DST-GNN: A Dynamic Spatiotemporal Graph Neural Network for Cyberattack Detection in Grid-Tied Photovoltaic Systems

Sha Peng, Mengxiang Liu, *Member, IEEE*, Li Chai, *Member, IEEE*, and Ruilong Deng, *Senior Member, IEEE*

***Abstract*—The increasing deployment of solar photovoltaic (PV) systems in the electric grid, aimed at addressing the energy crisis and surging power demands, has expanded the potential vulnerability to cyberattacks due to the inter-networking of the grid-connected power electronics converters. In this paper, we propose a dynamic spatiotemporal graph neural network (*DST-GNN*) for cyberattack detection in grid-tied PV systems. Specifically, to exploit the inherent graph topology of the grid-tied PV system, we start by employing a GNN with a dynamic weighted adjacency matrix to capture the latent *spatial* correlations within signal data. Then, a one-dimensional convolution neural network (1D-CNN) is utilized to extract the underlying *temporal* patterns. Notably, we leverage the system dynamics to determine the *dynamic* graph weights and the number of graph convolution layers, while the hyper-parameters of 1D-CNN are designed based on the periodicity of input signals. Finally, the integration of the priori physical system knowledge further enhances the interpretability and improves the detection performance of *DST-GNN*. To the best of our knowledge, this is the first work that embeds the grid-tied PV system into a graph structure for cyberattack detection. The effectiveness of *DST-GNN* is evaluated through comprehensive case studies on a hardware-in-the-loop (HIL) grid-tied PV testbed, and numerical results demonstrate its superiority over baseline methods.**

***Index Terms*—Cyber security, attack detection, graph neural network, machine learning, grid-tied photovoltaic systems.**

Manuscript received 8 February 2024; revised 21 June 2024; accepted 9 August 2024. Date of publication 16 August 2024; date of current version 26 December 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62293503, Grant 62293502, Grant 62293500, and Grant 62073285; in part by the Natural Science Foundation of Zhejiang Province under Grant LR23F030001 and Grant LZ24F030006; in part by the Xiaomi Foundation; in part by the Fundamental Research Funds for the Central Universities (Zhejiang University NGICS Platform); and in part by the Key Lab of CS&AUS of Zhejiang Province. Paper no. TSG-00217-2024. *(Corresponding author: Ruilong Deng.)*

Sha Peng and Li Chai are with the State Key Laboratory of Industrial Control Technology and the College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: pengsha@zju.edu.cn; chaili@zju.edu.cn).

Mengxiang Liu is with the Department of Automatic Control and Systems Engineering, University of Sheffield, S1 3JD Sheffield, U.K. (e-mail: mengxiang.liu@sheffield.ac.uk).

Ruilong Deng is with the State Key Laboratory of Industrial Control Technology and the College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China, and also with the Huzhou Institute of Industrial Control Technology, Huzhou 313000, China (e-mail: dengruilong@zju.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2024.3445113.

Digital Object Identifier 10.1109/TSG.2024.3445113

I. Introduction

With the increasing environmental concern and power demand, the electric grid is under its rapid transformation towards the sustainable and decarbonized future. This transformation is driven by the widely integrated renewable energy sources like solar, wind, hydro, etc. In 2023, the global annual renewable capacity additions surged nearly 50%, reaching about 510 GW, where solar constitutes 75% of the total additions [1]. The transformation has also boosted the adoption of distributed energy resources (DERs), marking a shift from large spinning generators to a grid primarily reliant on DERs. In 2021, nearly 50% of DERs are photovoltaic (PV) systems [2]. Benefiting from the advanced power electronics converters as well as information and communications technology (ICT), PV systems can produce the software-driven and digitally controlled outputs. This enhancement greatly improves the power supply's flexibility and efficiency, enabling participation in grid-support functions to maintain grid reliability [3].

As highlighted in a recent report from National Renewable Energy Laboratory [4], the widely adopted advanced ICT will enlarge the cyberattack risks in the meantime. An attacker could intrude into PV systems by compromising sensors, controllers, or communication networks [5], leading to hazardous voltage violations [6] or even blackouts [7], [8]. It is thus crucial to deploy appropriate cyber-resilient defensive methods to counter against these cyberattacks [9], [10]. Various PV-specific studies have emerged in the topic of cyber-resilience enhancement, which can be broadly categorized into attack prevention and intrusion detection.

- The *attack prevention* strategies focus on securing critical information infrastructure to thwart the adversary from hacking into and destroying the data availability, integrity, and confidentiality [11]. IEEE 1547 Std. [3] defines a series of trust and cryptography features for the communication protocols widely used in the interconnection of DERs with associated grid interfaces, including data encryption, access authentication, and key management. Nevertheless, powerful adversaries can still exploit zero-day vulnerabilities such as Stuxnet to bypass the employed preventive mechanisms [12].
- Hence, it is equally important to adopt the *intrusion detection* strategies in the during-attack phase, which can timely alarm system operators by perceiving anomalies

through the host-, network- or physics-related characteristics. This paper focuses on the physical-based detection strategies using physical signals like control commands and power readings to observe the anomalous physical statuses resulting from malicious data modification, which is usually regarded as the last intrusion detection line behind the host- and network-based intrusion detection systems [9].

Generally, the physical-based detection strategies fall into model-based and data-driven categories.

- The *model-based* methods develop detection strategies by validating the data integrity through the physical dynamics inherent in the electric circuit and controller. To detect cyberattacks in grid-tied PV systems, Zhang et al. [13] devised a harmonic state space transfer function (HSS-TF) based attack detector that can achieve high detection accuracy. Zhang and Ye [14] designed a residual-based attack detector utilising the principle of Kalman filter. Yan et al. [15] presented an unknown input observer to distributively detect and isolate the misbehaving agent in microgrid systems. By proactively perturbing the primary control gains, a converter-based moving target defense strategy was proposed in [16], [17] to defend against deception attacks in DC microgrids. Sahoo et al. [18] proposed a cooperative vulnerability factor detection framework for detecting cyberattacks in each DER system, where the factor is derived by the consensus-based secondary control algorithm. Nevertheless, these model-based methods rely heavily on the physical model's accuracy, which are extremely challenging to be obtained due to the power electronic device's essentially high complexity and non-linearity as verified in the comparative study [19].
- In contrast, the *data-driven* methods explore and utilise the characteristics behind data while not relying on an explicit physical model, thereby eliminating the potential adverse impact resulting from an inaccurate physical model. The works in [19], [20] detected cyberattacks in grid-tied PV systems by using a multi-layer long short-term memory (MLSTM) network to extract temporal patterns of electric waveform data. To detect and identify cyber and physical attacks, Guo et al. [21] extracted time-frequency domain features from a single waveform sensor and then employed an LSTM network/convolutional neural network (CNN). Li et al. [22] inputted the feature matrix of waveform data to a vector autoregressive model, which assumes a linear dependency among variables. To detect and mitigate cyberattacks in DC microgrids that are formed by parallel DC/DC converters, the work in [23] developed a method based on model predictive control and artificial neural networks. However, these data-driven methods do not fully consider the nonlinear spatiotemporal correlations beneath signal data and lack the incorporation of physical knowledge of PV systems, which, nevertheless, is greatly beneficial for the detection performance.

The grid-tied PV systems involve a variety of signal data, including (i) analog signals that offer continuous and precise information about physical systems, and (ii) digital signals that can be synchronized with other processes, enabling coordinated actions within the control system. Due to the underlying interactions and interconnections within the electric circuit and controller, different time series signals form multivariate time series data, exhibiting nonlinear dependency in a complex topological manner. In other words, each signal not only relies on its historical values but also has relations with other signals. For example, in grid-tied PV systems, Kirchhoff's law dictates that the grid-side current will be jointly influenced by its historical values as well as the inverter-side current, LCL capacitor voltage, and grid-side voltage. Consequently, the signals of PV systems show an underlying graph structure in non-Euclidean space. Since graphs can be irregular, with neighbors of a graph node being unordered and variable in size, certain crucial operators, such as the convolution, which are easy to implement in Euclidean space, face challenges when applying in non-Euclidean space with graph structure.

In recent years, the graph neural network (GNN) has emerged as a powerful tool in handling graph-structured data due to its permutation-invariance, local connectivity, and compositionality [24]. GNN propagates information through the graph structure, enabling each node to become aware of its neighborhood, and this iterative propagation effectively captures the global patterns and relationships within the graph. GNN has been widely used in the electric grid related problems, such as voltage regulation [25], false data injection attack detection [26], and frequency control of AC microgrids [27]. It follows that modeling multivariate time series signals in grid-tied PV systems using GNN can be a promising way to uncover the dynamic spatial correlations beneath signals.

While the conventional GNN is effective for handling static graph-structured data, it cannot characterize structured data with dynamic spatial and temporal correlations, which are quite common in grid-tied PV systems. Therefore, in this paper, we propose a dynamic spatiotemporal GNN for detecting cyberattacks in grid-tied PV systems (named *DST-GNN*). *DST-GNN* eliminates the complex process of constructing a mathematical model and effectively captures signals' spatial relationships and temporal patterns while integrating prior system knowledge. Firstly, to effectively capture the *spatial* relationships of signals, we incorporate a graph convolution module to address the dynamic inter-dependencies among time series while preserving temporal trajectory of signals. Secondly, a temporal convolution module, employing a one-dimensional convolution neural network (1D-CNN), is implemented to extract the *temporal* patterns. Furthermore, to enhance the interpretation and detection performance, *DST-GNN* is integrated with the priori physical knowledge of the grid-tied PV system. Especially, in the spatial dependency modeling, the system dynamics is utilized to determine the *dynamic* graph weights and the number of graph convolution layers; While in the temporal dependency modeling, the filter size and the number of 1D-convolution layers are designed based on the periodicity of input signals, which depends on

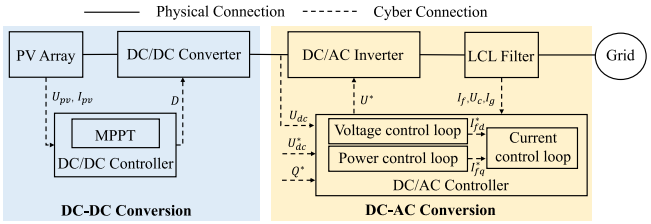


Fig. 1. The diagram of a typical two-stage grid-tied PV system.

