage, D., Hampson, S. & y Arcas, B.A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, pp. 1273–1282.

Moustafa, N. & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. 2015 Military Communications and Information Systems Conference (MilCIS), pp. 1–6. doi:10.1109/MilCIS.2015.7348942

Patel, R., Li, X. & Singh, D., 2022. Federated learning-based LSTM intrusion detection for cyber-physical systems. Future Generation Computer Systems, 129, pp. 453–465. DOI: https://doi.org/10.1016/j.future.2021.11.021

Pokhrel, S.R. & Choi, J. (2020). Federated learning with blockchain for autonomous vehicles: Analysis and design challenges. IEEE Transactions on Communications, 68(8), pp. 4734–4746. doi:10.1109/TCOMM.2020.2990666

Shokri, R. & Shmatikov, V. (2015). Privacy-preserving deep learning. Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, pp. 1310–1321. doi:10.1145/2810103.2813687

Stouffer, K., Pillitteri, V., Lightman, S., Abrams, M. & Hahn, A. (2015). Guide to Industrial Control Systems (ICS) Security: NIST SP 800-82 Rev. 2. National Institute of Standards and Technology. doi:10.6028/NIST.SP.800-82r2

Takahashi, T., Chen, L. & Xu, H., 2022. Anomaly detection for industrial CPS: A state-of-the-art survey and new perspectives. Computers & Security, 115, 102605. DOI: https://doi.org/10.1016/j.cose.2022.102605

Tao, F., Zhang, H., Liu, A. & Nee, A.Y.C. (2019). Digital twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics, 15(4), pp. 2405–2415. doi:10.1109/TII.2018.2873186

Yang, A. (2023). Application of meta-learning in cyberspace security: A survey. Computers & Security. doi:10.1016/j.cose.2023.103557

Zhang, C., Xie, Y., Bai, H. & Yu, B. (2022). A survey on federated learning for cybersecurity: Concepts, applications, and challenges. ACM Computing Surveys, 55(6), pp. 1–36. doi:10.1145/3501296

Zhao, Z., Zhang, Y., Qin, Z. & Liu, T., 2021. Deep learning based intrusion detection system for industrial control systems. IEEE Transactions on Industrial Informatics, 17(9), pp.5810–5818. DOI: https://doi.org/10.1109/TII.2020.3033584

Zheng, Y., Lai, S., Liu, Y., Yuan, X., Yi, X. & Wang, C. (2022). Aggregation service for federated learning: An efficient, secure, and more resilient realization. Future Generation Computer Systems, 137, pp. 346–357. doi:10.1016/j.future.2022.07.006

8. Appendix: Algorithm: Federated_LSTM_Training

Inputs:

N ← Number of edge nodes / clients

E ← Local epochs per client

B ← Mini-batch size

η ← Learning rate

G ← Number of global rounds

D_i ← Local dataset at client i

Model ← Initial LSTM model parameters

Outputs:

Global_Model ← Trained global LSTM model

Procedure:

```
1: Initialize Global_Model
2: for round = 1 to G do
3:    Selected_Clients ← Select subset of clients for this round
4:    Client_Updates ← []
5:    for each client i in Selected_Clients do
6:       Local_Model ← copy(Global_Model)
7:       for epoch = 1 to E do
8:          for each batch in D_i with size B do
9:             Gradients ← ComputeGradients(Local_Model, batch)
10:            Local_Model ← UpdateModel(Local_Model, Gradients, η)
11:         end for
12:      end for
13:      ΔModel_i ← Local_Model − Global_Model
14:      Client_Updates.append(ΔModel_i)
15:   end for
16:   Global_Model ← Aggregate(Client_Updates)
17: end for
18: return Global_Model
```

Function: Aggregate(Client_Updates)

```
Sum_Updates ← 0
for each ΔModel_i in Client_Updates do
   Sum_Updates ← Sum_Updates + ΔModel_i
end for
Global_Update ← Sum_Updates / length(Client_Updates)
```

return Global_Model + Global_Update

Policy Description:

- Client Selection Policy: Can be random, weighted by dataset size, or based on system reliability.
- Local Training: Each client trains the model independently using its private dataset.
- Aggregation Policy: Standard FedAvg is used, averaging updates across all participating clients.
- Privacy Consideration: Only model updates are shared; raw data never leaves the client.
- Stopping Criterion: Training stops after G global rounds, but can also include convergence thresholds.