TU Delft

Delft University of Technology

Document Version
Final published version

Licence
Dutch Copyright Act (Article 25fa)

Citation (APA)
Presekal, A., Semertzis, I., Goyel, H., Palensky, P., & Stefanov, A. (2026). Intrusion Detection System for Digital Substations Using Semi-Supervised Learning and Traffic Distance Similarity Clustering. *IEEE Transactions on Smart Grid*, *17*(1), 576-589. https://doi.org/10.1109/TSG.2025.3611345

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

*This work is downloaded from Delft University of Technology.*

**Green Open Access added to TU Delft Institutional Repository**
**as part of the Taverne amendment.**

More information about this copyright law amendment
can be found at https://www.openaccess.nl.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.

# Intrusion Detection System for Digital Substations Using Semi-Supervised Learning and Traffic Distance Similarity Clustering

Alfan Presekal, *Member, IEEE*, Ioannis Semertzis, *Graduate Student Member, IEEE*, Himanshu Goyel, *Graduate Student Member, IEEE*, Peter Palensky, *Senior Member, IEEE*, and Alexandru Ştefanov, *Member, IEEE*

***Abstract*—Cyber attacks on power grids are imminent and potentially have a severe impact, as evidenced by the cyber attacks in Ukraine in 2015, 2016, and 2022. In response to this challenge, machine learning-based Intrusion Detection Systems (IDS) have become more prevalent as a potential mitigation owing to their alignment with the latest advances in artificial intelligence. However, existing anomaly detection methods for power grid Operational Technology (OT) are often inadequate, as they primarily focus on detecting power grid physical anomalies at the later attack stages and suffer from the scarcity of available data for supervised machine learning. To address these limitations, we propose a novel semi-supervised IDS specifically for digital substations of the power system. The proposed detection method identifies the distinctive distance similarity of digital substation OT communication traffic using a Convolutional Neural Network and Chebyshev distance of packet payloads, and Kolmogorov-Smirnov of packets' interarrival time using Fast Fourier Transform amplitude. Subsequently, these traffic features are combined into a vector and classified using a novel hybrid semi-supervised Self-Organizing Map (SOM) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). Results indicate that the proposed method can identify zero-day attacks and achieve accuracy and F1 above 95%.**

***Index Terms*—Anomaly detection, Chebyshev distance, convolutional neural network, cyber attack, cyber security, digital substation, intrusion detection system, FFT, interarrival time, Kolmogorov Smirnov, power grids, self-organizing map.**

## Nomenclature

| | |
|---|---|
| $x$ | All types of packet payload. |
| $n$ | Normal packet payload. |
| $w$ | Convolutional kernel. |
| $k$ | Kernel size. |
| $c_{in}$ | Input channels. |
| $c$ | Index of input channel for the kernel. |
| $i$ | Vertical position (height index). |
| $j$ | Horizontal position (width index). |
| $z$ | Output from a convolutional operation. |
| $T$ | Packet interarrival time. |
| $t$ | Packet arrival time. |
| $\sigma$ | Sigmoid function. |
| $N$ | Total number of samples. |
| $\mathbb{N}$ | Total packet quantities. |
| $P$ | Fast Fourier Transform (FFT) amplitude. |
| $P_B$ | Baseline FFT amplitude. |
| $\eta$ | Frequency bin FFT. |
| $D$ | Maximum absolute difference. |
| EDF | Empirical distribution function. |
| $\lambda$ | Asymptotic approximation. |
| $p_{value}$ | Probability of Kolmogorov Smirnov (KS). |
| $\alpha$ | Convolutional neural network outputs. |
| $\beta$ | Variable Chebyshev distance. |
| $\tau$ | KS from the interarrival time. |
| $\gamma$ | KS from the FFT interarrival time. |
| $\psi = \langle \alpha, \beta, \tau \rangle$ | KS from the interarrival vector. |
| $\Phi = \langle \alpha, \beta, \gamma \rangle$ | KS from the FFT interarrival vector. |
| $g$ | Self-Organizing Map (SOM) model. |
| $w_g$ | SOM model weight. |
| $\Phi \rightarrow \Omega_{g^*}$ | Vector mapping using SOM. |
| $\mathbb{C}$ | All data clusters. |
| $\mathbb{C}_B$ | Baseline data cluster. |
| DBSCAN | Density-Based Spatial Clustering of Applications with Noise clustering. |
| $q$ | Number of DBSCAN clusters. |
| $[CB,\ I,\ V]_t$ | Vector from power system measurements. |

Received 7 January 2025; revised 16 May 2025 and 12 August 2025; accepted 9 September 2025. Date of publication 17 September 2025; date of current version 23 December 2025. This work was supported in part by the EU Horizon Europe Cooperative Cyber Protection for Modern Power Grids (COCOON) Project under Grant 101120221 and in part by the Dutch Research Council's Resilience and Cyber Security of Integrated Cyber-Physical Energy Systems (RESCUE) Project under Grant NWO ESI.2019.006. Paper no. TSG-02240-2024. *(Corresponding author: Alexandru Ştefanov.)*

Alfan Presekal is with the Department of Electrical Sustainable Energy, Delft University of Technology, 2628 CD Delft, The Netherlands, and also with the Department of Electrical Engineering, Universitas Indonesia, Depok, Jawa Barat 16424, Indonesia (e-mail: presekal@ui.ac.id).

Ioannis Semertzis, Himanshu Goyel, Peter Palensky, and Alexandru Ştefanov are with the Department of Electrical Sustainable Energy, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: I.Semertzis@tudelft.nl; H.Goyel@tudelft.nl; P.Palensky@tudelft.nl; A.I.Stefanov@tudelft.nl).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2025.3611345.

Digital Object Identifier 10.1109/TSG.2025.3611345

## I. Introduction

CYBER-PHYSICAL Power Systems (CPPS) are critical infrastructures that have experienced an increasing number of cyber attacks in recent years. In December 2015, a highly coordinated cyber attack had a major impact on the Ukrainian power grid, resulting a power outage for several hours [1]. In 2016, another cyber attack on Ukraine's power grid delivered a lower degree of impact compared with the 2015 attack [2]. The Ukrainian power grid also experienced a power outage in October 2022 due to the disruption caused

1949-3053 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISON WITH RELATED RESEARCH

| Ref. | Year | Protocol | Methods | Parameters |
|---|---|---|---|---|
| [15] | 2023 | GOOSE | Rule based | stNum, sqNum, packet payload |
| [9] | 2010 | IEC 61850 | Rule based | Payload |
| [16] | 2016 | IEC 61850 | Rule based | Payload |
| [17] | 2021 | GOOSE | SVM | stNum, sqNum, packet payload |
| [18] | 2024 | IEC 61850 | KNN | Payload |
| [19] | 2022 | GOOSE | DBSCAN and Autoencoder | Payload |
| Our | 2025 | GOOSE | CNN, FFT, SOM, and DBSCAN | Payload and Packets Interarrival Time |

by the Sandworm malware in its Operational Technology (OT) [3]. These incidents emphasize the imminent threat of cyber attacks on power grids.

Intrusion Detection Systems (IDSs) have emerged as a prominent solution for detecting anomalies in power grids [4], [5], [6], [7], with some of them focusing on the power grid's digital substations [6], [7], [8], [9], [10], [11]. Recent research highlights a growing interest in machine learning-based IDSs due to their superior performance [11], [12], [13]. However, machine learning models are predominantly based on supervised learning, which requires large quantities of data to train effectively and achieve optimal performance. This requirement often contrasts with the limited availability of data [12], especially for zero-day attacks. Considering this limitation, a fully supervised machine learning model may not be the optimal choice. Therefore, in this study, we implement a semi-supervised learning strategy to detect anomalies in digital substations of power grids and overcome the challenge of having limited data available. The technique leverages the advantages of the homogeneous characteristics of OT network traffic parameters generated from automated processes of machine-to-machine communications [14].

In this research, we address the aforementioned challenges of limited data availability for supervised learning and detection of zero-day attacks. To overcome this limitation, we propose an IDS that employs a semi-supervised learning approach based on the characteristics of communication traffic. The IDS analyzes traffic using distance-based similarity metrics, taking into account both packet payloads and interarrival times. In addition, we evaluate and validate the proposed IDS in a Hardware-in-the-Loop (HIL) digital substation environment using real Intelligent Electronic Devices (IEDs). Our results demonstrate that the proposed approach enhances situational awareness and effectively detects anomalies associated with zero-day attack scenarios.

The paper is structured as follows. Section II elaborates on the related works and contributions. Section III describes the method for semi-supervised IDS for digital substations. Section IV presents the experimental results, and Section V presents the conclusion of the research.

## II. RELATED WORKS AND CONTRIBUTIONS

There is existing research on intrusion detection systems that focuses on IEC 61850, particularly the Generic Object Oriented Substation Event (GOOSE) protocol. A summary of the related state-of-the-art research is provided in Table I. Most IDS solutions for IEC 61850 are implemented using rule-based strategies [9], [15], [16], which limit their ability to detect zero-day attacks, as they rely on predefined patterns. Other studies employ data clustering techniques using algorithms such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), DBSCAN, and Autoencoder [9], [17], [18], [19]. While these approaches are closer to unsupervised learning, they still fail to detect masquerading packets in spoofing attacks. This limitation arises because they rely primarily on packet payload, which provides insufficient distinction between legitimate and spoofed packets. Some approaches incorporate sequence number (sqNum) and status number (stNum) analysis for anomaly detection [15], [17]. However, these parameters become ineffective when an attacker employs an adaptive incremental numbering strategy, making the spoofed packets appear consistent with normal behavior. To address this challenge, our proposed method integrates both packet payload and interarrival time to characterize GOOSE traffic and improve classification accuracy.

There are many parameters of network traffic that can be utilized as input features for an IDS [20]. Among various parameters, some IDSs quantify the traffic parameters as a distance similarity for anomaly detection [21]. In [22], anomaly detection was performed using the quantified distance similarity of packet payload. However, the distance similarities derived from packet payloads are inadequate for mitigating spoofing attacks due to insignificant differences in packet anomalies when compared to legitimate packets. As an alternative to payload-based anomaly detection, traffic interarrival time parameters have been extensively studied and implemented for IDS applications [23], [24], [25], [26]. Traffic interarrival time is a relevant approach considering the homogeneous characteristics of digital substation traffic. Therefore, in this research, we proposed hybrid distance similarity parameters of digital substation network traffic based on the combination of packet payload and packet interarrival time.

