Federated Deep Packet Inspection for Intrusion Detection in Interconnected Power Grids

May Myat Thwe*, Alfan Presekal†, Alexandru Ştefanov*

*Department of Electrical Sustainable Energy
Delft University of Technology
Delft, The Netherlands
M.MyatThwe@tudelft.nl, A.I.Stefanov@tudelft.nl

†Faculty of Computer Science
University of Indonesia
Jakarta, Indonesia
presekal@ui.ac.id

***Abstract*— The integration of Information Technology (IT) - Operational Technology (OT) communication networks has enhanced automation and efficiency in power systems but has also expanded their attack surface. Advanced Persistent Threats (APTs) can coordinate cyber attacks across multiple power grid operators. Machine Learning (ML)-based Deep Packet Inspection (DPI) methods have been used to detect intrusions in OT networks. However, centralized training of ML-based DPI models using data from multiple power grid operators can raise data-sharing concerns while isolated local models fail to capture cross-border attack patterns. This paper proposes a Federated Deep Packet Inspection (FedGridDPI) method to analyze OT communication traffic from digital substations and control centers of various power grid operators and detect wide-area intrusions in interconnected power systems. It integrates principal component analysis with a gated recurrent unit classifier and employs federated averaging to aggregate distributed DPI models for collaborative intrusion detection. Experimental results demonstrate that FedGridDPI achieves 97.5% detection accuracy using OT network traffic from IEEE 39-bus simulations and public OT datasets, enabling privacy-preserving, large-scale cyber attacks detection.**

***Index Terms*—Deep Packet Inspection, Federated Learning, Operational Technology, Power Systems, Intrusion Detection.**

I. INTRODUCTION

The increasing integration of Information Technology (IT) and Operational Technology (OT) communication networks enhances power system automation, monitoring, and control. However, the IT-OT convergence has also made the power grid more vulnerable to cyber attacks. The Supervisory Control and Data Acquisition (SCADA) systems, Remote Terminal Units (RTUs), and Intelligent Electronic Devices (IEDs) in digital substations are the backbone of OT networks. For protection and control functions, they depend on communication protocols, e.g., IEC 61850, IEC 104, DNP3, and Modbus. Attackers are increasingly targeting these protocols by exploiting their weak security measures to disrupt power system operation [1]. Machine Learning (ML)-based intrusion detection techniques for OT networks have been researched in recent years. Signature and rule-based techniques are examples of traditional methodologies [2]. Signature-based approaches are effective at identifying known attacks, but they are ineffective against emerging or zero-day threats. Rule-based ones are not flexible to adjust to changing OT situations and need to be updated manually on a regular basis. On the other hand, adaptive, data-driven methods can detect unknown cyber attack patterns using ML and Deep Learning (DL) techniques [2]. Deep Packet Inspection (DPI) has been used to examine communication packet payload content to identify anomalies or malicious activity. It can detect minute variations in control and command payloads by utilizing the content-level semantics of OT data [3], [4]. Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) architectures are examples of DL-based DPI models that have demonstrated good results in the classification of OT protocol packets.

Interconnected power grids are operated by multiple Transmission System Operators (TSOs) and Distribution System Operators (DSOs). Cyber attacks are no longer limited to a single utility or national boundary. Advanced Persistent Threats (APTs) increasingly exploit the interdependencies between constituent systems in the System of Systems (SoS) [4][5]. The collaborative nature of SoS inherently introduces complex interconnections and dependencies that can be leveraged by attackers to violate system-level security requirements [5]. In the power system context, these interdependencies enable coordinated, multi-stage attacks, which can evolve and propagate across multiple system operators or countries. APTs can coordinate cyber attacks across several TSOs and DSOs, targeting geographically dispersed digital substations, control centers, and wide-area monitoring, protection and control systems of interconnected power grids. OT communication network events that appear legitimate within individual substations may collectively form the signature of a large-scale, cross-border cyber intrusion when correlated across multiple system operators. Consequently, power system-wide visibility and collaborative learning across several control areas are necessary for efficient wide scale cyber attacks detection. While isolated DPI independently deployed in control centers and substations can detect local abnormalities in OT protocol payloads, it has drawbacks. The capability to identify and capture cross-border APT patterns across large-

This work has received funding from the European Commission in the Horizon Europe TwinEU project under Grant Agreement 101136119. Corresponding author: Alexandru Ştefanov (a.i.stefanov@tudelft.nl).