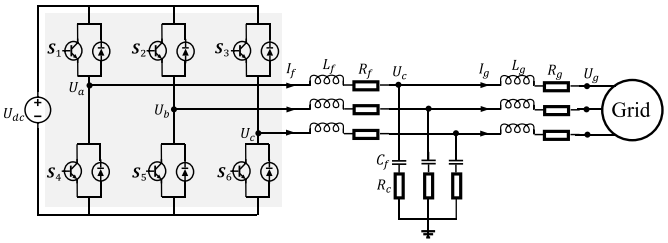


Fig. 2. The physical modeling of the DC-AC conversion.

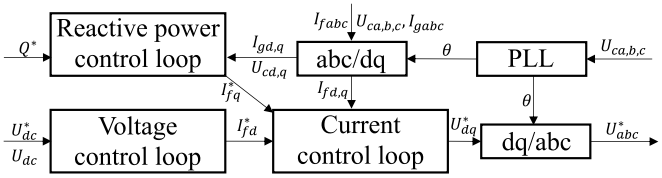


Fig. 3. The cyber modeling of the DC-AC conversion.

the grid frequency and signal sampling frequency. Overall, the main contributions of this paper are summarized as follows:

- To the best of the authors' knowledge, it is the first work that embeds the physical system and controller part of grid-tied PV systems into a graph structure.
- Based on the synthesised PV-specific graph, we propose a dynamic spatiotemporal GNN to extract inter-series spatial correlations and intra-series temporal relationships of multivariate time series signals for cyberattack detection.
- We further integrate the priori physical knowledge of grid-tied PV systems, including the system dynamics and periodicity of input signals, to improve the interpretation and detection performance.
- We conduct systematical experiments on a hardware-in-the-loop (HIL) grid-tied PV testbed to verify the effectiveness of *DST-GNN* and its superiority over existing methods.

The remainder of this paper is organized as follows. The detailed modeling of the grid-tied PV system is illustrated in Section II. In Section III, the cyberattack model and problem formulation are discussed. Section IV elaborates *DST-GNN* which is consisted of the dynamic graph construction, graph convolution and temporal convolution. Experiments are presented in Section V, and Section VI concludes this paper.

II. Grid-Tied PV System Modeling

A. Overview

Fig. 1 shows a typical two-stage grid-tied PV system.

- The *DC-DC conversion* stage includes a PV array and a DC/DC converter. The DC/DC controller takes the PV array voltage U_{pv} and current I_{pv} as inputs and uses the maximum power point tracking (MPPT) algorithm to optimize the power extraction. The DC/DC converter maintains the operating voltage at the maximum power point using the duty cycle D.
- The *DC-AC conversion* stage comprises a DC/AC inverter and an LCL filter. In the DC/AC controller, the voltage control loop maintains the DC link voltage U_{dc} and generates one reference current I_{fd}^*. The reactive power control loop derives the other reference current I_{fq}^* and generates the required reactive power. The current control loop ensures the inverter output current I_f tracks the current reference point. Since our work mainly focuses on the DC-AC conversion stage, we detail its physical and cyber parts in the following.

B. Physical Modeling

The physical topology of the DC-AC conversion is illustrated in Fig. 2. According to Kirchhoff's law, the DC/AC inverter and the LCL filter can be modeled as follows:

$$\begin{cases} I_f &= \frac{1}{L_f}\int [U - U_c - I_f R_f - (I_f - I_g)R_c] \\ U_c &= \frac{1}{C_f}\int (I_f - I_g) \\ I_g &= \frac{1}{L_g}\int [U_c + (I_f - I_g)R_c - I_g R_g - U_g] \end{cases}, \tag{1}$$

where I_f and I_g are the inverter-side and grid-side currents of LCL, respectively; U, U_c and U_g are the inverter-side, LCL capacitor and grid-side voltages, respectively; L_f and L_g are the inverter-side and grid-side inductance in LCL, respectively; R_f and R_g are the corresponding equivalent series resistance; C_f is the capacitor in LCL; and R_c is the shunt resistance.

C. Cyber Modeling

The controller inside the DC-AC conversion is treated as a cyber system and designed to convert the power from the DC circuit to the AC grid. Fig. 3 illustrates the basic blocks of the DC/AC controller.

1) The abc-to-dq/dq-to-abc Transformation: The abc-to-dq transformation converts three-phase quantities (*abc*) to direct-quadrature (*dq*) axis components, facilitating the designing of a simple control strategy like a proportional-integral (PI) controller. The transformation can be written as

$$\begin{bmatrix} d \\ q \end{bmatrix} = \frac{2}{3}\begin{bmatrix} \sin(\theta) & \sin\left(\theta - \frac{2\pi}{3}\right) & \sin\left(\theta + \frac{2\pi}{3}\right) \\ \cos(\theta) & \cos\left(\theta - \frac{2\pi}{3}\right) & \cos\left(\theta + \frac{2\pi}{3}\right) \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}, \tag{2}$$

where a, b and c are the components of the three-phase system; d and q are the components in the rotating reference frame; and θ is the synchronization phase angle obtained from the phase locked loop (PLL) block. Note that the dq-to-abc transformation is the inverse process of the abc-to-dq transformation, which can be trivially derived from Eq. (2).

2) Phase Locked Loop: PLL is utilized for grid synchronization by consistently monitoring the grid frequency and inverter output frequency, which can be written as

$$\begin{bmatrix} U_{cd} \\ U_{cq} \end{bmatrix} = |U_c| \begin{bmatrix} cos(\theta_g - \theta) \\ sin(\theta_g - \theta) \end{bmatrix}, \tag{3}$$

where $|U_c|$ and θ_g are the magnitude and fundamental phase of U_c, respectively; and θ is the synchronization phase angle. The PI controller of PLL tries to equal θ_g with θ to achieve synchronization.

3) Control Loops: The voltage control loop and the reactive power control loop maintain the DC voltage stability and generate the required reactive power through deriving the current reference I_{fd}^* and I_{fq}^*, respectively, which can be expressed as follows:

$$\begin{cases} I_{fd}^* = k_{pv}\left(U_{dc}^* - U_{dc}\right) + \int k_{iv}\left(U_{dc}^* - U_{dc}\right) \\ I_{fq}^* = k_{pq}(Q^* - Q) + \int k_{iq}(Q^* - Q) \end{cases}, \tag{4}$$

where k_{pv}, k_{iv}, k_{pq} and k_{iq} are the PI parameters; U_{dc}^* and Q^* are the references of the DC link voltage and reactive power, respectively; and Q represents the reactive power, which can be calculated as

$$Q = U_{cq} * I_{gd} - U_{cd} * I_{gq}. \tag{5}$$

In the current control loop, two separate PI controllers force the inverter-side current $I_{fd,q}$ to track the reference current $I_{fd,q}^*$, which can be expressed as

$$\begin{cases} U_d^* = U_{cd} + k_{pc}(I_{fd}^* - I_{fd}) + \int k_{ic}(I_{fd}^* - I_{fd}) - \omega L_f I_{fq} \\ U_q^* = U_{cq} + k_{pc}(I_{fq}^* - I_{fq}) + \int k_{ic}(I_{fq}^* - I_{fq}) - \omega L_f I_{fd} \end{cases}, \tag{6}$$

where k_{pc} and k_{ic} are the PI parameters, ω is the angular frequency, and $U_{d,q}^*$ is the controller signal.

III. Cyberattack Model and Problem Formulation

A. Cyberattack Model

The hierarchical control structure of a typical grid-tied PV system and its associated vulnerabilities are depicted in Fig. 4. Cyber threats can disrupt system operations by targeting sensors, controllers, or communication networks [13]. For instance, an attacker could manipulate control commands between the power plant control and device control or power grid control by compromising communication networks, employing methods such as man-in-the-middle (MITM) attacks and denial-of-service (DoS) attacks [28]. Cyberattacks can be executed through communication networks or physical access by exploiting firmware updates. Noninvasive attacks could pose a threat to grid operations by compromising PV converter sensors [29]. Numerous studies have been conducted to address cyberattacks targeting the networks and software of PV systems. For example, a blockchain-based method is proposed in [30] to detect MITM attacks in PV systems. A firmware over the air scheme based on blockchain technology is introduced in [31] to guarantee a secure, resilient, reliable, and auditable procedure of sending updates and/or patches to smart inverters. A distributed self-triggered mechanism for multiple PVs is developed in [32] to maintain control performance under DoS attacks while simultaneously reducing

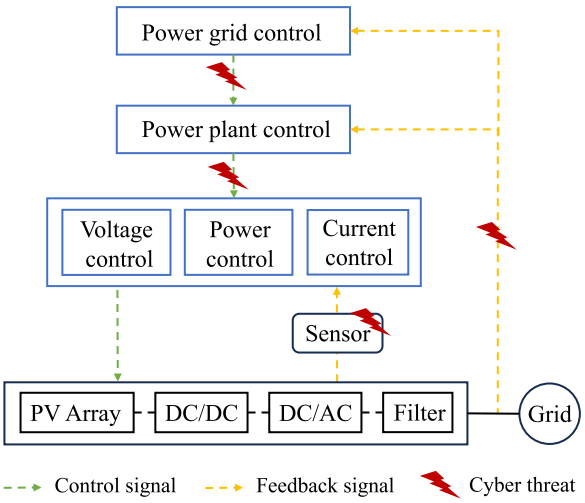


Fig. 4. Hierarchical control and vulnerabilities of a typical PV system.

data flow in the communication network. However, there is limited research focused on defending against cyberattacks that compromise sensor data in grid-tied PV systems [13], [21]. Once an attacker compromises a sensor, the device-level controller is immediately affected, leading to erroneous control commands and potentially disrupting the operation of PV converters.

