Attack Graph Model for Cyber-Physical Power Systems Using Hybrid Deep Learning

Presekal, A.; Stefanov, Alexandru; Subramaniam Rajkumar, Vetrivel; Palensky, P.

DOI
10.1109/TSG.2023.3237011

Publication date
2023

Document Version
Final published version

Published in
IEEE Transactions on Smart Grid

Attack Graph Model for Cyber-Physical Power Systems Using Hybrid Deep Learning

Alfan Presekal, *Member, IEEE*, Alexandru Ştefanov, *Member, IEEE*, Vetrivel Subramaniam Rajkumar, *Student Member, IEEE*, and Peter Palensky, *Senior Member, IEEE*

***Abstract*—Electrical power grids are vulnerable to cyber attacks, as seen in Ukraine in 2015 and 2016. However, existing attack detection methods are limited. Most of them are based on power system measurement anomalies that occur when an attack is successfully executed at the later stages of the cyber kill chain. In contrast, the attacks on the Ukrainian power grid show the importance of system-wide, early-stage attack detection through communication-based anomalies. Therefore, in this paper, we propose a novel method for online cyber attack situational awareness that enhances the power grid resilience. It supports power system operators in the identification and localization of active attack locations in Operational Technology (OT) networks in near real-time. The proposed method employs a hybrid deep learning model of Graph Convolutional Long Short-Term Memory (GC-LSTM) and a deep convolutional network for time series classification-based anomaly detection. It is implemented as a combination of software defined networking, anomaly detection in communication throughput, and a novel attack graph model. Results indicate that the proposed method can identify active attack locations, e.g., within substations, control center, and wide area network, with an accuracy above 96%. Hence, it outperforms existing state-of-the-art deep learning-based time series classification methods.**

***Index Terms*—Anomaly detection, cyber-physical system, graph neural network, network security, software defined networking, throughput, time series analysis.**

NOMENCLATURE

G	Graph
V	Known vertices/ nodes
E	Edges
$A_{i,j}$	Adjacency matrix with where i and j represent the node index numbers
$\widehat{A}$	Modified adjacency matrix where $\widehat{A} = A + I$ (identity matrix)
GCN_t^k	Graph convolutional equation for each k hop and time t
k	Number of neighbor hops in the graph
W_{gcn}	Weight of graph convolutional neural network
$\odot$	Hadamard product multiplication operator
$\{s_1, s_2, \ldots, s_n\} \in S$	S as all observable substations, and each individual substation s_n
$X_t, X \in s_n$	Data of network traffic for each time t for all nodes
$\{x_1, x_2, \ldots, x_n\} \in X$	Individual node traffic data
i_t	Input gate for long short-term memory
f_t	Forget gate for long short-term memory
o_t	Output gate for long short-term memory
c_t'	Internal cell state for long short-term memory
c_t	Transferable state for long short-term memory
h_t	Hidden state for long short-term memory
$W_i, W_f, W_o, W_c,$ U_i, U_f, U_o, U_c	Set of weights for long short-term memory
b_i, b_f, b_o, b_c	Set of biases for long short-term memory
σ	Sigmoid function
$tanh$	Hyperbolic tangent function
y_i^l	Convolution operation output for each l layers and i element
$ReLU$	Rectifier linear unit function
$\sum^{m-1} wy_{(i)}^{l-1} + b$	Convolution operation for layer l and element i with filter size (m), weight (w), and bias (b)
Λ	Attack graph
$\{a_i, \overline{a}_i\} \in V \in \Lambda$	Normal nodes (a_i), anomalous nodes ($\overline{a}_i$)
$\{u_i\} \notin V;\ u_i \in \Lambda$	Unidentified nodes (u_i)
G_{mean}	Geometric mean function

List of Acronyms

CNN	Convolutional Neural Network
CPS	Cyber-Physical System
DDoS	Distributed Denial-of-Service
EI	Expected Improvement
FCN	Fully Convolutional Neural Network

Manuscript received 12 August 2022; revised 30 November 2022; accepted 4 January 2023. Date of publication 16 January 2023; date of current version 23 August 2023. This work was supported by the Designing Systems for Informed Resilience Engineering (DeSIRE) Program of the 4TU Center for Resilience Engineering (4TU.RE). Paper no. TSG-01168-2022. *(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; V.SubramaniamRajkumar@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.2023.3237011.

Digital Object Identifier 10.1109/TSG.2023.3237011

GC-LSTM	Graph Convolutional Long Short-Term Memory
GCN	Graph Convolutional Network
GNN	Graph Neural Networks
IED	Intelligent Electronic Device
IT	Information Technology
LSTM	Long Short-Term Memory
MLP	Multi-Layer Perceptron
MU	Merging Unit
OT	Operational Technology
ROC	Receiver Operating Characteristic
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SDN	Software Defined Networking
TDG	Traffic Dispersion Graph
TSC	Time Series Classification
WAN	Wide Area Network

I. Introduction and Related Work

CYBER attacks on power grids are high-impact and low-frequency disturbances with a wide range of consequences. These could include but are not limited to, equipment damage, loss of load, and power system instability. In the worst-case scenario, cyber attacks and advanced persistent threats may cause system-wide cascading failures and a blackout. Therefore, cyber attacks on power grids are severe threats and have already been identified in the real world. For example, on December 23, 2015, a cyber attack was conducted on the power grid in Ukraine that resulted in a power outage, affecting 225,000 customers [1]. A more sophisticated cyber attack followed on December 17, 2016, resulting in a power outage in the distribution network, where 200 MW of load was left unsupplied [2]. The attackers employed several attack strategies and steps to achieve their objectives. These can be mapped with the seven stages of the cyber kill chain for an in-depth analysis of such an advanced persistent threat, i.e., reconnaissance, weaponization, delivery, exploitation, installation, command and control, and action on objectives [3] as depicted in Fig. 1. However, existing detection methods for cyber attacks on power grids are limited. Most of them are based on power system measurement anomalies that occur when an attack is successfully executed at the later stages of the cyber kill chain, e.g., false data injection [4], [5], [6], [7], [8], [9], [10], [11]. In contrast, in the aforementioned cyber attacks in Ukraine, the cyber kill chain lasted for more than six months between the reconnaissance and command and control stages. The latter caused power outages in a matter of minutes [1], [2], [12]. Hence, this highlights the urgency of timely early-stage attack detection through Information Technology-Operational Technology (IT-OT) system anomalies. Physical measurement-based anomaly detection is only valid for later stages in the cyber kill chain, i.e., command and control and actions on objectives. Therefore, in this research, we propose an early-stage anomaly detection method for OT systems. It is implemented in the control center to detect cyber attacks at the early stages of the cyber kill chain, based on throughput anomalies in OT communication traffic power system wide.

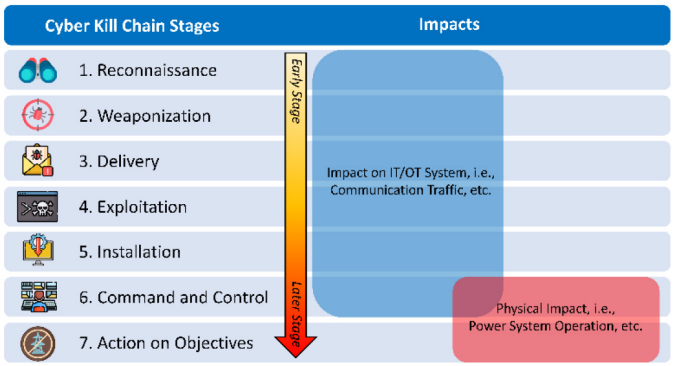


Fig. 1. Cyber kill chain stages and impacts.

Cyber attack detection on power grids have been extensively studied in recent years. Nonetheless, the majority of the existing research is focused on the identification of cyber attacks on power grids under False Data Injection (FDI) attack scenarios. These scenarios focus on analyzing power system measurements to identify anomalies in power grids [4], [5], [6], [7], [8], [9], [10], [11]. However, in the real-world cyber attacks on power grids reported in [1], [2], [12] adversaries did not perform FDI attacks. Instead, in the early stages of the cyber kill chain, attackers targeted the IT-OT communications. Therefore, in this research, we omit power system measurements under FDI attack scenarios and focus on the OT communication traffic anomalies.

There are four major methods reported in the literature for power grid communication traffic anomaly detection, i.e., signature-based [13], sequence-based [14], rule-based [15], [16], [17], and machine learning-based [18], [19], [20]. Recent research shows that machine learning-based methods are gaining increased attention and provide superior performance for anomaly detection [21], [22], [23]. Therefore, in this work, we focus on machine learning-based communication traffic anomaly detection. Our proposed model is based on a semi-supervised learning. It does not use signatures, sequences nor rules for detection and classification. The proposed model classifies OT network traffic into two categories, i.e., normal and anomalous, based on the network traffic throughput. Previous research in this area is discussed in [18], [20]. In [18], the authors used labeled communication packets from UNSW-NB15 and IDE2012/16 datasets as inputs to predict the Distributed Denial of Services (DDoS) attacks. Meanwhile, in [20], the authors use traffic data logs from Snort to create a sequence-based anomaly detection technique. However, both machine learning implementations do not use traffic throughput data, which is our research focus. Furthermore, the vast majority of machine learning-based anomaly detection methods only focus on IT systems [21], [22], [23], [24]. Even though the IT and OT systems of a utility are integrated, the traffic characteristics are distinct. The network traffic in OT systems is generated from automated processes with deterministic and homogenous behavior, whilst the IT system traffic consists of user-generated data with a stochastic behavior [25]. Hence, the implementation of traffic-based anomaly detection

for OT systems is fundamentally different from that of IT systems.