979-8-3315-8138-1

scale interconnected power grids is limited since each system operator only monitors a local subset of the overall OT communication network behavior. In addition, data privacy requirements of regulatory and governance frameworks restrict cross-operator, cross-country data exchange, making the sharing of raw OT communication traffic information between system operators impractical [6]. Therefore, centralized training of ML-based DPI models using OT data from multiple TSOs and DSOs may not be feasible. This raises the question on how to enable joint intrusion detection and cyber attack situational awareness over wide areas, e.g., at the European level, while maintaining the locality of critical OT data for each substation, control center or power system operator.

To overcome these challenges, this paper proposes a Federated Deep Packet Inspection (FedGridDPI) method that analyzes OT communication traffic from digital substations and control centers of various power grid operators and detects wide-area intrusions in interconnected power systems. By using Federated Learning (FL), the method enables collaborative, distributed deep learning-based DPI in digital substations and control centers of various TSOs and DSOs without exchanging raw OT communication network data. Each local DPI model sends only model updates to a cloud aggregation server. The server uses the Federated Averaging (FedAvg) algorithm to create a global model, while system operators keep their OT traffic information locally. FedGridDPI preserves system operator autonomy and data privacy of sensitive operational information, allowing the learning of large-scale, cross-border APT patterns that would be unobservable in isolated local models. The dataset used in this study consists of OT network traffic derived from cyber-physical system simulations using IEEE 39-bus and public OT datasets. It includes both normal communication and cyber attack scenarios such as Denial-of-Service (DoS), network scanning, malware injection, and exploitation attacks [8]-[14]. The data includes multiple OT communication protocols, i.e., IEC 61850, IEC 104, DNP3, and Modbus, with each packet preprocessed into numerical features. Principal Component Analysis (PCA) is applied to reduce dimensionality while preserving dominant traffic characteristics. The Gated Recurrent Unit (GRU)-based classifier was implemented locally on distributed clients. Federated training and evaluation are performed by simulating 27 federated clients, representing 27 substations in the IEEE-39 bus system. The main contributions of this work are as follows:

- A federated deep packet inspection method for privacy-preserving and collaborative intrusion detection in power system OT communication networks through federated learning.
- A domain-adapted DPI that combines PCA for dimensionality reduction and a GRU-based packet classifier to learn OT protocol traffic. It achieves similar detection accuracy as a centralized approach without requiring raw OT traffic data sharing.
- A cross-silo FL environment is simulated and configured to aggregate model updates from participating clients. It enables the learning of global attack patterns across heterogeneous OT networks, enhancing generalization to unseen attacks. According to our best knowledge, FedGridDPI is the first federated DPI for power grid OT domain.

The remainder of this paper is organized as follows. Section II provides background on DPI and FL for intrusion detection in power grids. Section III details the FedGridDPI method. Section IV describes the experimental setup. Section V discusses the experimental results and concludes the paper in Section VI.

II. Related research

A. Deep Packet Inspection in OT Networks

DPI focuses on analyzing the payload content of network packets to detect anomalies, malicious behavior, or policy violations. In OT networks, where deterministic communication patterns are common, payload analysis is especially effective for distinguishing legitimate control messages from crafted attack packets. The study in [4] proposed a spatio-temporal anomaly detection approach using an enhanced Graph-Convolutional LSTM (GC-LSTM) model to identify APT behaviors across substations. Their results demonstrated that payload-level and temporal feature extraction can improve accuracy in detecting evolving cyber threats [7].

B. Federated Learning for Intrusion Detection in Power Systems

FL has recently gained attention for privacy-preserving distributed intrusion detection [15]. A FL-based Intrusion Detection System (FL-IDS) enables multiple entities to collaboratively train an ML model to detect intrusions without data sharing in power grids [16]. FL improves scalability and computational efficiency by allowing parallel local updates, thus reducing overall training time [17]. By exchanging only model parameters, FL mitigates data privacy concerns while enhancing global situational awareness. This is important in power systems as OT network traffic collected from different substations cannot be centralized due to operational and regulatory constraints. Unlike isolated or centralized models, FL enables each substation to learn collaboratively from diverse OT network behaviors, improving detection of cross-border, coordinated cyber attacks. In power system OT networks, where substation traffic often contains sensitive control and device information, FL allows local training while only exchanging model parameters with a central aggregator.