Digital substation traffic under normal operating conditions is used as a reference point for anomaly detection. Based on the reference point, the packet payloads are quantified using a Convolutional Neural Network (CNN) and Chebyshev distance. The application of CNN and Chebyshev distance for the intrusion detection system has been proven separately in [11], [27] and [28], [29]. The traffic interarrival time statistical features, i.e., mean and standard deviation, have been used as input features for anomaly detection in [24], [25], [26],

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

and [30]. However, these statistical features are unable to adequately discriminate between normal and anomalous traffic due to the insignificant distinctions between them. Therefore, in this research, we introduce a novel packet interarrival time signature based on Fast Fourier Transform (FFT) and Kolmogorov-Smirnov (KS). FFTs have been implemented for anomaly detection owing to their ability to identify anomalies within both the time and frequency domains of data [31], [32], [33]. Meanwhile, the KS method has been implemented for anomaly detection based on the statistical features of the data [34], [35]. In this research, the FFT converts the interarrival times of traffic into FFT amplitudes. The FFT amplitudes of a particular traffic are compared to the FFT amplitudes of a baseline traffic using the KS in order to calculate the p-value, which indicates the statistical differences between the two. Subsequently, the parameters from CNN, Chebyshev distance, and KS p-value are integrated into a three-dimensional vector representing traffic distance similarities.

The vectors representing the normal and anomalous traffic signatures are then used as input for semi-supervised classification. In this research, we proposed a classifier based on Self-Organizing Map (SOM) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The application of SOM and DBSCAN for the unsupervised intrusion detection system has been proven independently in [36], [37], and [38]. Instead of using the clustering model independently, we proposed a novel hybrid classifier model for improving the classifier's performance. A hybrid machine learning model is an emerging approach for improving the stand-alone models. With a hybrid model, the implementation of an algorithm can be strengthened through the advantages of other algorithms [39]. Because of some partially labeled data included in this process, we consider our hybrid unsupervised method as a semi-supervised one. SOM can be implemented for unsupervised classification to reduce data dimensionality and complexity [36], [37]. Subsequently, the DBSCAN is implemented to enhance the complex data clustering process. The hybrid combination of SOM and DBSCAN aims to improve classification performance for normal and anomalous traffic in digital substations.

The scientific contributions of this paper are summarized as follows:

1. We propose a novel frequency domain interarrival time traffic characterization based on the Fast Fourier Transform and Kolmogorov-Smirnov. This method enhances the statistical-based methods that are unable to adequately discriminate between normal and anomalous traffic due to the insignificant distinctions between them. Compared to a statistical-based interarrival time, the combination of Fast Fourier Transform and Kolmogorov-Smirnov is able to improve the accuracy by 26% and F1 score by 41 %.
2. We propose a novel traffic distance similarity vector of operational technology communication traffic. The vector is derived from the packet payload and interarrival time. The vector quantifies the packet payload based on the Convolutional Neural Network and Chebyshev distance, and packet interarrival time using Fast Fourier Transform and Kolmogorov-Smirnov.
3. We propose a novel hybrid semi-supervised classification model based on Self-Organizing Map and DBSCAN. The hybrid combination of them aims to improve clustering performance and address the imbalanced dataset. Results indicate that the proposed method is able to identify zero-day attacks and achieve accuracy and F1 above 95 %.
4. We propose a digital substation state transition model based on historical records of traffic distance similarity vectors and power system measurements. This method correlates the historical records of communication traffic and physical features from power system measurements. The historical state transition is visualized and analyzed to discriminate against anomalies due to cyber attacks and physical disturbances. The visualization helps power system operators track the state transition of digital substation traffic and discriminate traffic anomalies due to faults, reclosure, and spoofing attacks.

## III. Semi-Supervised Intrusion Detection System for Digital Substation

This section explains the implementation of IDS for the power system digital substation and the methodology for the proposed hybrid semi-supervised IDS for the digital substation. Fig. 1 summarizes the model architecture of a digital substation, and Fig. 2 summarizes the overall architecture of the proposed methods, including preprocessing of the dataset from IEC 61850 packets into vectors and clusters. A more detailed explanation of the digital substation, proposed method, and corresponding processes is provided in the following subsections.

### A. Digital Substation Architecture and Cyber Threat Model

The power grid's OT communication network consists of several network segments, including the control center, wide area network, and digital substation [22], [40]. There are two primary network elements in the digital substation, i.e., the process bus and station bus, depicted in Fig. 1. The process bus connects the high/medium-voltage equipment in the field to the protection, control, and monitoring systems. In the process bus, Merging Units (MUs) digitize the analog signals from the current transformers (CTs) and voltage transformers (VTs). These signals are communicated with the IEDs in the station bus. The IEDs monitor the status of circuit breakers, transformers, and other assets, and they also act on protective relays to isolate faults. Additional systems are present in the digital substation environment, e.g., Firewall, Gateway, GPS clock, Supervisory Control and Data Acquisition (SCADA) database, and Human Machine Interface (HMI).

Digital substations retain a critical role due to their ability to directly control physical power grid equipment. Consequently, they represent a valuable target for adversaries aiming to disrupt power grids. In this subsection, we present a cyber threat model for digital substations. Digital substations are constructed using standardized communication protocols, such as

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

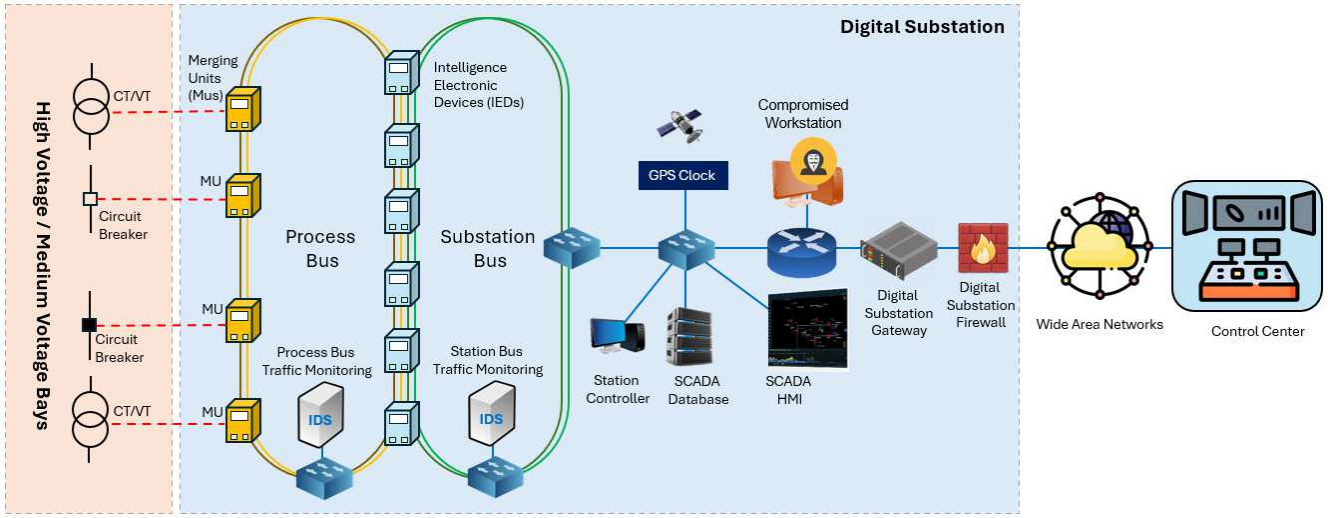


Fig. 1. Digital substation architecture.

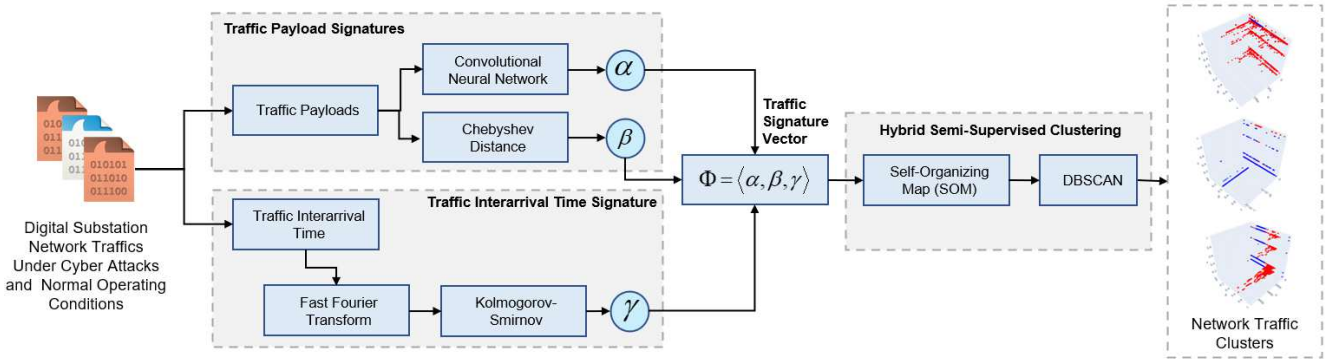


Fig. 2. Summary of the proposed hybrid semi-supervised intrusion detection systems.

IEC 61850, which were initially designed for operational efficiency and interoperability, rather than cyber security. These protocols frequently lack fundamental security mechanisms such as encryption, authentication, and integrity checking. As a result, once an adversary gains access to the network, they can eavesdrop on and masquerade as packets within the network. This vulnerability poses significant risks in digital substations, where messages such as GOOSE and Sampled Values (SV) directly govern protection functions and system stability.

The threat landscape in digital substations is shaped by several key factors described in the advanced persistent threat kill chain for cyber-physical power systems [41]. Before reaching the final objective, the adversaries perform preliminary actions, including reconnaissance, weaponization, exploitation, lateral movements, and other tactics. A prominent real-world example is the Ukrainian power grid cyber attack in 2016, where adversaries gained initial access via the Information Technology (IT) network and eventually executed malicious commands to IEDs in the substation, causing power outages [1], [2].

In our threat model, the attack begins with malware delivery through a phishing email. An unsuspecting user within the organization interacts with a malicious attachment or link, resulting in the installation of malware on an engineering laptop that is also used for configuration of OT devices connected to the digital substation network. This malware establishes a remote access backdoor, allowing the attacker to maintain persistent control over the compromised system. Through this foothold, the attacker is able to move laterally within the substation OT communication network, identify critical communication paths, and ultimately launch attacks against substation devices, including IEDs. This attack vector was used in real cyber attacks on power grids [1], [2]. In this research, we focus specifically on the later stages of this intrusion, where the adversary has already compromised a device within the substation OT communication network. At this point, the attacker has the capability to both passively monitor the OT network traffic and actively inject malicious packets. This reflects a strong internal threat with full access to substation communications.

The primary attack vector considered is a GOOSE frame spoofing attack. The attacker takes advantage of the GOOSE protocol's reliance on multicast messaging and absence of an integrity checking mechanisms. By passively eavesdropping on network traffic, the attacker gathers legitimate GOOSE data frames to understand the format, timing, and control parameters used. To perform the spoofing attack, the adversary constructs masqueraded GOOSE packets that appear structurally and temporally similar to legitimate messages. These packets are crafted by modifying key payload fields,

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

particularly those indicating circuit breaker status. Additionally, protocol-specific control fields such as the stNum and sqNum are incremented adaptively to maintain consistency with ongoing communication, thus bypassing integrity checks within the IEDs.