Amongst the machine learning-based traffic anomaly detection methods, most recent works use deep learning models that provide a better performance [22], [26]. In [27], the authors propose a deep reinforcement learning-based method for traffic flow matching control. They focus on detection of DDoS attacks that systematically trigger considerable anomalies in traffic throughput. Therefore, this method is not suitable to detect infinitesimally small changes in OT network traffic throughput, e.g., caused by stealthy attacks [27]. In [28], the authors used Convolutional Neural Network (CNN) for communication traffic classification. However, the CNN method cannot detect unknown cyber attacks because it depends on preliminary traffic data for the training. To address this gap, instead of using specific labeled data for each attack category, we use the quantitative anomaly. The quantitative anomaly detection uses the throughput of the OT communication traffic. The throughput is quantified as a time series to generate a unique waveform pattern as shown in [29], [30], [31]. Therefore, instead of classifying specific attack types or sequences, in this work we classify the time series traffic flow into two categories, i.e., normal and anomalous. In other related work, time series-based anomaly detection and classification were studied in [32], [33], [34], [35]. The state-of-the-art Time Series Classification (TSC) methods are based on deep learning models, as described in [34], [35]. However, based on our experiments, they do not perform well in the detection of stealthy attacks due to infinitesimally small changes in the traffic throughput. Additionally, these methods do not perform well due to imbalanced data that is indicated in their F1 and Geometric mean scores. Therefore, to address these challenges, we propose a novel hybrid deep learning model for anomaly detection in power grid OT network traffic. The hybrid model uses Graph Neural Networks (GNN), Long Short-Term Memory (LSTM), and CNN. It employs unsupervised learning to learn the complex behavior of OT network traffic throughput and supervised learning to classify the OT traffic.

GNN-based deep learning models have been implemented for various applications, e.g., residential load forecasting [36], detection of false data injection [37], road traffic prediction [38], and road traffic anomaly detection [39]. LSTM has been used to detect anomalies in Supervisory Control and Data Acquisition (SCADA) systems [40]. This method can detect anomalies based on temporal features of time series data. CNN has been proposed to detect anomalies in power system data [41]. It has advantages in learning spatial features and correlations of the datasets. In this research, we propose the application of a Graph-Convolutional Long Short-Term Memory (GC-LSTM) to preprocess the data of OT network traffic and generate traffic predictions. The output from the GC-LSTM is then used as an input for the CNN-based time-series classification. We generate an attack graph to identify in near real-time the active cyber attack locations in the power grid.

The attack graph provides topological information on the possible attack paths for a specific cyber attack on a given network. Hence, the attack graph is an important method to identify vulnerabilities in the system [42]. The knowledge about the attack path is also crucial to prevent and mitigate cyber attacks. At current, the attack graphs are mostly constructed based on vulnerability information obtained from network elements [43], [44]. This type of attack graph is not flexible, because it heavily depends on system vulnerability data. However, in this research, we propose an alternative attack graph map generation model, based on the online traffic monitoring in the OT networks of power grids. This is made possible through the wide deployment of an emergent technology, i.e., Software Defined Networking (SDN). SDN is a networking paradigm based on network virtualization and segregation of data and control planes [45]. In the SDN architecture, as seen in Fig. 2, there are three abstraction layers present, i.e., data plane, control plane, and management plane. The data plane represents locations of conventional communication networks, while control plane provides controllability over the data plane. Additionally, the management plane in SDN allows the deployment of network applications, e.g., attack graph model. Although SDN is an emergent paradigm in the field of computer networking, earlier research has investigated its implementation in cyber-physical power systems [46], [47], [48], [49], [50]. Earlier research has used SDN for anomaly detection based on traffic flow information [27], [51]. However, these works are not designed to detect anomalies triggered by cyber attacks in OT networks. In this research, we use SDN to monitor the network traffic in real-time, originating from the data plane of the OT Wide Area Network (WAN) for power systems. In summary, a critical examination of related state-of-the-art methods for communication traffic anomaly detection reveals the following. (1) Existing SDN applications for cyber-physical systems are not focused on cyber security of OT networks [27], [47], [48], [49], [50], [51]. (2) They are solely based on packet flow rules [51]. (3) They overlook the cyber kill chain and do not address any type of stealthy cyber attacks [27], [51].

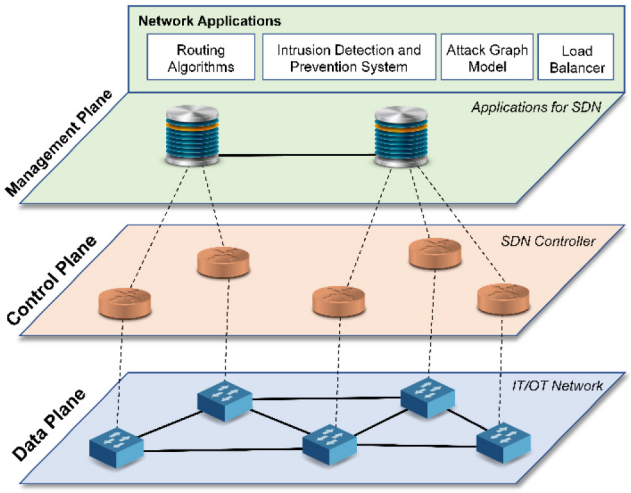


Fig. 2. Abstraction layers of SDN architecture.

The scientific contributions of this paper are as follows:

1) To the best knowledge of the authors, we propose the first known SDN-based online cyber attack situational awareness

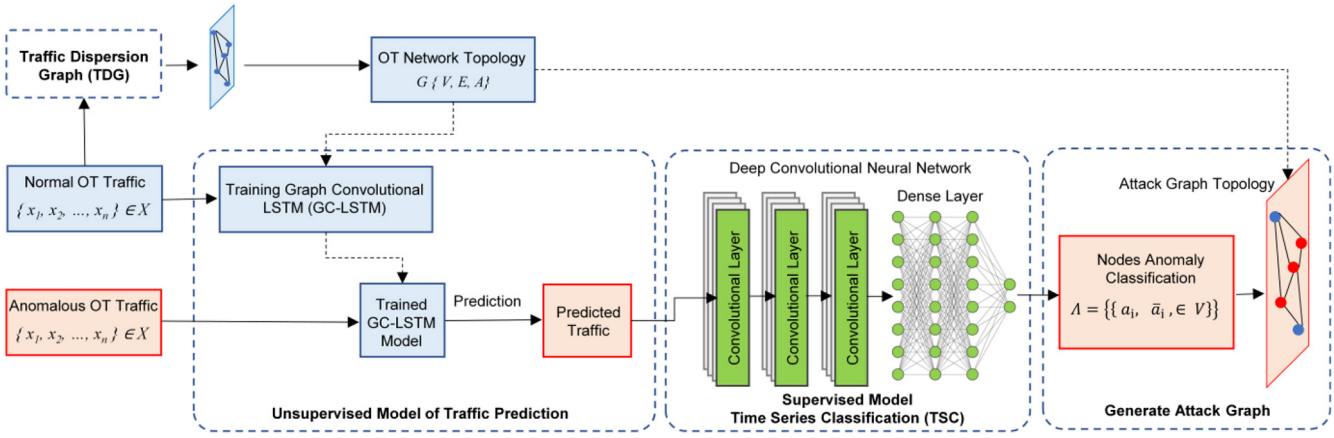


Fig. 3. Attack graph creation using CyResGrid method.

method, i.e., Cyber Resilient Grid (CyResGrid). It is specifically designed for anomaly detection using communication traffic throughput in OT networks for stealthy cyber attacks during the early stages of the cyber kill chain, e.g., network reconnaissance. Therefore, CyResGrid aids operators to locate and identify power system-wide cyber attacks in near real-time through an attack graph map.

2) We propose a hybrid deep learning model to classify the OT network traffic throughput as anomalous or normal. The model combines GC-LSTM and a deep convolutional network to detect OT network anomalies caused by cyber attacks. It outperforms existing state-of-the-art deep learning-based time series classifiers [34], [35], as indicated by Geometric mean and F1 scores. To achieve this, we use GC-LSTM for traffic normalization. Subsequently, to detect the anomaly, we design a deep convolutional network by tuning the hyperparameters through Bayesian optimization. Based on the network throughput monitoring and anomaly detection, we create an attack graph map of power system-wide cyber attacks, in near real-time.

3) As there is a strong need for synthetic Cyber-Physical System (CPS) datasets for research [52], we create the first synthetic dataset of OT communication traffic throughput, which is generated through a cyber-physical power system model. To the best of our knowledge, the majority of the existing datasets are not suitable for cyber security [53], [54], [55], [56], [57], [58], [59]. A cyber-physical system dataset was proposed in [60], [61] for intrusion detection. However, the OT traffic data is only in the form of signature-based logs without detailed traffic information [60], [61]. Therefore, in this research, we employ a CPS model of the power grid consisting of the physical system and associated OT communication networks. The model is used to co-simulate the power grid and OT network, from substations up to the control center. It also has cyber range capabilities to simulate various cyber attack scenarios. Based on this model, we generate a synthetic dataset of OT communication traffic throughput for cyber-physical power system operation under cyber attacks.

The paper is structured as follows. Section I is the introduction and Section II describes the methodology proposed in this paper, including cyber-physical system model, Traffic Dispersion Graph (TDG), GC-LSTM, TSC for anomaly detection, and the attack graph model. Section III provides the experimental results. Section IV presents the conclusions and future work.

II. Anomaly Detection and Attack Graph Model

In this section, the proposed methods for anomaly detection and attack graph modeling are introduced. Furthermore, we also elaborate on the cyber-physical model that serves as the basis for the aforementioned methods. Fig. 3 summarizes the methodology of anomaly detection and attack graph creation. The method consists of four steps as follows.

