Stealthy Data Integrity Attacks Against Grid-tied Photovoltaic Systems

Sha Peng, Mengxiang Liu, Ke Zuo, Wei Tan, and Ruilong Deng

***Abstract*—Under the transformation of electric grid towards sustainability and decarbonization, a large number of distributed energy resources including solar photovoltaic (PV) farms are expected to penetrate the grid. As one of the critical state infrastructures, the cybersecurity of PV systems has attracted numerous attention especially with the standardization of grid support services. Various data-driven and model-based intrusion detection systems (IDSs) have emerged for the cybersecurity issue of grid-tied PV systems, among which the *stealthy* data integrity attacks (DIA) are rarely mentioned. In this paper, we propose a generation scheme of stealthy DIAs, which can bypass two recently proposed (almost state-of-the-art) data-driven and model-based IDSs simultaneously. The attack stealthiness is guaranteed by compromising the sensor measurements cooperatively conforming the physical dynamics of the grid-tied PV system, and meanwhile the attack vector needs to change with an imperceptible speed to avoid steep and observable increase/decrease. Systematical HIL experiments are conducted to verify the stealthiness of the designed stealthy DIA and evaluate its attack impact on PCC voltages.**

***Index Terms*—Stealthy Data Integrity Attack, Grid-tied Photovoltaic Systems, Intrusion Detection System**

I. Introduction

The current electric grid is undergoing significant and rapid transformation to a sustainable and decarbonized electric grid [1]. The deployment of variable generation, primarily wind and solar, is leading this transformation and is associated with the move from the physics of large spinning generation to power systems dominated by power electronics enabled resources [2]. It is expected that the deployment of distributed energy resources (DER) will be approximately 280 gigawatts (GW) by 2025, where nearly half of DER in 2021 are solar photovoltaic (PV) systems [3]. Owing to the rapid development of power electronics converters and information communication technologies, PV systems can generate software-driven and digital-controlled output powers according to participated grid support services like reactive power capability and voltage/power control [4], whose configuration parameters can be remotely tuned by system operators. Several technical reports from Sandia National Laboratories pointed out that the cyber vulnerabilities in the DER dominated electric grid come from a wide range of factors including communication protocols, software updates and patches, supply chain, insider, social engineering, etc [5]–[7]. Once the attacker intrudes into the network of a PV system, the information comprising the meter readings, monitoring/diagnostics data, control loop commands/measurements/parameters, etc [8]. could be manipulated to cause economical and operational losses to the electric grid like voltage violation [9], line failure, and blackout [10].

Although IEEE 1547 Std [4] has proposed trust and cryptography features including data encryption, access authentication, and key management for communication protocols (Modbus, IEC 61850, DNP3, etc,) that are widely used in the interconnection of DERs with associated area electric grid, it is still possible for powerful adversaries to bypass the *prevention* security strategies by exploiting zero-day vulnerabilities like the Stuxnet accident [11]. It thereby has great importance to deploy intrusion detection systems (IDSs) in PV systems, namely the second defense line, which can perceive the existence of adversaries through either signature of known attacks or observed host, network, and physical anomaly induced by attacks [12]. In this paper, we focus on the physical-based IDS that utilizes physical measurements such as point of common voltage/current, output active/reactive power, etc. to determine the anomaly caused by adversaries, which has attracted great attention in the power and control societies. According to the detection techniques, physical-based IDSs can be classified into data-driven and model-based IDSs.

The data-driven IDS can adopt machine learning methods, statistical patterns, and specifications to identify anomaly data from normal behaviours. Li *et al.* proposed an adaptive hierarchical cyberattack detection and localization for active distribution systems with DERs using the electrical waveform, where the multi-layer long short-term memory (MLSTM) network is utilized to classify anomaly from normal behaviours [13], [14]. Guo *et al.* presented a detection and diagnosis framework for power electronics converter enabled PV farms via single waveform sensor to distinguish between normal conditions, open-circuit faults, and cyberattacks, where innovative frequency-domain magnitude-based and time-domain mean current vector-based features are proposed and LSTM and convolutional neural network are used for classification [15]. Mustafa *et al.* proposed an attack detection mechanism

This work was supported in part by the National Natural Science Foundation of China under Grant 62293503, 62073285, in part by the Natural Science Foundation of Zhejiang Province under Grant LR23F030001, LZ21F020006, in part by the Xiaomi Foundation, and in part by the Fundamental Research Funds for the Central Universities (226-2022-00120).