While both DPI and FL have advanced power system cyber security, their integration remains unexplored. Conventional DPI systems rely on centralized datasets, which is impractical for geographically distributed OT networks. This work bridges that gap by proposing FedGridDPI, which performs federated learning of DL-based DPI models across multiple TSO and DSO substations and control centers.

III. Federated deep packet inspection for OT Networks in Interconnected Power Grids

This section presents FedGridDPI method for intrusion detection in OT networks. It combines the FedAvg method for cross-silo collaborative model training, a GRU-based

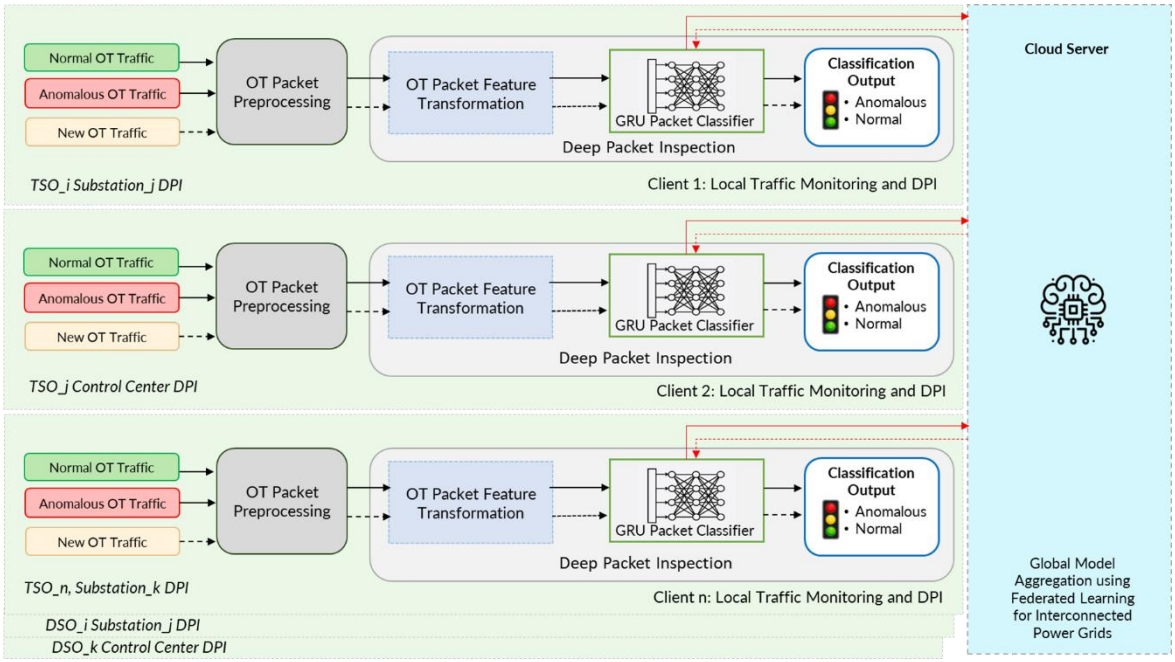

Figure 1. FedGridDPI - Overall architecture of federated DPI approach for intrusion detection in OT networks.

classifier, and PCA for feature transformation. The approach is designed for DPI across dispersed substations and control centers while maintaining privacy without sending private OT traffic information to a single location. Fig. 1 illustrates the FedGridDPI architecture.

A. OT Traffic Collection and Dimensionality Reduction

The dataset includes publicly accessible industrial protocol datasets such as IEC 104, IEC 61850, DNP3, and Modbus, as well as simulated OT traffic. Additionally, Industroyer malware traffic samples are included. Using a Root Mean Square (RMS) dynamic model of the IEEE 39-bus test system in DIgSILENT PowerFactory, the cyber-physical power system is simulated in real-time. We also created extra network traffic using a Mininet simulation, resulting in both normal and malicious TCP/IP communication that is indicative of OT control flows. An estimated 7.71 GB of.pcap data were gathered and examined from all sources. A structured numerical representation of each packet capture was created, with each row denoting a single packet instance and columns represent numerical features extracted from protocol headers and payload bytes. Attack categories were encoded as integer labels, and zeros were used to replace missing data. Min-Max scaling is used to normalize all attributes. Following normalization, the dataset is divided into training and testing sets and dispersed among the FL setup's substation clients. As a feature representation and dimensionality reduction stage, PCA is used because of the high dimensionality and redundancy of packet-level features. It preserves the majority of the information content by projecting the original feature space into a lower-dimensional subspace. The covariance matrix of the training data $X \in \mathbb{R}^{N\times d}$ is defined as:

$$C = \frac{1}{N-1}(X-\bar{X})^{\top}(X-\bar{X}), \qquad (1)$$

where $\bar{X}$ is the mean feature vector. After eigen-decomposition $Cu_i = \lambda_i u_i$, the top m principal components are retained such that:

$$\frac{\sum_{i=1}^{m}\lambda_i}{\sum_{i=1}^{d}\lambda_i} \geq \tau, \quad \tau \in (0,1), \qquad (2)$$

where τ denotes the variance retention threshold. The resulting reduced feature matrix $Z = XU_m \in \mathbb{R}^{N\times m}$, serves as the input to the GRU-based classifier. This transformation reduces feature redundancy, enhances computational efficiency and accelerates convergence under the FL framework.

B. GRU-Based Packet Classifier

The GRU network acts as the packet-level classifier within the FedGridDPI. It performs supervised classification for packet payload from the OT communication network into normal and anomalous. Each PCA-transformed packet feature vector $Z \in \mathbb{R}^m$ is fed into the GRU network to learn nonlinear inter-feature dependencies. GRU captures feature-level correlations within a single packet. The GRU employs two gating mechanisms, the reset gate and update gate, to regulate the flow of information between successive hidden representations as described in (3). Given an input vector z_t and the previous hidden state h_{t-1}, the internal updates are defined as:

$$\begin{aligned} r_t &= \sigma(W_r z_t + U_r h_{t-1} + b_r), \\ u_t &= \sigma(W_u z_t + U_u h_{t-1} + b_u), \\ \tilde{h}_t &= \tanh(W_h z_t + U_h(r_t \odot h_{t-1}) + b_h), \\ h_t &= (1-u_t)\odot h_{t-1} + u_t \odot \tilde{h}_t, \end{aligned} \qquad (3)$$

where $\sigma(\cdot)$ denotes the sigmoid activation, and $\odot$ represents element-wise multiplication. The final hidden state h_t encodes the learned feature representation and is passed through fully connected layers with a ReLU activation, followed by a

SoftMax layer to generate class probability distributions for multi-class traffic classification.

C. Cross-Silo Federated Learning

The proposed FedGridDPI architecture uses a cross-silo FL paradigm to protect data across substations. Each silo represents an individual operator's control center or substation. They work together to train a common global intrusion detection DPI model while simultaneously gathering and storing their own local OT traffic data. The silos have different packet distributions that reflect their different operational settings and protocol usage, although sharing a common feature space. After the local PCA transformation step, each client $k \in \{1,2,\dots,K\}$ maintains a dataset $D_k = \{(Z_k, y_k)\}$, where $Z_k \in \mathbb{R}^{N_k \times m}$ denotes the matrix of PCA-reduced packet features and y_k the corresponding class labels. The local GRU-based packet-level classifier parameterized by weights w_k^t is trained on D_k for E local epochs, minimizing the cross-entropy loss as in (4):

$$L_k = -\frac{1}{N_k}\sum_{i=1}^{N_k}\sum_{c=1}^{C} y_{i,c}\log\left(\hat{y}_{i,cl}\right), \qquad (4)$$

where $\hat{y}_{i,c} = f_{\mathrm{GRU}}(z_i;\ w_k^t)$ is the model's predicted probability for class cl. When local training is completed, each client transmits its updated model parameters w_k^t (not raw data) to the central server. The server then applies the FedAvg algorithm to integrate these local updates into a new global model by computing a weighted mean of all received model parameters. The global model parameters after the communication round t are given by (5).

$$w^{(t+1)} = \sum_{k=1}^{K}\frac{n_k}{n}w_k^t, \qquad (5)$$

where n_k is the number of local training samples at the client $k, n = \sum_{k=1}^{K} n_k$ and K is the total number of participating silos. The updated global model $w^{(t+1)}$ is then redistributed to all clients for the next training round. This process repeats for T communication rounds until convergence, progressively aligning local models toward a shared decision boundary that generalizes across all substations.

IV. Experimental Setup

A. Federated DPI Configuration and System Assumptions

To simulate a cross-silo environment that is reflective of power system, the FedGridDPI is developed using the Flower federated AI framework. While the broader concept targets large-scale collaboration among multiple TSOs/DSOs across countries, the experimental setup simulates this through a representative case of multiple substations under a single operator. It is assumed that every client is a substation node with sufficient computation power to run a local GRU model on its own data. To aggregate model updates and redistribute the updated global model to participants, all clients securely communicate with a central coordinating server. The IEEE 39-bus cyber-physical system has 27 substations that act as federated clients. Every training round, a subset of substations participates and performs local training epochs.