- *Step 1:* *GC-LSTM training and TDG.* The normal OT traffic is used to train the GC-LSTM model for traffic prediction. The process generates a trained GC-LSTM model. Additionally, the normal OT traffic is used to generate the OT network topology using a TDG.
- *Step 2:* *Deep CNN training.* The trained GC-LSTM model is used to predict the OT network traffic. The prediction is then used to train a Deep Convolutional Neural Network for TSC. This process generates a trained Deep CNN model for OT traffic classification.
- *Step 3:* *Online node classification.* This step monitors the online OT traffic as input for node classification. The trained GC-LSTM and Deep CNN are used sequentially to classify the nodes as normal or anomalous.
- *Step 4:* *Attack graph generation.* The node classification results from step 3 in conjunction with OT graph data from step 1 are used to generate the attack graph visualization.

A more detailed explanation of the method in each step is provided in the following subsections.

A. Cyber-Physical System Model

Detailed CPS models are needed for research on cyber security of power grids. They are used to simulate the power systems along with their associated IT-OT communication

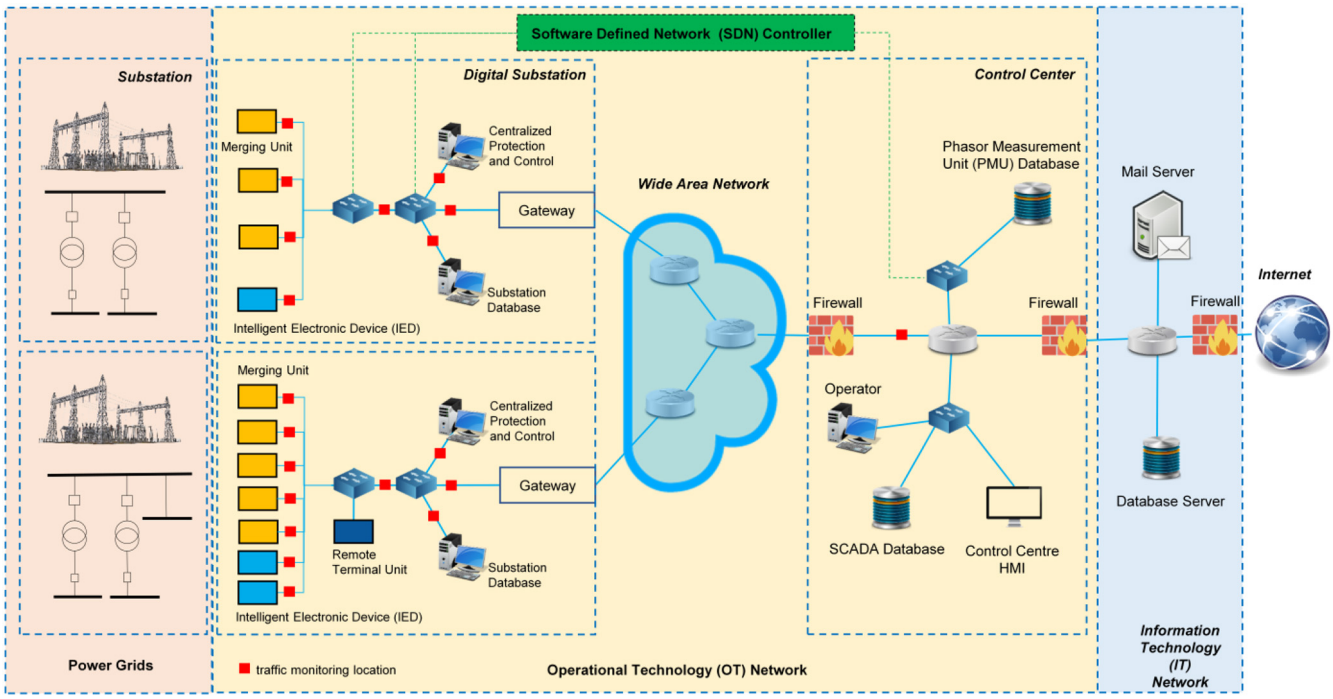


Fig. 4. Cyber-physical system model of the power grid with IT-OT communication networks.

networks and cyber events. The state-of-the-art in smart grid modeling and simulations is discussed in [62], [63], [64], [65], [66], [67], [68]. Hence, as part of our CPS model, we perform a co-simulation of the power grid and IT-OT systems, as depicted in Fig. 4.

The CPS model provides time-domain measurement data from substation bays, e.g., buses, lines, and generators, in the form of active and reactive power, voltage, and current measurements. All measurement data is then delivered from the substation to the control center via a WAN as SCADA telemetry. The SCADA data is also stored in local databases located in substations and the control center. For the cyber system, every node in the OT network is emulated using operating system-level virtualization. The network connectivity between substations, WAN, and control center is realized through network virtualization and SDN. With this configuration, the developed CPS architecture can model and simulate realistic OT network traffic for the power system.

The OT network is modeled based on custom functions for every device in the communication network. The measurement devices represent components, such as Merging Units (MUs), Remote Terminal Units (RTUs), and Intelligent Electronic Devices (IEDs). These devices perform data acquisition from the power grid, with a SCADA sampling rate of one sample per second. Legitimate control commands from the control center modify the set points for power grid controllers in real-time. For example, a control command can set a circuit breaker to open or close, set values for voltage, and active power set points of generator 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.

The CPS model is integrated with SDN capability that creates network virtualization using virtual switches. Based on Fig. 4, the OT and IT networks are present in the data plane layer of the SDN. Meanwhile, the control and management plane are represented by the SDN controller. Network virtualization allows the SDN controller to monitor and control traffic and run custom network applications. Fig. 4 depicts how the SDN controller is applied to the typical SCADA architecture. SDN improves the OT network monitoring and control by collecting OT communication traffic reports in the control center. The traffic observation points are visualized as red squares, which are distributed across the substations and control center. Using these points, we observe real-time OT network traffic from the control center to detect traffic anomalies for each observation location and create a power system wide attack graph.

B. *Traffic Dispersion Graph*

The TDG is an analytical model for communication traffic monitoring and analysis. The core idea for TDG is derived from the social behavior of hosts in a network [69]. Therefore, the flow of OT network traffic is analyzed based on the interactions between all hosts in the communication network. Based on this analysis, information related to communication sources and destinations is extracted. Furthermore, TDG represents nodal information using graph structures. Every host in a network is represented by a single node in a graph. On the other hand, communication between hosts is represented by connectivity between nodes, i.e., graph edges. Fig. 5 shows the TDG generation processes. Firstly, information on the IP address source and destination from flowing packets in

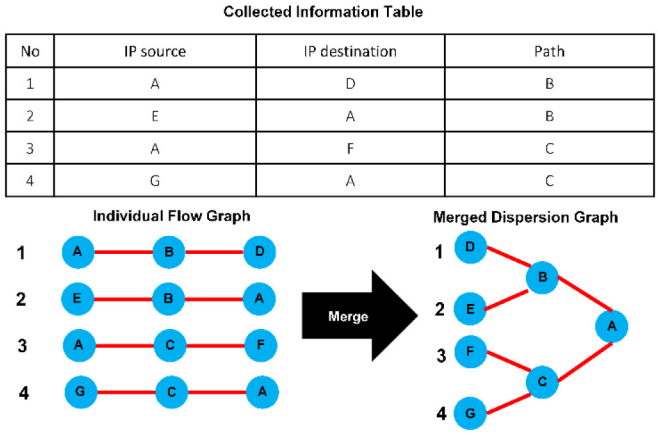


Fig. 5. Traffic Dispersion Graph (TDG) processes.

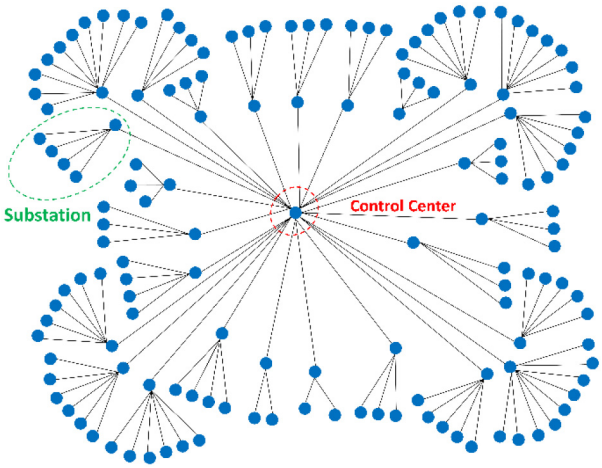


Fig. 6. Traffic dispersion graph of 27 substations.

the network is in the collected information table. Information about the path between two IP addresses is added based on prior knowledge of the network topology. The information in the table is then used to create an individual flow graph. Finally, all individual graph is converged into a dispersion graph which provides an overall topology of the network.

The TDG has previously been used to analyze communication network patterns. For example, a research proposed an application of TDG for anomaly detection based on the degree distribution values of a graph [70]. In our research, the CyResGrid method uses TDG to generate graph structures of the power system OT network. This includes a graphical representation of the OT network topology between the control center and substations. The anomalous nodes in the graph are then detected based on OT network traffic anomalies. In our model, the CPS topology of a power grid possesses a tree-like network structure. Fig. 6 illustrates the TDG of the OT network that is used in our model, containing a total of 27 substations and one control center. Every substation consists of OT devices, e.g., MUs, IEDs, RTUs, etc., and a communication gateway, e.g., router/firewall, that communicates with the control center.

In this research, the nodes represent traffic observation locations, while edges represent communication links between nodes. The traffic observation locations are situated in the Ethernet ports of virtual SDN switches that are directly connected to a host. All measurement data from each substation is sent to the control center via SCADA protocols, e.g., IEC 104 and DNP3. Thereby, this traffic flow allows the control center to gain a complete overview of the entire OT network. Using observation locations in the control center, the dispersion graph determines the nodes that actively communicate measurements. Also, the dispersion graph can determine unusual behavior, i.e., when a node is not sending measurement data or sending an abnormal quantity of traffic. In this research, anomaly detection works based on the total volume of observed network traffic, i.e., throughput, measured in KiloBytes per second (KBps). Furthermore, the dispersion graph can also identify unknown nodes with unidentified or unknown sources and destinations of IP Addresses or MAC Addresses.