S. Peng, K. Zuo and R. Deng 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, kezuo, dengruilong}@zju.edu.cn)

M. Liu is with Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: m.liu22@imperial.ac.uk)

W. Tan is with Shanghai Hanxiang Intelligent Technology Co.,Ltd., Shanghai 201600 China. (e-mail:wei.tan@hanxiang-tech.com)

using a Kullback-Liebler (KL) divergence-based criterion for each DER to detect any misbehavior on its neighboring DERs, where the KL divergence is a non-negative measure of the relative entropy between two probability distributions [16]. Beg *et al.* developed signal temporal logic detection of two major types of cyberattacks, namely false data injection and denial-of-service attacks in DC microgrids, where STL is a formalism to monitor the output voltages and currents of DC microgrids against the defined specifications, such as operational bounds, over time [17].

The model-based IDS focuses on verifying the data anomaly employing the physical dynamics involving the electric circuit and controller. Zhang *et al.* presented a physics-data-based detection method to detect a variety of cyber-attacks in PV farms using the frequency-domain power electronics-enabled harmonic state space (HSS) models, which requires less sensor measurements compared to Kalman Filter-based methods and can achieve comparable performance [18]. Gallon and Liu *et al.* proposed a distributed detection scheme that can be directly applied to DC microgrids by synthesising a Luenberger observer and a bank of unknowninput observers, realizing improved detection performance that cannot be achieved by either observer module [19]–[21]. Sahoo *et al.* proposed a novel cooperative vulnerability factor detection framework against false data injection attacks for each DER system, where the factor is derived insipred from the consensus-based secondary control algorithm and will be nontrivial only under the presence of attacks [22].

One critical issue of the physical-based IDSs in the existing literature is that the threat of powerful adversaries like the state-sponsored actor is overlooked especially in the scenario of PV systems. The recently proposed IDSs in [13], [18] merely consider multiply- and add-based biases, while the possibility for the adversary to design *stealthy* attack vectors that conform to the physical dynamics and deviate the states from normal ones slowly is ignored. Given the lesson learned from the Stuxnet accident, the state-sponsored adversary can not only exploit zero-day vulnerabilities but also steal, buy, infer critical information including the system topology, electrical parameters, control algorithm, and detection scheme. Therefore, it has great importance to consider *stealthy* attacks in the critical state infrastructure, namely the grid-tied PV farm, which is becoming common as the global trend towards decarbonization. To fill this gap, in this paper, we propose a generation scheme of the *stealthy* data integrity attack (DIA) that can bypass two typical data-driven [13] and model-based [18] IDSs. The contributions are as follows:

- We design an automatic stealthy DIA generation scheme, only requiring that the adversary can obtain local static parameters like electrical and control related parameters;
- To bypass the model-based IDS, the sensor measurements are compromised simultaneously conforming the physical dynamics; To bypass the date-driven IDS, the injected bias changes with an imperceptible speed avoiding steep and observable increase/decrease.
- We conduct systematical experiments on a hardware-in-the-loop (HIL) testbed to validate the stealthiness of generated DIA vectors and evaluate their attack impacts on grid-tied PV systems.

The remainder of this paper is as follows: Section II presents the grid-tied PV system modeling, and Section III introduces the model-based and data-driven IDSs. Section IV details the design of the stealthy DIA and Section V demonstrates the experimental results. Section VI concludes this paper and provides possible defensive strategies in future works.

II. Grid-tied PV System Modeling

The block diagram of a grid-connected PV system is shown in Fig. 1, which includes two main stages, i.e., the DC-DC and DC-AC conversion stages. As for the DC-DC conversion stage, the PV array output voltage U_{pv} and current I_{pv} are interfaced with the DC/DC control. The DC/DC controller employs the Maximum Power Point Tracking (MPPT) algorithm to extract the maximum power from the PV panels. The DC/DC controller output duty cycle D is given as input to the DC/DC converter that maintains the operating voltage at the maximum power point. In the DC-AC conversion stage, the DC/AC inverter is mainly responsible for changing the DC power the first stage produces to AC power. As shown in Fig. 1, the DC/AC controller consists of the voltage control loop, reactive control loop and current control loop. The voltage control loop maintains DC link voltage U_{dc} and derives the optimal inverter output current value in d frame I^*_{fd}. The other optimal value I^*_{fq} is determined by the reactive control loop used to generate the required reactive power. The current control loop is designed to force the DC/AC inverter output current I_f to track the reference set points I^*_{fd} and I^*_{fq}.

