TU Delft

Delft University of Technology

Spatio-Temporal Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems using Enhanced GC-LSTM

Presekal, A.; Stefanov, Alexandru; Semertzis, I.; Palensky, P.

DOI
10.1109/TSG.2024.3474039

Publication date
2024

Document Version
Final published version

Published in
IEEE Transactions on Smart Grid

Citation (APA)
Presekal, A., Stefanov, A., Semertzis, I., & Palensky, P. (2024). Spatio-Temporal Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems using Enhanced GC-LSTM. *IEEE Transactions on Smart Grid*, *16*(2), 1654-1666. https://doi.org/10.1109/TSG.2024.3474039

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

Copyright
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.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

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.

Spatio-Temporal Advanced Persistent Threat Detection and Correlation for Cyber–Physical Power Systems Using Enhanced GC-LSTM

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

***Abstract*—Electrical power grids are vulnerable to cyber attacks, as seen in Ukraine in 2015, 2016, and 2022. These cyber attacks are classified as Advanced Persistent Threats (APTs) with potential disastrous consequences such as a total blackout. However, state-of-the-art intrusion detection systems are inadequate for APT detection owing to their stealthy nature and long-lasting persistence. Furthermore, they are ineffective as they focus on individual anomaly instances and overlook the correlation between attack instances. Therefore, this research proposes a novel method for spatio-temporal APT detection and correlation for cyber-physical power systems. It provides online situational awareness for power system operators to pinpoint system-wide anomaly locations in near real-time and preemptively mitigate APTs at an early stage before causing adverse impacts. We propose an Enhanced Graph Convolutional Long Short-Term Memory (EGC-LSTM) by using sequential and neural network filters to improve APT detection, correlation, and prediction. Control center and substation communication traffic is used to determine cyber anomalies using semi-supervised deep packet inspection and software-defined networking. Power grid circuit breaker status is used to determine physical anomalies. Cyber-physical anomalies are correlated in cyber-physical system integration matrix and EGC-LSTM. The EGC-LSTM outperforms existing state-of-the-art spatio-temporal deep learning models, achieving the lowest mean square error.**

***Index Terms*—Advanced persistent threat, anomaly correlation, anomaly detection, cyber-physical system, graph neural network, intrusion detection system, software-defined networking.**

I. Introduction

CYBER-PHYSICAL Power Systems (CPPS) are critical infrastructures that have been targeted by a growing number of cyber attacks in recent years. Some of the notable cyber attacks on power grids are the cyber attacks in Ukraine in 2015 [1], [2], 2016 [3], and 2022 [4]. These incidents highlight the imminent threat of cyber attacks on power grids, which had patterns resembling to Advanced Persistent Threats (APTs). The detection of APTs poses significant challenges owing to their stealthy nature and long-lasting persistence [5]. The majority of the existing research on APTs focuses on individual anomaly instances and overlooks the correlation between them [5], [6], [7], [8]. Those studies have highlighted the necessity of anomaly correlation, but there is still a shortage of research in this area. Furthermore, according to literature studies, existing research on APTs only focuses on the cyber system [5], [6], [7], [8] and omits the APTs on Cyber-Physical Systems (CPS).

The literature review lists four main methods for detecting power grid communication traffic anomalies, i.e., signature-based [9], sequence-based [10], rule-based [11], [12], and machine learning-based [13], [14]. Machine learning-based anomaly detection methods have gained popularity due to their superior performance [15], [16]. However, machine learning models need large amounts of data to learn and perform well. Meanwhile, cyber attack data in CPPS is scarce [15], especially for zero-day attacks. Given this constraint, a fully supervised machine learning model may not be the best option. Therefore, in this research, we employ semi-supervised Deep Packet Inspection (DPI) to identify anomalies in Operational Technology (OT) communication traffic of CPPS. The technique leverages the advantages of the homogeneous characteristics of OT network traffic generated from automated processes [17].

Semi-supervised classifiers can be constructed by combining Convolutional Neural Network (CNN) and Hamming Distance (HD). CNN usually solves supervised classification problems, i.e., intrusion detection [18]. Integration of the CNN classifier with the HD addresses data dependency in supervised learning. The HD application for distance metric learning has been proposed in [19]. CNN and HD generate Gaussian Mixture Model (GMM) vectors for semi-supervised classification with partial labeling. This GMM classification strategy improves classifier robustness with scarce labeled data [20], [21], [22]. Therefore, this classification method is suitable for zero-day attacks.

Along with a semi-supervised classifier for anomaly instance detection, system-wide monitoring is needed to correlate anomalies and track APT propagation. The state-of-the-art system-wide intrusion detection graphs are only focused on

Received 12 April 2024; revised 8 August 2024; accepted 27 September 2024. Date of publication 4 October 2024; date of current version 21 February 2025. This work was supported in part by the EU Horizon Europe COCOON Project under Grant 101120221, and in part by the Dutch Research Council's RESCUE Project under Grant NWO ESI.2019.006. Paper no. TSG-00615-2024. *(Corresponding author: Alfan Presekal.)*

The authors are with the Department of Electrical Sustainable Energy, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: A.Presekal@tudelft.nl; A.I.Stefanov@tudelft.nl; I.Semertzis@tudelft.nl; P.Palensky@tudelft.nl).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2024.3474039.

Digital Object Identifier 10.1109/TSG.2024.3474039

1949-3053 © 2024 IEEE. 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 March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

cyber anomalies and omit physical anomalies [14], [23], [24]. Meanwhile, as demonstrated in [25], combining cyber and physical anomalies would provide better cyber attack detection on CPS. Our literature review shows that cyber and physical system-wide anomaly detections in power systems are not integrated. Existing methods track anomalies using cyber graphs [14] and power system graphs [26], [27]. In [28], the authors proposed Long Range Memory (LRM) to correlate anomalies and use this knowledge to predict future attack trends. In [29], the authors proposed an Artificial Intelligence (AI) generative model for addressing limited OT traffic and estimating the CPPS vulnerabilities and potential intrusion likelihood based on anomaly correlation. Align with our research objectives, these works highlight the necessity of spatial and temporal correlation for cyber attacks mitigation. Therefore, anomalies must be integrated and correlated to provide a system-wide visibility for spatio-temporal APT events.

Spatio-temporal correlation for APTs can determine the correlation of the anomalies based on spatial and temporal data. Spatio-temporal correlation based on a dynamic heterogeneous graph network has been proposed to detect and correlate APTs in [30]. Graph representation and natural language processing were used to detect spatio-temporal APTs [31]. However, these APT spatio-temporal correlation methods only use IT system logs and are insufficient for the CPPS. Spatio-temporal graph modelling was proposed in [32] to correlate spatial and temporal features from sensor network measurement data. This research only focused on sensor measurement anomalies and did not consider cyber anomaly detection. According to the literature review, graph-based methods are used to develop state-of-the-art spatial correlations, and temporal machine learning models like Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), and Long Short-Term Memory (LSTM) are used to build temporal correlations. The LSTM is the most advanced temporal model and performs best. However, LSTM has limitations when it comes to long-term memory preservation [33]. Therefore, the LSTM is not optimal for capturing the temporal correlation of APTs with non-deterministic temporal windows.

In this paper, we propose a novel spatio-temporal APT detection, correlation, and prediction in cyber-physical power systems. It allows power system operators to locate system-wide anomalies in near real-time from control centers and mitigate APTs early before they cause adverse impacts. At substations and control centers, distributed semi-supervised DPI classifiers monitor OT communication traffic using Software Defined Networking (SDN)-enabled switch. The summary of the proposed architecture is presented in Fig. 1. They communicate with the SDN controller at the control center to construct a cyber anomaly graph. This is generated based on the DPI classification results using a Traffic Dispersion Graph (TDG) with SDN [14]. The power system graph is constructed based on the energized power lines in accordance with the status of Circuit Breakers (CBs) [34], [35]. The cyber-physical anomaly graph is input into a Cyber-Physical System Integration Matrix (CPSIM) for spatio-temporal correlation. Subsequently, an Enhanced Graph-Convolutional Long Short-Term Memory (EGC-LSTM) model with sequential and neural network filters is used to predict APTs in CPPS. Furthermore, to identify zero-day APT patterns, we propose a resilient associative method based on vector databases and K-Nearest Neighbor (KNN). The method employs a CPPS log comparator function to verify and differentiate between circuit breakers opened by operators, faults, and cyber attacks. The overall processes from the proposed methods are presented in Fig. 2. The scientific contributions of this paper are summarized as follows:

1. We propose a novel semi-supervised deep packet inspection method for OT communication network traffic utilizing the OT homogeneous characteristics. The method uses a combination of CNN and Hamming distance to generate vectors. The method identifies zero-day attacks by utilizing semi-supervised clustering on the baseline OT traffic vectors using a Gaussian mixture model with partial labeling. In addition, the proposed method is also integrated with software defined networking and traffic dispersion graph to facilitate power system-wide OT communication traffic monitoring in the control center and substations.

2. We propose a cyber-physical system integration matrix that constructs a topological correlation of cyber and physical system anomalies in CPPS. Control center and substation OT communication network traffic is used to construct a cyber anomalies graph. The circuit breaker status is used to construct a power system graph. The CPSIM matrix serves as the primary data for the APT spatio-temporal correlation and prediction processes.