C. Graph Convolutional Long Short-Term Memory

GC-LSTM aims to learn the traffic behavior of the OT network. Two machine learning models are applied in GC-LSTM, i.e., Graph Convolutional Network (GCN) and LSTM. GCN processes the OT network topological information expressed as a graph, along with localized features from neighboring communication nodes in the spatial domain. Subsequently, LSTM performs temporal learning based on time-series data of observed OT network traffic. The combination of GCN and LSTM has the advantage of learning from both the spatial and temporal domains. Various applications using graph-based spatial and temporal models were proposed in [36], [37], [38], [39]. In this research, we propose a novel method for nodal feature prediction based on communication network topology and features of neighboring nodes. CyResGrid proposes an innovative application of GC-LSTM to model the OT network traffic of the power system. It uses a hybrid combination of unsupervised and supervised models for OT traffic anomaly detection. The former is based on GC-LSTM which learns the complex behavior of OT network data and topology. Subsequently, the GC-LSTM generates traffic for the supervised predictions of the TSCs. The OT traffic model is then integrated with deep convolutional network-based TSC to generate an attack graph based on observed anomalies in the communication network traffic.

The graph structure of the OT network topology serves as the main input for GC-LSTM method. This graph structure is obtained from the TDG. It can be represented as $G = (V, E)$ where G is the graph, V represents the vertices/nodes and E represents the edges/links. The connection between the nodes in the graph is represented by the adjacency matrix A. Elements of the adjacency matrix are represented by $A_{i,j}$ where i and j represent the node index numbers, such that $A_{i,j}= 1$ when two nodes are connected, and $A_{i,j}= 0$ otherwise.

$$GCN_t^k \leftarrow \left(W_{gcn} \odot A^k\right) X_t \tag{1}$$

$$f_t = \sigma\left(\left(W_f GCN_t^k\right) + \left(U_f h_{t-1}\right) + b_f\right) \tag{2}$$

$$i_t = \sigma\left(\left(W_i GCN_t^k\right) + \left(U_i h_{t-1}\right) + b_i\right) \tag{3}$$

$$o_t = \sigma\left(\left(W_o GCN_t^k\right) + (U_o h_{t-1}) + b_o\right) \quad (4)$$

$$c_t' = \tanh\left(\left(W_c GCN_t^k\right) + (U_c h_{t-1}) + b_{c'}\right) \quad (5)$$

$$c_t = (f_t \odot c_{t-1}) + \left(i_t \odot c_t'\right) \quad (6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (7)$$

The GCN function is used to obtain the nodal features as described in (1). GCN operates based on the Hadamard product multiplication ($\odot$) of the weight matrix (W_{gcn}), adjacency matrix (A), and node features from the observed traffic data (X_t). The adjacency matrix captures information related to the OT network topology. The adjacency matrix (A) is added with the identity matrix (I) to form a modified adjacency matrix ($\widehat{A}$). The data set (X_t) is represented as a time series, where the equation considers the single time instant (t) and total number of time observations, T. The node feature matrix (X) contains individual nodal information (x_i), where the total number of nodes is represented by (n). The equation also considers the number of hops from a communication node to neighboring nodes, i.e., k as an exponent of $\widehat{A}$, as explained in [38], [71]. This research uses the maximum number of hops between each substation and the control center being two, i.e., $k = 2$.

After obtaining the spatial features from the graph convolutional operation, LSTM is then used to analyze the temporal / time-series features. The LSTM functions and processes inside an LSTM cell are described in (2 - 7). 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 previously calculated nodal features output (GCN_t^k) serves as the input for the LSTM cell.

In this work, we consider each substation to have unique characteristics. Given the communication network traffic data from all nodes that are present in a substation as (X), Algorithm 1 describes how an independent process is performed for each substation to provide the independent set GC-LSTM models for every substation (s_i). During the training process, this output is compared with the real OT traffic data (X_{t+1}) to update the weight values in GCN and LSTM. The final output of LSTM predicts the OT traffic in corresponding nodes represented by (h_t). This output serves as input for the TSC in the following stage.

Algorithm 1 CyResGrid Attack Graph Generation

Inputs: $S\{s_1, s_2, \ldots s_n\}$; $X \in$ sn: Substations traffic data
$\{x_1, x_2, \ldots, x_n\} \in X$: Nodes traffic data
Outputs: $\Lambda = \{\{a_i, \bar{a}_i \in V\}\}$: Nodes classification as attack graph

Iteration for each substation
for s_i in S **do**
  for t = 1 to T do
    Traffic prediction
    $GCN_t^k \leftarrow \left(W_{gcn} \odot \widehat{A}^k\right) X\{x_1, x_2, \ldots, x_n\}_t$
    $h_t, c_t = LSTM(X\{x_1, x_2, \ldots, x_n\}_t, GCN_t^k, h_{t-1}, C_{t-1})$
    Iteration for each node a in V
    for a in V
      Node classification
      $\bar{a}_i = \sum^{m-1} w h_{t,(i)}^{l-1} + b$
    end for
  end for
end for
return: $\Lambda = \{\{a_i, \bar{a}_i \in V\}\}$

D. Time Series Anomaly Detection

TSC for anomaly detection was studied in [32], [33], [34], [35]. In this research, we propose a new method using TSC to detect anomalies in the OT communication network traffic throughput for power systems. As a benchmark, we focus on state-of-the-art deep learning-based anomaly detection techniques, i.e., ResNets [72], Inception [35], Fully Convolutional Neural Networks (FCN) [73], and Multi-Layer Perceptron (MLP) [74]. Meanwhile, in our research, we propose CyResGrid; a hybrid of method for unsupervised and supervised OT traffic anomaly detection. The unsupervised learning application for time series data was studied in [75]. We specifically use an unsupervised GC-LSTM model to learn the complex behavior of OT network data and topology. Subsequently, the GC-LSTM generates traffic predictions as inputs to TSCs.

$$y_i^l = \mathrm{ReLU}\left(\sum^{m-1} w y_{(i)}^{l-1} + b\right) \quad (8)$$

$$x^* = \underset{x}{\operatorname{argmax}} f(x) \quad (9)$$

We propose a supervised deep convolutional neural network for TSC-based anomaly detection. The deep convolutional network is based on a multi-layer one-dimensional convolutional with the ReLU activation function as shown in (8). In (8), we consider the number of layers (l), filter size (m), weight (w), and bias (b). This model is trained to optimize the performance of classification based on the previous GC-LSTM output. To formulate our hybrid deep learning model, we perform hyperparameter tuning based on the number of layers, filters, and kernel size. Bayesian optimization [76] is used to optimize the deep learning model. The objective function maximizes the deep learning performance as described in (9). Bayesian optimization works based on the surrogate model and acquisition function. The surrogate model is a Gaussian process that quantifies the uncertainty of the unobservable region. To achieve the optimum value of the objective function, we use the Expected Improvement (EI) as the acquisition function. Bayesian optimization performs iterations to obtain a function with the best performance. From the iterative process, we obtain the best performing deep convolutional network that has 3 layers, 64 filters, and 3 kernel sizes. Fig. 7 shows the architecture of CyResGrid hybrid deep learning model that consists of a GC-LSTM layer, three layers of convolutional neural network, and one layer of fully connected neural network (dense).

E. Attack Graph Model

An attack graph is a method to model CPS vulnerabilities and potential exploits. Since a successful exploit of a vulnerability may lead to a partial or even a total failure of the CPS,

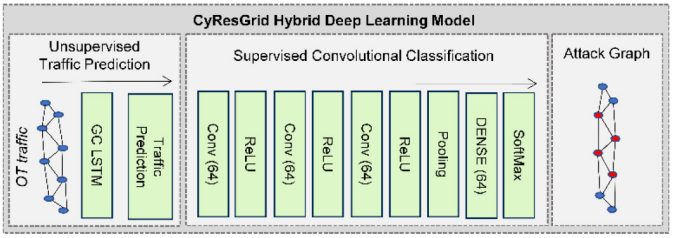


Fig. 7. CyResGrid – hybrid deep learning model.

an attack graph is an important tool for vulnerability analysis and mitigation strategies. Meanwhile, in a communication network, there are many hosts that may become vulnerable. As a result, the cyber security of the entire CPS cannot only rely on the security of a single host. Therefore, it is important to locate and identify all vulnerable nodes/hosts in a communication network as a set of potential threats in the CPS. Subsequently, in this research, we propose the observation and analysis of anomalous OT traffic behavior to detect nodes potentially compromised by cyber attacks. The information regarding anomalous nodes is then used to construct an online attack graph in near real-time for the entire OT network of the power grid.

Algorithm 1 explains the process of attack graph generation. The OT network traffic (X) is the input for the algorithm. The network traffic from each substation (X_n) is used to predict the OT traffic using GC-LSTM. The GC-LSTM model provides a set of traffic predictions (h_t) as outputs. The output from the prediction is then used as input for the TSC-based CNN. The time series-based anomaly detection is performed for each node (a) in V. The classifier labels each node as anomalous or normal based on the input OT traffic prediction. This information is then used to construct the attack graph.

$$\Lambda = \{\{a_i, \overline{a}_i, \in V\}\} \tag{10}$$

$$\Lambda = \{\{a_i, \overline{a}_i, \in V\}, \{u_i \notin V\}\} \tag{11}$$

There are two types of attack graphs as described through equations (10) and (11). The attack graph type I in (10) is constructed based on prior knowledge of the OT network topology and node classification results. Meanwhile, the attack graph type II in (11) considers unidentified nodes based on the TDG. There are two elements of attack graph (Λ) type I as indicated in (10), i.e., normal nodes (a_i), and anomalous nodes ($\overline{a}_i$). Both of the nodes are elements of the known nodes (V). In contrast attack graph (Λ) type II as indicated in (11) contains one extra element of unidentified nodes (u_i). The unidentified nodes are considered as anomalous since these nodes are not elements of the known nodes (V).