B. Dataset Preparation and Client Partitioning

The experiment uses both normal and anomalous packet-level OT network traffic produced by IEEE 39-bus system simulations [8]. Malware, exploit scenarios, DoS attacks, and network scanning are types of attacks included. To represent protocol diversity, the dataset includes other OT protocols such as IEC 61850, IEC 104, DNP3, and Modbus [9]–[14]. For intrusion detection, every packet that is collected is converted into a numerical feature vector encoding header and payload properties. Two scenarios are simulated across 27 clients in order to assess the consequences of various local data distributions (how data varies between clients):

- Independent and Identically Distributed (IID): The dataset is randomly and uniformly divided, guaranteeing that every client receives a fair distribution of all attack and traffic types.
- Non-IID (Non-Independent and Identically Distributed): This type of dataset is divided using quantity and label skew to simulate actual substation situations. In addition to regular traffic, each substation receives samples of two distinct attack types, which leads to an unequal distribution of classes and data sizes. This setup mimics real OT networks, where substations are vulnerable to attacks and encounter localized traffic as a result of different protocols.

Each client uses PCA to reduce dimensionality before training, keeping 95% of the variation. This procedure stabilizes local learning on high-dimensional data, removes redundancy, and lowers communication overhead during aggregation.

C. Federated DPI Process

The central server in FedGridDPI employs the FedAvg algorithm to coordinate local model updates from substation clients. Ten clients are chosen at random for local training in each round, and the server provides the most recent GRU-based model parameters to these clients. They employ the Adam optimizer with 10 local epochs, a learning rate of 0.001, and a momentum of 0.9. This method simulates power system networks, in which maintenance, delay, or bandwidth constraints prevent all substations from being available simultaneously. Clients send updated model weights to the central server after local training, and FedAvg aggregates them. 20 clients simultaneously assess the combined model using their validation data without sending gradients. PyTorch 1.13.1, Python 3.9.20, and the Flower framework with Ray for parallel execution are used for all of the experiments. An Intel Xeon W-2245 CPU (8 cores, 16 threads), 64 GB of RAM, and Windows 11 are used for the simulations. We measure the performance of the FedGridDPI in terms of accuracy, precision, recall, and F1-score. These metrics assessed FedGridDPI's learning capabilities in both IID and non-IID configurations.

V. Results

As depicted in Fig. 2, the effectiveness of FedGridDPI was assessed in two data distribution scenarios: IID and non-IID. The centralized method outperforms the centralized DPI in

Figure 2. Performance of FedGridDPI under IID and non-IID settings.

terms of accuracy (97.61%), precision (99.46%), recall (93.72%), and F1 score (96.2%), due to access to the complete dataset. Each client in the IID scenario received a representative and well-balanced combination of attack and normal packets. After 10 communication rounds, it still attains nearly the same accuracy of 97.5% (with less training time), precision of 96%, recall of 83%, and an F1-score of 88%. These results show that the global model effectively learns discriminative features of OT traffic by utilizing the GRU's ability to capture inter-feature correlations when client data distributions are uniform. The non-IID setup, on the other hand, restricted each substation to a specific subset of attack types. With an accuracy of 95.57%, a precision of 52.07%, a recall of 58.69%, and an F1-score of 54.62%, it led to a decrease in performance. Non-IID distributions cause this performance degradation, as they cause client drift. Inconsistent gradient updates during aggregation are the result of local models becoming biased toward their particular traffic profiles. This makes it difficult for the global model to generalize to a variety of attack behaviors that are not evenly distributed across substations.

VI. Conclusion

This study introduced FedGridDPI, a privacy-preserving federated deep packet inspection method for intrusion detection in OT networks of interconnected power grids. It combines a GRU-based packet classifier with federated averaging and PCA for feature compression. FedGridDPI allows cross-border TSO and DSO substations and control centers to train models without exchanging raw OT traffic data. FedGridDPI successfully detects cyber attacks across several OT protocols while maintaining data privacy. Experimental results show strong detection performance and consistent convergence under IID settings. On the other hand, client drift and inconsistent gradient updates led to a decrease in model performance for non-IID data distributions, which are caused by unbalanced and varied attack types among clients. The inherent challenge of FL in heterogeneous OT environments, where substations encounter different traffic and protocol distributions, is highlighted by this performance disparity between IID and non-IID. FedGridDPI demonstrated its viability for distributed deployment in power system substations and control centers by maintaining steady local training and effective performance. To address non-IID issues, future research will focus on developing various aggregation and customization techniques. FedGridDPI's scalability and communication effectiveness will be validated with OT traffic generated from a synthetic Dutch cyber-physical power system model.