Previous studies demonstrated IEC 61850 attacks using a single counterfeit packet [8], [42], [43]. In our experimental setup involving a real industrial IED, a single counterfeit packet was not successfully executed. A single counterfeit GOOSE packet may not be sufficient to affect the circuit breaker operation, due to the continuous presence of legitimate traffic. For the spoofing attack to be successful, the counterfeit messages must be transmitted at a rate comparable to or higher than the legitimate GOOSE frame frequency. This repetition increases the likelihood that a forged packet will be processed and acted upon by the target IED. In our experiment, a single counterfeit packet may successfully affect the IED if the attacker can fully block the legitimate traffic. The full methodology and implementation details of the spoofing attack are described in our previous work [44].

In this research, we did not focus on single counterfeit packet detection. We acknowledge that our detection method, which relies on packet interarrival time analysis, is better suited for detecting patterns over multiple packets rather than single isolated events. This reliance on multiple packet observations is a deliberate design choice aimed at enhancing detection robustness and reducing false positives in realistic operational conditions. Consequently, the inability to detect single-packet attacks represents a trade-off in favor of minimizing false positives in detecting sustained or patterned attacks. We focus on the detection strategy for masquerade attacks instead of performing the attack. The detection strategy aims to address the limitation of the traditional intrusion detection systems based on packet payloads or signatures. This limitation is due to the insignificant difference between a masqueraded attack packet and a legitimate packet. To address this challenge, we propose a novel strategy based on the similarity of traffic interarrival distances and hybrid semi-supervised clustering.

To detect such intrusions, traffic monitoring workstations are deployed in both the process bus and station bus. These workstations passively observe all network traffic through span ports and analyze communication patterns. Our proposed IDS leverages these monitoring points to identify anomalies and detect spoofed GOOSE messages in real time, enhancing the cyber security of digital substations.

### *B. Traffic Distance Similarity Vectors in Digital Substations*

OT traffic poses a distinct characteristic in comparison to IT traffic. IT traffic typically originates from the human user with non-deterministic behavior, which makes it more heterogeneous. Meanwhile, OT traffic is generated by automated processes and machine-to-machine communications and is characterized by a greater level of homogeneity. While the OT traffic tends to be homogeneous, it still exhibits a certain degree of variation due to traffic seasonality. Therefore, relying solely on statistical parameters for OT traffic characterization is inadequate. In this research, we proposed a novel OT traffic distance similarity vector based on packet payloads and packet interarrival time.

Packet payload contains data that is communicated between nodes in a network. The packet payload of IEC 61850 includes various features such as Generic Object Identifier (goID), stNum, and sqNum. Among the various IEC 61850 implementations and datasets, they are technically not the same. This distinction arises from variations in the configuration and operational settings of the digital substations. To address this issue, we implement a generalization approach for the payload to ensure the method's applicability across different implementations and dataset variants. This generalization is feasible because our methodology does not rely on the explicit selection of specific features from the packet payload, thereby maintaining adaptability across diverse IEC 61850-based packets. Instead, we utilized a feature-agnostic approach for payload characterization to enhance the generalizability of the proposed method with various data. Furthermore, the integration of explicit feature selection would require prior knowledge of the data, consequently diminishing the method's adaptability, especially in the context of potential zero-day attack scenarios.

The overall process of transforming IEC 61850 packet data into a vector representation is illustrated in Fig. 2. Initially, the packet payload is converted into a two-dimensional array, which serves as input for both a CNN and a Chebyshev distance algorithm. This research evaluates various distance measurement techniques to determine the most effective method for our application. The assessed methods, ordered from highest to lowest performance, comprise Chebyshev, Minkowski, Euclidean, Hamming, and Jaccard. The metrics were evaluated according to their performance on the tested dataset. Among them, the Chebyshev distance consistently yielded the best results. Therefore, it was selected as the preferred distance metric for this research. The outputs generated by the CNN and the Chebyshev distance, along with the traffic interarrival time, are then integrated to form a traffic signature vector. These signature vectors are subsequently utilized as input for a clustering algorithm to identify patterns and group similar traffic behaviors.

This work characterized the packet payload using CNN and Chebyshev distance. The CNN classified the traffic into two classes, i.e., normal and anomalous. The packet payload from the nominal operating condition serves as a normal class, and the payload from other packets serves as an anomalous class. Considering the possibility of unknown payloads from a zero-day attack, in this research, we used a one percent anomalous payload for the supervised CNN training.

CNN learns spatial hierarchies of features from input data and uses convolutional layers with filters to detect patterns. Eq. 1 shows the convolutional operation used in CNN. Parameter $x$ represents input data from the packet payload. Parameter $w$ represents the weight of the CNN filters that are learned during training to detect specific features in the input data. These weights slide over the input to compute weighted sums, allowing the model to extract meaningful patterns at different spatial locations. In CNN, $b$ represents the bias of the convolution operation to enable the model to fit the data

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

more flexibly by shifting the activation function. The $c_{in}$ represents the input channel dimension, while $c$ represents the input channel index. This research uses a two-dimensional (2D) CNN to capture a 2D representation of packet payloads. Therefore, in Eq. 1, there are two parameters, $k_h$ and $k_w$, that represent the kernel from a 2D space. The 2D space also applies to index locations ($i$ and $j$) and for kernel size ($m$ and $n$). The output from the convolution operation ($z$) serves as an input for the sigmoid function ($\sigma$) in Eq. 2. The CNN operation in Eq. 1 and Eq. 2 produced a vector variable alpha ($\alpha$). Eq. 3 shows the Chebyshev distance equation to obtain a vector variable $\beta$. This process is based on the input from the packet payload $x$ and the average value of the normal traffic payload ($n$).

$$z[i,j] = \sum_{m=0}^{k_h-1}\sum_{n=0}^{k_w-1}\sum_{c=0}^{c_{in}-1} x[i+m, j+n, c].w[m,n,c] + b \quad (1)$$

$$\alpha = \sigma(z) = \frac{1}{1+e^{-z}} \quad (2)$$

$$\beta = D_{Chebyshev}(x,n) = \max|x_i - n_i| \quad (3)$$

Packet interarrival time ($T$) is calculated based on the individual packet arrival time ($t$) for all packet quantities ($\mathbb{N}$) depicted in Eq. 4. The interarrival time serves as an input for Eq. 5. Eq. 5 represents the calculation of the amplitude spectrum ($P$) using the FFT. Specifically, it computes the squared magnitude $P$ of the Discrete Fourier Transform (DFT) at frequency bin $\eta$. The parameter $N$ in Eq. 5 represents the total number of samples, and the exponential term represents the complex sinusoid basis functions used in the FFT. This information is obtained from the sliding window size for all interarrival data ($\mathbb{N}$). The FFT amplitude ($P$) serves as an input for the KS equations in Eqs. 6, 7, and 8. In Eq. 6, the maximum absolute difference ($D$) between the two Empirical Distribution Functions ($EDF$) is based on the calculated FFT amplitude ($P$) and the baseline FFT amplitude ($P_B$). The value of $D$ is then used to calculate $\lambda$, which represents an asymptotic approximation formula in Eq. 7. Subsequently, $\lambda$ is then used to calculate vector variables $\gamma$ based on $p_{value}$ in Eq. 8. Finally, all vector variables are combined into a traffic distance similarity vector ($\Phi$) depicted in Eq. 9.

$$T_{i=1}^{\mathbb{N}} = t_{i+1} - t_i \quad (4)$$

$$P = \left|\sum_{i=0}^{N-1} T[i].e^{-j\frac{2\pi}{N}\eta N}\right|^2 \quad (5)$$

$$D = \max_i \left|EDF_P(i) - EDF_{P_B}(i)\right| \quad (6)$$

$$\lambda = \left(\sqrt{\frac{N_P \times \mathbb{N}_{PB}}{N_P + \mathbb{N}_{PB}}} + 0.12 + \frac{0.11}{\sqrt{N_P + \mathbb{N}_{PB}}}\right) \times D \quad (7)$$

$$\gamma = 1 - p_{value} = 1 - \left(2\sum_{i=1}^{\infty}(-1)^{k-1} e^{-2i^2\lambda^2}\right) \quad (8)$$

$$\Phi = \langle \alpha, \beta, \gamma \rangle \quad (9)$$

### C. Hybrid Semi-Supervised Intrusion Detection System

The primary objective of the proposed IDS is to cluster traffic data according to distance similarity vectors $\Phi$.

TABLE II
COMPARISON OF GOOSE TRAFFIC DATA

| Data | No. of IEDs | No. of Packets | ∑.pcap File Size |
|---|---|---|---|
| A[42] | 2 | 27259 | 5.31 MB |
| B[43] | 18 | 328017 | 56.5 MB |
| C[44] | 9 | 1048576 | 895 MB |

A clustering-based approach is utilized because it does not necessitate labeled data, unlike supervised learning techniques. This attribute improves the system's robustness in detecting previously unrecognized traffic patterns, including those linked to zero-day attacks. This research employs a combination of two foundational unsupervised learning techniques, SOM and DBSCAN, to develop an effective and adaptable IDS.

The decision to combine SOM and DBSCAN in this research is based on a thorough evaluation of their individual performances. Both SOM and DBSCAN have been independently validated as effective techniques for unsupervised intrusion detection systems, as demonstrated in prior studies [36], [37], and [38]. According to the literature, these methods represent recent advancements in unsupervised learning approaches that align closely with the objectives of this study. As such, SOM and DBSCAN were initially selected as primary benchmarks. Building upon their individual strengths, we proposed a hybrid approach that integrates SOM and DBSCAN. Experimental results presented in Table III of Section IV indicate that the hybrid method outperforms the standalone models. Therefore, we adopt the combined SOM-DBSCAN approach as the core of our proposed Hybrid Semi-Supervised IDS. Furthermore, we also benchmark the method with the Gaussian Mixture Model (GMM) and KNN that we implemented in our previous research in [22].