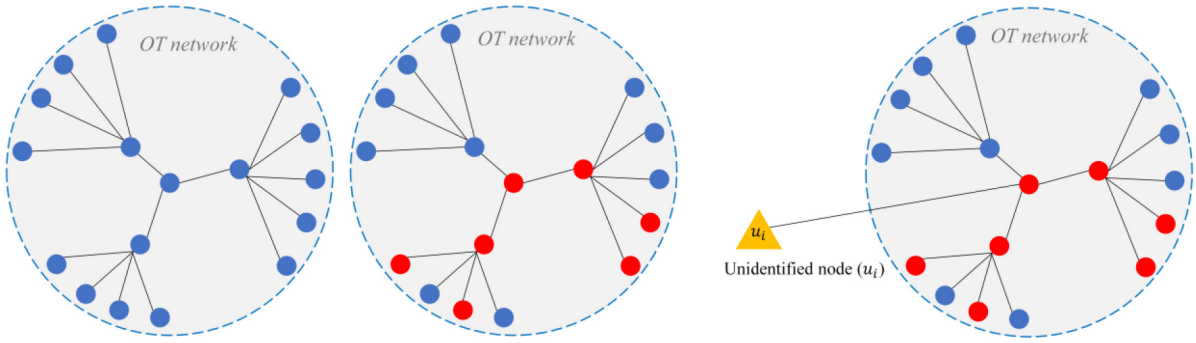
Fig. 8 depicts an example comparison of attack graph representations of the OT network under normal network traffic conditions in Fig. 8(a) and anomalous traffic in Figs. 8(b) and 8(c). The anomalous network traffic conditions are determined based on observed abnormal node behavior shown in red. Subsequently, these nodes are combined to form an attack graph (Λ). There are three elements in the attack graph, i.e., normal nodes (a_i), anomalous nodes ($\overline{a}_i$), and unidentified nodes (u_i). The attack graph type I from Fig. 8(b) only classifies nodes as anomalous based on observed traffics from all known nodes. This notion is represented by a set of attack graphs (Λ) and described through (10). On the other hand, the attack graph type II in Fig. 8(c) also considers all unidentified nodes for the classification of anomalous behavior, as described in (11). The unidentified nodes (u_i) are determined based on unknown sources or destinations address obtained from the TDG. The unknown nodes (u_i) are assumed to indicate an active cyber attack, originating from an unlisted host in the known OT network (V).

III. Experimental Results

A. Experimental Setting

All experiments in this paper are conducted using the previously discussed CPS model of the power grid represented in Fig. 4. The power system is simulated in real-time using a Root Mean Square (RMS) dynamic model of the IEEE 39-bus test system in DIgSILENT PowerFactory. The CPS model employs OPC UA implemented through Python to interface the time domain simulation of the power grid and emulated OT communication network. The OT network emulation is based on Mininet, which uses the operating-system-level virtualization. The entire emulated OT network runs on 10 virtual servers and consists of 27 user-defined substations, 118 measurement devices, and over 800 data points for the entire simulated power system. SCADA device functionality within the OT network is realized through custom Python code. Therefore, we generate SCADA traffic from substations and the control center. All OT network traffic is captured using the Linux *bwm-ng* tool and used as the main dataset for this research. The OT network traffic is measured in KBps. The observed OT network traffic data under nominal operating conditions is used to train the GC-LSTM model.

We collect OT network traffic data during various cyber attack scenarios. Two types of cyber attacks are considered, i.e., DDoS and active reconnaissance, i.e., OT network scanning. The DDoS attack is launched to target multiple substations and aims to disrupt the power system operation with a malicious increase of the OT network traffic loading. To this end, we use the well-known Syn Flood cyber attack vector that exploits vulnerabilities in the TCP/IP packets to target network hosts [77]. This attack vector is chosen as it can flood the OT network and cause the targeted hosts to crash. The DDoS attack is executed using the Linux *hping3* tool. The second examined cyber attack scenario is based on OT network scanning. This attack aims to enumerate active hosts within the OT network. Network scanning targets IP addresses and ports within a specified range. It is typically performed during reconnaissance at the early stages of a cyber attack kill chain. In this work, we conduct a six-level network scanning using *nmap*, i.e., paranoid, sneaky, polite, normal, aggressive, and insane. The first two scanning levels are stealthy and used to evade intrusion detection systems [78]. The scanning intensity determines the number of packets delivered to the network. For all cyber attack scenarios and simulations, we collect the observed OT network traffic data into a labeled dataset for deep learning applications.



(a) Normal graph, $\{a_i, \in V\}$ (b) Attack graph type I, $\Lambda = \{a_i, \bar{a}_i, \in V\}$ (c) Attack graph type II, $\Lambda = \{\{a_i, \bar{a}_i, \in V\}, \{u_i \notin V\}\}$

Fig. 8. Attack graph representation for normal and anomalous traffic: a) Normal graph, b) Attack graph type I which contains normal and anomalous nodes, and c) Attack graph type II which contains normal, anomalous and unidentified nodes.

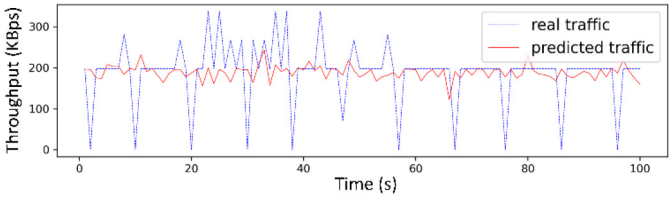


Fig. 9. Comparison of real and predicted traffic under normal conditions.

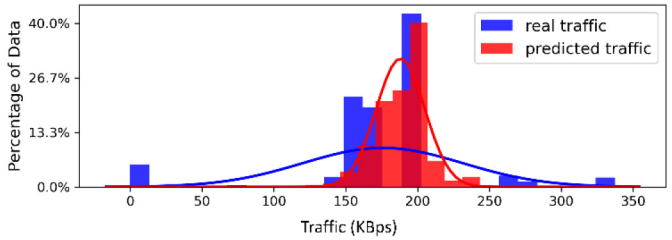


Fig. 10. Histogram of real and predicted traffic under normal conditions.

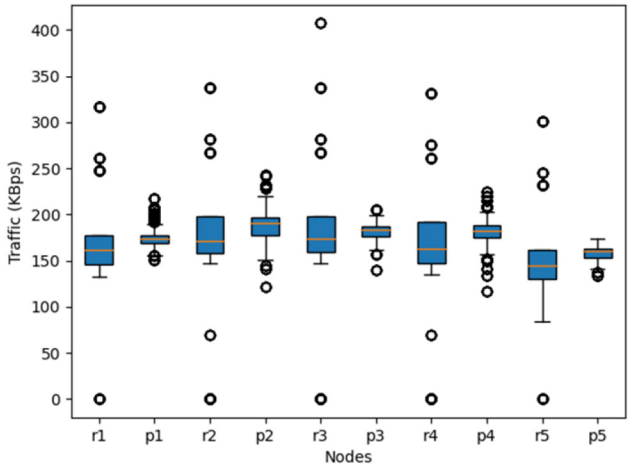


Fig. 11. Statistical comparison of real (r) and predicted traffic (p).

B. Network Traffic Prediction

In this research, the training of the GC-LSTM model is performed using the simulated OT network traffic dataset. This dataset consists of operational data for 27 substations, resulting in a total of 146 columns and 25 x 10^4 rows. The number of columns represents the total number of traffic observation points in the OT network. On the other hand, the number of rows in the dataset represents the temporal observations. The sampling rate for all observations is 1 sample/second. Therefore, the dataset for normal OT traffic is collected for a total duration of 25 x 10^4 seconds. The training was performed using a computer with the following specifications: Intel Xeon CPU 3.60GHz, 64 GB of RAM, and an NVIDIA Quadro RTX 4000 graphics processing unit. During the training process, the OT observation points are further classified for each individual substation to create 27 independent models of traffic predictions. The total training time for all 27 substations is 26.5 hours.

Fig. 9 shows the comparison of the real OT traffic under normal conditions and GC-LSTM predicted traffic in node 2, substation 7. The observed traffic rate is around 197 KBps. However, occasionally, the real OT traffic slightly increases or drops to zero but we cannot consider this situation as an anomaly. In distributed communication systems, the zero-value and variability happen because of the latency and delay that lead to variations in the packet arrival time. These factors are common phenomena for distributed communications, which have been studied in [79]. The zero value in Fig. 9 represents zero in Fig. 13. On average, the observed OT traffic data contains 3.6% of zeroes.

Fig. 10 presents the histogram and probability distribution of the real and predicted OT traffic in node 2, substation 7. Fig. 10 shows that the predicted OT traffic is more concentrated. We also compare the normal and predicted OT traffic for nodes 1 to 5 in substation 7 as represented in Fig. 11. The box plot in Fig. 11 shows the statistical summary from the traffic data including the minimum, median, maximum, first quartile, and third quartile. The box plot also indicates the variability, spread, and skewness of the data. The circles in the plot indicate the outlier data. Based on the plots in Fig. 9–11, the predicted OT traffic has a more concentrated value and fewer outliers compared to the real data. Therefore, the GC-LSTM performs as a filter to normalize and reduce the variability and outliers traffic.

Fig. 12 shows the comparison of the real and predicted OT traffic during a sneaky cyber attack. The cyber attack triggers

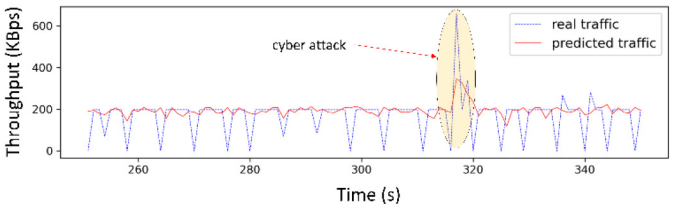


Fig. 12. Comparison of throughput between real and predicted OT traffic for sneaky network scanning cyber attack scenario.