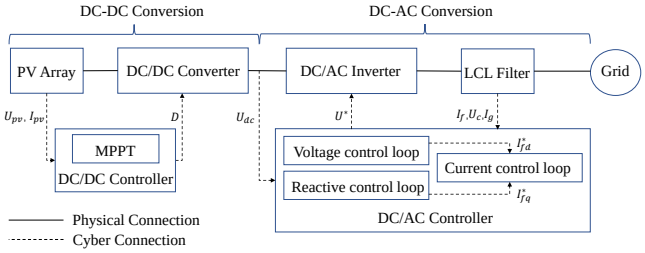


Fig. 1: The block diagram of a grid-connected PV system.

Fig. 2 shows a detailed DC-AC conversion stage. The LCL filter can be represented by:

$$\begin{cases} \dot{I_{fd}} = \dfrac{[U_d - I_{fd}R_f - U_{cd} - (I_{fd} - I_{gd})\,R_c]}{L_f} + \omega I_{fq} \\ \dot{I_{fq}} = \dfrac{[U_q - I_{fq}R_f - U_{cq} - (I_{fq} - I_{gq})\,R_c]}{L_f} - \omega I_{fd} \\ \dot{U_{cd}} = \dfrac{(I_{fd} - I_{gd})}{C_f} + \omega U_{cq} \\ \dot{U_{cq}} = \dfrac{(I_{fq} - I_{gq})}{C_f} - \omega U_{cd} \\ \dot{I_{gd}} = \dfrac{[U_{cd} + (I_{fd} - I_{gd})R_c - I_{gd}R_g - U_{gd}]}{L_g} + \omega I_{cq} \\ \dot{I_{gq}} = \dfrac{[U_{cq} + (I_{fq} - I_{gq})R_c - I_{gq}R_g - U_{gq}]}{L_g} - \omega I_{cd} \end{cases} \tag{1}$$

where $I_{fd,q}$ is the inverter side current of LCL, U_{dq} is the inverter side voltage, $U_{cd,q}$ is the LCL capacitor, $I_{d,q}$ is the grid side current of LCL and $U_{gd,q}$ is the grid side voltage; ω is the system frequency; R_f,L_f,R_c,C_f,R_g and L_g are LCL filter parameters. Let $\gamma_{d,q}$ denotes the state of the inner current control loop, which can be expressed as:

$$\begin{cases} \dot{\gamma_d} = I_{fd}^* - I_{fd} \\ \dot{\gamma_q} = I_{fq}^* - I_{fq} \end{cases}, \tag{2}$$

$$\begin{cases} U_{id}^* = U_{cd} + k_p(I_{fd}^* - I_{fd}) + k_i \int (I_{fd}^* - I_{fd}) - \omega L_f i_{fq} \\ U_{iq}^* = U_{cq} + k_p(I_{fq}^* - I_{fq}) + k_i \int (I_{fq}^* - I_{fq}) + \omega L_f i_{fd} \end{cases}. \tag{3}$$

Assuming $U_{id,q}^* = U_{id,q}$, the model of PV converter and current control loop can be written as

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} = C\mathbf{x} \end{cases}, \tag{4}$$

where

$$A = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ \frac{k_i}{L_f} & 0 & -\frac{R_f+k_p+R_c}{L_f} & 0 & 0 & 0 & \frac{R_c}{L_f} & 0 \\ 0 & \frac{k_i}{L_f} & 0 & -\frac{R_f+k_p+R_c}{L_f} & 0 & 0 & 0 & \frac{R_c}{L_f} \\ 0 & 0 & \frac{1}{C_f} & 0 & 0 & \omega & -\frac{1}{C_f} & 0 \\ 0 & 0 & 0 & \frac{1}{C_f} & 0 & 0 & \omega & -\frac{1}{C_f} \\ 0 & 0 & \frac{R_c}{L_g} & 0 & \frac{1}{L_g} & 0 & -\frac{R_g+R_c}{L_g} & \omega \\ 0 & 0 & 0 & \frac{R_c}{L_g} & 0 & \frac{1}{L_g} & -\omega & -\frac{R_g+R_c}{L_g} \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 & \frac{k_p}{L_f} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{k_p}{L_f} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{L_g} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{L_g} \end{bmatrix}^T,$$