3. We propose a novel EGC-LSTM model with sequential and neural network filters to predict subsequent anomalies resulting from APT attacks. The proposed EGC-LSTM uses the Sequential and Neural Network filter to minimize the Mean Square Error (MSE). Standalone implementation of the Sequential and Neural Network (NN) filter reduces the MSE by 31% and 35%, respectively. Meanwhile, the integration of both filters reduces MSE by 97%.

4. We propose a resilient associative method based on vector databases and KNN to improve the resilience of EGC-LSTM for detection of zero-day attack scenarios. The vector database of CPSIM allows the proposed model to associate zero-day attack scenarios with the known attacks using the KNN search.

5. We propose a CPPS log comparator to corelate CPPS information, i.e., operator activities, OT communication network traffic, COMmon format for TRAnsient Data Exchange (COMTRADE) information from protective relays, power system circuit breaker (CB) status, and CPSIM. The log comparator enables system operators to verify and differentiate between physical power system anomalies caused by cyber attacks and physical power system disturbances.

The paper is structured as follows. Section II explains the CPPS and cyber threat model. Section III describes the method for spatio-temporal anomaly detection, correlation, and prediction. Section IV provides the experimental results. Section V presents the conclusions and future work.

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

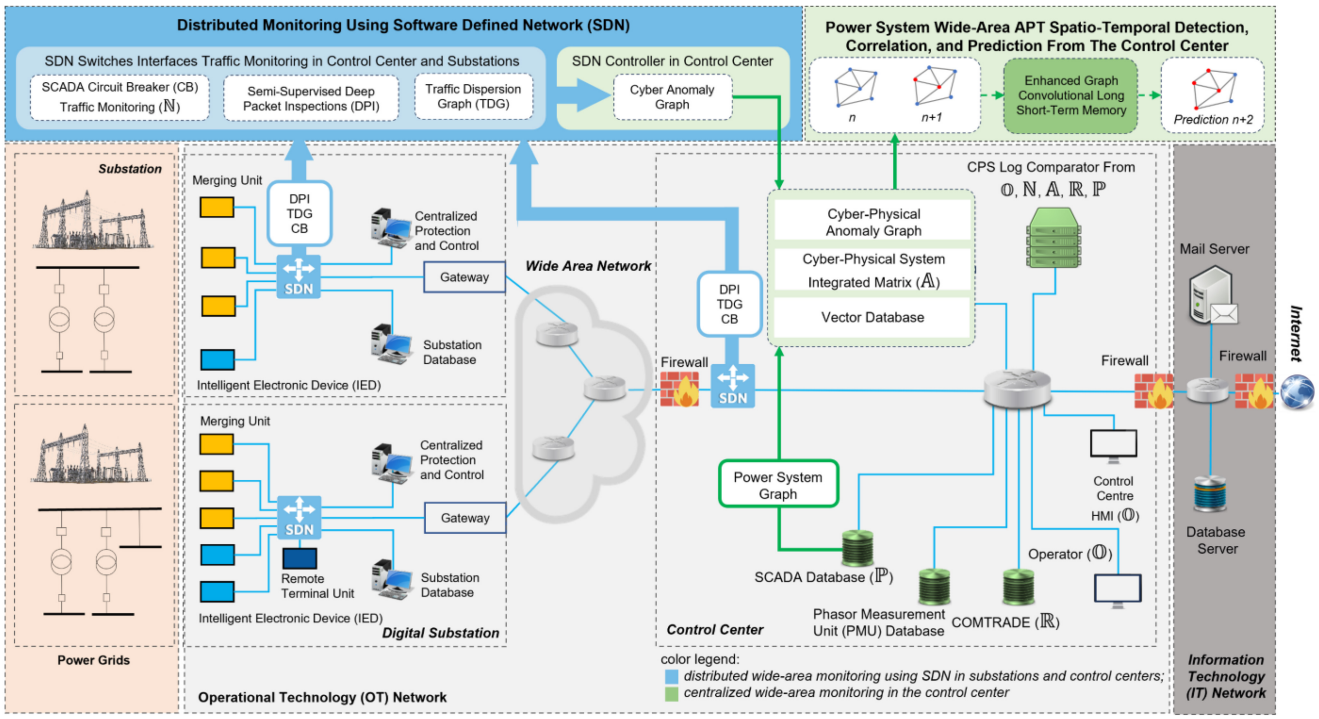

Fig. 1. Cyber-physical system model of the power grid with IT-OT communication networks.

II. Cyber-Physical Power System and Cyber Threat Model

A. Cyber-Physical Power System Model

CPPS models are essential for conducting research on power system cyber security. Therefore, we model the power system integrated with IT-OT communication networks as depicted in Fig. 1. The CPPS model incorporates SDN functionality to establish OT communication network virtualization through SDN switches in the substations and control center and SDN controller in the control center. SDN has three abstraction layers, i.e., data plane, control plane, and management plane. The data plane forwards the OT network traffic, which is controlled by the control plane. In the management plane, SDN allows the deployment of custom network applications. The model is built based on our previous research in [14]. Compared to the previous research, we improved the CPPS model with new SDN management and control functionalities, i.e., monitor the traffic of Supervisory Control and Data Acquisition (SCADA) measurements and CB status, collect the summary from the semi-supervised DPI, and deploy TDG.

The CPPS model is used to compute time-domain simulations and generate measurement data from substation bays, such as busbars, power lines, transformers, and generators. This data includes measurements of active and reactive power, voltage, current, and circuit breaker status. The measurements are communicated from the substations to the control center via a wide area communication network as SCADA telemetry. The SCADA data is kept in local databases situated within substations as well as the control center. The CPPS architecture emulates the OT communication network traffic for power system monitoring and control.

The OT communication network consists of customized functionalities for each OT device within the communication network. The measurement devices include Merging Units (MUs), Remote Terminal Units (RTUs), and Intelligent Electronic Devices (IEDs). These devices collect data from the power grid using SCADA with a sampling rate of one sample per second. The control center uses control commands to dynamically adjust the set points for power grid controllers in real-time. For example, control commands are used to either open or close circuit breakers for power lines, and change set points for automatic voltage regulators and governors. The measurement values and control set points are communicated across the OT network using Transmission Control Protocol/Internet Protocol (TCP/IP) packets.

B. Cyber Threat Model for Cyber-Physical Power System

A cyber threat model is a systematic representation of potential security threats and an analysis of the techniques and pathways that attackers may employ to exploit communication network vulnerabilities. In this research, the cyber threat model is constructed based on the cyber attacks on the Ukrainian power grid in 2015 [2], 2016 [3], and 2022 [4]. These attacks resemble APT's strategies from the early phase of the intrusions until the power outages in the later stages. In the Ukrainian power grid attack in 2015, the adversary used spear phishing emails as an attack vector against the distribution system operators. The phishing emails contained a Microsoft Excel file attachment that was infected with the BlackEnergy3 malware. Subsequently, adversaries performed stealthy operations in the Information Technology (IT) and OT communication networks while preparing for the final

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

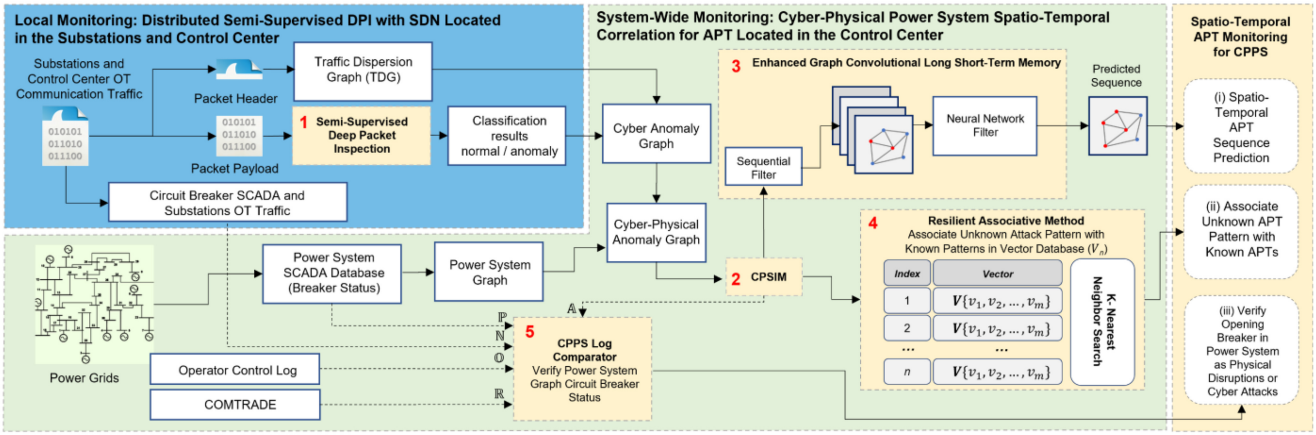

Fig. 2. Integrated processes for spatio-temporal anomaly detection, correlation, and prediction.

phase of the attack. During the early attack phase, the adversaries conducted several malicious activities to intrude from the IT communication network into the control center and substations, i.e., reconnaissance, exploit, lateral movement, firmware modification, and command and control. These activities inevitably caused anomalies in the IT-OT communication network traffic. However, the absence of an early detection mechanism rendered these activities imperceptible to the distribution system operators.