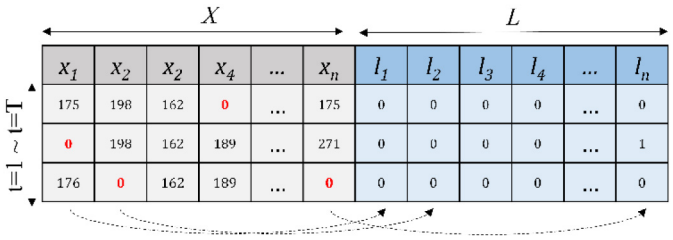


Fig. 13. Dataset for time series classification.

TABLE I
CYBER ATTACK SCENARIOS

Attack Type	Intensity	Tool	Time Duration (s)
DDoS	High	*hping3*	30,000
	Medium	*hping3*	30,000
	Low	*hping3*	30,000
Reconnaissance	Paranoid	*nmap*	75,000
	Sneaky	*nmap*	50,000
	Polite	*nmap*	40,000
	Normal	*nmap*	30,000
	Aggressive	*nmap*	30,000
	Insane	*nmap*	30,000

a higher spike in OT traffic. The time series-based anomaly detection is then expected to distinguish the spikes due to traffic variability and cyber attacks. Therefore, the GC-LSTM-based prediction is important to normalize the OT traffic and reduce data variability on the predicted traffic. This is then used to improve the anomaly detection accuracy of TSC.

C. Anomaly Detection

To perform anomaly detection on the OT traffic, we generate a dataset with network traffic (X) and labels (L) for univariate TSC. This is depicted in Fig. 13. Each column (x_n) in the observed traffic data has one associated label column (l_n). A label value of zero corresponds to the normal operation, while one represents anomalous OT traffic. We simulate two types of cyber attacks to generate anomalous traffic, i.e., DDoS and OT network scanning during the reconnaissance stage of the cyber kill chain. The attack scenarios are summarized in Table I. There are nine variations in the intensity of the communication network scanning amongst the scenarios. In total, the cyber attacks run for 345,000 seconds, and data is collected every second to create the dataset, as represented in Fig. 13, from t = 1 until t = 345,000. This dataset is then used to train 70% and test 30% the TSC algorithm.

Using the same generated dataset, we compare our proposed CyResGrid method with four state-of-the-art deep learning-based TSC techniques for anomaly detection, i.e., ResNets [72], Inception [35], FCN [73], and MLP [74]. These deep learning models are chosen as they address the general time series classification problem and are not domain specific. This makes them suitable for benchmarking and comparison of various TSC methods. Additionally, we also combine them with the proposed GC-LSTM method and test their performances, as summarized in Table II.

$$G_{mean} = \sqrt{true\ positive\ rate\ *\ truenegativerate} \quad (12)$$

$$F1 = \frac{2 * precision * recall}{precision + recall} \quad (13)$$

In Table II, we classify the cyber attacks into two scenarios. The first is for all combined attacks, i.e., no. 1-9, and the second only focuses on stealthy attack scenarios, i.e., paranoid and sneaky attacks no. 10-16. We consider the test dataset as imbalanced because, for the combined attacks, only 6.4% of the data is labeled as an anomaly. Meanwhile, for the stealthy attacks, only 2.7% of the data is labeled as an anomaly. Therefore, to evaluate the anomaly detection performance, we use as metrics the Geometric mean (G mean) in Equation (12) [80] and F1 score in Equation (13) [81], [82]. From Table II, it is clearly seen that for the combined attack scenario, CyResGrid provides the best performance with the highest scores in the Area Under The Curve (AUC), accuracy, G mean, and F1. Meanwhile, for the stealthy attack dataset, we ignore the MLP method due to its lower performance. For this scenario, Inception seems to provide the best AUC and accuracy. However, its true positive rate is significantly low. Furthermore, its F1 and G mean score are amongst one the lowest. Therefore, we can still conclude that CyResGrid provides the most balanced performance, even for stealthy attack detection.

Table II also indicates that GC-LSTM hybrid models can significantly improve the performance of deep learning-based classification, as indicated in row number 5, 6, 7, 8, 13, 14, and 15. The performance comparisons are also shown in Figs. 14 and 15. The Receiver Operating Characteristic (ROC) curve shows the performance of the classifier. The hybrid classification integrated with GC-LSTM provides improved result, as seen in Fig. 15, in comparison to the one without GC-LSTM in Fig. 14. According to Figs. 7–9, the actual OT traffic data is noisier compared to the predicted one. This condition leads to better anomaly detection using the hybrid model, as described above.

D. Attack Graph Generation and Analysis

As discussed in Section II-D, the attack graph is modeled by comparing the normal and anomalous OT traffic. The result of this comparison is then used to determine the nodal abnormality. The attack graph classifies nodes into two categories, i.e., normal and anomalous. Anomalous nodes ($\overline{a}_i$) are indicated by red, while normal nodes (a_i) are highlighted in blue.

Fig. 16 illustrates the entire attack graph map for online cyber attack identification and visualization. Fig. 16 (a) depicts

TABLE II
PERFORMANCE COMPARISON OF ANOMALY DETECTION METHODS

No	Methods	AUC	TN	FP	FN	TP	Accuracy	F1	G mean	Time (s)
					Combined attack scenarios					
1	ResNet	0.849	82.27	11.32	3.49	2.92	85.19	28.29	15.50	633
2	Inception	0.961	93.50	0.20	4.10	2.31	95.71	51.76	14.68	976
3	FCN	0.955	88.16	5.43	3.92	2.49	90.65	34.76	14.81	1016
4	MLP	0.758	72.22	21.37	4.86	1.55	73.77	10.55	10.57	113
5	GC-LSTM + Resnet	0.974	93.29	0.31	3.27	3.14	96.42	63.77	17.12	1056
6	GC-LSTM + Inception	0.976	92.10	1.49	3.35	3.06	95.16	55.87	16.79	1409
7	GC-LSTM + FCN	0.972	92.28	1.30	3.68	2.73	95.01	52.26	15.87	1342
8	GC-LSTM + MLP	0.937	93.40	0.19	6.13	0.28	93.68	8.14	5.12	765
9	**CyResGrid**	**0.984**	**93.47**	**0.13**	**3.42**	**2.99**	**96.45**	**65.03**	**17.16**	**714**
					Stealthy attack scenarios					
10	ResNet	0.8637	86.94	12.02	0.96	0.08	87.02	1.26	2.69	91
11	Inception	0.9887	98.93	0.02	1.04	0.0004	98.93	0.09	0.22	224
12	FCN	0.9833	87.82	11.13	1.01	0.02	87.85	0.47	1.58	240
13	GC-LSTM + Resnet	0.9524	89.93	9.02	0.95	0.09	90.02	1.87	2.92	226
14	GC-LSTM + Inception	0.9489	89.96	8.99	0.95	0.10	90.05	1.87	2.92	303
15	GC-LSTM + FCN	0.9491	89.96	8.99	0.95	0.10	90.05	1.87	2.92	304
16	**CyResGrid**	**0.9243**	**91.15**	**7.81**	**0.94**	**0.111**	**91.25**	**2.32**	**3.08**	**138**

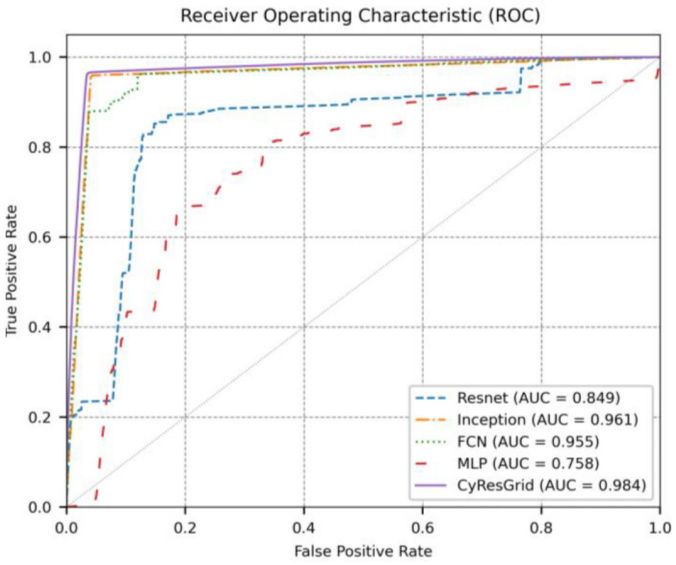


Fig. 14. ROC comparison of the deep learning-based TSC.

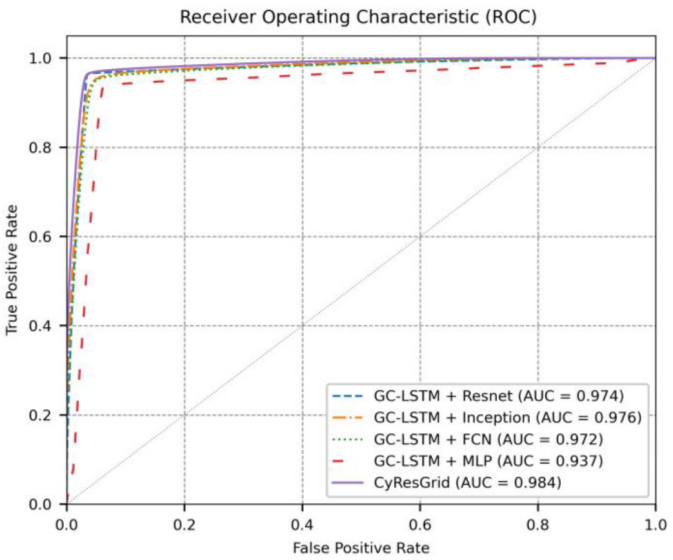


Fig. 15. ROC comparison of the hybrid deep learning-based TSC.