**Algorithm 1** Hybrid Semi-Supervised Intrusion Detection System

**Inputs:** $\Phi = \langle \alpha, \beta, \gamma \rangle$ // set of *traffic distance similarity vectors*
$q = 2$ // *number of output clusters*

**Outputs:** $\mathbb{C}_{q=2}$ // *clustering result*

*SOM train to get the model g Best Matching Unit (BMU)*
$g* = \arg\min_g \left\|\Phi - w_g\right\|_2$
*Model G is used to reduce the dimensionality* $\Phi \rightarrow \Omega_{g^*}$
*DBSCAN performs clustering* $\mathbb{C}_q = DBSCAN(\Omega)$
if $q > 2$:
    $\mathbb{C}_q \xrightarrow{\mathbb{C}_B} \mathbb{C}_{q=2}$
return: $\mathbb{C}_{q=2}$ // *clustering result*

The process for the hybrid semi-supervised intrusion detection system is summarized in Algorithm 1. The traffic characterization vector ($\Phi$) serves as input for the hybrid semi-supervised IDS. In this research, we proposed a hybrid sequential clustering process using SOM and DBSCAN. Initially, SOM trains a model ($g$) to find the Best Matching Unit (BMU) depicted in Eq. 10. The best model is represented by the weight ($w_g$) that is closest to the vector $\Phi$. The model with the BMU ($g$) is used to perform a mapping from the three-dimensional vector $\Phi$ into a two-dimensional vector $\Omega$

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

TABLE III
PERFORMANCE COMPARISON OF CLUSTERING METHODS

| Methods | Data A | | | Data B | | | Data C | | | Average | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Acc. | F1 | FPR | Acc. | F1 | FPR | Acc. | F1 | FPR | Acc. | F1 | FPR |
| input $\psi = \langle \alpha, \beta, \tau \rangle$ | | | | | | | | | | | | |
| KNN | 0.5665 | 0.4959 | 0.3957 | 0.5689 | 0.4981 | 0.3657 | 0.5755 | 0.5076 | 0.3621 | 0.5703 | 0.5005 | 0.3745 |
| Agglomerative | 0.5815 | 0.5814 | 0.3128 | 0.5522 | 0.4543 | 0.3927 | 0.5876 | 0.5024 | 0.3453 | 0.5738 | 0.5127 | 0.3503 |
| GMM | 0.5065 | 0.4935 | 0.4684 | 0.5061 | 0.4831 | 0.4671 | 0.5062 | 0.4836 | 0.4623 | 0.5063 | 0.4867 | 0.4660 |
| SOM | 0.8166 | 0.5649 | **0.1584** | 0.8151 | 0.5543 | **0.1893** | 0.8171 | 0.5837 | **0.1672** | 0.8163 | 0.5676 | **0.1716** |
| DBSCAN | **0.8459** | 0.4895 | 0.2106 | **0.8319** | 0.4842 | 0.2235 | **0.8403** | 0.4917 | 0.1923 | **0.8394** | 0.4885 | 0.2088 |
| SOM DBSCAN | 0.6401 | **0.6283** | 0.3809 | 0.6288 | **0.6184** | 0.2921 | 0.6497 | **0.6363** | 0.2554 | 0.6395 | **0.6277** | 0.3095 |
| input $\Phi = \langle \alpha, \beta, \gamma \rangle$ | | | | | | | | | | | | |
| KNN | 0.5615 | 0.4595 | 0.4231 | 0.5046 | 0.4176 | 0.4821 | 0.7368 | 0.7052 | 0.2433 | 0.6010 | 0.5274 | 0.3828 |
| Agglomerative | 0.5477 | 0.4332 | 0.4327 | 0.8908 | 0.8899 | 0.1135 | 0.7675 | 0.7541 | 0.2129 | 0.7353 | 0.6924 | 0.2530 |
| GMM | 0.6011 | 0.5994 | 0.2981 | 0.9936 | 0.9936 | 0.0051 | 0.7697 | 0.7799 | 0.1938 | 0.7881 | 0.7910 | 0.1657 |
| SOM | 0.8471 | 0.7582 | 0.1602 | **0.9977** | **0.9951** | **0.0023** | **0.9999** | 0.6985 | **0.0021** | 0.9482 | 0.8173 | 0.0549 |
| DBSCAN | 0.8424 | 0.4819 | 0.2237 | 0.9961 | 0.9086 | 0.0237 | 0.9999 | 0.6653 | 0.0072 | 0.9461 | 0.6853 | 0.0849 |
| SOM DBSCAN | **0.8903** | **0.8823** | **0.1089** | 0.9942 | 0.9942 | 0.0049 | 0.9976 | **0.9976** | 0.0025 | **0.9607** | **0.9580** | **0.0388** |

depicted in Eq. 11. The SOM application aims to reduce data dimensionality and complexity into a lower-dimensional space while preserving the higher-dimensional structure of the data. Therefore, despite the reduction in dimensionality, the new vector $\Omega$ inherits the features of the payload and interarrival time traffic characteristics.

$$g* = \arg\min_{g} \left\| \Phi - w_g \right\|_2 \tag{10}$$

$$\Phi \rightarrow \Omega_{g^*} \tag{11}$$

$$\mathbb{C}_q = DBSCAN(\Omega) \tag{12}$$

$$\mathbb{C}_q \xrightarrow{\mathbb{C}_B} \mathbb{C}_{q=2} \tag{13}$$

Subsequently, the output vector $\Omega$ from SOM is used as an input for the DBSCAN clustering algorithm shown in Eq. 12. The DBSCAN classifies the data into clusters ($\mathbb{C}$) with the number of clusters $q$. The proposed IDS is intended to categorize the traffic characteristics into two classes, i.e., normal and anomalous. Therefore, the value of $q$ equals two. However, the DBSCAN algorithm can potentially generate more than two clusters, which does not align with the intended outcome of two classes expected by the IDS. In this case, the number of DBSCAN clusters ($q$) is reduced into two cluster categories through a mapping process depicted in Eq. 13. Our experiment comprises both normal and anomalous traffic data. Consequently, the minimum number of clusters will be two.

The mapping in Eq. 13 is performed based on the baseline data cluster from the normal traffic ($\mathbb{C}_B$). Therefore, if a cluster predominantly contains values associated with a normal class in partially labeled data, it will be mapped into a normal class. Otherwise, it will be mapped to an anomalous class. In Algorithm 1, the step in Eq. 13 is only performed when the number of clusters from DBSCAN is greater than 2. It is also possible that the DBSCAN directly obtains 2 clusters. In this case, the process in Eq. 13 is not performed. Cluster reduction is possible in DBSCAN because the number of clusters is not predefined. In DBSCAN, the number of clusters is determined by the number of neighbors and the radius. Therefore, by adjusting the parameters, it is possible to obtain DBSCAN results with two clusters. Furthermore, because of some partially labeled data included in this process, we consider our hybrid method as a semi-supervised clustering. Based on the aforementioned steps, the hybrid SOM and DBSCAN process is then used to classify traffic distance similarity vectors into two classes normal and anomalous traffic.

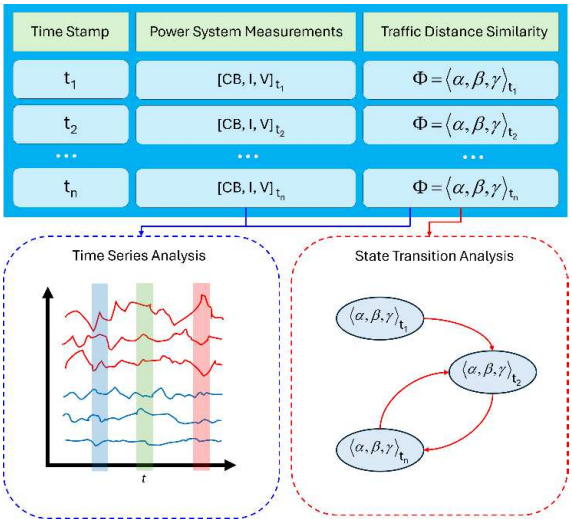


Fig. 3. Digital substation cyber-physical system state transition.

### D. Digital Substation Traffic State Transition Model

To comprehensively analyze the cyber-physical system state of digital substations, we collect historical information on traffic distance similarity vectors and power system measurements depicted in Fig. 3. The collected information is then analyzed on a state transition model based on a time series plot and a 3-dimensional vector plot. The traffic state transition model is intended to oversee the digital substation's traffic state, denoted by a vector $\Phi$, alongside real-time power system

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

measurements. The model improves situational awareness in the digital substation environment by correlating traffic data with measurement information. Furthermore, it facilitates the timely monitoring of dynamic phenomena, recording state transitions that incorporate both cyber and physical parameters. This integrated method enables differentiation between traffic anomalies induced by cyber attacks and those arising from physical disruptions. In this research, we classify the anomalies represented by the state transition vector in digital substations into faults, reclosures, and spoofing attacks.

In Fig. 3, the time series plot presents time series data from distance similarity vectors and power system measurements. Using the time series plot, this method can show the anomaly correlation between digital substation traffic and power system measurements. Meanwhile, the 3-dimensional vector plot presents the distance similarity vector transition in the digital substation. These visualizations aim to help power system operators track the state transition of digital substation traffic and discriminate traffic anomalies. Furthermore, the continuous tracking of information provides valuable insights into the dynamic behavior and interactions between various system states.

## IV. Experimental Results and Discussion

In this section, we present the experimental results of the proposed methods, including a detailed overview of the dataset used, findings from traffic payload and interarrival time distance similarity, the performance of a hybrid semi-supervised IDS, and state transition analysis. The following subsections provide a more detailed explanation of the experimental results.

### A. Experimental Setting and Dataset

Fig. 4 depicts the architecture of a Hardware in the Loop (HIL) digital substation utilized in our experiments. The digital substation simulation consists of a power system simulator and OT networks. The power system simulations were implemented using a Real-Time Digital Simulator (RTDS). The OT communications are implemented using several IEDs and computers. The IEDs implement the IEC 61850 standard, which ensures that the devices can utilize GOOSE messaging and employ SV for measurements. In this research, we are focusing on the GOOSE protocol, which represents control functionality to open or close circuit breakers in the simulated power system. In the simulated cyber attack, the adversaries gained control of a computer in the substation and utilized it to carry out the cyber attack. From the computer, the adversary performs several cyber attack scenarios, i.e., network scanning, simulating malware traffic, and GOOSE spoofing attacks. During the attack, the IDS monitors the network traffic from the span port of a switch in the HIL setup. The monitored traffic is used for the implementation of the proposed hybrid semi-supervised IDS.

The experiment utilizes data derived from the simulation conducted in the HIL simulation setup depicted in Fig. 4. This data is denoted as data A [45] and includes the normal operating conditions of IEC61850 GOOSE traffic from two

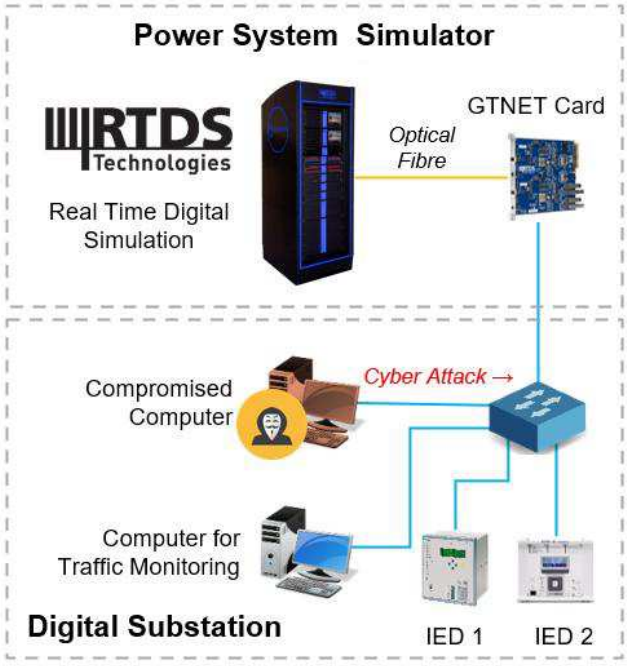


Fig. 4. Digital substation co-simulation architecture with HIL.