References

[1] M. Willett, “The cyber dimension of the Russia–Ukraine war,” Survival, pp. 7–26, Oct.–Nov. 2022.

[2] Q. Liu, V. Hagenmeyer, and H. B. Keller, “A review of rule learning-based intrusion detection systems and their prospects in smart grids,” IEEE Access, vol. 9, pp. 57542–57564, 2021.

[3] N. Sahani, R. Zhu, J. H. Cho, and C. C. Liu, “Machine learning-based intrusion detection for smart grid computing: A survey,” ACM Transactions on Cyber-Physical Systems, vol. 7, no. 2, pp. 1–31, 2023.

[4] A. Presekal, A. Ştefanov, I. Semertzis and P. Palensky, "Spatio-Temporal Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems Using Enhanced GC-LSTM," in IEEE Transactions on Smart Grid, vol. 16, no. 2, pp. 1654-1666, March 2025.

[5] M. M. Thwe, Z. M. Belay, E. Jee and D. -H. Bae, "Cybersecurity Vulnerability Identification in System-of-Systems using Model-based Testing," 2022 17th Annual System of Systems Engineering Conference (SOSE), Rochester, NY, USA, 2022, pp. 317-322.

[6] European Data Protection Supervisor, “Federated learning,” [Online]. Available:https://www.edps.europa.eu/press-publications/publications/techsonar/federated-learning_en

[7] C. Wang, W. Zhang, H. Hao, and H. Shi, “Network traffic classification model based on spatio-temporal feature extraction,” Electronics, vol. 13, no. 7, p. 1236, 2024.

[8] A. Presekal, A. Ştefanov, V. S. Rajkumar and P. Palensky, "Cyber Forensic Analysis for Operational Technology Using Graph-Based Deep Learning," 2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Glasgow, United Kingdom, 2023, pp. 1-7, doi: 10.1109/SmartGridComm57358.2023.10333922.

[9] S. M. S. Hussain, C. Yaohao, M. M. Roomi, D. Mashima, and E. C. Chang, “An open-source framework for publishing/subscribing IEC 61850 R-GOOSE and R-SV,” SoftwareX, vol. 23, Jul. 2023.

[10] P. Matoušek, O. Ryšavý, and P. Grofčík, “ICS dataset for smart grid anomaly detection,” Dataset, IEEE Dataport. [Online]. Available: https://ieee-dataport.org/documents/ics-dataset-smart-grid-anomaly-detection

[11] P. R. Grammatikis, V. Kelli, T. Lagkas, V. Argyriou, and P. Sarigiannidis, “DNP3 intrusion detection dataset,” Dataset, IEEE Dataport. [Online]. Available: https://ieee-dataport.org/documents/dnp3-intrusion-detection-dataset

[12] T. Ardley. “ICS security tools,” GitHub. 2023. [Online]. Available: https://github.com/ITI/ICS-Security-Tools/

[13] H. Kang, “IoT network intrusion dataset,” Dataset, IEEE Dataport. [Online]. Available: https://ieee-dataport.org/open-access/iot-network-intrusion-dataset

[14] E. Hjelmvik. “Industroyer IEC 104 analysis.” NETRESCEC. 2022.

[15] M. Wen, Y. Yang, D. Wang, J. Yu, Y. Xiang and H. Jin, "FedDetect: A novel privacy-preserving federated learning framework for energy theft detection in smart grid," IEEE Internet Things J., vol. 9, no. 8, pp. 6069–6080, 2021.

[16] M.M.Thwe. P. Palensky, A. Stefanov, "Clustered Federated Learning for Early Stage Cyber Attacks Detection in Power Systems," in Proc. IEEE ISGT2025 Europe, Valletta, Malta, 2025.

[17] Q. Xia, W. Ye, Z. Tao, J. Wu, and Q. Li, “A survey of federated learning for edge computing: Research problems and solutions,” High-Confidence Computing, vol. 1, no. 1, p. 100008, 2021.