$\mathbf{x} = [\gamma_d, \gamma_q, I_{fd}, I_{fq}, U_{cd}, U_{cq}, I_{gd}, I_{gq}]^T$ is the state of system, $\mathbf{u} = [I_{fd}^*, I_{fd}^*, U_{gd}, U_{gq}]^T$ is the control input, $\mathbf{y}$ is the output vector, C is the output matrix, and k_p and k_i are the PI controller parameters.

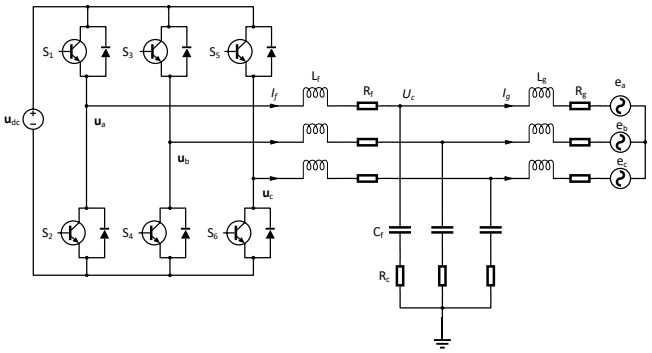


Fig. 2: The DC-AC conversion stage.

III. Model-based and Data-driven IDSs

A. Model-based IDS

In the system model described in eq. (4), the signal $\mathbf{x}$, $\mathbf{u}$ and $\mathbf{y}$ are represented by $\mathbf{x}(t)$, $\mathbf{u}(t)$ and $\mathbf{y}(t)$, respectively. With the HSS modeling, these time domain signals can be transformed to the frequency domain using the fast Fourier transform (FFT). Then the exponentially modulated periodic (EMP) signal as the kernel function (e^{-st}) is used to describe the dynamic performance in the time and frequency domains. Finally, the HSS equation can be expressed as

$$\begin{cases} (s + jm\omega) X_n = \sum_{n=-\infty}^{\infty} A_{n-m} X_m + \sum_{n=-\infty}^{\infty} B_{n-m} U_m \\ Y_n = \sum_{n=-\infty}^{\infty} C_{n-m} X_m \end{cases}, \tag{5}$$

which is equivalent to

$$\begin{cases} s\mathbf{X} = (A_T - N)\mathbf{X} + B_T\mathbf{U} \\ \mathbf{Y} = C_T\mathbf{X} \end{cases}, \tag{6}$$

where

$$\begin{aligned} &\mathbf{X} = [\cdots, X_{-2}, X_{-1}, X_0, X_1, X_2, \cdots]^T, \\ &\mathbf{U} = [\cdots, U_{-2}, U_{-1}, U_0, U_1, U_2, \cdots]^T, \\ &\mathbf{Y} = [\cdots, Y_{-2}, Y_{-1}, Y_0, Y_1, Y_2, \cdots]^T, \\ &N = blockdiag(\cdots, -j2\omega I, -j\omega I, \mathbf{0}, j\omega I, j2\omega I, \cdots). \end{aligned}$$