Therefore, the main focus of this paper is the cyber threat on sensor measurements in the grid-tied PV system. Following the references [13], [21], the cyberattack model is defined as:

$$\mathbf{M_A} = \boldsymbol{\alpha} \cdot \mathbf{M} + \boldsymbol{\beta}. \tag{7}$$

Here, $\mathbf{M_A}$ denotes the attacked measurements which are eventually utilized by the controller. $\mathbf{M}$ represents the original measurements. We assume that the sensors operate reliably, therefore $\mathbf{M}$ does not include errors. $\boldsymbol{\alpha}$ is a multiplicative factor matrix and $\boldsymbol{\beta}$ is a vector, that both define unknown attack weights, with compatible dimensions. Specifically, for the DC/AC controller, $\mathbf{M}$ is a 10-dimensional vector including measurements of the DC link voltage, three-phase inverter-side current of LCL, three-phase LCL capacitor voltage, and three-phase grid-side current of LCL, which can be represented as

$$\mathbf{M} = \left[U_{dc}, I_{fa,b,c}, U_{ca,b,c}, I_{ga,b,c}\right]^T. \tag{8}$$

The measurements can be obtained using waveform measurement units (WMUs) [33], such as Hall sensors [29]. When these measurements are compromised as described in Eq. (7), they provide incorrect state information to the controller as per Eq. (4)-(6). This leads to the generation of erroneous control signals and ultimately degrades system performance. In our experiments, we employ this cyberattack model to simulate false data in the sensors, creating a dataset for network testing.

B. Problem Formulation

PV systems exhibit diverse operational states, including normal conditions and various underattack states with distinct attack types. It is challenging to collect a sufficiently large and varied set of cyberattack samples and to define clear

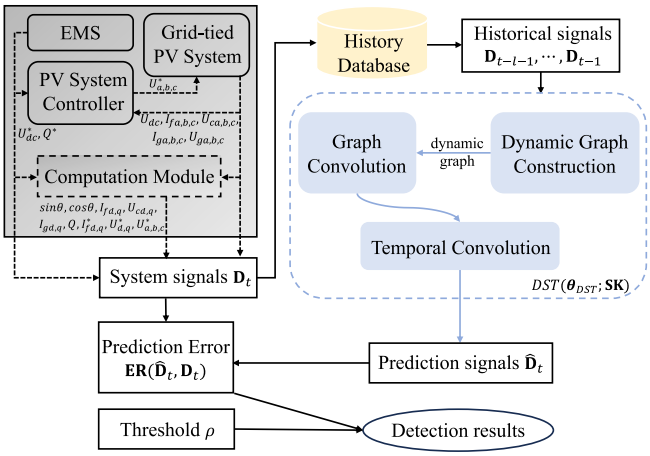


Fig. 5. The general structure for detecting cyberattacks.

attack characteristics. Since two-class detection requires a training dataset that includes both normal and attacked data, we formulate the cyberattack detection in PV systems as a multivariate time series one-class detection problem. Training a one-class detection model requires only normal data, making it advantageous when dealing with a potentially large number of different attack scenarios [19].

Let $\mathbf{D}$ represent the set of the following signals $[U_{dc}, I_{fa,b,c}, U_{ca,b,c}, I_{ga,b,c,}, U_{ga,b,c}, U^*_{dc}, Q^*, sin\theta, cos\theta, I_{fd,q}, U_{cd,q}, I_{gd,q}, Q, I^*_{fd,q}, U^*_{d,q}, U^*_{a,b,c}]^T$ and $\mathbf{SK}$ denote the set of the system parameters $[L_f, C_f, L_g, R_f, R_g, R_c, k_{pv}, k_{iv}, k_{pq}, k_{iq}, k_{pc}, k_{ic}, \omega]^T$. Note that the signal values in $\mathbf{D}$ are instantaneous values, in which the physical measurements can be monitored directly by WMUs, and the references of the DC link voltage and reactive power are provided by the facility's energy management systems (EMS). The remaining signals computed by the DC/AC controller can be obtained through an independent computation module without posing excessive communication burdens on the actual controller in practice. Given the observed data signals of $l-1$ sampling times $\mathbf{D}_{t-l-1}, \ldots, \mathbf{D}_{t-2}, \mathbf{D}_{t-1}$ and the priori system knowledge $\mathbf{SK}$, the aim of the formulated problem is to detect cyberattacks through predicting the value of $\mathbf{D}_t$ and comparing the prediction error with a threshold.

The general structure for cyberattack detection is illustrated in Fig. 5, where the blue part represents the trained neural network used for prediction. Specifically, let *DST* denote the trained network and $\boldsymbol{\theta}_{DST}$ represent its parameters learned through offline training. We firstly predict the value of $\mathbf{D}_t$ by

$$\hat{\mathbf{D}}_t = DST(\mathbf{D}_{t-l-1}, \ldots, \mathbf{D}_{t-2}, \mathbf{D}_{t-1}; \mathbf{SK}; \boldsymbol{\theta}_{DST}). \tag{9}$$

Then, we calculate the prediction error $\mathbf{ER}\ (\hat{\mathbf{D}}_t, \mathbf{D}_t)$ and compare it with the threshold ρ, where $\mathbf{ER}(\cdot)$ denotes the error calculating function and ρ is determined empirically through a validation process. The detection results can be written as a sign function $\mathrm{sgn}(\cdot)$:

$$\mathrm{sgn}(\hat{\mathbf{D}}_t, \mathbf{D}_t; \rho) := \begin{cases} 1 & \text{if } \mathbf{ER}\left(\hat{\mathbf{D}}_t, \mathbf{D}_t\right) \geq \rho \\ 0 & \text{if } \mathbf{ER}\left(\hat{\mathbf{D}}_t, \mathbf{D}_t\right) < \rho \end{cases}. \tag{10}$$

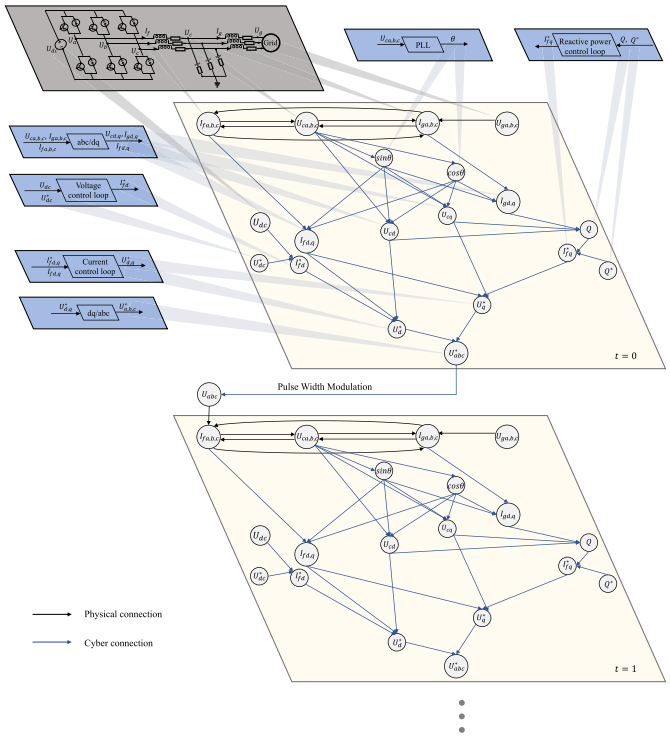


Fig. 6. The directed graph of the DC-AC conversion over time. For clearer illustration, the larger circle denotes combination of multiple nodes, while the smaller circle denotes a single node.

Here, if $\mathrm{sgn}(\hat{\mathbf{D}}_t, \mathbf{D}_t; \rho) = 1$, it indicates the existence of cyberattacks; otherwise, the system is under normal states.

IV. DST-GNN for Cyberattack Detection in Grid-Tied PV Systems

DST-GNN aims to extract both inter-series spatial relationships and intra-series temporal correlations of multivariate time series signals and then detect cyberattacks from the learned patterns. In this section, we firstly introduce the dynamic graph construction module, elaborating on how to embed the DC-AC conversion into a directed graph and derive the dynamic weighted adjacency matrix for GNN. Then, we design the graph convolution module and the temporal convolution module, which aim to capture the spatial and temporal features, respectively. Finally, we present the overall architecture of *DST-GNN*.

A. Dynamic Graph Construction

In order to better capture the dynamic correlations among multivariate time series signals and bypass the extraordinarily complicated modeling process, the DC-AC conversion is approximated utilizing a directed graph. The graph can be expressed as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ having a set of nodes $\mathcal{V}$, a set of edges $\mathcal{E}$, and a weighted adjacency matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$, where $N = |\mathcal{V}|$ is the number of nodes. Fig. 6 shows the representation of the DC-AC conversion as a graph structure. The goal of the dynamic graph construction module is to learn the latent spatial relationships among the data points of $\mathbf{D}$, so the 31 data points of $\mathbf{D}$ provide the nodes of the graph. For a

clear presentation of Fig. 6, we group the three-phase variables and depict them as a single node. We also group I_{fd}, I_{fq} and I_{gd}, I_{gq} and depict them as $I_{fd,q}$ and $I_{gd,q}$, respectively. However, in the proposed *DST-GNN*, we denote each of these terms as a single node. As depicted in Fig. 6, we do not incorporate the node $U_{a,b,c}$ directly into the graph structure but position it between the graph structures of two consecutive time steps. This decision is due to the absence of $U_{a,b,c}$ measurements in real-world systems. Additionally, including the node $U_{a,b,c}$ would significantly increase the computational complexity, as the relationship between $U_{a,b,c}$ and $U^*_{a,b,c}$ involves the pulse width modulation module, making it highly intricate. Generally, the set of edges $\mathcal{E}$ and weighted adjacency matrix $\mathbf{W}$ can be constructed by the priori physical knowledge or automatically inferred from the data. Here, we generate $\mathcal{E}$ and $\mathbf{W}$ according to the physical and cyber modeling of the DC-AC conversion described in Section II. Leveraging additional insights from the modeling could enhance the effectiveness of the spatial characteristics extraction within the data. It is worth to notice that graphs inferred from the data might be noise-sensitive because signals from practical situations may contain noises of random fluctuations and artifacts. This can lead to the misinterpretation of noise as meaningful patterns, resulting in inaccuracies. In contrast, graphs constructed from priori physical knowledge are based on robust principles, making them less susceptible to the inconsistencies and errors introduced by noisy data.