OT network scanning, originating from the control center to an OT device in substation 7. Consequently, this leads to the control center, substation 7 gateway, and targeted OT device to be flagged as anomalous, as shown in red. Fig. 16 (b) depicts a DDoS attack targeting substations 1-7 that originates from the control center. The DDoS attack on multiple substation targets triggers widespread traffic anomalies in substations 1-7, as indicated in red. It is considerably easier to detect a DDoS attack, as it results in notably increased OT network traffic volume, in comparison to a network scanning attack. Figs. 16 (c) and (d) depict attack graphs for cyber attacks originating from other sources than the control center. In Fig. 16 (c), we highlight OT network scanning performed by a compromised OT device located in substation 7. The scanning attacks lead to all nodes in substation 7 being classified as anomalous, except the router gateway. This scenario is explained as a local cyber attack that occurs in a substation. Finally, Fig. 16 (d) shows OT network scanning by an unidentified node, as indicated by an orange triangle. The attack source is classified as unidentified because it is not included on the list of known nodes in the OT network.

IV. CONCLUSION AND FUTURE WORK

With the ever-increasing threat of cyber attacks on power grids, it is now crucial to improve attack detection capabilities in OT systems. In this work, we proposed CyResGrid, a hybrid model of GC-LSTM and a deep convolutional network for anomaly detection in OT communication networks for power grids. It helps power system operators to localize and identify cyber attacks in near real-time. GC-LSTM creates OT traffic predictions based on the spatial and temporal features of the input data. Through its predictions, the data variability and outliers are reduced. GC-LSTM also serves as a mechanism to improve the anomaly detection performance of TSCs.

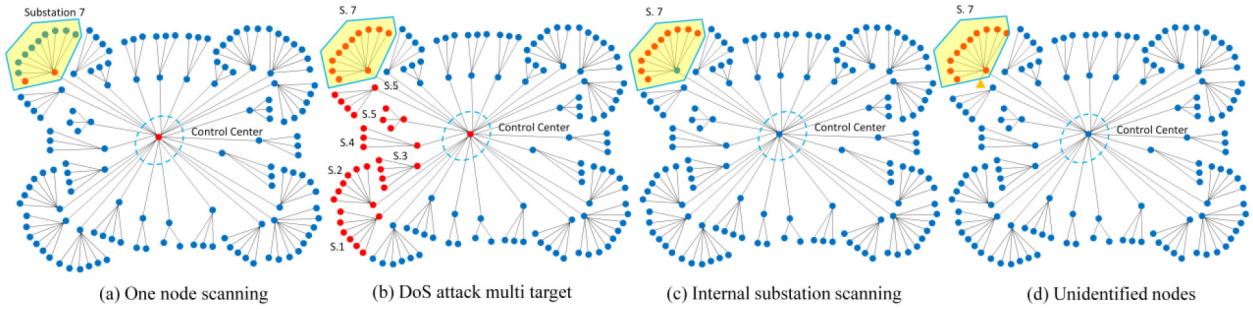


Fig. 16. Attack graph maps to identify and visualize cyber attack locations.

Furthermore, the deep convolutional network in CyResGrid is designed based on the hyperparameter tunning using Bayesian optimization. Hence, CyResGrid outperforms the state-of-the-art deep learning-based TSC. It provides the best detection performance, with the highest accuracy of 96.45%, F1 score of 65.03%, and G mean of 17.16%, and the lowest false positive rate of 0.13%. Additionally, for stealthy cyber attack scenarios, i.e., paranoid and sneaky attacks, CyResGrid provides the best performance indicated by the highest F1 score of 2.32% and G mean score of 3.08%. Other methods seem to provide higher accuracy and AUC. However, they have a lower performance to detect anomalies as indicated by the lower True Positive (TP), F1, and G mean scores. This classification is then used to generate an attack graph that serves as an online tool for power system operators to identify and localize active cyber attacks in OT networks of power systems.

In a future work, we will focus on augmenting the proposed CyResGrid method with prevention capabilities, in addition to the existing detection features. Subsequently, it can be integrated with an intrusion detection and prevention system. The developed method is equally applicable to different OT networks and CPS topologies, besides other cyber attack vectors, such as malware-based and privilege escalation attacks. Moreover, the performance of the detection algorithm can further be improved to detect more variations of cyber attacks with infinitesimally small changes to OT network traffic intensity and frequency of occurrences.

ACKNOWLEDGMENT

DeSIRE is funded by the 4TU-program High Tech for a Sustainable Future (HTSF). 4TU is the federation of the four technical universities in the Netherlands.

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. Prot. Relay Eng.*, College Station, TX, USA, Apr. 2017, pp. 1–8.

[2] M. J. Assante, R. M. Lee, and T. Conway, "ICS defense use case no. 6: Modular ICS malware," Electricity Inf. Sharing Center (E-ISAC), Washington, DC, USA, Rep. 6, Aug. 2017, vol. 2.

[3] E. M. Hutchins, M. J. Cloppert, and R. M. Amin, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," Lockheed Martin Corp., Bethesda, MD, USA, Rep. 1, 2011. Accessed: Jul. 5, 2022. [Online]. Available: https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf

[4] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.

[5] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—Attacks, impacts, and defense: A survey," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.

[6] A. S. Musleh, G. Chen, and Z. Y. Dong, "A survey on the detection algorithms for false data injection attacks in smart grids," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2218–2234, May 2020.

[7] H. T. Reda, A. Anwar, and A. Mahmood, "Comprehensive survey and taxonomies of false data injection attacks in smart grids: Attack models, targets, and impacts," *Renew. Sustain. Energy Rev.*, vol. 163, pp. 1–24, Jul. 2022.

[8] A. Sayghe et al., "Survey of machine learning methods for detecting false data injection attacks in power systems," *IET Smart Grid*, vol. 3, no. 5, pp. 581–595, Oct. 2020.

[9] H. Zhang, B. Liu, and H. Wu, "Smart grid cyber-physical attack and defense: A review," *IEEE Access*, vol. 9, pp. 29641–29659, 2021.

[10] U. Inayat, M. F. Zia, S. Mahmood, H. M. Khalid, and M. Benbouzid, "Learning-based methods for cyber attacks detection in IoT systems: A survey on methods, analysis, and future prospects," *Electronics*, vol. 11, no. 9, pp. 1–20, Jan. 2022.

[11] A. S. Musleh, H. M. Khalid, S. M. Muyeen, and A. Al-Durra, "A prediction algorithm to enhance grid resilience toward cyber attacks in WAMCS applications," *IEEE Syst. J.*, vol. 13, no. 1, pp. 710–719, Mar. 2019.

[12] SANS ICS, "Analysis of the cyber attack on the Ukrainian power grid," Electricity Inf. Sharing Center (E-ISAC), Washington, DC, USA, Rep. 2, Mar. 2016, vol. 388.

[13] 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.

[14] 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.

[15] 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.

[16] 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 Syst.*, vol. 23, no. 3, pp. 831–844, Aug. 2008.

[17] Y. Yang et al., "Intrusion detection system for network security in synchrophasor systems," in *Proc. IET Int. Conf. Inf. Commun. Technol. (IETICT)*, Beijing, China, 2013, pp. 246–252.

[18] S. Ali and Y. Li, "Learning multilevel auto-encoders for DDoS attack detection in smart grid network," *IEEE Access*, vol. 7, pp. 108647–108659, 2019.

[19] 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.

[20] M. Panthi, "Anomaly detection in smart grids using machine learning techniques," in *Proc. 1st Int. Conf. Power Control Comput. Technol. (ICPC2T)*, Raipur, India, Jan. 2020, pp. 220–222.

[21] 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.

[22] H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," *Appl. Sci.*, vol. 9, no. 20, pp. 1–28, Oct. 2019.

[23] 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.

[24] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 686–728, 1st Quart., 2019.

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

[26] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," Jan. 2019, *arXiv:1901.03407*.

[27] T. V. Phan, T. G. Nguyen, N.-N. Dao, T. T. Huong, N. H. Thanh, and T. Bauschert, "DeepGuard: Efficient anomaly detection in SDN with fine-grained traffic flow monitoring," *IEEE Trans. Netw. Service Manag.*, vol. 17, no. 3, pp. 1349–1362, Sep. 2020.

[28] R.-H. Hwang, M.-C. Peng, C.-W. Huang, P.-C. Lin, and V.-L. Nguyen, "An unsupervised deep learning model for early network traffic anomaly detection," *IEEE Access*, vol. 8, pp. 30387–30399, 2020.

[29] X. Guan, T. Qin, W. Li, and P. Wang, "Dynamic feature analysis and measurement for large-scale network traffic monitoring," *IEEE Trans. Inf. Forensics Security*, vol. 5, pp. 905–919, 2010.

[30] A. Kind, M. P. Stoecklin, and X. Dimitropoulos, "Histogram-based traffic anomaly detection," *IEEE Trans. Netw. Service Manag.*, vol. 6, no. 2, pp. 110–121, Jun. 2009.

[31] K. Xu, Z.-L. Zhang, and S. Bhattacharyya, "Internet traffic behavior profiling for network security monitoring," *IEEE/ACM Trans. Netw.*, vol. 16, no. 6, pp. 1241–1252, Dec. 2008.

[32] H.-S. Wu, "A survey of research on anomaly detection for time series," in *Proc. 13th Int. Comput. Conf. Wavelet Active Media Technol. Inf. Process. (ICCWAMTIP)*, Chengdu, China, Dec. 2016, pp. 426–431.

[33] K. Shaukat et al., "A review of time-series anomaly detection techniques: A step to future perspectives," in *Proc. Future Inf. Commun. Conf.*, Vancouver, BC, Canada, Apr. 2021, pp. 865–877.

[34] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: A review," *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, Jul. 2019.

[35] I. Fawaz et al., "InceptionTime: Finding AlexNet for time series classification," *Data Min. Knowl. Discov.*, vol. 34, no. 6, pp. 1936–1962, Sep. 2020.