Here, $\mathbf{X}$, $\mathbf{U}$ and $\mathbf{Y}$ are harmonics vectors. The harmonics vector is the Fourier coefficients, and the vector element subscript denotes the harmonic order. A_T, B_T and C_T are the Toeplitz matrix. Since A, B and C are constants in the PV system model descried in Eq. (4), $A_T = blockdiag(A)$, $B_T = blockdiag(B)$ and $C_T = blockdiag(C)$. Thus, the harmonic transfer function is

$$H_T = C_T(sI + N - A_T)^{-1}B_T, \tag{7}$$

The HSS transfer function (HSS-TF)-based detector is a residual-based detection algorithm consisting of a voltage control loop detector (VLD) and a current control loop detector (CLD). As for VLD, the residual at time instant k is calculated as

$$R_{U_{dc}}(k) = \frac{\sum_{j=k-N_d+1}^{k} (|U_{dc}^* - U_{dc}(j)|)}{N_d U_{dc}^*}, \tag{8}$$

where U_{dc}^* is the DC-link voltage reference, $U_{dc}(j)$ is the sensor measurement at time j and N_d is the window length. The VLD detection result is

$$VLD(k) = \begin{cases} \text{Abnormal}, R_{U_{dc}}(k) \geq \gamma_u \\ \text{Normal}, R_{U_{dc}}(k) < \gamma_u \end{cases}, \tag{9}$$

where γ_u is the predefined threshold. In CLD, at time instant k, the residual can be expressed as

$$R_{I_{gd}}(k) = \frac{\sum_{j=k-N_d+1}^{k} \left(\left|\mathbf{y}_{I_{gdest}}(j) - \mathbf{y}_{I_{gd}}(j)\right|\right)}{N_d |\mathbf{y}_{I_{gdest}}(k)|}, \tag{10}$$

where $\mathbf{y}_{I_{gdest}}$ is the estimate of I_{gd} obtained by the harmonic transfer function H_T and $\mathbf{y}_{I_{gd}}$ is the measurement data. The detection result about I_{gd} is

$$CLD_d(k) = \begin{cases} \text{Abnormal, } R_{I_{gd}}(k) \geq \gamma_i \\ \text{Normal }, R_{I_{gd}}(k) < \gamma_i \end{cases} . \quad (11)$$

Here, γ_i is the predefined threshold. The detection result about I_{gq} can be similarly calculated.

B. Data-driven IDS

The MLSTM-based detector aims to detect cyberattacks using the PCC node's time series 3-phase voltage and current measurements. LSTM belongs to the gate-controlled recurrent neural networks that show excellent ability in dealing with time series data. MLSTM is a stacked LSTM architecture consisting of multiple LSTM layers. The sequence input of the LSTM layer below is the sequence output of the LSTM layer above. MLSTM utilizes previous time series measurements to model complex nonlinear temporal dependencies of the system and predict the future measurements the system would produce. Abnormal data can be detected by comparing the prediction results with the sensor measurements. The MLSTM-based detector result at time instant k can be expressed as

$$R(k) = \begin{cases} \text{Abnormal, } |\mathcal{G}^*(M(k-1)) - m(k)|_2 \geq \alpha \\ \text{Normal }, |\mathcal{G}^*(M(k-1)) - m(k)|_2 < \alpha \end{cases} \quad (12)$$

where $m(k)$ denotes the measurement vector at time instant k, $M(k-1)$ denotes a set of measurement vectors from $k - N_d$ to $k-1$ and $\mathcal{G}^*$ is the trained MLSTM model.

IV. Stealthy DIA Model

If an attacker gains knowledge of static parameters of the PV model described in eq. (4), the HSS-TF-based and the MLSTM-based detectors mentioned in Section III have a shared vulnerability that allows the attacker to circumvent them without being detected. The detectors assume that the observed measurements will deviate from the estimates when cyberattacks occur. However, when an attacker knows PV model parameters, he/she can delicately design the stealthy attack such that the above assumption is violated, thus introducing disturbance to the PV system state without being detected by the detectors.

The assumptions on the attacker's capabilities in the proposed stealthy DIA model are listed as follows:

- ***Assumption 1:*** The attacker knows the static parameters of the model, i.e., the attacker can calculate the matrices of A, B and C in eq. (4).
- ***Assumption 2:*** The attacker can manipulate a subset of sensor measurements before the controllers access them. Specifically, the attacker can tamper with I_{fd}, I_{fq}, U_{cd}, U_{cq}, I_{gd} and I_{gq}, simultaneously.

The key to our attack design is that, by calculating the state model parameters A, B, and C, the attacker can further obtain the stable transfer function $H = C(-A)^{-1}B$. By adding an attack signal which is the linear combination of H, we ensure that the attack signal cannot be detected by the transfer function based detector.