As expressed in Eq. (1)-(6), the 31 data points have underlying interactions and interconnections due to the electricity flow and control signal flow. Hence, as shown in Fig. 6, the physical and cyber connections among all these nodes are depicted as the edges of the graph. Furthermore, the physical and cyber modeling of the DC-AC conversion also gives an insight into the construction of $\mathbf{W}$. For the edge's weight from the source node to the destination node at time point t, we compute the first-order partial derivative of the destination node's value at t with respect to the source node's value at t according to Eq. (1)-(6). The first-order partial derivatives offer an approximate measure of the impact of the source node on the destination node while keeping other variables constant. This concept closely aligns with the edges' weight in the graph, representing the strength of the connection between adjacent nodes. Following this insight, we adopt the first-order partial derivatives as the edge weights and construct $\mathbf{W}$. Noted that $\mathbf{W}$ is dynamic since some nodes' first-order partial derivatives are related to variables and thus change with time.

1) Physical Connections: For the edge weights of the physical connections described by Eq. (1), we need to discretize the dynamics to calculate the first-order partial derivatives at time point t. This is because that Eq. (1) involves the integral which range from the start time to the current time, but we only require the relationship at current time t. Besides, the sampled data signals of the continuous variable signals are discrete. Therefore, we approximate the continuous integral by the summation of discrete signals and calculate the first-order partial derivatives at time t. This technique aligns with the established practices, i.e., right Riemann sum, in numerical approximation, allowing for a more tractable analysis. As a result, to derive the physical connections' weights, we firstly approximate Eq. (1) as follows:

$$\begin{cases} I_f(t) \approx \frac{1}{L_f+R_c+R_f}\{U(t) - U_c(t) + I_g(t)R_c + \sum_{i=1}^{t-1} \\ [U(i) - U_c(i) - I_f(i)R_f - \big(I_f(i) - I_g(i)\big)R_c]\} \\ U_c(t) \approx \frac{1}{C_f}[I_f(t) - I_g(t) + \sum_{i=1}^{t-1}\big(I_f(i) - I_g(i)\big)] \\ I_g(t) \approx \frac{1}{L_g+R_c+R_g}\{U_c(t) + I_f(t)R_c - U_g(t) + \sum_{i=1}^{t-1}[U_c(i) \\ \quad +\big(I_f(i) - I_g(i)\big)R_c - I_g(i)R_g - U_g(i)]\} \end{cases}. \quad (11)$$

Recall that we only consider the first-order partial derivatives of the destination node's value at time t with respect to the source node's value at time t, so we set the sum of all previous time points as a constant. Let $\mathbf{W}(a, b)$ denote the edge's weight from the source node a to the destination node b. Therefore, we have $\mathbf{W}(U_c, I_f) = -\frac{1}{L_f+R_c+R_f}$, $\mathbf{W}(I_g, I_f) = \frac{R_c}{L_f+R_c+R_f}$, $\mathbf{W}(I_f, U_c) = \frac{1}{C_f}$, $\mathbf{W}(I_g, U_c) = -\frac{1}{C_f}$, $\mathbf{W}(U_c, I_g) = \frac{1}{L_g+R_c+R_g}$, $\mathbf{W}(I_f, I_g) = \frac{R_c}{L_g+R_c+R_g}$ and $\mathbf{W}(U_g, I_g) = -\frac{1}{L_g+R_c+R_g}$. Note that these edge weights are applicable to all the three phase quantities.

2) Cyber Connections:

a) The abc-to-dq/dq-to-abc transformation: Based on Eq. (2), we can directly compute the edge weights from the source nodes I_f, U_c, and I_g in the dq reference frame to the destination nodes $sin\theta$, $cos\theta$ and the corresponding *abc* axis components using the first-order partial derivatives. Conversely, we can also derive the edge weights from the source nodes I_f, U_c, and I_g in the *abc* frame to the destination nodes $sin\theta$, $cos\theta$ and the corresponding *dq* axis quantities.

b) Phase locked loop: To compute the edge weights inside PLL, we assume that the PI controller of PLL has a higher frequency which can ensure θ_g equals to θ in omittable short time [29]. Under this assumption, we have $sin(\theta_g - \theta) = 0$, which derives $U_{cq} = 0$ in Eq. (3). Therefore, applying the abc-to-dq transformation, we can calculate the edge weights from the source node U_c in the *abc* reference frame to the destination nodes $sin\theta$, $cos\theta$ based on the following equation:

$$U_{cq} = \frac{2}{3}\left(cos\theta U_{ca} + cos\left(\theta - \frac{2\pi}{3}\right)U_{cb} + cos\left(\theta + \frac{2\pi}{3}\right)U_{cc}\right) = 0. \quad (12)$$

From Eq. (12), we have

$$sin^2\theta = \frac{\left(U_{ca} - \frac{1}{2}U_{cb} - \frac{1}{2}U_{cc}\right)^2}{\left(U_{ca} - \frac{1}{2}U_{cb} - \frac{1}{2}U_{cc}\right)^2 + \frac{3}{4}(U_{cc} - U_{cb})^2}. \quad (13)$$

According to the chain rule, we further have

$$\frac{\partial sin^2\theta}{\partial U_c} = 2sin\theta\frac{\partial sin\theta}{\partial U_c}. \quad (14)$$

Combining Eq. (13) and (14), we can calculate the first-order partial derivatives of $sin\theta$ with respect to U_{ca}, U_{cb} and U_{cc}. Similarly, the edge weights from the source nodes U_{ca}, U_{cb} and U_{cc} to the destination node $cos\theta$ can be computed accordingly.

c) Control loops: As expressed in Eq. (4) and Eq. (6), the voltage control loop, reactive power control loop and current control loop share the integral operation. Therefore, we approximate the integral in Eq. (4) and Eq. (6) with right Riemann sum as described in Section IV-A1). Then we can calculate the edge weights between the destination nodes $I^*_{fd,q}$, $U^*_{d,q}$ and their corresponding source nodes. In addition, based on Eq. (5), we have $\mathbf{W}(U_{cq}, Q) = I_{gd}$, $\mathbf{W}(I_{gd}, Q) = U_{cq}$, $\mathbf{W}(U_{cd}, Q) = -I_{gq}$ and $\mathbf{W}(I_{gq}, Q) = -U_{cd}$.

B. Graph Convolution

In the grid-tied PV system, the physical connections and control signal flows result in the highly correlated multivariate data signals, revealing a non-Euclidean underlying graph structure. Therefore, we incorporate a graph convolution module to capture these spatial correlations. This module analyzes data from the graph structure in non-Euclidean space, aiming to model the message propagation process and integrate a node's information with its neighbors to address spatial dependencies in the graph. From the spatial perspective, the state variables in the PV system interact, and this inter-dependence exhibits a notably dynamic nature. The weighted adjacency matrix $\mathbf{W}$ adjusts the weights of two connected nodes based on the temporal inputs. It can reflect the system dynamics implicitly, providing effective information for capturing dynamic spatial relationships among the multivariate data signals.

The graph convolution module consists of two phases - the information propagation (IP) phase and the information selection (IS) phase [34]. Let $\mathbf{W}_t$ represent the weighted adjacency matrix at the sampling time t, which is constructed according to the priori system knowledge $\mathbf{SK}$ and data signal $\mathbf{D}_t$. Given an adjacency matrix $\mathbf{W}_t$, we have the mathematical form of the k-hop IP phase at time t:

$$\mathbf{H}_t^{(k)} = \gamma \mathbf{H}_t^{in} + (1-\gamma)\mathbf{W}_t \mathbf{H}_t^{(k-1)}, \tag{15}$$

where γ is a hyper-parameter that governs the ratio of preserving the original states of the root node, $\mathbf{H}_t^{(k)} \in \mathbb{R}^{N \times D_{in}}$ is the k-th layer node representation matrix at time t, and $\mathbf{H}_t^{in}$ is the input node states at time t. Set $\mathbf{H}_t^{(0)} = \mathbf{H}_t^{in}$. In this paper, we employ per unit normalization [35] for all quantities in $\mathbf{D}_t$ except $sin\theta$ and $cos\theta$, serving the normalized $\mathbf{D}_t$ as $\mathbf{H}_t^{in}$. This normalization facilitates the analysis of the grid-tied PV system with different voltage levels. Furthermore, it ensures that all variables in $\mathbf{D}$ have a similar scale and prevents specific data points from dominating others in magnitude, which is a crucial consideration for the network convergence. It's important to highlight that the resistance, capacitor and inductance value in $\mathbf{SK}$ also require pre-processing through per-unit normalization for constructing the weighted adjacency matrix applicable to the normalized data signals. Base on Eq. (15), we can calculate the node information for each input time point along with the associated weighted adjacency matrix separately and obtain the k-th layer node representation matrix $\mathbf{H}^{(k)} \in \mathbb{R}^{l \times N \times D_{in}}$ for the multivariate time series of length l. Then, the IS phase is introduced to filter out crucial information, which is defined as follows:

$$\mathbf{H}^{out} = \sum_{k=0}^{K} \mathbf{H}^{(k)} \boldsymbol{\theta}_g^{(k)}, \tag{16}$$

where K is the depth of propagation, $\mathbf{H}^{out} \in \mathbb{R}^{l \times N \times D_{out}}$ is the output node states and $\boldsymbol{\theta}_g \in \mathbb{R}^{(K+1) \times D_{in} \times D_{out}}$ is learnable parameter matrix, serving as a feature selector. Notably, the outputs of the IP phase of different time points share the same parameter matrix in the IS phase, which offers several advantages, including the reduced network size, improved generalization and enhanced training efficiency. In this paper, we select the graph convolution module with only one layer to learn the hidden correlations of the data signals. This is because that in dynamic systems, changes in system variables are predominantly associated with directly relevant state variables outlined in the dynamic equations. The impact on system variables through indirect means is less significant and can be dismissed. Besides, unlike CNN which usually exhibits enhanced performance with deeper architectures, GNN typically necessitates fewer than 3 layers. Given these considerations, we establish the graph convolution module with a single layer.