In addition to the aforementioned threat posed by external adversaries, there is also the possibility of an attack from internal actors, known as an insider threat. Insider threats have different characteristics from external threats. The external threat required a lateral movement to reach its final objective in a timely fashion. These scenarios provide an opportunity for the early identification of external threats. However, insider threat potentially has direct access to the substations and control centers and has the potential to cause an immediate severe impact. There is also a possibility when the external and insider threats are combined into more sophisticated and coordinated attack scenarios. However, modelling the insider threat behavior and integration with the test simulation has been identified as a notable challenge for anomaly detection [36]. Therefore, in this research, we omit the insider threat constraint in our CPPS threat model.

Using the aforementioned CPPS co-simulation, our research simulates the early phase of a cyber attack in the simulated substations and control center, which includes reconnaissance, command injection, and malware traffic. The normal and anomalous communication traffic is then used to train the semi-supervised deep packet inspection. Using a traffic dispersion graph, anomaly detection also tracks the sources and destinations of anomalous packets. The graph representation of anomalies is able to track lateral movement processes within the OT network from the entry point to the end device that has direct control of the power grid components. This information is also combined with the CB status of the power lines and transformers to track anomalies in power system-wide. Subsequently, the cyber and physical anomalies are used for the spatio-temporal anomaly correlation and prediction.

III. Spatio-Temporal Anomaly Detection, Correlation and Prediction

The cyber-physical power system architecture integrating the power grid, IT-OT communication networks, and SDN is depicted in Fig. 1. The wide-area network monitoring is enabled based on data collected in near real-time at substations and control center, i.e., OT communication network traffic and CB status. Fig. 2 shows the integrated processes of the proposed method for spatio-temporal anomaly detection, correlation, and prediction. Their implementation in CPS is represented in Fig. 1. The OT communication traffic is monitored locally in all substations and control center on the SDN-enabled switches. The OT traffic is classified using semi-supervised DPI to determine whether an individual packet is normal or anomalous. This information is combined with TDG to generate and update in near real-time as a system-wide cyber anomaly graph. A power system graph is updated in near real-time based on the CB status. It is combined with the cyber anomaly graph into CPSIM. The EGC-LSTM runs continuously to predict subsequent anomalies according to the input from the last four anomalies in CPSIM. To identify zero-day attacks, the resilient associative method associates the zero-day CPSIM with the known CPSIM scenarios using a KNN-based search on the vector database. The CPPS log comparator runs in near real-time to verify the CB status and distinguish between a CB opened by cyber attacks and physical power system disturbances.

The proposed method provides three main results, i.e., (i) spatio-temporal APT detection, correlation and sequence prediction, (ii) identification of zero-day attacks, and (iii) identification of circuit breakers opened by cyber attacks. A detailed description of the proposed method and corresponding processes are provided in the following subsections.

A. Semi-Supervised Learning for Deep Packet Inspection

The DPI uses supervised CNN, HD, and semi-supervised learning based on GMM with partial labeling. The CNN model performs supervised classification for packet payload from OT communication network into normal and anomalous. The packet payload is converted into a 2-Dimensional (2D) data

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

representation as an image. Eq. (1) shows the convolution function from the 2D CNN layer. The * denotes convolution operation, f is the filter size mxn and, g is the input data size i,j. Bayesian optimization [37] is used to optimize the CNN model. Between the convolutional layers, CNN uses Rectified Linear Unit (ReLU) activation function and pooling layer. An activation function is applied to introduce non-linearity into the model, and pooling layers are used to reduce the size. In the end part of CNN layers, there are flattening processes and fully connected neural networks. The flattening layer converts the 2D data into one-dimensional data. The fully connected neural network allows weight adjustment during the training to produce the best-fit output. After the fully connected neural networks, a SoftMax function in Eq. (2) is used to classify the output according to a probability distribution.

Typical neural networks use SoftMax to determine the class categories. However, in our model, we use the SoftMax probability score to generate a vectorΦ. The vector Φ in Eq. (4) is generated from the combination of SoftMax (σ) in Eq. (2) and Hamming Distance ($\mathbb{H}$) in Eq. (3). The HD computes the distance between the average normal traffic payload in OT and other traffic payloads. For normal traffic, the hamming distance will be close to zero (HD $\approx$ 0). Meanwhile, anomalous traffic tends to result in a larger HD score. The combination of outputs from CNN and HD is then used to generate 2D vectorsΦ for semi-supervised learning.

$$(f * g)(i,j) = \sum_m \sum_n f(m,n).g(i-m, j-n) \tag{1}$$

$$\sigma = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{2}$$

$$\mathbb{H} = \frac{1}{N} \sum_{i=1}^{i=N} |a_i - b_i| \tag{3}$$

$$\Phi = \langle \mathbb{H}, \sigma \rangle \tag{4}$$

$$GMM_p(\Phi) = \sum_{k=1}^{K} \pi_k \aleph \left(\Phi | \mu_k, \sum_k \right) \tag{5}$$

GMM with partial labeling is used as a semi-supervised learning technique to classify the 2D vectorsΦ. It uses both labeled and unlabeled data to fit the model. The presence of labeled data facilitates the learning process by enhancing the accuracy of the estimation of GMM parameters. Eq. (5) shows the equation for GMM probability where$\Phi$is the 2D vector data points, K is the number of Gaussian distributions in the mixture, π_k represents the mixing coefficient, and $\aleph$ is the probability density function. The presence of a limited number of labels in the data denotes the presence of a limited dataset that can be used for anomaly detection. Unlabeled data represents the zero-day attack traffic. By employing this semi-supervised learning strategy, our model can effectively identify zero-day attack traffic.

B. Cyber-Physical System Integration Matrix

In order to integrate the cyber and physical components of the CPPS as an integrated graph, we construct the Cyber-Physical System Integration Matrix (CPSIM). CPSIM

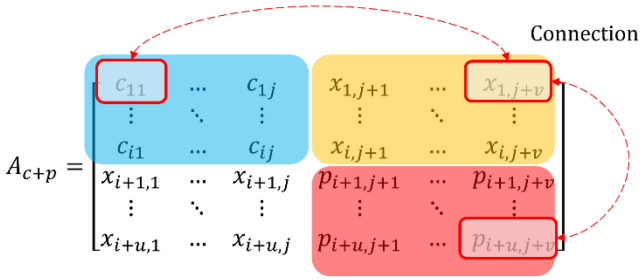

Fig. 3. Cyber-physical system integration matrix.

is formed by combining the adjacency of the OT network topology (c) with the power system topology (p). Fig. 3 shows the representation of CPSIM integration. The cyber adjacency matrix shown in the blue area is represented by $c_{i,j}$, where i and j indicate the element of the matrix. Meanwhile, the power system adjacency matrix shown in the red area is represented by $p_{u,v}$ where u and v indicate the element of the matrix. The CPSIM $\mathbb{A}_{c+p}$ is a combination of cyber and physical elements with dimensions $\mathbb{A}_{i+u,j+v}$. Other than the cyber and physical elements, we introduce a connection matrix in the yellow area represented by x. This area represents the functional connectivity between cyber and physical systems. For example, a node in the cyber element is able to change the physical state of a node in the physical element. The connection matrix is constructed based on the prior information of control function configuration from cyber into the physical system. All information from cyber and physical topology and connectivity information are integrated into a single adjacency of CPSIM.

The adjacency matrix from CPSIM (A) serves as a main reference for the entire cyber and physical system state in CPSIM ($\mathbb{A}$). In the CPSIM, the anomalous elements are indicated by $\mathbb{A}_{i+u,j+v}$=1, and $\mathbb{A}_{i+u,j+v}$=0 otherwise. This reference is then used to track anomalies in both cyber and power systems. Our model identifies the power system graph by analyzing the energized lines based on breaker status information. When the circuit breaker is closed, it indicates a normal condition (0) in the CPSIM. Alternatively, when the breaker is in the open position, it indicates a potential anomaly state (1) in the CPSIM. Meanwhile, to identify the anomaly on the OT network, we use the semi-supervised DPI and TDG. The TDG is utilized to determine the location of OT communication traffic anomalies. The TDG utilizes graph structures to depict nodal information. Every node in a graph represents a distinct host in the communication network. The transfer of information between hosts is shown by the interconnectedness of nodes, specifically, the edges of a graph [38]. By combining DPI and TDG, we identify the location of anomalies in the OT communication networks. In the CPSIM record, the cyber anomalous element is recorded as $\mathbb{A}_{i,j}$=1 in the CPSIM, and $\mathbb{A}_{i,j}$=0 otherwise.