IEDs depicted in Fig. 4. Other than data A, we also use other GOOSE attack data from the IEC61850 Security [46] and Power Duck [47]. In this study, these two datasets are denoted as data B and C. The data A, B, and C have similarities in the GOOSE communication protocols for both normal and anomaly conditions. Table II summarizes the comparison of all data based on the number of IEDs, the number of GOOSE packets, and the total size of all.pcap files. All data primarily represents the GOOSE traffic under normal and anomalous conditions. Therefore, in this research, we focus on the characterization of the GOOSE under normal and anomaly conditions. Other than the GOOSE traffic data, we incorporate data from other types of cyber attacks, including Denial of Service (DoS) and network reconnaissance. The DoS and reconnaissance data are collected from the attack simulation using the tools hping3 and Nmap. In addition, we also incorporate sample attack traffic from publicly available data in [45]. The total size for all cyber attacks is 1.5 GB.

### B. Traffic Payload Distance Similarities

The traffic from the three datasets has different average sizes. The average size from the data A, B, and C is 169.58, 165.31, and 217.02 Bytes. From all data, the maximum traffic size is 250 Bytes, originating from data C. The cyber attack data has an average size of 304.74 Bytes. The different sizes of cyber attacks and the GOOSE lead to inconsistencies in the size of the payload characterization processes. Therefore, in the experiment, we decided to standardize the packet size to 256 Bytes to cover all possible GOOSE data sizes. When cyber attack traffic exceeds 256 bytes, the extra bytes are discarded. Conversely, when the GOOSE or cyber attack traffic is less than 256 bytes, the remaining spaces are filled with zeros.

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

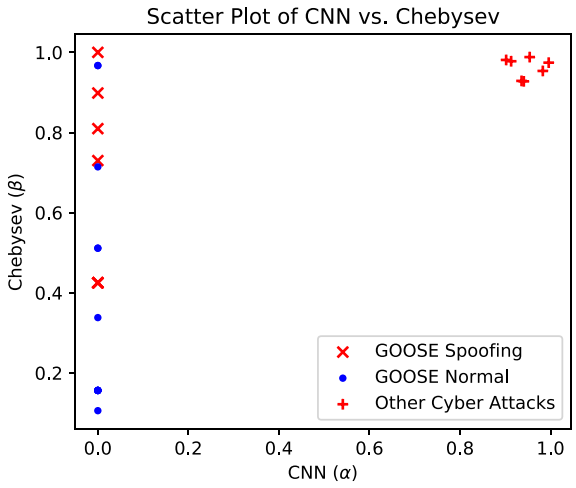


Fig. 5. 2D scatter plot based on the CNN ($\alpha$) and Chebyshev distance ($\beta$) for GOOSE normal and spoofing, and other cyber attacks.

This process ensures the consistency of the traffic payload characterization using CNN and Chebyshev distance.

Alternative distance measurement methods for anomaly detection exist beyond Chebyshev distance [28], [29]. In the experiment, we also evaluate alternative distance measurement methodologies, including Minkowski [49], Euclidean [50], Hamming [51], and Jaccard distances [52]. Our experiment evaluates these distance methods across all datasets utilizing feature selection through a Random Forest [53] and ANOVA test [54]. Upon assessing five distance metrics, the average ranks are as follows: Chebyshev (1.92), Minkowski (2.19), Euclidean (2.32), Hamming (4.74), and Jaccard (4.83). The overall result shows that Chebyshev poses the best average rank. Based on this result, our method uses Chebyshev as the best distance parameter.

Fig. 5 shows the scatter plot based on the CNN ($\alpha$) and Chebyshev distance ($\beta$) for three types of traffic samples. The plot shows that other types of cyber attacks are significantly different from the GOOSE packets, with a higher $\alpha$ and $\beta$ score. However, the GOOSE normal and GOOSE spoofing are relatively the same, with zero values for $\alpha$ score and scattered values for $\beta$ score. Therefore, the $\alpha$ and $\beta$ scores are unable to discriminate between GOOSE normal and GOOSE spoofing. This is because the spoofing attack originated from the normal one, and the payloads from both packet categories exhibit similarities. Based on the plot, it appears that using CNN and Chebyshev distance could be effective in discriminating cyber attacks in digital substations. Consequently, CNN and Chebyshev distance are potentially sufficient to discriminate against other types of cyber attacks in digital substations, including traffic from zero-day attack cases. Fig. 5 presents the results based on zero-day attacks that were not performed over the GOOSE protocol. This result is constructed based on prior knowledge of the limited protocols in the digital substation as an OT network. Therefore, any traffic not associated with the digital substation may be regarded as anomalous. However, the CNN and Chebyshev application is inadequate for addressing zero-day attacks related to GOOSE. Consequently, this study introduces additional parameters derived from the interarrival time of the GOOSE protocol.

Based on the experimental results in Fig. 5, the dataset includes no normal traffic aside from GOOSE. This is because our experiment aimed to differentiate between standard GOOSE messages and GOOSE-based spoofing attacks. Consequently, only these two traffic types are depicted, alongside other traffic classified as anomalies associated with cyber attacks. In situations where other protocols exist, particularly if they correspond with those employed in attack traffic, our approach requires retraining the CNN model to identify anomalies associated with those protocols. This protocol-specific detection strategy corresponds with the practical implementation of OT protocols in digital substations, where communication is generally regulated by a limited set of protocols. However, we recognize that this methodology may encounter more challenges in an IT environment, where protocol diversity is substantially higher and more complex.

### C. Traffic Interarrival Distance Similarities

Packet interarrival time is the duration between the arrivals of consecutive data packets in a network. It is an essential measure to analyze the patterns of network traffic, enabling network performance evaluation, congestion, and the effectiveness of data transfer. In this research, we use the traffic interarrival time to capture the communication signature and identify traffic anomalies. Fig. 6 depicts the plot for visualizing traffic characteristics from data A, B, and C. The blue area and blue line represent normal traffic. Otherwise, the red area and red line represent anomalous traffic.

The top row in Fig. 6 shows the probability density distribution with the normal distribution curve for interarrival time measured in seconds (s). The red and blue area represents the probability density distribution in a bar chart. Meanwhile, the red and blue lines represent the bell curve of the data's normal distribution. Based on the plot, the interarrival time of the GOOSE attack tends to be smaller than the normal one. The bottom row in Fig. 6 depicts the FFT amplitude for the traffic interarrival time. The processes of obtaining the FFT power amplitudes are based on Eq. 4 and Eq. 5. Subsequently, this information is then used to obtain $\gamma$ value based on the KS process depicted in Eqs. 6, 7, and 8.

### D. Hybrid Semi-Supervised Classifier

The hybrid semi-supervised traffic classifier uses the vector $\Phi$ as an input. The visualization of the 3D vector plot for data A, B, and C is depicted in Fig. 7. The blue node represents normal GOOSE traffic, and the red node represents anomalous GOOSE traffic. Every axis in the plot represents the vector element of $\alpha$, $\beta$ and $\gamma$ respectively. As shown in Fig. 7, the normal data is not always concentrated. There is also a possibility where the normal data is scattered, e.g., data B and C. The data B and C pose more scattered normal data due to more variability in the normal GOOSE data. For example, data B considers normal variable loading in the GOOSE traffic, which significantly makes the normal data plot more scattered. The plot in Fig. 7 also shows that the normal GOOSE

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

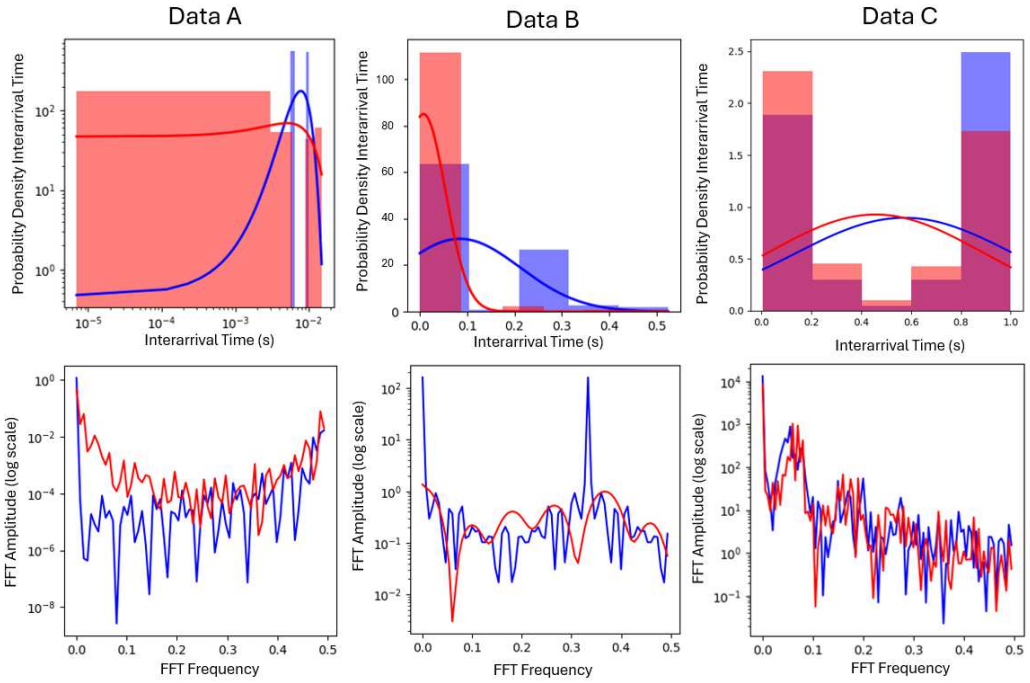


Fig. 6. Characteristic of interarrival time from sample data A, B, and C. The top row plots show the probability density distribution with the normal distribution curve of interarrival time. The bottom row plots show the FFT amplitude for the traffic interarrival time.

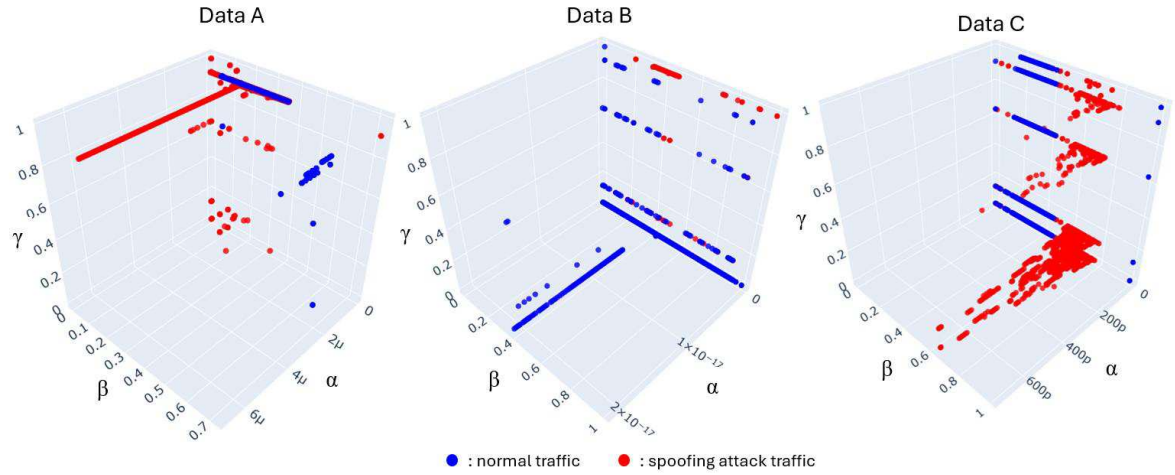


Fig. 7. 3D scatter plots for data A, B, and C, with the blue nodes representing normal GOOSE and the red nodes representing anomalous GOOSE.

traffic signature from different digital substations is unique. Therefore, we propose that this distinctive signature can serve as a reference point for comparing the anomalous GOOSE traffic and use this distance similarity parameter to perform a hybrid semi-supervised traffic classifier. Table III shows the performance comparison of several clustering methods including K Nearest Neighbor (KNN) [55], Agglomerative clustering [56], Gaussian Mixture Model (GMM) [22], [57], SOM [36], [37], DBSCAN [38], and our proposed hybrid SOM and DBSCAN.