C. Temporal Convolution

Data signals of the grid-tied PV system exhibit complicated nonlinear temporal characteristics due to the nature of the control strategies, operating conditions and functionalities of the converter. Hence, we utilize a 1D-CNN in the temporal convolution module to learn these latent temporal patterns.

CNN stands out as one of the most widely used deep neural networks. It employs shared local filters and hierarchical information processing to extract hidden data features, taking a comprehensive global perspective. 1D-CNN is a type of CNN designed for processing sequences of data arranged along a single dimension. It moves 1D convolution kernels along the input data in one direction. These kernels process the input sequence in a localized manner and allow the network to derive the hidden representations from different slices of the input sequence. This makes 1D-CNN suitable for analyzing sequential data, such as time series, audio signals and text. Furthermore, 1D-CNN is computationally efficient and can process large sequences faster than other methods, like recurrent neural networks.

In this paper, the temporal convolution module incorporates multiple 1D convolution layers and a flattening layer. The mathematical form of the convolution layer can be formulated as follows:

$$\boldsymbol{c}^{j,q} = \sigma\left(\sum_k \boldsymbol{c}^{k,q-1} * \boldsymbol{\theta}_c^{k,j,q} + \boldsymbol{b}^{j,q}\right), \tag{17}$$

where $\boldsymbol{c}^{j,q}$ and $\boldsymbol{b}^{j,q}$ denote the j-th feature map and bias at the q-th convolution layer, respectively; the learnable weight parameter of k-th feature map of the j-th filter kernel at the q-th convolution layer is denoted as $\boldsymbol{\theta}_c^{k,j,q}$; $*$ denotes the convolution operation; and $\sigma(\cdot)$ denotes an nonlinear activation function. In this paper, the leaky rectified linear

unit (LeakyReLU) serves as the nonlinear activation function for each convolution layer. Before each LeakyReLU, a 1D batch-normalization is implemented to enhance the detection performance and expedite the training. Following a series of convolution layers, a flattening layer is employed to transform the extracted feature maps into a singular vector, facilitating the subsequent modeling tasks.

Generally, the filter size and the number of convolution layers are hyper-parameters. The filter size determines each filter's receptive field, and selecting the appropriate one poses a challenging task for convolution networks. It may be either too small - hindering the effective extraction of long-term temporal features, or too large - making capturing the subtle short-term temporal features difficult. Inspired by [36], we determine the filter size according to the inherent physical characteristics of the input time series. Due to the grid synchronization, the temporal signals of the grid-tied PV system show periodicity, which is determined by the grid frequency. Combined with the sampling frequency, we can further compute the periodicity in the sampled signals. We denote m_i as the filter size at the i-th convolution layer, and then

$$\sum_{i=1}^{K} m_i - K + 1 = T, \tag{18}$$

where K is the number of the convolution layers and T is the periodicity in the sampled signals. This enables us to extract the temporal information over the entire period. For example, we can use a 1×9 filter in the first convolution layer and a 1×7 filter in the second to represent the period 15.

D. Overall Architecture

Fig. 7 depicts the overall architecture of *DST-GNN*, with the orange and blue circles representing the operation dimensions of graph convolution and temporal convolution, respectively. The procedure is summarized as follows:

1) The multivariate time series signals are fed into the dynamic graph construction module to generate the weighted adjacency matrices $\mathbf{W}_t$ at all time points.
2) In the graph convolution module, for each time point, the IP phase takes the data signals and the corresponding weighted adjacency matrix as inputs to update each specific node's representation by incorporating its neighbors' information. After processing all time points, the output of the IP phase at each time point serves as an input to the IS phase to filter out the important information and derive the latent spatial correlations among data signals separately for each time point. The spatial dependency modeling part integrates with the priori physical knowledge of the grid-tied PV system to construct the dynamic weighted adjacency matrix and determine the number of graph convolution layers, effectively extracting the dynamic spatial features.
3) The temporal convolution module utilizes a 1D-CNN to further capture the hidden temporal patterns of each univariate time series signal. The filter size and the number of 1D-convolution layers are designed according to the grid frequency and the sampling frequency of input signals. Integrating the input time series' physical characteristics makes 1D-CNN more effective for learning the temporal features.
4) The learned spatiotemporal features are fed into a final CNN to output the predicted value.
5) The prediction error is calculated and compared with the empirically determined threshold to detect cyberattacks.

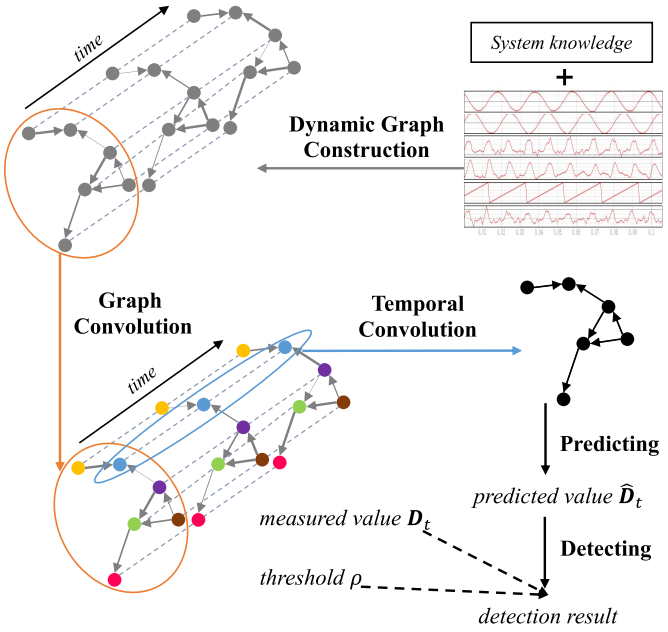


Fig. 7. The overall architecture of *DST-GNN*. Note that the graph convolution operates separately for each time point, while the temporal convolution operates separately for each node.

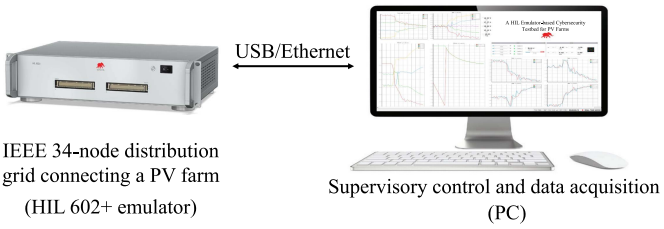


Fig. 8. Overview of the HIL-based grid-tied PV testbed [38].

V. Experimental Validations

In this section, we perform the hardware-in-the-loop (HIL) experiments using a real-time Typhoon HIL testbed [37] to analyze the effectiveness of *DST-GNN*. As illustrated in Fig. 8, we construct the IEEE 34-node distribution grid, which links to a power electronic converter-enabled PV farm, within the HIL 602+ emulator. Besides, a connection between the emulator and a personal computer (PC) is also established to facilitate the supervisory control and data acquisition (SCADA). In the HIL experiments, the system frequency is set to 50 Hz.

A. Datasets

During the training process, we define various types of normal cases by specifying different irradiance and temperature conditions. As shown in Table I, in addition to scenarios with constant irradiance and temperature (N1∼N5), we also

TABLE I
DEFINITION OF NORMAL CASES

Type	Irradiance (w/m^2)	Temperature (°C)	No.
Constant irradiance and temperature	800	25	N1
	900	15	N2
	900	25	N3
	900	35	N4
	1000	25	N5
Changing irradiance and temperature	900	25 $\sim$ 30	N6
	900	25 $\sim$ 20	N7
	800 $\sim$ 1000	25	N8
	1000 $\sim$ 800	25	N9

TABLE II
DEFINITION OF ATTACK CASES

Type	Target	$[\alpha, \beta]$	No.
Single Attack	I_f (3 phase)	[1.5, 0]	A1
	I_f (1 phase)	[0.8, 0]	A2
	U_{dc}	[0.1, 0]	A3
Coordinated Attack	I_f (1 phase), U_{dc}	[0.8, 0]	A4
	I_f, U_c (3 phase)	[0.5, 0]	A5
	I_f, U_c (1 phase)	[1.2, 0]	A6
Replay Attack	I_f (3 phase)	—	A7

Note: Refer to [13, 19] for setting α, β. Here 1 phase refers to phase a. For A7, fake I_f is captured under 800w/m^2, 15°C.

include scenarios with gradual changes in the irradiance and temperature (N6∼N9). For each scenario, data is captured over a period of 10 seconds, with a selected sampling frequency of 1.3 kHz. Thus, 117,000 signal time points are simulated for this normal scenario dataset.

Following the cyberattack model in Eq. (7), different cyberattacks, including the single attack, coordinated attack and replay attack are formulated, as illustrated in Table II.

- To select the threshold and test the effectiveness of *DST-GNN*, we separately conduct these cyberattacks under 900w/m^2 irradiation and 25°C. In each attack scenario, data is collected over 10 seconds with a sampling frequency of 1.3 kHz. The normal condition spans 5 seconds, followed by the attack duration of another 5 seconds. We refer to this dataset as multi-type attack scenario dataset.
- Furthermore, we test the effectiveness of *DST-GNN* under the stealthy attack proposed in [38]. The stealthy attack scenario lasts 80 seconds, in which the first and the last 20 seconds are normal conditions, while the attack lasts 40 seconds. Detailed information regarding the setting of attack parameters can be found in [38].