C. APT Spatio-Temporal Correlation

Methods exist in the literature for spatio-temporal correlation, i.e., Graph Convolutional Gated Recurrent Unit (GConvGRU) [39], Temporal Graph Convolutional Network

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

Algorithm 1 Sequence Filter Algorithm

Inputs: $\mathbb{A} = []$ // *Initialize the log storage for CPSIM*
$n = 0$ // *Number of element in* $\mathbb{A}$
$CPSIM_t$ // *CPSIM matrix stream for every time t*

Outputs: $\mathbb{A} = [CPSIM_1, \ldots., CPSIM_n]$ // *CPSIM matrix log*

Iteration for every CPSIM stream
  if $CPSIM_t \neq \mathbb{A}[n]$:
    $\mathbb{A}[n+1] == CPSIM_t$
    $n = n + 1$
else:
  Continue
return: $\mathbb{A} = [CPSIM_1, \ldots., CPSIM_n]$

(TGCN) [40], and Graph Convolution embedded LSTM (GC-LSTM) [41]. These methods can capture the spatial correlation of data using graph convolution. However, they are not the best-fit solution to achieve optimal performance in terms of temporal correlation for APTs. The APTs exemplify non-deterministic temporal characteristics and typically endure extended time intervals between attack stages. Meanwhile, the time-series models, i.e., GRU and LSTM, have the limitation to address long-term temporal correlation [33]. Therefore, in this paper we propose an Enhanced Graph Convolutional LSTM (EGC-LSTM) to address the issue arising from the spatio-temporal correlation of APTs. There are three main improvements in the EGC-LSTM, i.e., Bayesian optimization, sequential filter, and NN filter. Bayesian optimization aims to optimize the GC-LSTM model architecture [37]. The sequential filter is implemented to reduce the recorded data from CPSIM. The filter selectively saves CPSIM data that have distinct values compared to the most recent data in CPSIM, instead of saving all data indiscriminately. This mechanism enables the GC-LSTM to prioritize the detection of anomaly changes instead of analyzing the entire data stream. This mechanism improves the temporal correlation performance. Algorithm 1 shows the pseudocode of the sequence filter algorithm. This algorithm aims to address the APT non-deterministic temporal windows of anomaly records in the CPSIM.

$$GCN_t^k \leftarrow \left(W_{gcn} \odot \hat{A}^k\right)\mathbb{A} \tag{6}$$

$$f_t = \sigma\left(\left(W_f GCN_t^k\right) + \left(U_f h_{t-1}\right) + b_f\right) \tag{7}$$

$$i_t = \sigma\left(\left(W_i GCN_t^k\right) + \left(U_i h_{t-1}\right) + b_i\right) \tag{8}$$

$$o_t = \sigma\left(\left(W_o GCN_t^k\right) + \left(U_o h_{t-1}\right) + b_o\right) \tag{9}$$

$$c'_t = \tanh\left(\left(W_c GCN_t^k\right) + \left(U_c h_{t-1}\right) + b_{c'}\right) \tag{10}$$

$$c_t = \left(f_t \odot c_{t-1}\right) + \left(i_t \odot c'_t\right) \tag{11}$$

$$h_t = o_t \odot \tanh(c_t) \tag{12}$$

$$\text{EGC} - \text{LSTM} = f(Wo_t + b) \tag{13}$$

The output from the sequence filter serves as input for GC-LSTM that combines Graph Convolutional Network (GCN) and LSTM. The GCN function is utilized to extract the nodal characteristics from CPPS elements in CPSIM $\mathbb{A}$ as described in Eq. (6). GCN operates based on the Hadamard product multiplication ($\odot$) of the weight matrix (W_{gcn}), adjacency matrix (A), and node features from CPSIM $\mathbb{A}$. The adjacency matrix (A) is augmented with the identity matrix (I) to create a modified adjacency matrix ($\hat{A}$). The equation incorporates the number of hops from a communication node to neighboring nodes, denoted as k. The temporal features are processed using LSTM subsequent to the acquisition of the spatial features through the GCN. The LSTM input originated from the last four CPSIM anomalies to predict subsequent anomalies in near real-time. The operations performed within an LSTM cell are described in Eq. (7)–(12). There are six main sub-equations in the LSTM process, including the forget gate (f_t), input gate (i_t), output gate (o_t), internal cell state ($c_{'t}$), transferable cell state (c_t), and hidden state (h_t). The predicted output from EGC-LSTM (o_t) serves as an input for the NN filter in Eq. (13). This EGC-LSTM and NN filter are optimized using a Bayesian optimization for hyperparameters tuning. The NN filter transforms the EGC-LSTM output into a binary value of 0 or 1 to enhance the prediction performance of the EGC-LSTM.

D. Resilient Associative Method

To enhance the resilience of EGC-LSTM in handling new or unknown APT patterns, we propose a resilient associative method by performing a KNN search on a vector database. Vector databases are specifically designed to store and handle vector data, which consists of data points defined by arrays or lists of values [42]. Fig. 4 represents the vector database search strategy with KNN. Vector database $\mathbb{A}_n$ represents known historical anomalies recorded in the CPSIM matrix. There are numerous potential combinations of anomalous events (U) that may not be included in the existing data $\mathbb{A}_n$, i.e., zero-day attacks. Therefore, to address event detection for zero-day attacks, the KNN algorithm is implemented to search for the most similar pattern from known data. By implementing this strategy, the model can identify zero-day anomalies (U) by associating this anomaly with the known one ($\mathbb{A}_n$).

E. CPPS Log Comparator

We introduce an innovative circuit breaker log comparator as a multi-log anomaly detection system for CPPS that specifically targets CB-related events. The breaker log comparator function is utilized to compare recorded log activities from CPSIM anomaly record ($\mathbb{A}$), CB of SCADA and substations traffic ($\mathbb{N}$), operator control log ($\mathbb{O}$), power system SCADA database CB status ($\mathbb{P}$), and relay COMTRADE ($\mathbb{R}$). Fig. 5 shows the diagram of the breaker log comparator. Data $\mathbb{A}$ originates from the spatio-temporal data of anomalies in CPSIM. Log $\mathbb{N}$ is generated by the observed CB control traffic in the OT network. Log $\mathbb{O}$ is produced through authorized control operations recorded by the power system operator. Log $\mathbb{P}$ denotes the current state of the CB in the physical power grid. Log $\mathbb{R}$ represents the relay COMTRADE that records transient events in the power system. COMTRADE data has been used to identify electrical disturbances in power systems based on

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

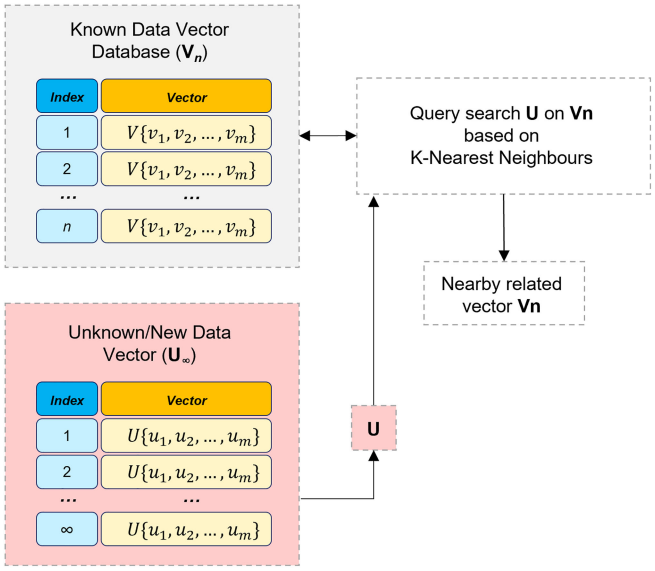

Fig. 4. Vector database query search strategy with KNN.

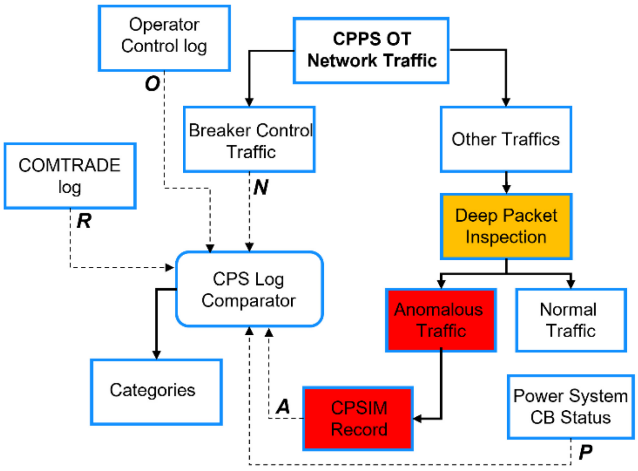

Fig. 5. Circuit breaker log comparator function.

TABLE I
CPS BREAKER LOG COMPARATOR CATEGORIES

$\mathbb{R}$	$\mathbb{O}$	$\mathbb{A}$	$\mathbb{N}$	$\mathbb{P}$	Causes	Categories
0	0	0	0	0	-	Normal operation
1	0	0	0	1	Physical disturbances	Direct response in BCU
1	0	0	1	1		Coordinated protection
0	1	0	1	1	Operator	Operator control action
0	1	1	1	1	Cyber attacks	Compromised operator control
0	0	1	1	1		Spoofing attack
0	0	1	0	1		Compromised BCU device
0	0	1	0	0		Preliminarily kill chain stages

transient waveforms [43]. The breaker log comparator utilizes the COMTRADE to differentiate anomalies in the power system caused by physical system events and cyber attacks.

Table I presents the log comparison categories corresponding to the five types of data logs. The value of 1 indicates the presence of an activity log or anomaly, while the value of 0 represents the normal operating condition. During normal operation, all log parameters are indicated as 0. Otherwise, elements of CPSIM ($\mathbb{A}$) records are indicated with 1. The anomalies in CPSIM are subsequently compared with other logs to identify different scenarios, i.e., system operator performing control actions, physical disturbances, activation of protection relays, and cyber attacks. For physical disturbances, the COMTRADE data serves as a primary indicator. Meanwhile, CPSIM ($\mathbb{A}$) serves as a primary indicator for cyber attacks.