[36] W. Lin, D. Wu, and B. Boulet, "Spatial-temporal residential short-term load forecasting via graph neural networks," *IEEE. Trans. Smart Grid*, vol. 12, no. 6, pp. 5373–5384, Nov. 2021.

[37] O. Boyaci, M. R. Narimani, K. R. Davis, M. Ismail, T. J. Overbye, and E. Serpedin, "Joint detection and localization of stealth false data injection attacks in smart grids using graph neural networks," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 807–819, Jan. 2022.

[38] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting," *IEEE Trans. Intell. Transp. Sys.*, vol. 21, no. 11, pp. 4883–4894, Nov. 2020.

[39] L. Deng, D. Lian, Z. Huang, and E. Chen, "Graph convolutional adversarial networks for spatiotemporal anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 6, pp. 2416–2428, Jun. 2022.

[40] H. Chen, H. Liu, X. Chu, Q. Liu, and D. Xue, "Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network," *Renew. Energy*, vol. 172, pp. 829–840, Jul. 2021.

[41] S. Basumallik, R. Ma, and S. Eftekharnejad, "Packet-data anomaly detection in PMU-based state estimator using convolutional neural network," *Int. J. Electr. Power Energy Syst.*, vol. 107, pp. 690–702, May 2019.

[42] X. Ou, W. F. Boyer, and M. A. McQueen, "A scalable approach to attack graph generation," in *Proc. ACM Conf. Comput. Commun. Security*, Oct. 2006, pp. 336–345.

[43] K. Kaynar and F. Sivrikaya, "Distributed attack graph generation," *IEEE Trans. Dependable Secure Comput.*, vol. 13, no. 5, pp. 519–532, Sep./Oct. 2016.

[44] S. Yoon, J.-H. Cho, D. S. Kim, T. J. Moore, F. Free-Nelson, and H. Lim, "Attack graph-based moving target defense in software-defined networks," *IEEE Trans. Netw. Service Manag.*, vol. 17, no. 3, pp. 1653–1668, Sep. 2020.

[45] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14–76, Jan. 2015.

[46] J. Wu, S. Luo, S. Wang, and H. Wang, "NLES: A novel lifetime extension scheme for safety-critical cyber-physical systems using SDN and NFV," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2463–2475, Apr. 2019.

[47] Y. Li, Y. Qin, P. Zhang, and A. Herzberg, "SDN-enabled cyber-physical security in networked microgrids," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1613–1622, Jul. 2019.

[48] X. Zhang, K. Wei, L. Guo, W. Hou, and J. Wu, "SDN-based resilience solutions for smart grids," in *Proc. Int. Conf. Softw. Netw. (ICSN)*, Jeju, South Korea, May 2016, pp. 1–5.

[49] A. Montazerolghaem and M. H. Yaghmaee, "Demand response application as a service: An SDN-based management framework," *IEEE Trans. Smart Grid*, vol. 13, no. 3, pp. 1952–1966, May 2022.

[50] M. H. Rehmani, F. Akhtar, A. Davy, and B. Jennings, "Achieving resilience in SDN-based smart grid: A multi-armed bandit approach," in *Proc. IEEE Conf. Netw. Softw. Workshop (NetSoft)*, Montreal, QC, Canada, Jun. 2018, pp. 366–371.

[51] P. Zhang et al., "Network-wide forwarding anomaly detection and localization in software defined networks," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 332–345, Feb. 2021.

[52] V. Krishnan et al., "Validation of synthetic U.S. electric power distribution system data sets," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4477–4489, Sep. 2020.

[53] X. Zheng et al., "A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids," *Nat. Sci. Data*, vol. 9, p. 359, Jun. 2022.

[54] S. Soltan, A. Loh, and G. Zussman, "A learning-based method for generating synthetic power grids," *IEEE Syst. J.*, vol. 13, no. 1, pp. 625–634, Mar. 2019.

[55] A. Venzke, D. K. Molzahn, and S. Chatzivasileiadis, "Efficient creation of datasets for data-driven power system applications," *Electr. Power Syst. Res.*, vol. 190, pp. 1–8, Jan. 2021.

[56] M. F. Elaha, M. Jin, and P. Zeng, "Review of load data analytics using deep learning in smart grids: Open load datasets, methodologies, and application challenges," *Int. J. Energy Res.*, vol. 45, no. 10, pp. 14274–14302, Aug. 2021.

[57] S. Tavakkoli, J. Macknick, G. A. Heat, and S. M. Jordaan, "Spatiotemporal energy infrastructure datasets for the United States: A review," *Renew. Sustain. Energy Rev.*, vol. 152, pp. 1–10, Dec. 2021.

[58] Y. Himeur, A. Alsalemi, F. Bensaali, and A. Amira, "Building power consumption datasets: Survey, taxonomy and future directions," *Energy Build.*, vol. 227, pp. 1–16, Nov. 2020.

[59] M. Naglic, *PMU Measurements of IEEE 39-Bus Power System Model*, IEEE DataPort, Piscataway, NJ, USA, Jun. 2019.

[60] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov. 2015.

[61] U. Adhikari, T. Morris, and S. Pan, "WAMS cyber-physical test bed for power system, cybersecurity study, and data mining," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2744–2753, Nov. 2017.

[62] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 847–855, Jun. 2013.

[63] M. H. Cintuglu, O. A. Mohammed, K. Akkaya, and A. S. Uluagac, "A survey on smart grid cyber-physical system testbeds," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 446–464, 1st Quart., 2017.

[64] B. B. Gupta and T. Akhtar, "A survey on smart power grid: Frameworks, tools, security issues, and solutions," *Ann. Telecommun.*, vol. 72, no. 9, pp. 517–549, Sep. 2017.

[65] C.-C. Sun, A. Hahn, and C.-C. Liu, "Cyber security of a power grid: State-of-the-art," *Int. J. Electr. Power Energy Syst.*, vol. 99, pp. 45–56, Jul. 2018.

[66] X. Zhou, X. Gou, T. Huang, and S. Yang, "Review on testing of cyber physical systems: Methods and testbeds," *IEEE Access*, vol. 6, pp. 52179–52194, 2018.

[67] M. Z. Gunduz and R. Das, "A comparison of cyber-security oriented testbeds for IoT-based smart grids," in *Proc. 6th Int. Symp. Digit. Forensic Security (ISDFS)*, Antalya, Turkey, Mar. 2018, pp. 1–6.

[68] J. Montoya et al., "Advanced laboratory testing methods using real-time simulation and hardware-in-the-loop techniques: A survey of smart grid international research facility network activities," *Energies*, vol. 13, no. 12, pp. 1–38, Jun. 2020.

[69] M. Iliofotou, P. Pappu, M. Faloutsos, M. Mitzenmacher, S. Singh, and G. Varghese, "Network monitoring using traffic dispersion graphs (TDGS)," in *Proc. 7th ACM SIGCOMM Conf. Internet Meas.*, San Diego, CA, USA, Oct. 2007, pp. 315–320.
[70] 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.*, Hanoi, Vietnam, Oct. 2011, pp. 36–41.
[71] J. Chen, X. Wang, and X. Xu, "GC-LSTM: Graph convolution embedded LSTM for dynamic network link prediction," *Appl. Intell.*, vol. 52, pp. 7513–7528, Sep. 2021.
[72] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[73] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 3431–3440.
[74] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
[75] M. Längkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognit. Lett.*, vol. 42, pp. 1–14, Jun. 2014.
[76] 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, Dec. 2012, pp. 1–9.
[77] R. Mohammadi, R. Javidan, and M. Conti, "SLICOTS: An SDN-based lightweight countermeasure for TCP SYN flooding attacks," *IEEE Trans. Netw. Service Manag.*, vol. 14, no. 2, pp. 487–497, Jun. 2017.
[78] T. Zitta et al., "Penetration testing of intrusion detection and prevention system in low-performance embedded IoT device," in *Proc. 18th Int. Conf. Mechtronics (ME)*, Brno, Czech Republic, Dec. 2018, pp. 1–5.
[79] J. M. Johansson, "On the impact of network latency on distributed systems design," *Inf. Technol. Manag.*, vol. 1, no. 3, pp. 183–194, Jan. 2000.
[80] R. Barandela, J. S. Sánchez, V. Garcıa, and E. Rangel, "Strategies for learning in class imbalance problems," *Pattern Recognit.*, vol. 36, no. 3, pp. 849–851, Mar. 2003.
[81] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, pp. 1–54, Dec. 2019.
[82] B. Kim, Y. Ko, and J. Seo, "Novel regularization method for the class imbalance problem," *Expert Syst. Appl.*, vol. 188, pp. 1–8, Feb. 2022.

Alfan Presekal (Member, IEEE) received the B.Eng. degree in computer engineering from Universitas Indonesia in 2014, and the M.Sc. degree in secure software system from the Department of Computing, Imperial College London, U.K., in 2016. He is an Assistant Professor of Computer Engineering with the Department of Electrical Engineering, Universitas Indonesia. He holds various cyber security certifications from EC Council, CompTIA, and CISCO. He is currently a Doctoral Researcher in Cyber Resilient Power Grids within Intelligent Electrical Power Grids with the Department of Electrical Sustainable Technology, 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 an Assistant Professor of Intelligent Electrical Power Grids with TU Delft, The Netherlands. He is the Director of the Control Room with 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.

Vetrivel Subramaniam Rajkumar (Student Member, IEEE) received the bachelor's degree in electrical engineering from Anna University, India, in 2013, and the M.Sc. degree in electrical power engineering from the Delft University of Technology, The Netherlands, in 2019, where he is currently a Doctoral Researcher with the Intelligent Electrical Power Grids Group, Department of Electrical Sustainable Technology. His current research interests include cyber security for power grids and impact analysis of cyber attacks on power systems.

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 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 active in international committees, such as ISO or CEN. His research interests include energy automation networks, smart grids, and modeling intelligent energy systems. He also serves as an IEEE IES AdCom 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.