Additionally, we define a special normal case which lasts 4 seconds, with a sudden irradiance change to test the robustness of *DST-GNN*. In the special normal scenario, the irradiance increases to 1000w/m^2 from 800w/m^2 within one second and then returns to 800w/m^2 in the following second. The temperature is maintained at 25°C, and the signals are sampled at a frequency of 1.3 kHz, consistent with other scenarios.

B. Experimental Setups

1) Baseline Methods: We evaluate the performance of *DST-GNN* in comparison with three recently proposed methods for cyberattack detection in grid-tied PV systems: MLSTM [19], [21], CNN [21] and HSS-TF [13].

- The MLSTM-based and CNN-based detectors use a dataset with only 6 data points, i.e., 3-phase voltage and current in the PCC node. Therefore, we train the MLSTM and CNN models using a dataset with 6 data points, denoted as MLSTM-1 and CNN-1, respectively. However, *DST-GNN* uses 31 data points as expressed in Section III-B. To facilitate a fair comparison and explore the impact of utilizing more data points, we also train the MLSTM and CNN models using a dataset with 31 data points, denoted as MLSTM-2 and CNN-2, respectively. Unlike these data-driven detection methods, the HSS-TF-based detector is model-based and does not involve a training process; it only requires fewer time series data of $I^*_{fd,q}$, $U_{gd,q}$, $I_{gd,q}$, U_{dc} and U^*_{dc}.
- In addition, two popular machine learning-based anomaly detection methods - multi-layer perceptron (MLP) and k-nearest neighbor (KNN) - are also implemented as baselines. They are both trained using a dataset with 31 data points. As they are not inherently designed for time series data, we reshape the two-dimensional input into a one-dimensional array for input purposes.

2) Evaluation Metrics: The standard metrics (i.e., recall, precision, F1-score and accuracy) are used in our experiments to evaluate the cyberattack detection performance. The formulas for these metrics are as follows:

- Recall $= \frac{\text{TP}}{\text{TP+FN}}$
- Precision $= \frac{\text{TP}}{\text{TP+FP}}$
- F1-score $= \frac{2\times\text{TP}}{2\times\text{TN+FP+FN}}$
- Accuracy $= \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}$

where TP, TN, FP and FN are the number of correctly predicted actual attacks, correctly predicted actual normals, actual attacks incorrectly predicted as normals and actual normals incorrectly predicted as attacks, respectively. To detect cyberattacks, we use the mean absolute error (MAE) [39] as the error calculating function $\mathbf{ER}(\cdot)$. Any sample with MAE over the empirically determined threshold will be regarded as an attack at the test time.

3) Training Settings: We implement the aforementioned machine learning networks in Python 3.7.11 using Pytorch 1.8.2 and Sklearn 1.0.2 libraries on a PC equipped with an Intel i5-12600KF CPU, an nVidia GeForce RTX 3060Ti GPU, and 16GB RAM. For *DST-GNN*, we set the input length of historical data to 49, which meaning *DST-GNN* predicts the value of the 50th data point and determines whether cyberattacks exist. In the graph convolution module, we employ the IP phase with one layer as mentioned in Section IV-B, set γ to 0.05 and set embedding dimension D_{out} to 4. Since the system frequency is 50 Hz and the sampling frequency is 1.3 kHz, we can calculate that the periodicity in the sampled signals is 26. In the temporal convolution module, based on Eq. (18) and the periodicity of the sampled signals, we set the number of convolution layers to 3 and use the 1×10, 1×9 and 1×9 filters to the first, second and third layers, respectively. Regarding the final CNN, the number of layers is set to 3, and the kernel size of each layer is set to 7. The LeakyReLU servers as the nonlinear activation function

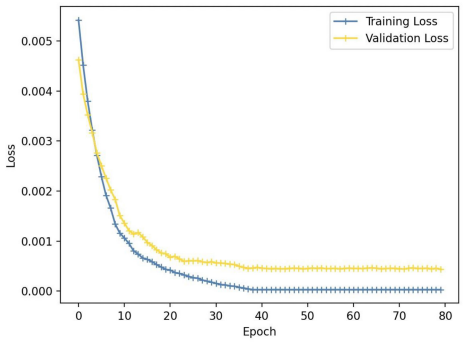


Fig. 9. The loss curves of the proposed network during the training process.

for each layer in the temporal convolution module and the final CNN. Before each LeakyReLU, a batch-normalization is implemented. We use the Adam optimizer for the training. The batch size is 64, and the training is conducted for 80 epochs. The learning rate is initialized by 0.001 and decayed with a rate of 0.5 after every 5 epochs to prevent overfitting.

During the training process, the normal scenario dataset is divided into a training set and a validation set-1. The training set is used to train the network with the mean squared error (MSE) [39] as the loss function, while validation set-1 is used to tune hyperparameters and detect overfitting. Additionally, 20% of the multi-type attack scenario dataset, maintaining a balanced normal/attack ratio, is used to determine the threshold. We manually select a threshold that balances the trade-off between precision and recall based on detection performance on the validation set-2.

C. Detection Performance

Fig. 9 displays the training and validating performances of the proposed network. Then, we perform experiments in two different attack scenarios, i.e., the multi-type attack scenario and the stealthy attack scenario, to test the effectiveness of *DST-GNN*. In the multi-type attack scenario, the dataset contains all seven attacks listed in Table II. Note that the multi-type attack scenario dataset used for testing does not include the data used to select the threshold. In contrast, in the stealthy attack scenario, the dataset includes the stealthy attack proposed in [38]. In addition, we test the robustness of *DST-GNN* in the special normal scenario.

The detection results of *DST-GNN* and the baseline methods in the multi-type attack scenario and the stealthy attack scenario are presented in Table III and IV, respectively.

- The metrics scores show that *DST-GNN* outperforms the baselines in both scenarios, achieving a high accuracy of 97.38% in the multi-type attack scenario and 91.96% in the stealthy attack scenario. In terms of F1-score, which is a critical metric for cyberattack detection, *DST-GNN* exhibits 2.01% and 0.77% higher F1-score than the second-best method in the respective scenarios.
- Furthermore, the results highlight the significance of using more data points for the effective cyberattack detection. In both scenarios, the methods utilizing 31 data points, i.e., *DST-GNN*, CNN-2, MLSTM-2, MLP and KNN, outperform those using fewer data points, i.e., CNN-1, MLSTM-1 and HSS-TF.
- In addition, CNN-1, MLSTM-1 and HSS-TF experience a notable decline in the detection performance in the stealthy attack scenario compared to the multi-type attack scenario, with decreases in the accuracy of 14.5%, 18.96% and 37.9%, respectively. This decline can be attributed to the fundamental principles underlying each method. The HSS-TF-based detector primarily assesses whether the data adheres to the physical dynamics, while CNN-1 and MLSTM-1 focus on the sudden changes in the PCC node voltage and current. However, the stealthy attack will manipulate sensor measurements to simultaneously conform to the physical dynamics and introduce the bias changes at an imperceptible speed, thereby avoiding the observable fluctuations in the PCC node voltage and current.
- Finally, among these baseline methods using 31 data points, CNN-2 exhibits the best detection performance in both scenarios. Its effectiveness lies in extracting the data features in two dimensions, while MLSTM-2 only focuses on the data correlations in the time dimension, and MLP and KNN use the flatted data inputs. This suggests that the method extracting features in both temporal and spatial dimensions is more effective in learning the latent relationships within multivariate time series signals in grid-tied PV systems.

TABLE III
MULTI-TYPE ATTACK SCENARIO: DETECTION RESULTS IN TERMS OF RECALL(%), PRECISION(%), F1-SCORE(%), AND ACCURACY(%)

Method	Recall	Precision	F1-score	Accuracy
CNN-1	90.12	90.10	90.11	91.05
CNN-2	95.11	95.07	95.09	95.56
MLSTM-1	90.03	89.94	89.99	90.94
MLSTM-2	94.89	94.99	94.94	95.42
HSS-TF	87.30	87.33	87.31	88.40
MLP	93.44	93.45	93.44	94.07
KNN	91.14	91.11	91.12	92.04
DST-GNN	**97.06**	**97.14**	**97.10**	**97.38**

TABLE IV
STEALTHY ATTACK SCENARIO: DETECTION RESULTS IN TERMS OF RECALL(%), PRECISION(%), F1-SCORE(%), AND ACCURACY(%)

Method	Recall	Precision	F1-score	Accuracy
CNN-1	76.58	76.58	76.58	76.55
CNN-2	91.35	91.08	91.21	91.19
MLSTM-1	71.96	72.04	72.00	71.98
MLSTM-2	90.02	89.63	89.24	89.79
HSS-TF	51.37	50.53	50.95	50.50
MLP	90.66	89.89	90.27	90.22
KNN	88.04	88.03	88.04	88.02
DST-GNN	**92.04**	**91.92**	**91.98**	**91.96**

TABLE V
ACCURACY (%) IN THE SPECIAL NORMAL SCENARIO

CNN-1	CNN-2	MLSTM-1	MLSTM-2
91.93	96.01	93.75	97.90
HSS-TF	MLP	KNN	*DST-GNN*
99.83	96.29	92.92	98.60

In a scenario of sudden irradiance change under special normal conditions, the performance of *DST-GNN* and baseline methods is detailed in Table V. Given the absence of attacked samples, we solely use the accuracy as the performance metric. The results indicate that the HSS-TF method demonstrates superior robustness against sudden irradiance changes. This can be attributed to the HSS-TF method's reliance entirely on system dynamics for detection. Conversely, despite incorporating system dynamics, our method remains dependent on latent data variations. Consequently, our method may misinterpret abrupt irradiance changes in a normal scenario as attacks. Nonetheless, our method exhibits an accuracy only slightly inferior to that of the HSS-TF method, demonstrating its robustness in normal scenarios with sudden irradiance changes.