There are four cyber attack scenarios. The first scenario is when the adversaries compromise the legitimate operator's control workstation. Therefore, the operator control log ($\mathbb{O}$) will be indicated by 1. The second scenario is characterized by the adversaries executing a spoofed remote control and disguising themselves as a legitimate operator. It does not originate from the legitimate operator's control workstation and is indicated by control log 0. The spoofed controls originate from a compromised device in the OT communication network that sends malicious breaker control commands. The third attack scenario occurs when the adversaries compromise a device in the Bay Control Units (BCU). Compared to the previous scenarios, this attack does not provide an indicator of a breaker control command in the network. This is possible due to the position of the BCU devices that have a direct connection with the power grids. The fourth attack scenario is a cyber anomaly, which refers to a situation where the cyber anomaly is recorded in CPSIM ($\mathbb{A}$) and does not have any impact on the power system. This category corresponds to the reconnaissance phase in the cyber kill chain.

IV. EXPERIMENTAL RESULTS

In this section, the experimental results of the proposed methods are presented. This section provides an overview of the experimental setup, including the cyber-physical power system co-simulation setup and dataset. Subsequently, two main results are presented in this section. First, the semi-supervised deep packet inspection is presented to quantitatively assess the capability of zero-day detection. The experimental results demonstrate the effectiveness of the proposed method in identifying unknown attacks despite having a limited amount of training data. This solution intends to address the problem of the limited availability of the dataset acquired from the OT communication traffic of power grids. Second, the spatio-temporal anomaly correlation and prediction demonstrate the prediction performance for subsequent anomalies resulting from APT attacks. The experimental results present the superiority of the proposed EGC-LSTM in comparison with the state-of-the-art graph spatio-temporal deep learning models. A more detailed explanation is provided in the following subsections.

A. Experimental Setup

The experiments in this work are performed using the CPPS model of the power grid represented in Fig. 1. The power system is simulated in real-time using a Root Mean Square (RMS) dynamic model of the IEEE 39-bus test system in

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

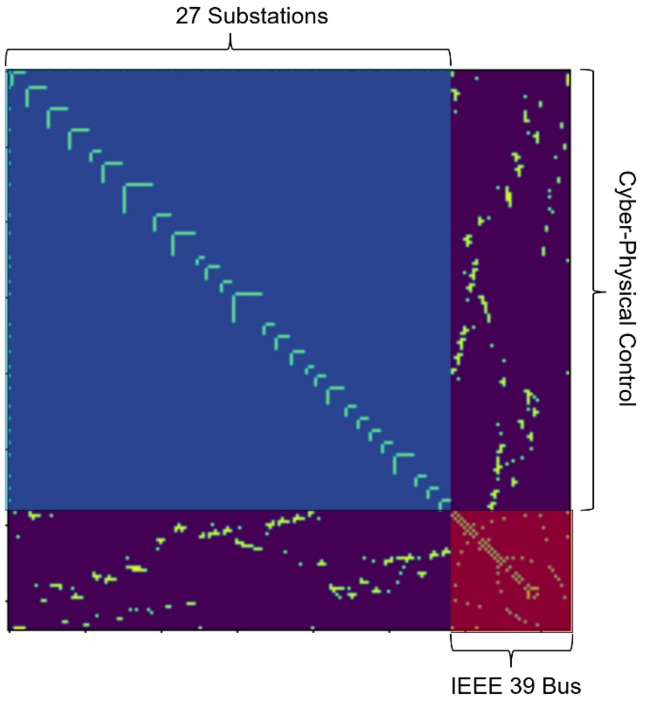

Fig. 6. CPPS Adjacency Matrix.

DIgSILENT PowerFactory. The CPS model employs OPC UA to establish a connection between the time domain simulation of the power grid and simulated IT-OT communication networks. The OT network emulation utilizes Mininet deployed on a total of 10 virtual servers. It consists of 27 user-defined substations, 118 measurement devices, and over 800 data points to emulate the OT communication network of IEEE 39-bus test power system. The CPPS model comprises a total of 185 nodes, consisting of 146 OT nodes and 39 physical nodes from the IEEE 39-bus system. Fig. 6 depicts the adjacency matrix representing the connection between 185 nodes of CPPS that are associated with CPSIM. The nodes that are connected are represented by the value 1 in the adjacency matrix, whereas the nodes that are not connected are represented by the value 0. The blue area corresponds to the OT adjacency matrix, while the red area corresponds to the IEEE 39-bus adjacency matrix. The cyber-physical control region illustrates the functional connectivity between the OT and power system, coupling the cyber and physical systems together.

The SCADA functionalities in the OT communication network are achieved by implementing customized Python code on each Mininet host. The OT devices include MUs, RTUs, IEDs, database server, gateway, human machine interfaces, and control center. The measurement values and control set points are communicated across the OT network using TCP/IP packets. In this research, we focus on the control traffic associated with CBs control. As shown in Fig. 1, the CPPS model is integrated with the SDN application to monitor OT traffic payload using SDN-enabled switch interfaces. With this capability, the CPPS model performs traffic monitoring in substations and control center.

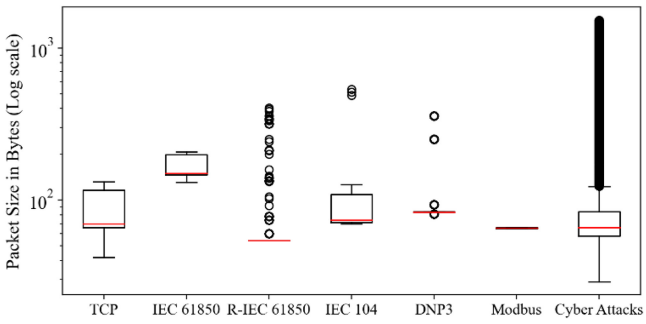

Fig. 7. Statistical box plot from normal traffic and cyber attacks.

From the CPPS co-simulation, the experiment collects two types of data. The first type of data is network traffic from the OT communication network collected as *.pcap* files. The traffic from multiple locations contains the source and destination addresses. This is processed using a traffic dispersion graph. In addition, the packet payload is classified as normal or anomalous using semi-supervised deep packet inspection. The second type of data is the CB status collected from the DIgSILENT PowerFactory simulation, which represents the status of the power system. Subsequently, both the cyber and physical system data are combined into the CPPS dataset in the CPSIM. The historical wide-area CPPS data from the CPSIM is used as input parameters for the EGC-LSTM. Based on this information, the EGC-LSTM performs spatio-temporal anomaly correlation and prediction. In this experiment, the ECG-LSTM model combines graph convolution and LSTM with hidden state vector parameters with the size of 32. Subsequently, the EGC-LSTM uses the ReLU as an activation function.

B. Semi-Supervised Deep Packet Inspection for OT Anomaly Detection

Our research uses OT traffic generated from the CPPS simulation to evaluate the performance of DPI. The simulation in Mininet produced TCP/IP traffic in the OT network. In the CPPS model, we test several cyber attacks scenarios, i.e., Denial of Service (DoS), network scanning, exploits, and malwares. In addition to the OT traffic generated by our simulation, we verify the model's performances by using open OT traffic datasets, i.e., IEC 61850 [44], Routable IEC 61850 [45], IEC 104 [46], DNP 3 [47], and Modbus [48]. Furthermore, we also incorporate samples of cyber attack datasets [49] and Industroyer malware traffic samples [50]. A total of 7.71 GB of *.pcap* data is collected from the OT traffic samples for the evaluation of semi-supervised DPI. Fig. 7 depicts the statistical distribution of packet size using a box plot across several OT traffic categories. Overall, the average size of the normal OT traffic from various protocols is 118.599 bytes, and 304.735 bytes for cyber attacks. In order to handle the size of the OT traffic, we use a 16x16 convolutional input with a total capacity of 256 bytes. When the traffic exceeds 256 bytes, the extra bytes are discarded. Conversely, when the traffic is less than 256 bytes, the remaining spaces are filled with zeros. The top part of Fig. 8 shows the image

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

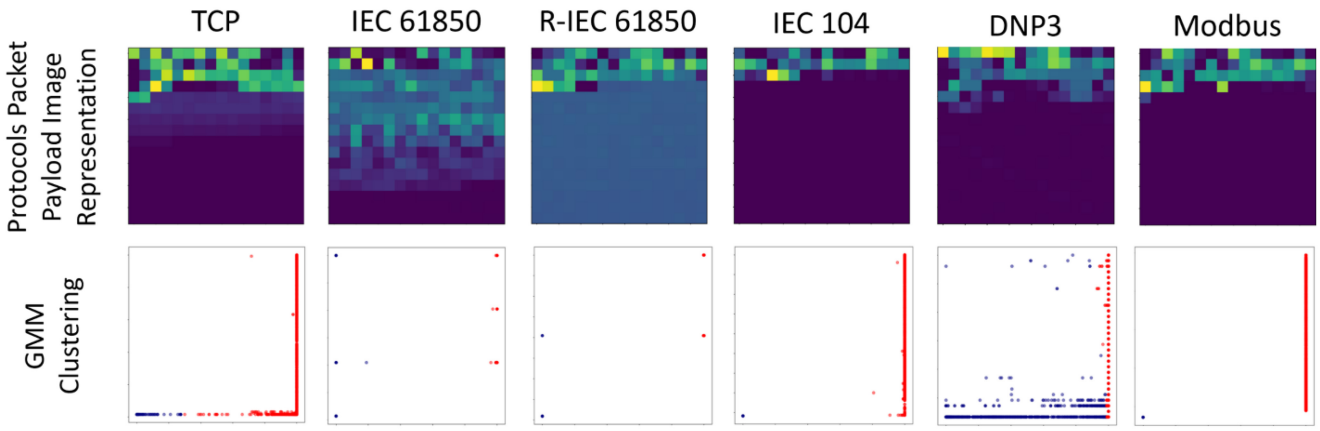

Fig. 8. OT traffic images representation and the result of Gaussian Mixture with partial labelling for each protocols.

representation from each tested OT traffic. This 2D data is used as input for the supervised CNN algorithms.