The parameter configurations for KNN, Agglomerative Clustering, Gaussian Mixture Models (GMM), Self-Organizing Maps (SOM), DBSCAN, and the hybrid SOM-DBSCAN method are selected in accordance with the overarching objective of partitioning the data into two distinct clusters (i.e., $q = 2$). Accordingly, the parameters for each clustering method are specifically adjusted to fulfill this requirement. For instance, in methods such as GMM and Agglomerative Clustering, the number of components or clusters is explicitly set to two. Similarly, in KNN-based clustering, the model is designed to assign data points to one of two neighboring classes. In the case of SOM and DBSCAN, the configurations are tailored to facilitate the identification of two principal groupings. These two clusters correspond to the research's main objective of distinguishing between normal and anomalous traffic in a digital substation.

The top of Table III shows the performance benchmarking for input $\psi = \langle \alpha, \beta, \tau \rangle$, where $\tau$ represents KS from the

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

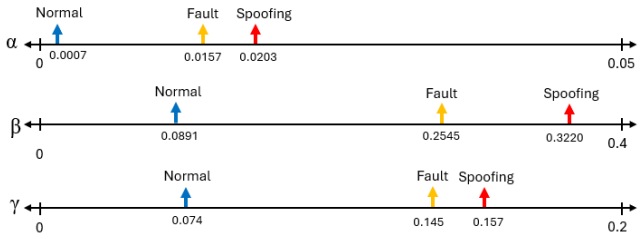


Fig. 8. Comparison of the average value of $\alpha, \beta, \gamma$ under normal, fault, and spoofing traffic.

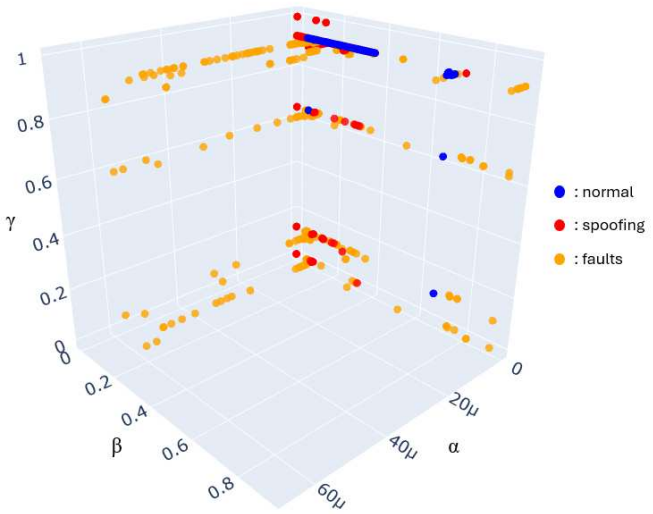


Fig. 9. 3D scatter plots for traffic under normal, spoofing, and faults.

interarrival time. Meanwhile, the bottom of Table III shows the performance benchmarking for the input $\Phi = \langle \alpha, \beta, \gamma \rangle$, where $\gamma$ represents the KS from the FFT amplitude of the interarrival time. From the result in Table III, the input $\psi$ has an average accuracy of 0.6576 and an average F1 of 0.5307. For the input $\Phi$ has an average accuracy of 0.8299 and an average F1 of 0.7452. Based on these results, the implementation of the KS from the FFT amplitude of interarrival time improves the accuracy by 26% and the F1 score by 41 %, and reduces False Positive Rate (FPR) by 48%.

We evaluate the performance based on the accuracy and F1 score. To effectively evaluate model performance under these imbalanced conditions, we employ the F1 score as the primary evaluation metric. The F1 score is particularly suitable for scenarios where class distributions are imbalanced, as it provides a balanced measure of a classifier's ability to detect minority class instances while minimizing false positives. Its significance and efficacy in network intrusion detection tasks involving highly imbalanced datasets have been demonstrated in multiple studies [58], [59], [60].

According to the result, some methods provide relatively high accuracy. However, the accuracy metric can be misleading when dealing with imbalanced datasets, as it may produce high accuracy if the model only correctly predicts the majority class and ignores the minority class. Therefore, in order to evaluate the overall performance, we determined that accuracy and F1 scores were equally important. Out of all the methods in Table III for the input $\Phi$, SOM achieves the highest accuracy and F1 score as a standalone method, while DBSCAN comes in second. However, neither of them can provide the best performance compared to our proposed method, the hybrid SOM and DBSCAN. On average, the hybrid method is able to result in an accuracy of 96%, F1 of 95%, and FPR of 3.9% for input $\Phi$. In Table III for the input $\psi$, the hybrid SOM and DBSCAN do not provide the best accuracy. However, it increases the F1 score. From Table III, overall, the hybrid is able to improve the F1 score. Therefore, the hybrid method is suitable for addressing an imbalanced dataset.

### *E. Digital Substation State Transition Analysis*

This section presents the state transition analysis of digital substations based on traffic distance similarity vectors and power system measurements. The analysis primarily highlights the state transition between traffic anomalies due to faults, reclosures, and spoofing attacks. In a digital substation, an anomaly of GOOSE traffic is also possibly caused by a fault in a power system. When a fault occurs in the power system, the IEDs detect the abnormal condition. Using the GOOSE protocol, these IEDs send an instantaneous trip command to the breaker. In this work, we also simulate GOOSE packets due to a fault in the power system. Fig. 8 depicts the comparison of the average value of $\alpha, \beta, \gamma$ under normal, fault, and spoofing traffic. Fig. 9 shows the 3D plots of traffic under normal, fault, and spoofing. Based on these plots, the traffic from normal, fault, and spoofing shows a distinct characteristic. We implemented the hybrid SOM DBSCAN to classify the data depicted in Fig. 7 into three classes. The result shows that the hybrid SOM DBSCAN can achieve an accuracy of 0.8889 and an F1 score of 0.8785. Although the result is not as good as the two-class classification, this result indicates that our proposed method is also capable of discriminating between fault and spoofing.

We conduct experiments for analyzing the performance of the methods under the state transition from normal, faults, reclosure, and spoofing attacks. The experiments are conducted using the HIL testbed, while the IEEE 5-bus system utilizes a real-time implementation and testing of virtualized controllers for software-defined IEC 61850 in digital substations. Fig. 10 and Fig. 11 show an overview of the transition from the simulated scenarios. Initially, the system runs under normal operating conditions. In step 1, at t = 1.4 s, a three-phase fault occurs in a connected transmission line. In step 2 at t = 1.5 s, the IED responds to the fault by sending the GOOSE trip command, based on the distance protection scheme. The trip command triggers a circuit breaker (CB) to open and change the CB status from 1 to 0. In a digital substation, the trip command is sent by a protection IED to CB circuit breakers when a fault is detected. This action isolates the fault, protects equipment, and ensures system stability. After the fault is cleared, the IED aims to restore the system state to the normal operating condition. Therefore, in step 3 at t = 5.7 s, the IED sends reclosure commands to close the CB and change the CB status from 0 to 1. After the reclosure, as shown in Fig. 10, the system returns to the nominal operating

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

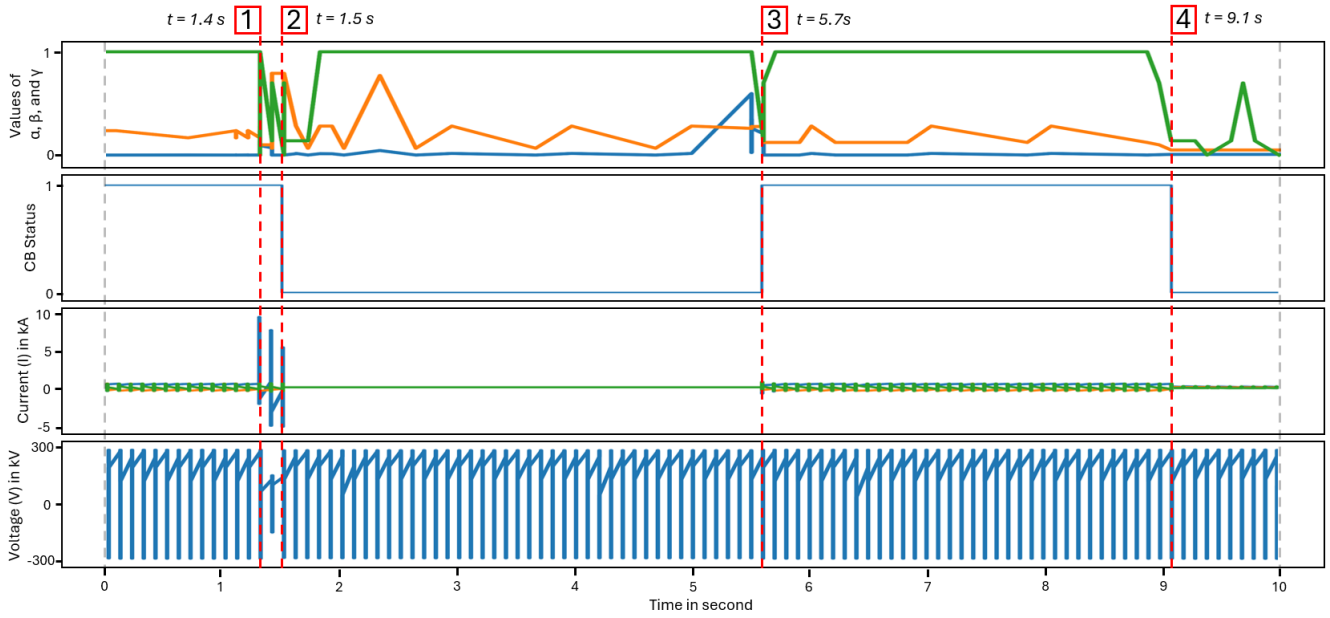


Fig. 10. Time series plot from the simulated events representing value of traffic vector $\Phi = \langle \alpha, \beta, \gamma \rangle$, circuit breaker status, current (kA) and voltage (kV).