In conclusion, *DST-GNN* outperforms all aforementioned baselines in both cyberattack scenarios, validating its effectiveness. Furthermore, utilizing more data points proves beneficial for cyberattack detection, as the methods using 31 data points outperform those with fewer data points. The top two methods, i.e., *DST-GNN* and CNN-2, capture both temporal patterns and spatial correlations, emphasizing the efficacy of the spatiotemoral features for detecting cyberattacks in grid-tied PV systems. Additionally, *DST-GNN* exhibits strong robustness against sudden irradiance changes. Although HSS-TF shows the best performance in the special normal scenario, it fails to defend against stealthy attacks that adhere to physical dynamics. Since the stealthy attack is a particular case, we employ the dataset of the multi-type attack scenario in the subsequent experiments.

D. Ablation Study

We perform the ablation study to investigate the effectiveness of the graph convolution module and the temporal convolution module in *DST-GNN*. Specifically, we gradually exclude these modules and observe the resulting degradation in the detection performance. *DST-GNN* without different modules is denoted as follows:

- *w/o GC: DST-GNN* without the graph convolution module. To assess the importance of extracting the spatial information among input signals, we remove the IP phase and IS phase and directly input the signals to the next module.
- *w/o TC: DST-GNN* without the temporal convolution module. To analyze the significance of the temporal convolution module in capturing the temporal patterns, we disable the 1D-CNN and directly apply the outputs of the graph convolution module as the inputs of the final CNN.

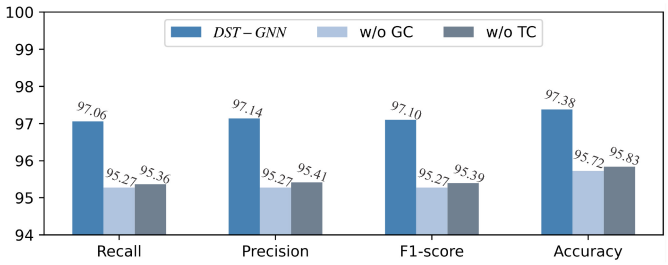


Fig. 10. Ablation study: detection results in terms of Recall(%), Precision(%), F1-score(%), and Accuracy(%).

The detection results for the ablation study are illustrated in Fig. 10. The results indicate that both two modules are indispensable. Although the final CNN can process data in two dimensions, the graph convolution module and the temporal convolution module, designed for learning the latent features in the spatial dimension and temporal dimension, respectively, are necessary for detecting cyberattacks in grid-tied PV systems. Furthermore, the introduction of the graph convolution module makes more significant contributions to cyberattack detection as it enables the capture of the spatial correlations in non-Euclidean space. The impact of the temporal convolution module is evident as well, validating the beneficial use of 1D-CNN for the temporal pattern extraction.

E. Interpretability Analysis

As stated before, *DST-GNN* integrates with the priori physical knowledge of the grid-tied PV system. Specifically, in the graph convolution module, the system dynamics is utilized to determine the graph weights and the number of graph convolution layers. In the temporal convolution module, the filter size and the number of 1D-convolution layers are designed based on Eq. (18). The integration of the priori physical knowledge can improve the interpretation and detection performance of *DST-GNN*. To validate this, we conduct a comparative analysis between *DST-GNN* and its variants, which involve variations in the weighted adjacency matrices, graph convolution layers and 1D-convolution filter sizes, specifically:

- *Self-attention:* the weighted adjacency matrix in *DST-GNN* is learned by the self-attention mechanism automatically [40].
- *Topology:* the weighted adjacency matrix in *DST-GNN* is determined by the grid-tied PV system topology, wherein the edge weight is set to 1 if two nodes are connected, and 0 otherwise.
- *Two-layer: DST-GNN* with a two-layer graph convolution module.
- *Three-layer: DST-GNN* with a three-layer graph convolution module.
- *Smaller filter:* the filter size of *DST-GNN*'s first, second and third 1D-convolution layers is set to 1×9.
- *Larger filter:* the filter size of *DST-GNN*'s first, second and third 1D-convolution layers is set to 1×10, 1×10 and 1×9, respectively.

The detection results are summarized in Table VI, which demonstrates the importance of incorporating the priori

TABLE VI
INTERPRETABILITY ANALYSIS: DETECTION RESULTS IN TERMS OF RECALL(%), PRECISION(%), F1-SCORE(%), AND ACCURACY(%)

Variation	Method	Recall	Precision	F1-score	Accuracy
Weighted adjacency matrices	Self-attention	95.10	95.27	95.18	95.65
	Topology	96.31	96.35	96.33	96.68
Graph convolution layers	Two-layer	96.15	96.15	96.15	96.52
	Three-layer	95.43	95.43	95.43	95.87
1D-convolution filter sizes	Smaller filter	96.03	96.04	96.03	96.41
	Larger filter	95.98	96.03	96.01	96.39
DST-GNN		**97.06**	**97.14**	**97.10**	**97.38**

physical knowledge about the grid-tied PV system in cyberattack detection.

- Firstly, among the variations of the weighted adjacency matrices, **Self-attention** exhibits the worst detection performance. It infers the latent correlations between multivariate time series through the self-attention mechanism without considering any physical and cyber connections between nodes. In contrast, **Topology**, which integrates the grid-tied PV system topology, demonstrates improved detection performance compared to **Self-attention**. Nevertheless, the topology information is static and lacks the specific strength indicators for connections between adjacent nodes. Consequently, **Topology** still falls short of the detection performance achieved by *DST-GNN*. The variations of the weighted adjacency matrices highlight the significance of leveraging both the physical and cyber modeling of the grid-tied PV system to construct the dynamic weighted adjacency matrix of *DST-GNN*.
- Secondly, as mentioned in Section IV-B, we choose the graph convolution module with a single layer for *DST-GNN*, given that the dynamics of state variables are primarily influenced by the directly relevant state variables. The detection results of the variations of graph convolution layers demonstrate the effectiveness of this design. Both **Two-layer** and **Three-layer** underperform *DST-GNN*, and the detection performance of the graph convolution module decreases with an increasing number of layers.
- Thirdly, in the variations of 1D-convolution filter sizes, **Smaller filter** and **Larger filter** lead to the left-hand side of Eq. (18) being smaller or larger than the right-hand side, respectively. As illustrated in Table VI, *DST-GNN*, which adheres to Eq. (18), outperforms both **Smaller filter** and **Larger filter**. This indicates the importance of determining the filter size and the number of 1D-convolution layers based on the periodicity in the input time series.

VI. CONCLUSION

In this paper, the *DST-GNN* is proposed to detect cyberattacks in grid-tied PV systems by capturing both inter-series spatial correlations and intra-series temporal relationships of multivariate time series signals. The latent spatial dependencies are captured in the graph convolution module, in which the dynamic graph and convolution layers are determined by the system topology and dynamics. The underlying temporal patterns are extracted by the temporal convolution module employing the 1D-CNN, in which the hyper-parameters (i.e., the filter size and the number of 1D-convolution layers) are designed based on the periodicity of input signals. Unlike existing cyberattack detection methods in grid-tied PV systems, *DST-GNN* not only eliminates the complex process of building the physical model but also leverages the priori system knowledge to improve its interpretation and detection performance. Experiments conducted on the HIL testbed demonstrate that *DST-GNN* shows excellent performance on cyberattack detection, even in the particular stealthy attack scenario. The ablation study confirms the necessity of the graph convolution module and the temporal convolution module. Furthermore, the significance of integrating the priori physical knowledge is verified by performing the comparative analysis between *DST-GNN* and its variants.

The proposed method is inherently adaptable to other systems, such as PV systems using grid-forming control and larger-scale power grids comprising multiple PV systems. This adaptability stems from the core concept of our approach, which enables creating a dynamic graph based on the physical and cyber connections among system signals. Once this dynamic graph is established, *DST-GNN* can be retrained to capture the specific interactions and dependencies of the new system. However, several challenges need to be addressed when applying this method to larger-scale power grids with multiple PV systems. First, training *DST-GNN* on a larger-scale graph poses challenges in computational complexity, memory usage, and convergence guarantee. Second, it is difficult to represent the cyber connections of a large number of PV systems with varying time scales on different operation levels. Our future work will focus on extending the current method to accommodate for large-scale systems more effectively.

REFERENCES

[1] "Renewables 2023." IEA. 2024. [Online]. Available: https://www.iea.org/reports/renewables-2023
[2] B. Kellison, (Wood Mackenzie, Edinburgh, U.K.). *The Next Five Years Will See Massive Distributed Energy Resource Growth*. 2021. [Online]. Available: https://www.woodmac.com/news/editorial/der-growth-united-states/
[3] *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces*, IEEE Standard 1547-2018 (Revision of IEEE Std 1547-2003), 2018.

[4] W. Hupp, D. Saleem, J. T. Peterson, and K. Boyce, "*Cybersecurity Certification Recommendations for Interconnected Grid Edge Devices and Inverter Based Resources*," Nat. Renew. Energy Lab., Golden, CO, USA, 2021.

[5] J. Ye et al., "A review of cyber–physical security for photovoltaic systems," *IEEE J. Emerg. Select. Topics Power Electron.*, vol. 10, no. 4, pp. 4879–4901, Aug. 2022.

[6] A. Teymouri, A. Mehrizi-Sani, and C.-C. Liu, "Cyber security risk assessment of solar PV units with reactive power capability," in *Proc. IECON 44th Annu. Conf. IEEE Ind. Electron. Soc.*, 2018, pp. 2872–2877.

[7] S. Soltan, P. Mittal, and H. V. Poor, "{BlackIoT}:{IoT} botnet of high wattage devices can disrupt the power grid," in *Proc. 27th USENIX Secur. Symp. (USENIX Security 18)*, 2018, pp. 15–32.

[8] Z. Zhang, R. Deng, and D. K. Yau, "Vulnerability of the load frequency control against the network parameter attack," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 921–933, Jan. 2024.

[9] M. Liu et al., "Enhancing cyber-resiliency of der-based smart grid: A survey," *IEEE Trans. Smart Grid*, early access, Mar. 5, 2024, doi: 10.1109/TSG.2024.3373008.

[10] J. Hou, C. Hu, S. Lei, and Y. Hou, "Cyber resilience of power electronics-enabled power systems: A review," *Renew. Sustain. Energy Rev.*, vol. 189, Jan. 2024, Art. no. 114036.