The outputs from CNN and HD generate vectors for GMM with partial labelling. The bottom section of Fig. 8 depicts the result from GMM with partial labelling for all protocols. K represents the number of classes in the GMM classifier. This parameter value is decided based on the number of classes in the tested dataset. The GMM uses the probability density function of multivariate Gaussian distribution with a full covariance matrix. The GMM implementation does not utilize an explicit distance metric, e.g., Euclidean distance. GMM uses Mahalanobis distance to calculate a point's likelihood for a given Gaussian component. The Mahalanobis distance quantifies the spatial separation between an individual data point and a given probability distribution.

In our scenarios, we create a pair of two classes from the normal/baseline OT protocols with cyber attack traffic. In this experiment, we evaluate the performance of the GMM with partial labelling with different proportion of training and test data. The labeled data proportion selection is carried out by running GMM with a variation of the labeled data proportion between 1% and 30%. As shown in Fig. 9, the majority of the tested dataset only required less than 5% labelled data to achieve the minimum MSE. However, for the DNP3, it required 19% labelled data to achieve the best performance. This is because of the DNP3 characteristics, which has more data variation, as shown in Fig. 8. In addition, compared to other datasets, the DNP3 dataset has a substantial amount of DNP3-modified attack packets, which resemble the normal DNP3traffic. Therefore, the clustering plot of DNP3 is different from that of the other protocols. Consequently, to achieve the best performance for all datasets, the experiment incorporates a 20% proportion of labeled data.

The x-axis represents the probability scores from the CNN, and the y-axis represents the HD scores. The red dots indicate the cyber attack traffic that is associated with a higher CNN anomaly probability and HD score close to one. Conversely, the blue dots indicate normal OT traffic that has a lower probability of CNN anomalies and a low HD score nearing zero. Fig. 10 depicts the Receiver Operating Characteristic (ROC) curve from all traffic categories and Area Under

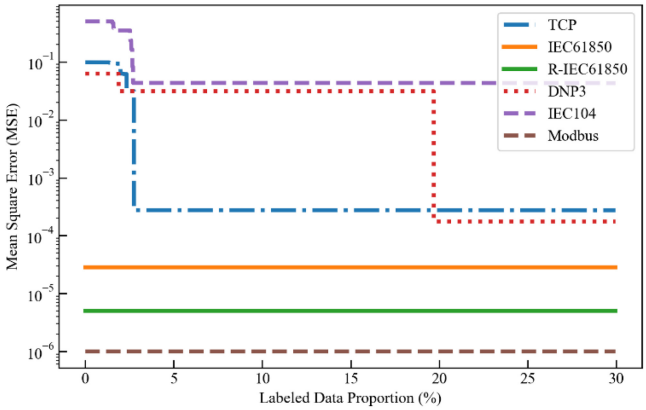

Fig. 9. Proportion of labeled data impact on the MSE for all protocols.

the Curve (AUC) score. This plot indicates that the semi-supervised DPI and GMM with partial labelling provides a good classification performance.

C. Spatio-Temporal Anomaly Correlation and Prediction

The anomaly detection result generated from the semi-supervised DPI is further processed using TDG and recorded in CPSIM, together with the CB's status retrieved from the power grid. During instances of cyber attacks, the CPSIM matrix will deviate from its normal state (all zeroes). As the attack progresses, particular elements of CPSIM are shifting to a value of 1. In the matrix, the value 1 corresponds to traffic anomaly or an open CB in the power grid. Based on this constraint, the transition on the CPSIM matrix will be varied depending on the cyber attack scenarios and location.

This research performs 220 cyber attacks scenarios with variation of location and methods. These scenarios serve as primary data to evaluate the performance of EGC-LSTM. In addition, we also perform benchmarking with the state-of-the-art graph-based spatio-temporal deep learning models, i.e., GConvLSTM and GConGRU [39], TGCN [40], and GC-LSTM [41]. Table II shows the performance comparison of the tested models based on MSE. A smaller MSE indicates superior prediction results. Our proposed strategy with sequence and NN filter reduces the MSE for all models.

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

TABLE II
MSE SCORES COMPARISON OF GRAPH-BASED SPATIO-TEMPORAL DEEP LEARNING MODELS

Combinations	Performance Parameters	Tested Models				
		GConvLSTM	GConvGRU	TGCN	GC-LSTM	EGC-LSTM
Original	MSE ± SDev	0.055±0.017	0.054±0.018	0.052±0.021	0.045±0.017	**0.035±0.014**
	Time ± SDev	528±112	371±91	253±67	304±84	**307±79**
+ Seq. Filter	MSE ± SDev	0.034±0.012	0.039±0.013	0.037±0.018	0.034±0.016	**0.021±0.011**
	Time ± SDev	532±118	381±97	259±64	308±81	**309±80**
+ NN Filter	MSE ± SDev	0.027±0.011	0.037±0.014	0.026±0.011	0.043±0.019	**0.023±0.012**
	Time ± SDev	542±145	378±94	258±66	311±81	**323±72**
+ Seq. Filter & NN Filter	MSE ± SDev	0.0026 ±0.0012	0.0011±0.0007	0.0016±0.008	0.0019±0.0009	**0.0003±0.0002**
	Time ± SDev	549±127	385±99	257±69	313±83	**324±84**

Seq. Filter = Sequential Filter, NN. Filter = Neural Network Filter, SDev = Standard Deviation, Time in milliseconds

TABLE III
PERFORMANCE COMPARISON OF GRAPH-BASED SPATIO-TEMPORAL DEEP LEARNING MODELS WITH SEQUENTIAL AND NEURAL NETWORK FILTERS

Parameter	Original	Seq. Filter	NN Filter	Seq. and NN Filter
Average MSE	0.0482 (+ 0 %)	0.033 (- 31%)	0.0312 (- 35%)	0.0015 (- 97%)
Average Time (ms)	352.6 (+ 0%)	357.8 (+1.5%)	362.4 (+ 2.8%)	365.6 (+ 3.7%)

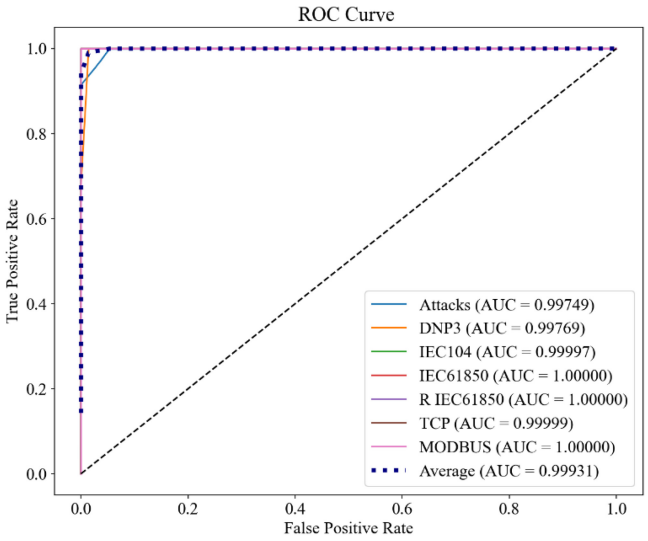

Fig. 10. ROC curve from all traffic categories.

Table III shows the average performance comparison of the five original graph-based spatio-temporal deep learning models and their variants with Algorithm 1 Sequential and NN filter. Table III quantifies the impact of the implementation of Seq. filter, NN filter, and combination of Seq. filter and NN filter. The standalone implementation of the Sequential and NN filter reduces the MSE by 31% and 35%, respectively. Meanwhile, the integration of both filters reduces MSE by 97%. The best MSE of 0.0003 is achieved in the proposed EGC-LSTM that implements Bayesian optimization, sequence filter, and NN filter. Besides reducing the MSE, as shown in Table III, the filters also increase the computing time by 1.5∼3.7%.