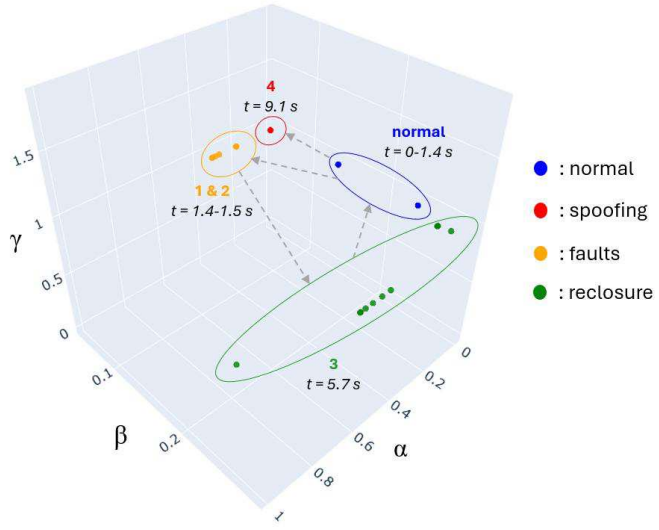


Fig. 11. 3D scatter plots for traffic transition from normal operation, faults (1 and 2), reclosure (3), and spoofing (4).

conditions. In step 4, at t = 9.1 s, GOOSE spoofing packets are injected into the digital substation. The spoofing GOOSE traffic instructs the CB to open. Consequently, this command led to the current anomaly. The measured current, obtained by the CT in the substation, is zero after the CB opens. Due to the direct connection of the substation to a generator, the voltage is equal to the generator's output voltage.

Based on the experiment, we analyze the state transition of traffic distance similarities from different steps shown in Fig. 11. The plots show that every step poses different vector representation values. Therefore, it indicates that our proposed method can distinguish between normal, faults, reclosure, and spoofing attacks. This capability is crucial for the power system operator to identify the anomalies due to a fault or a cyber attack. To improve the confidence of accuracy, the fault can also be detected through measurement anomaly, as shown in Fig. 10. Meanwhile, the confidence level of cyber attack detection can be improved using anomalies from cyber attack traffic depicted in Fig. 5. Based on the cyber kill chain and the real cyber attack in Ukrainian power grids, the adversaries potentially trigger traffic anomalies in the early stage of the cyber kill chain.

The early-stage anomaly detection strategies are based on our previous research in [22] and [40]. In those works, we implement early-stage anomaly detection for a wide-area cyber-physical power system based on packet payload and throughput. However, the early-stage anomaly detections have limitations when addressing spoofing attacks, as shown in Fig. 5. Therefore, in this work, we aim to address the limitations by considering the interarrival time. In future work, the integration of our previous and current research is expected to collaboratively address the challenges of cyber attack detection in cyber-physical power systems from the early to the final stages of the cyber kill chain.

## V. Conclusion

In light of the increasing threat that comes from cyber attacks targeting power grids, it is critical to enhance the capabilities for detecting such attacks in the power grids' OT systems. It is necessary to acknowledge that we are now living in a world where AI is playing an expanding role, particularly with the development of advanced AI models such as deep learning, physics-informed models, and generative AI models. However, existing AI-based IDS for power grids are limited in their ability to achieve optimal performance because of the limited availability and access to the data from the power grid's OT.

In this study, we introduced an innovative hybrid semi-supervised method to detect anomalies in power grid digital substations. The proposed method aims to address limited data availability for AI model training, especially for zero-day attacks. This method works based on the novel digital

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

substation distance similarity vector, which consists of traffic payload and traffic interarrival time distance similarities. The experimental results demonstrate that our hybrid SOM and DBSCAN algorithm outperforms other clustering methods, achieving accuracy and F1 score levels exceeding 95%, respectively. In addition, the implementation of the KS from the FFT amplitude of interarrival time improves the accuracy by 26% and the F1 score by 41 %, and reduces FPR by 48%. The method can also classify normal faults and spoofing with an accuracy of 0.8889 and an F1 score of 0.8785. In future works, our proposed methods can be enhanced by incorporating power system measurement data. This will enable a more thorough evaluation of the cyber-physical power system. In addition, the traffic distance similarity vector and hybrid semi-supervised model can also be implemented in other OT communication protocols beyond IEC61850.

## REFERENCES


[1] D. E. Whitehead, K. Owens, D. Gammel, and J. Smith, "Ukraine cyber-induced power outage: Analysis and practical mitigation strategies," in *Proc. 70th Annu. Conf. Protective Relay Eng. (CPRE)*, Apr. 2017, pp. 1–8.

[2] D. U. Case, "Analysis of the cyber attack on the Ukrainian power grid," *Electr. Inf. Sharing Center (E-ISAC)*, vol. 388, no. 1, pp. 1–29, Mar. 2016.

[3] K. Proska et al.. (Aug. 2024). *Sandworm Disrupts Power in Ukraine Using a Novel Attack Against Operational Technology*. [Online]. Available: https://www.mandiant.com/resources/blog/sandworm-disrupts-power-ukraine-operational-technology

[4] A. Carcano, A. Coletta, M. Guglielmi, M. Masera, I. N. Fovino, and A. Trombetta, "A multidimensional critical state analysis for detecting intrusions in SCADA systems," *IEEE Trans. Ind. Inf.*, vol. 7, no. 2, pp. 179–186, May 2011.

[5] P. Nader, P. Honeine, and P. Beauseroy, "$L_{\text{p}}$-norms in one-class classification for intrusion detection in SCADA systems," *IEEE Trans. Ind. Inf.*, vol. 10, no. 4, pp. 2308–2317, Nov. 2014.

[6] J. Hong, R. F. Nuqui, A. Kondabathini, D. Ishchenko, and A. Martin, "Cyber attack resilient distance protection and circuit breaker control for digital substations," *IEEE Trans. Ind. Inf.*, vol. 15, no. 7, pp. 4332–4341, Jul. 2019.

[7] G. Elbez, K. Nahrstedt, and V. Hagenmeyer, "Early attack detection for securing GOOSE network traffic," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 899–910, Jan. 2024.

[8] S. E. Quincozes, C. Albuquerque, D. Passos, and D. Mossé, "A survey on intrusion detection and prevention systems in digital substations," *Comput. Netw.*, vol. 184, Jan. 2021, Art. no. 107679.

[9] U. K. Premaratne, J. Samarabandu, T. S. Sidhu, R. Beresh, and J.-C. Tan, "An intrusion detection system for IEC61850 automated substations," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2376–2383, Oct. 2010.

[10] C.-W. Ten, J. Hong, and C.-C. Liu, "Anomaly detection for cybersecurity of the substations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 865–873, Dec. 2011.

[11] A. A. Nassar and W. G. Morsi, "Detection of cyber-attacks and power disturbances in smart digital substations using continuous wavelet transform and convolution neural networks," *Electric Power Syst. Res.*, vol. 229, Apr. 2024, Art. no. 110157.

[12] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: Techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1–22, Dec. 2019.

[13] A. Aldweesh, A. Derhab, and A. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," *Knowl. Based Syst.*, vol. 189, Feb. 2019, Art. no. 105124.

[14] R. R. R. Barbosa, R. Sadre, and A. Pras, "Difficulties in modeling SCADA traffic: A comparative analysis," in *Proc. Int. Conf. Passive Act. Meas.*, Mar. 2012, pp. 126–135.

[15] S. Hussain et al., "A novel hybrid methodology to secure GOOSE messages against cyberattacks in smart grids," *Sci. Rep.*, vol. 13, no. 1, p. 1857, Feb. 2023.

[16] Y. Yang, H.-Q. Xu, L. Gao, Y.-B. Yuan, K. McLaughlin, and S. Sezer, "Multidimensional intrusion detection system for IEC 61850-based SCADA networks," *IEEE Trans. Power Del.*, vol. 32, no. 2, pp. 1068–1078, Apr. 2017.

[17] T. S. Ustun, S. M. S. Hussain, A. Ulutas, A. Onen, M. M. Roomi, and D. Mashima, "Machine learning-based intrusion detection for achieving cybersecurity in smart grids using IEC 61850 GOOSE messages," *Symmetry*, vol. 13, no. 5, p. 826, May 2021.

[18] V. E. Quincozes, S. E. Quincozes, D. Passos, C. Albuquerque, and D. Mossé, "Towards feature engineering for intrusion detection in IEC-61850 communication networks," *Ann. Telecommun.*, vol. 79, nos. 7–8, pp. 537–551, Aug. 2024.

[19] D. Jay, H. Goyel, U. Manickam, and G. Khare, "Unsupervised learning based intrusion detection for GOOSE messages in digital substation," in *Proc. 22nd Nat. Power Syst. Conf. (NPSC)*, New Delhi, India, Dec. 2022, pp. 242–247.

[20] K. He, D. D. Kim, and M. R. Asghar, "Adversarial machine learning for network intrusion detection systems: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 538–566, 1st Quart., 2023.

[21] D. J. Weller-Fahy, B. J. Borghetti, and A. A. Sodemann, "A survey of distance and similarity measures used within network intrusion anomaly detection," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 70–91, 1st Quart., 2015.

[22] 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," *IEEE Trans. Smart Grid*, vol. 16, no. 2, pp. 1654–1666, Mar. 2025.

[23] P. Kokoszka, H. Nguyen, H. Wang, and L. Yang, "Statistical and probabilistic analysis of interarrival and waiting times of internet2 anomalies," *Stat. Methods Appl.*, vol. 29, no. 4, pp. 727–744, Dec. 2020.

[24] X. Kong, Y. Zhou, Y. Xiao, X. Ye, H. Qi, and X. Liu, "IDetector: A novel real-time intrusion detection solution for IoT networks," *IEEE Internet Things J.*, vol. 11, no. 19, pp. 31153–31166, Oct. 2024.

[25] T. E. T. Djaidja, B. Brik, S. Senouci, A. Boualouache, and Y. Ghamri-Doudane, "Early network intrusion detection enabled by attention mechanisms and RNNs," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 7783–7793, 2024.

[26] N. Wang, Y. Chen, Y. Xiao, Y. Hu, W. Lou, and Y. T. Hou, "MANDA: On adversarial example detection for network intrusion detection system," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 2, pp. 1139–1153, Mar. 2023.

[27] L. Mohammadpour, T. C. Ling, C. S. Liew, and A. Aryanfar, "A survey of CNN-based network intrusion detection," *Appl. Sci.*, vol. 12, no. 16, p. 8162, Aug. 2022.

[28] G. Baldini and I. Amerini, "Online distributed denial of service (DDoS) intrusion detection based on adaptive sliding window and morphological fractal dimension," *Comput. Netw.*, vol. 210, pp. 1–13, Jun. 2022.

[29] X. Liu et al., "Secure computation protocol of Chebyshev distance under the malicious model," *Sci. Rep.*, vol. 14, no. 1, p. 17115, Jul. 2024.