[11] J. Dai, J. Yang, Y. Wang, and Y. Xu, "Blockchain-enabled cyber-resilience enhancement framework of microgrid distributed secondary control against false data injection attacks," *IEEE Trans. Smart Grid*, vol. 15, no. 2, pp. 2226–2236, Mar. 2024.

[12] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security Privacy*, vol. 9, no. 3, pp. 49–51, Jun. 2011.

[13] J. Zhang, L. Guo, and J. Ye, "Cyber-attack detection for photovoltaic farms based on power-electronics-enabled harmonic state space modeling," *IEEE Trans. Smart Grid*, vol. 13, no. 5, pp. 3929–3942, Sep. 2022.

[14] J. Zhang and J. Ye, "Cyber-attack detection for active neutral point clamped (ANPC) photovoltaic (PV) converter using Kalman filter," in *Proc. IEEE Appl. Power Electron. Conf. Expo. (APEC)*, 2022, pp. 1939–1944.

[15] J. Yan, F. Guo, and C. Wen, "Attack detection and isolation for distributed load shedding algorithm in microgrid systems," *IEEE J. Emerg. Select. Topics Ind. Electron.*, vol. 1, no. 1, pp. 102–110, Jul. 2020.

[16] M. Liu, C. Zhao, Z. Zhang, R. Deng, P. Cheng, and J. Chen, "Converter-based moving target defense against deception attacks in DC microgrids," *IEEE Trans. Smart Grid*, vol. 13, no. 5, pp. 3984–3996, Sep. 2022.

[17] M. Liu, C. Zhao, J. Xia, R. Deng, P. Cheng, and J. Chen, "PDDL: Proactive distributed detection and localization against stealthy deception attacks in DC microgrids," *IEEE Trans. Smart Grid*, vol. 14, no. 1, pp. 714–731, Jan. 2023.

[18] S. Sahoo, S. Mishra, J. C.-H. Peng, and T. Dragičević, "A stealth cyber-attack detection strategy for DC microgrids," *IEEE Trans. Power Electron.*, vol. 34, no. 8, pp. 8162–8174, Aug. 2019.

[19] F. Li et al., "Detection and diagnosis of data integrity attacks in solar farms based on multilayer long short-term memory network," *IEEE Trans. Power Electron.*, vol. 36, no. 3, pp. 2495–2498, Mar. 2021.

[20] Q. Li, J. Zhang, J. Zhao, J. Ye, W. Song, and F. Li, "Adaptive hierarchical cyber attack detection and localization in active distribution systems," *IEEE Trans. Smart Grid*, vol. 13, no. 3, pp. 2369–2380, May 2022.

[21] L. Guo, J. Zhang, J. Ye, S. J. Coshatt, and W. Song, "Data-driven cyber-attack detection for PV farms via time-frequency domain features," *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1582–1597, Mar. 2022.

[22] F. Li et al., "Detection and identification of cyber and physical attacks on distribution power grids with PVs: An online high-dimensional data-driven approach," *IEEE J. Emerg. Select. Topics Power Electron.*, vol. 10, no. 1, pp. 1282–1291, Feb. 2022.

[23] M. R. Habibi, H. R. Baghaee, F. Blaabjerg, and T. Dragičević, "Secure MPC/ANN-based false data injection cyber-attack detection and mitigation in DC microgrids," *IEEE Syst. J.*, vol. 16, no. 1, pp. 1487–1498, Mar. 2022.

[24] X. Ma et al., "A comprehensive survey on graph anomaly detection with deep learning," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12012–12038, Dec. 2023.

[25] Y. Wang, D. Qiu, Y. Wang, M. Sun, and G. Strbac, "Graph learning-based voltage regulation in distribution networks with multi-microgrids," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 1881–1895, Jan. 2024.

[26] S. Peng, Z. Zhang, R. Deng, and P. Cheng, "Localizing false data injection attacks in smart grid: A spectrum-based neural network approach," *IEEE Trans. Smart Grid*, vol. 14, no. 6, pp. 4827–4838, Nov. 2023.

[27] R. Yan, Y. Xu, and R. Zhang, "Graph attention network based reinforcement learning method for optimal distributed frequency control of an islanded AC microgrid," in *Proc. IEEE Power Energy Soc. General Meet. (PESGM)*, 2023, pp. 1–5.

[28] S. Zhao, J. Xia, R. Deng, P. Cheng, Q. Yang, and X. Jiao, "Dual-triggered adaptive torque control strategy for variable-speed wind turbine against denial-of-service attacks," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 3072–3084, Jul. 2023.

[29] A. Barua and M. A. Al Faruque, "Hall spoofing: A non-invasive DoS attack on grid-tied solar inverter," in *Proc. 29th USENIX Secur. Symp. (USENIX Secur.)*, 2020, pp. 1273–1290.

[30] J. Choi, B. Ahn, G. Bere, S. Ahmad, H. A. Mantooth, and T. Kim, "Blockchain-based man-in-the-middle (MITM) attack detection for photovoltaic systems," in *Proc. IEEE Design Methodol. Conf. (DMC)*, 2021, pp. 1–6.

[31] R. Akkaoui, A. Stefanov, P. Palensky, and D. H. Epema, "Resilient, auditable and secure IoT-enabled smart inverter firmware amendments with blockchain," *IEEE Internet Things J.*, vol. 11, no. 5, pp. 8945–8960, Mar. 2024.

[32] S. Weng, D. Yue, J. Chen, X. Xie, and C. Dou, "Distributed resilient self-triggered cooperative control for multiple photovoltaic generators under denial-of-service attack," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 1, pp. 226–237, Jan. 2023.

[33] H. Mohsenian-Rad and W. Xu, "Synchro-waveforms: A window to the future of power systems data analytics," *IEEE Power Energy Mag.*, vol. 21, no. 5, pp. 68–77, Oct. 2023.

[34] F. Li et al., "Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution," *ACM Trans. Knowl. Disc. Data*, vol. 17, no. 1, pp. 1–21, 2023.

[35] H. Mohsenian-Rad, *Smart Grid Sensors: Principles and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2021.

[36] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang, "Connecting the dots: Multivariate time series forecasting with graph neural networks," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mini.*, 2020, pp. 753–763.

[37] M. Liu, Z. Jin, J. Xia, M. Sun, R. Deng, and P. Cheng, "Demo abstract: A HIL emulator-based cyber security testbed for DC microgrids," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2021, pp. 1–2.

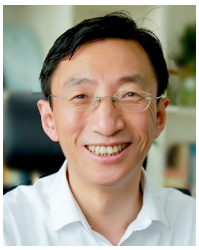
[38] S. Peng, M. Liu, K. Zuo, W. Tan, and R. Deng, "Stealthy data integrity attacks against grid-tied photovoltaic systems," in *Proc. IEEE 6th Int. Conf. Ind. Cyber-Phys. Syst. (ICPS)*, 2023, pp. 1–7.

[39] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006.

[40] D. Cao et al., "Spectral temporal graph neural network for multivariate time-series forecasting," in *Proc. 34th Conf. Neural Inf. Process. Syst.*, 2020, pp. 17766–17778.

Sha Peng received the B.S. degree from the College of Computer Science and Technology, Jilin University, Changchun, China, in 2020. She is currently pursuing the Ph.D. degree with the College of Control Science and Engineering, Zhejiang University, Hangzhou, China. She is a Visiting Scholar with the National University of Singapore, Singapore, from 2023 to 2024. Her current research interests include data analytics, cyber security, and smart grid.

Mengxiang Liu (Member, IEEE) received the B.Sc. degree in automation from Tongji University, Shanghai, in 2017, and the Ph.D. degree in cyberspace security from Zhejiang University, Hangzhou, in 2022. He is currently a Research Associate of Power Systems Engineering with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K. His research interests include cyber resiliency, smart grid, active defense, and cyber-physical co-simulation.



Li Chai (Member, IEEE) received the B.S. degree in applied mathematics and the M.S. degree in control science and engineering from Zhejiang University, China, in 1994 and 1997, respectively, and the Ph.D. degree in electrical engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2002.

From August 2002 to December 2007, he was with Hangzhou Dianzi University, China. He worked as a Professor with the Wuhan University of Science and Technology, China, from 2008 to 2022. In August 2022, he joined Zhejiang University, China, where he is currently a Full Professor with the College of Control Science and Engineering. He has been a Postdoctoral Researcher or a Visiting Scholar with Monash University, Newcastle University, Australia, and Harvard University, USA. He is the Expert of State Council Special Allowance, China. He is the recipient of the Distinguished Young Scholar of the National Science Foundation of China. He has published over 100 fully refereed papers in prestigious journals and leading conferences. His research interests include distributed optimization, graph signal processing, filter banks, and networked control systems. He has been an Associate Editor of IEEE TRANSACTIONS ON CIRCUIT AND SYSTEMS–II: EXPRESS BRIEFS, and *Journal of Control and Decision*.

Ruilong Deng (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, Zhejiang, China, in 2009 and 2014, respectively.

He was a Research Fellow with Nanyang Technological University, Singapore, from 2014 to 2015; an AITF Postdoctoral Fellow with the University of Alberta, Edmonton, AB, Canada, from 2015 to 2018; and an Assistant Professor with Nanyang Technological University from 2018 to 2019. He is currently a Professor with the College of Control Science and Engineering, Zhejiang University. His research interests include smart grid, cyber security, and control systems. He serves/served as an Associate Editor for IEEE TRANSACTIONS ON SMART GRID, IEEE POWER ENGINEERING LETTERS, IEEE/CAA JOURNAL OF AUTOMATICA SINICA, and IEEE/KICS JOURNAL OF COMMUNICATIONS AND NETWORKS, and a Guest Editor for IEEE TRANSACTIONS ON CLOUD COMPUTING, IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING, IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN INDUSTRIAL ELECTRONICS, and *IET Cyber-Physical Systems: Theory and Applications*. He also serves/served as the Symposium Chair for IEEE SmartGridComm, IEEE ICPS, and IEEE GLOBECOM.