Fig. 11 depicts the sample prediction result from EGC-LSTM for cases 76 and 218. In case 76, the cyber attack started from substation 9 and compromised merging unit 9.1. This MU has the capability to control the CB of the power line between Bus 6 and Bus 7. During the state n, EGC-LSTM can predict incoming events in $n+1$ before they actually happen. In state $n+1$, the method can predict the circuit breaker that will be affected after the power line between Bus 6 and 7 is disconnected. For case 218, the cyber attack is starting from substation 27. Compared to case 76, this scenario shows different highlighted anomalous locations. In the majority of cases, the breaker opening attack will not trigger other breakers to open. However, in cases 76 and 218, the opening a few breakers will trigger more breakers to open due to protection schemes implemented in the IEEE-39 bus model.

D. Detection for Zero Day Attack Scenarios

Considering the complexity of CPPS topology, there are various possibilities of cyber attacks scenarios. To address this concern, we implement the resilient associative method of vector database search using KNN as depicted in Fig. 4. To evaluate this method, we generate 20 new scenarios and test several vector search strategies, i.e., KNN [51], Euclidean Distance (ED), K Decision Tree (KDT) [52], Hierarchical Navigable Small World (HNSW) [53], K Means (KM) [54], and Locality Sensitive Hashing (LSH) [55]. Fig. 12 shows a computation time comparison with variety of data quantities. Compared to the tested methods, KNN provides the most stable computational performance. Methods such as KDT, HNSW, and LSH provide a faster search time. However, these methods need preliminary computation to preprocess or generate the hash map. Therefore, these methods may not be suitable for fast pace changing data. Based on the search evaluation, the tested methods find the most related

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

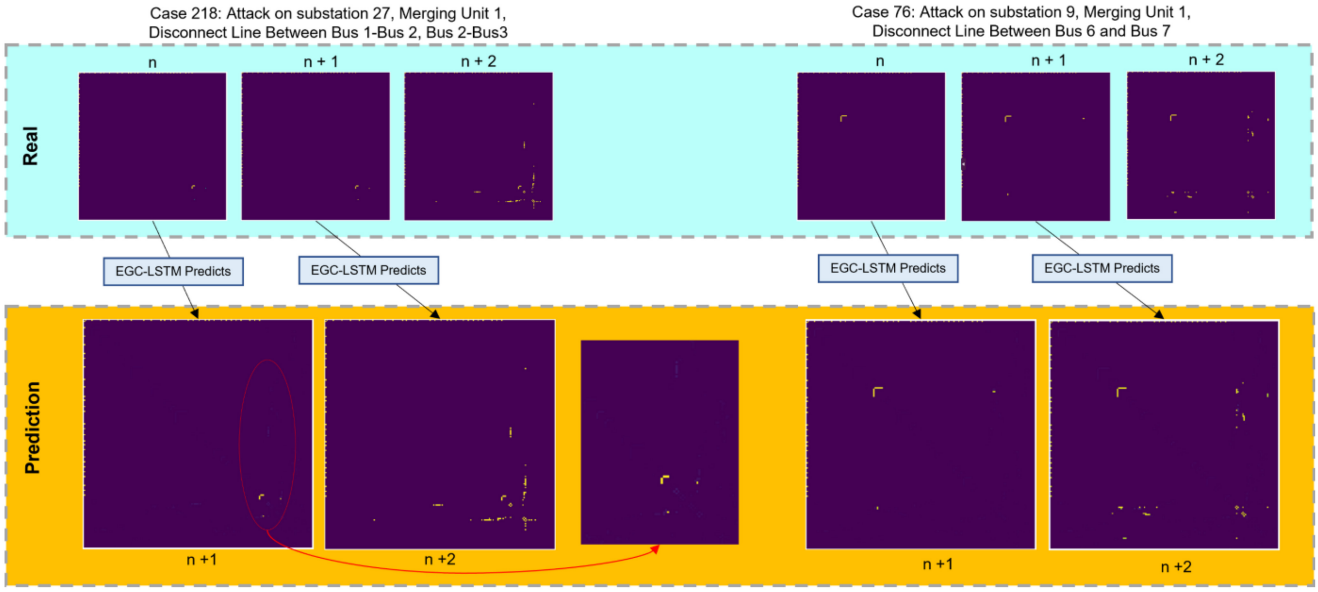

Fig. 11. Sample EGC-LSMT Prediction Results from case 76 and 218.

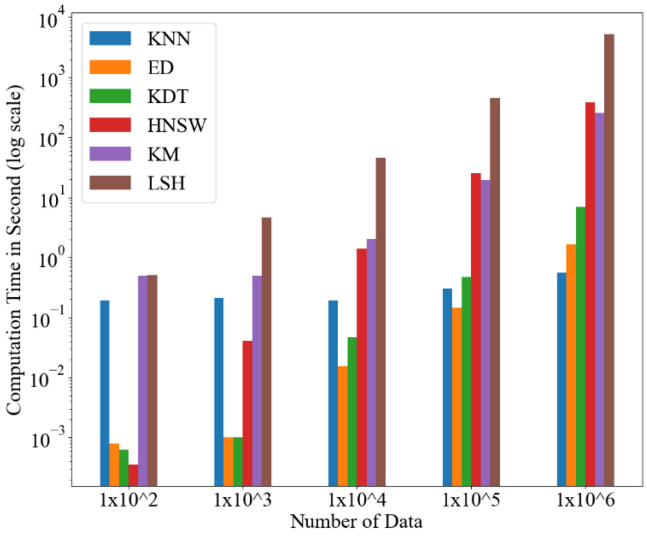

Fig. 12. Search algorithms computation time comparison.

scenarios from the known scenarios. This strategy will serve as a resilient mechanism in identifying new possible APT cyber attack scenarios, i.e., zero-day attacks.

V. Conclusion and Future Works

With the growing threat of cyber attacks on power grids, it is now more critical than ever to strengthen the attack detection capabilities in OT communication networks. It is important to note that, from 2024 onward, we will be living in a world where AI plays an increasing role alongside the advancement of AI models, i.e., deep learning, physic-informed, and generative AI models. In this context, our research aligns with this trend by proposing AI-based spatio-temporal APTs detection, correlation, and prediction in power systems. The implementation of deep learning for intrusion detection systems is becoming increasingly crucial to address the sophisticated and evolving nature of APTs. The proposed methods comprise of semi-supervised DPI, CPSIM, EGC-LSTM, a resilient associative method, and CPPS log comparator. The EGC-LSTM outperforms the state-of-the-art graph-based spatio-temporal deep learning model with the lowest MSE score of 0.0003. The proposed methods are also capable of identifying zero-day attacks, locating anomalous elements in the CPPS, and predicting the potential impact of anomalies. In contrast to most research that emphasizes the physical anomalies that occur during the later stages of a cyber attack on power systems, the proposed methods have the potential to detect cyber attacks during the early phases of the cyber kill chain. In addition, AI-based intrusion detection provides an online situational awareness for power system operators to pinpoint system-wide anomaly locations in near real-time and preemptively mitigate APTs at an early stage before causing adverse impacts.

In this work, the methods primarily focus on cyber anomalies that originate from external threat actors. The external APT required a lateral movement to reach its final objective in a timely fashion. These scenarios provide an opportunity for the early identification of the APT. However, there is also possibly an insider threat that can cause an immediate impact on the CPPS. Currently, the insider threat constraint is omitted from our objectives. Therefore, insider threat detection can become a potential future research direction, along with external threat detection. Furthermore, to enhance the methodology, it is essential to conduct comprehensive testing of the OT communication traffic by including a wider range of APT scenarios.

References

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

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

[2] M. J. Assante, R. M. Lee, and T. Conway, '*ICS Defense Use Case no. 6: Modular ICS Malware*, Electr. Inf. Sharing Center (E-ISAC), Washington, DC, USA, Aug. 2017.

[3] SANS ICS, "Analysis of the cyber attack on the Ukrainian power grid," Electr. Inf. Sharing Center (E-ISAC), Washington, DC, USA, White Paper, Mar. 2016.

[4] K. Proska et al. "Sandworm disrupts power in Ukraine using a novel attack against operational technology." Nov. 2023. [online]. Available: https://www.mandiant.com/resources/blog/sandworm-disrupts-power-ukraine-operational-technology

[5] A. Alshamrani, S. Myneni, A. Chowdhary, and D. Huang, "A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1851–1877, 2nd Quart., 2019.

[6] A. Ahmad, J. Webb, K. C. Desouza, and J. Boorman, "Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack," *Comput. Secur.*, vol. 86, pp. 402–418, Sep. 2019.

[7] B. Stojanović, K. Hofer-Schmitz, and U. Kleb, "APT datasets and attack modeling for automated detection methods: A review," *Comput. Secur.*, vol. 92, May 2020, Art. no. 101734.

[8] B. Tang et al., "Advanced Persistent Threat intelligent profiling technique: A survey," *Comput. Electr. Eng.*, vol. 103, Oct. 2022, Art. no. 108261.

[9] 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

[10] Q. Wang, X. Cai, Y. Tang, and M. Ni, "Methods of cyber-attack identification for power systems based on bilateral cyber-physical information," *Int. J. Electr. Power Energy Syst.*, vol. 125, pp. 1–12, Feb. 2021.

[11] R. Mitchell and I.-R. Chen, "Behavior-rule based intrusion detection systems for safety critical smart grid applications," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1254–1263, Sep. 2013.

[12] G. M. Coates, K. M. Hopkinson, S. R. Graham, and S. H. Kurkowski, "Collaborative trust-based security mechanisms for a regional utility Intranet," *IEEE Trans. Power System*, vol. 23, no. 3, pp. 831–844, Aug. 2008.