[30] L. Yang, Y. Zhai, Y. Zhang, Y. Zhao, Z. Li, and T. Xu, "A new methodology for anomaly detection of attacks in IEC 61850-based substation system," *J. Inf. Secur. Appl.*, vol. 68, pp. 1–12, Aug. 2022.

[31] M. Herrera, Y. Proselkov, M. Pérez-Hernández, and A. K. Parlikad, "Mining graph-Fourier transform time series for anomaly detection of internet traffic at core and metro networks," *IEEE Access*, vol. 9, pp. 8997–9011, 2021.

[32] L. Sui and Y. Jiang, "Argo data anomaly detection based on transformer and Fourier transform," *J. Sea Res.*, vol. 198, Apr. 2024, Art. no. 102483.

[33] S. Kanarachos, S.-R.-G. Christopoulos, A. Chroneos, and M. E. Fitzpatrick, "Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform," *Expert Syst. Appl.*, vol. 85, pp. 292–304, Nov. 2017.

[34] O. S. Ohunakin, E. U. Henry, O. J. Matthew, V. U. Ezekiel, D. S. Adelekan, and A. T. Oyeniran, "Conditional monitoring and fault detection of wind turbines based on Kolmogorov–Smirnov non-parametric test," *Energy Rep.*, vol. 11, pp. 2577–2591, Jun. 2024.

[35] B. Cherkaoui, M.-A.-E. Houssaini, M. Kasri, A. Beni-Hssane, and M. Erritali, "Kolmogorov–Smirnov based method for detecting black hole attack in vehicular ad-hoc networks," *Proc. Comput. Sci.*, vol. 236, pp. 177–184, May 2024.

[36] X. Qu et al., "A survey on the development of self-organizing maps for unsupervised intrusion detection," *Mobile Netw. Appl.*, vol. 26, no. 2, pp. 808–829, Apr. 2021.


Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.

[37] C. S. Wickramasinghe, K. Amarasinghe, and M. Manic, "Deep self-organizing maps for unsupervised image classification," *IEEE Trans. Ind. Inf.*, vol. 15, no. 11, pp. 5837–5845, Nov. 2019.
[38] S. Pitafi, T. Anwar, I. D. M. Widia, and B. Yimwadsana, "Revolutionizing perimeter intrusion detection: A machine learning-driven approach with curated dataset generation for enhanced security," *IEEE Access*, vol. 11, pp. 106954–106966, 2023.
[39] B. F. Azevedo, A. M. A. C. Rocha, and A. I. Pereira, "Hybrid approaches to optimization and machine learning methods: A systematic literature review," *Mach. Learn.*, vol. 113, no. 7, pp. 4055–4097, Jan. 2024.
[40] A. Presekal, A. Ştefanov, V. S. Rajkumar, and P. Palensky, "Attack graph model for cyber-physical power systems using hybrid deep learning," *IEEE Trans. Smart Grid*, vol. 14, no. 5, pp. 4007–4020, Sep. 2023.
[41] A. Presekal, A. Ştefanov, V. S. Rajkumar, I. Semertzis, and P. Palensky, "Advanced persistent threat kill chain for cyber-physical power systems," *IEEE Access*, vol. 12, pp. 177746–177771, 2024.
[42] S. E. Quincozes, C. Albuquerque, D. Passos, and D. Mossé, "ERENO: A framework for generating realistic IEC–61850 intrusion detection datasets for smart grids," *IEEE Trans. Dependable Secure Comput.*, vol. 21, no. 4, pp. 3851–3865, Jul. 2024.
[43] J. G. Wright and S. D. Wolthusen, "Stealthy injection attacks against IEC61850's GOOSE messaging service," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Eur. (ISGT-Europe)*, Oct. 2018, pp. 1–6.
[44] V. S. Rajkumar, M. Tealane, A. Stefanov, A. Presekal, and P. Palensky, "Cyber attacks on power system automation and protection and impact analysis," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur. (ISGT-Eur.)*, Oct. 2020, pp. 247–254.
[45] N. Cibin et al., "Cyber-physical power system dataset for cyber security of digital substations," Zenodo, Tech. Rep., May 2025.
[46] P. P. Biswas, H. C. Tan, Q. Zhu, Y. Li, D. Mashima, and B. Chen, "A synthesized dataset for cybersecurity study of IEC 61850 based substation," in *Proc. IEEE SmartGridComm*, Beijing, China, 2019, pp. 1–7.
[47] S. Zemanek et al., "PowerDuck: A GOOSE data set of cyberattacks in substation," in *Proc. 15th Work. Cyber Secu. Exp. Test*, Feb. 2022, pp. 1–5.
[48] H. Kang et al., "IoT network intrusion dataset," IEEE Dataport, Tech. Rep., Sep. 27, 2019.
[49] S. Asif and Qurrat-Ul-Ain, "A fuzzy Minkowski distance-based fusion of convolutional neural networks for gastrointestinal disease detection," *Appl. Soft Comput.*, vol. 158, no. 5, 2024, Art. no. 111595.
[50] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, "Multivariate correlation analysis technique based on Euclidean distance map for network traffic characterization," in *Information and Communication Security*. Berlin, Germany: Springer, Nov. 2011, pp. 388–398.
[51] R. K. Pandey and T. K. Das, "Anomaly detection for industrial control networks using Hamming distance," in *Proc. Int. Conf. Inf. Syst. Manag. Sci.*, Jun. 2021, pp. 280–290.
[52] Z. Yao, P. Mark, and M. Rabbat, "Anomaly detection using proximity graph and pagerank algorithm," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1288–1300, Aug. 2012.
[53] M. A. M. Hasan, M. Nasser, S. Ahmad, and K. I. Molla, "Feature selection for intrusion detection using random forest," *J. Inf. Secur.*, vol. 7, no. 3, pp. 129–140, Apr. 2016.
[54] P. Dhal and C. Azad, "A comprehensive survey on feature selection in the various fields of machine learning," *Int. J. Speech Technol.*, vol. 52, no. 4, pp. 4543–4581, Mar. 2022.
[55] Y. Liao and V. R. Vemuri, "Use of K-nearest neighbor classifier for intrusion detection," *Comput. Secur.*, vol. 21, no. 5, pp. 439–448, Oct. 2002.
[56] F. A. Mazarbhuiya, M. Y. AlZahrani, and L. Georgieva, "Anomaly detection using agglomerative hierarchical clustering algorithm," in *Proc. ICISA*, Apr. 2018, pp. 475–484.
[57] H. Zhang, L. Huang, C. Q. Wu, and Z. Li, "An effective convolutional neural network based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset," *Comput. Netw.*, vol. 177, Aug. 2020, Art. no. 107315.
[58] T.-T.-H. Le, Y. Shin, M. Kim, and H. Kim, "Towards unbalanced multiclass intrusion detection with hybrid sampling methods and ensemble classification," *Appl. Soft Comput.*, vol. 157, May 2024, Art. no.111517.
[59] R. Ahsan, W. Shi, and J.-P. Corriveau, "Network intrusion detection using machine learning approaches: Addressing data imbalance," *IET Cyber-Phys. Syst., Theory Appl.*, vol. 7, no. 1, pp. 30–39, Mar. 2022.
[60] A. O. Widodo, B. Setiawan, and R. Indraswari, "Machine learning-based intrusion detection on multi-class imbalanced dataset using SMOTE," *Proc. Comput. Sci.*, vol. 234, pp. 578–583, Mar. 2024.

**Alfan Presekal** (Member, IEEE) received the bachelor's degree in computer engineering from the Universitas Indonesia in 2014, the master's degree in secure software systems from the Department of Computing, Imperial College London, U.K., in 2016, and the Ph.D. degree in cyber-resilient power grids from Delft University of Technology in 2025. He is a Guest Researcher at TU Delft and an Assistant Professor of computer engineering with the Department of Electrical Engineering, Universitas Indonesia. Previously, he was a Researcher on cyber-resilient power grids with the Intelligent Electrical Power Grids Section, TU Delft, contributing to European Horizon projects HVDC-WISE and Cooperative Cyber Protection for Modern Power Grids (COCOON). He is a Board Member of Indonesia Cyber Awareness and Resilience Center (idCARE.UI). In 2025, one of his papers published in IEEE TRANSACTIONS ON SMART GRID was recognized as one of the top five outstanding papers. His research interests include cybersecurity, cyber-physical systems, and artificial intelligence.

**Ioannis Semertzis** (Graduate Student Member, IEEE) received the Diploma degree in electrical and computer engineering from the Democritus University of Thrace, Greece, in 2019, and the M.Sc. degree in electrical power engineering from Delft University of Technology, Delft, The Netherlands, in 2021, where he is currently pursuing the Ph.D. degree with the Department of Electrical Sustainable Energy. His main research interests include cyber security, cyber-physical power systems, power system stability, and artificial intelligence for power system applications.

**Himanshu Goyel** (Graduate Student Member, IEEE) received the bachelor's degree in electrical engineering from the University of Mumbai, and the Master of Science degree in electrical engineering from Indian Institute of Technology Madras, Chennai, India. He is currently pursuing the Ph.D. degree in cybersecurity for power systems with the Technische Universiteit Delft, Delft, The Netherlands. He has professional experience, as an Engineer with Grid-Sentry and Rakuten Mobile Inc., Tokyo. His research interests include power systems optimization, cybersecurity, digital substation, machine learning, and smart power grids.

**Peter Palensky** (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. and Habilitation degrees from Vienna University of Technology, Austria, in 1997, 2001, and 2015, respectively. He has co-founded Envidatec, a German startup on energy management and analytics. In 2008, he joined the Lawrence Berkeley National Laboratory, Berkeley, CA, USA, as a Researcher, and the University of Pretoria, South Africa. In 2009, he was appointed as the Head of the Business Unit, Austrian Institute of Technology (AIT), in sustainable building technologies, where he was the first Principal Scientist of complex energy systems. In 2014, he was appointed as a Full Professor in intelligent electric power grids with TU Delft, The Netherlands. His research interests include energy automation networks, smart grids, and modeling intelligent energy systems. He is active in international committees, such as ISO and CEN. He serves as an IEEE IES Ad Com Member-at-Large in various functions for IEEE. He is the past Editor-in-Chief of *IEEE Industrial Electronics Magazine* and an associate editor of several other IEEE publications and regularly organizes IEEE conferences.

**Alexandru Ştefanov** (Member, IEEE) received the M.Sc. degree from the University POLITEHNICA of Bucharest, Romania, in 2011, and the Ph.D. degree from University College Dublin, Ireland, in 2015. He is Associate Professor in intelligent electrical power grids with the Department of Electrical Sustainable Energy, TU Delft, The Netherlands. He is the Director of Control Room of the Future (CRoF) Technology Centre. He is leading the Cyber Resilient Power Grids (CRPG) Research Group. His research interests include cyber security of power grids, resilience of cyber-physical systems, and next generation grid operation. He holds the professional title of Chartered Engineer from Engineers Ireland.

Authorized licensed use limited to: TU Delft Library. Downloaded on January 12,2026 at 09:46:24 UTC from IEEE Xplore. Restrictions apply.