[13] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1773–1786, Aug. 2016.

[14] 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.

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

[16] A. Aldweesh, A. Derham, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: A survey taxonomy and open issues," *Knowl. Based Syst.*, vol. 189, pp. 1–19, Feb. 2020.

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

[18] 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.

[19] M. Norouzi, D. J. Fleet, and R. Salakhutdinov, "Hamming distance metric learning," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1070–1078.

[20] Y. Zhou, J. He, and H. Gu, "Partial label learning via Gaussian processes," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4443–4450, Dec. 2017.

[21] H.-C. Yan, J.-H. Zhou, and C. K. Pang, "Gaussian mixture model using semisupervised learning for probabilistic fault diagnosis under new data categories," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 4, pp. 723–733, Apr. 2017.

[22] M. S. Yang, C.-Y. Lai, and C.-Y. Lin, "A robust EM clustering algorithm for Gaussian mixture models," *Pattern Recognit.*, vol. 45, no. 11, pp. 3950–3961, Nov. 2012.

[23] T. Bilot, N. E. Madhoun, K. A. Agha, and A. Zouaoui, "Graph neural networks for intrusion detection: A survey," *IEEE Access*, vol. 11, pp. 49114–49139, 2023.

[24] C. Wang and H. Zhu, "Wrongdoing monitor: A graph-based behavioral anomaly detection in cyber security," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 2703–2718, 2022.

[25] S. Kim, K.-J. Park, and C. Lu, "A survey on network security for cyber–physical systems: From threats to resilient design," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1534–1573, 3rd Quart., 2022.

[26] S. Li, A. Pandey, B. Hooi, C. Faloutsos, and L. Pileggi, "Dynamic graph-based anomaly detection in the electrical grid," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 3408–3422, Sep. 2022.

[27] Z. Wang, W. Jiang, J. Xu, Z. Xu, A. Zhou, and M. Xu, "Grid2Vec: Learning node representations of digital power systems for anomaly detection," *IEEE Trans. Smart Grid*, vol. 15, no. 5, pp. 5031–5042, Sep. 2024.

[28] X. Ling, Y. Rho, and C.-W. Ten, "Predicting global trend of cybersecurity on continental honeynets using vector autoregression," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur.*, Bucharest, Romania, 2019, pp. 1–5.

[29] Z. Yang, S. Zhang, C.-W. Ten, T. Liu, X. Pang, and H. Sun, "Implementation of risk-aggregated substation testbed using generative adversarial networks," *IEEE Trans. Smart Grid*, vol. 14, no. 1, pp. 677–689, Jan. 2023.

[30] P. Gao et al., "Detecting unknown threat based on continuous-time dynamic heterogeneous graph network," *Wireless Commun. Mobile Comput.*, vol. 2022, no. 1, pp. 1–10, Jul. 2022.

[31] C. K. Chen, S. C. Lin, S. C. Huang, Y. T. Chu, C. L. Lei, and C. Y. Huang, "Building machine learning-based threat hunting system from scratch," *Digit. Threats Res. Pract.*, vol. 3, no. 3, pp. 1–21, Sep. 2022.

[32] J. Yang and Z. Yue, "Learning hierarchical spatial-temporal graph representations for robust multivariate industrial anomaly detection," *IEEE Trans. Ind. Informat.*, vol. 19, no. 6, pp. 7624–7635, Jun. 2023.

[33] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, vol. 379, no. 2194, pp. 1–14, Apr. 2021.

[34] M. Prais and A. Bose, "A topology processor that tracks network modifications," *IEEE Trans. Power Syst.*, vol. 3, no. 3, pp. 992–998, Aug. 1988.

[35] S. N. Talukdar, E. Cardozo, and T. Perry, "The operator's assistant–an intelligent, expandable program for power system trouble analysis," *IEEE Trans. Power Syst.*, vol. 1, no. 3, pp. 182–187, Aug. 1986.

[36] S. Yuan and X. Wu, "Deep learning for insider threat detection: Review, challenges and opportunities," *Comput. Secur.*, vol. 104, May 2021, Art. no. 102221.

[37] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1–9, Dec. 2012.

[38] D. Q. Le, T. Jeong, H. E. Roman, and J. W.-K. Hong, "Traffic dispersion graph based anomaly detection," in *Proc. 2nd Symp. Inf. Commun. Technol.*, Oct. 2011, pp. 36–41.

[39] Y. Seo, M. Defferrard, P. Vandergheynst, and X. Bresson, "Structured sequence modeling with graph convolutional recurrent networks," in *Proc. Int. Conf. Neural Inf. Process.*, 2018, pp. 362–373.

[40] L. Zhao et al., "T-GCN: A temporal graph convolutional network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Sep. 2020.

[41] J. Chen, X. Xu, Y. Wu, and H. Zheng, "GC-LSTM: Graph convolution embedded LSTM for dynamic network link prediction," *Appl. Intell.*, vol. 52, pp. 7513–7528, May 2022.

[42] J. Pan, J. Wang, and G. Li, "Survey of vector database management systems," Oct. 2023, *arXiv.2310.14021*.

[43] A. J. Wilson, D. R. Reising, R. W. Hay, R. C. Johnson, A. A. Karrar, and T. Daniel Loveless, "Automated identification of electrical disturbance waveforms within an operational smart power grid," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4380–4389, Sep. 2020.

[44] A. Presekal, A. Ştefanov, V. S. Rajkumar, and P. Palensky, "Cyber forensic analysis for operational technology using graph-based deep learning," in *Proc. IEEE Int. Conf. Commun.*, Glasgow, U.K., Oct. 2023, pp. 1–7.

[45] 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, Art. no. 101415.

[46] P. Matoušek, O. Ryšavý, and P. Grofčík, Mar. 16, 2022, "ICS dataset for smart grid anomaly detection," Dataset, IEEE Dataport. [Online]. Available: https://ieee-dataport.org/documents/ics-dataset-smart-grid-anomaly-detection

[47] P. R. Grammatikis, V. Kelli, T. Lagkas, V. Argyriou, and P. Sarigiannidis, Nov. 22, 2022, "DNP3 intrusion detection dataset," Dataset, IEEE Dataport. [Online]. Available: https://ieee-dataport.org/documents/dnp3-intrusion-detection-dataset

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.

[48] T. Ardley. "ICS security tools." GitHub. 2023. [Online]. Available: https://github.com/ITI/ICS-Security-Tools/
[49] H. Kang et al., Sep. 27, 2019, "IoT network intrusion dataset," Dataset, IEEE Dataport. [Online]. Available: https://ieee-dataport.org/open-access/iot-network-intrusion-dataset
[50] E. Hjelmvik. "Industroyer IEC 104 analysis." NETRESCEC. 2022. [Online]. Available: https://www.netresec.com/?page=Blog&month=2022-04&post=Industroyer2-IEC-104-Analysis
[51] Y. Zhang, J. Wu, J. Wang, and C. Xing, "A transformation-based framework for KNN set similarity search," *IEEE Trans. on Know. and Data Eng.*, vol. 32, no. 3, pp. 409–423, Mar. 2020.
[52] J. L. Bentley, "*K*-d trees for semidynamic point set," in *Proc. 6th Annu. Symp. Comput. Geom.*, May 1990, pp. 187–197.
[53] Y. A. Malkov and D. A. Yashunin, "Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 824–836, Apr. 2020.
[54] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.
[55] A. Gionis, P. Indyk, and R. Motwani, "Similarity search in high dimensions via hashing," in *Proc. 25th VLDB Conf.*, vol. 99, 1999, pp. 1–12.

Alfan Presekal (Member, IEEE) received the bachelor's degree in computer engineering from Universitas Indonesia in 2014, and the master's degree in secure software systems from the Department of Computing, Imperial College London, U.K., in 2016. He worked as an Assistant Professor of Computer Engineering with the Department of Electrical Engineering, Universitas Indonesia. He is currently a Researcher of Cyber Resilient Power Grids within Intelligent Electrical Power Grids with the Department of Electrical Sustainable Energy, Delft University of Technology. His main research interest includes cyber security, cyber–physical systems, and artificial intelligence.

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 of Intelligent Electrical Power Grids with the Department of Electrical Sustainable Energy, TU Delft, The Netherlands. He is the Director of the Control Room of the Future Technology Centre. He is leading the Cyber Resilient Power Grids 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.

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 the 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.

Peter Palensky (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. and Habilitation degrees from the Vienna University of Technology, Austria, in 1997, 2001, and 2015, respectively. He 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 became appointed as the Head of the Business Unit, Austrian Institute of Technology in sustainable building technologies, where he was the first Principal Scientist of Complex Energy Systems. In 2014, he was appointed as a Full Professor of Intelligent Electric Power Grids with TU Delft, The Netherlands. He is currently the Head of the Department of Electrical Sustainable Energy, Delft University of Technology. His research interests include energy automation networks, smart grids, and modeling intelligent energy systems. He is active in international committees, such as ISO or CEN. 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. He also serves as an IEEE IES adcom member-at-large in various functions for IEEE.

Authorized licensed use limited to: TU Delft Library. Downloaded on March 03,2025 at 11:52:14 UTC from IEEE Xplore. Restrictions apply.