Let $\tilde{\mathbf{z}}$ denote the sequential observations that may be attacked. At the time instant k, $\tilde{\mathbf{z}}$ is:

$$\begin{aligned} \tilde{\mathbf{z}}(k) &= \mathbf{z}(k) + \mathbf{z}_{att}(k) \\ \mathbf{z}(k) &= (\mathbf{y}(k - N_d + 1), \cdots, \mathbf{y}(k)) \end{aligned} \quad (13)$$

where N_d is the length of the sliding window, and $\mathbf{z}_{att}(k)$ is a sequential malicious data. The $\mathbf{z}_{att}(k)$ is a zero matrix if no attack is implemented.

To implement an attack, the attacker can choose any arbitrary nonzero matrix as $\mathbf{z}_{att}(k)$. However, the attack matrix should satisfy the following conditions to accomplish a stealthy attack.

1) Firstly, the attack vectors, which are the column vectors of $\mathbf{z}_{att}(k)$, should be linear combinations of the column vectors of stable transfer function H, i.e., $\mathbf{z}_{att}(k) = H \cdot w(k)$ where $w(k)$ is some constant matrix that is a N_d duplicate of its column vector. We refer $w(k)$ as the attack weight matrix, and its column vectors as the attack weight vector;
2) Secondly, since the detection criteria are based on the fraction of residual and estimation, the attack vector should increase the observation on $I_{gd,q}$ to decrease the criterion;
3) Thirdly, the attack vector should change gradually to ensure a steady transient process that could bypass MLSTM based detector. The rapidity of change is irrelevant to whether the change happens in milliseconds or seconds, and the gradual change means that the slope of the attack vector should not exceed this system's normal range.

We specifically illustrate the characteristic of matrix $w(k)$. If $w(k)$ lies between changes in the attack signal, each column vector of matrix $w(k)$ is the same. Suppose $w(k)$ contains the instant that the attack signal changes, we can still approximately consider that the matrix $w(k)$ has the same column vectors since the attack signal changes slowly and gradually. Therefore, we deem each column vector of $w(k)$ the same in the following derivation.

To begin with, we prove that the proposed method could bypass the CLD detector. At time instant k, the attacked observation sequence would be: $\tilde{\mathbf{z}}(k) = \mathbf{z}(k) + \mathbf{z}_{att}(k) = \mathbf{z}(k) + Hw(k)$, where H is the stable transfer function and $Hw(k)$ is the attack matrix. Since in stable condition, $\mathbf{y}(k) = H\mathbf{u}(k)$, the effect of adding the proposed attack signal is equivalent to adding an attack weight matrix on the control input as

$$\tilde{\mathbf{u}}(k) = \mathbf{u}(k) + \mathbf{u}_{att}(k) = \mathbf{u}(k) + w(k). \quad (14)$$

The CLD-detector would use the Fourier coefficient of $\tilde{\mathbf{u}}(k)$ and harmonic transfer function H_T to calculate the Fourier coefficient of the following estimated observation:

$$\tilde{\mathbf{u}}(k) = \mathcal{F}(\tilde{\mathbf{u}}(k)) = \mathcal{F}(\mathbf{u}(k) + w(k)), \quad (15)$$

$$\begin{aligned}\tilde{\mathbf{Y}}(k) &= H_T\tilde{\mathbf{u}}(k) = H_T\mathcal{F}(\mathbf{u}(k)+w(k)) \\ &= H_T\mathcal{F}(\mathbf{u}(k)) + H_T\mathcal{F}(w(k))\end{aligned}. \tag{16}$$

Note that the second term is the effect of the attack signal. The detector then uses inverse Fourier transform to obtain the estimated signal of observation as:

$$\begin{aligned}\mathbf{y}_{est}(k) &= \mathcal{F}^{-1}(\tilde{\mathbf{Y}}(k)) = \mathcal{F}^{-1}(H_T\mathcal{F}(\mathbf{u}(k)) + H_T\mathcal{F}(w(k))) \\ &= \mathcal{F}^{-1}(H_T\mathcal{F}(\mathbf{u}(k))) + \mathcal{F}^{-1}(H_T\mathcal{F}(w(k)))\end{aligned} \tag{17}$$

According to the first condition of the attack matrix, the attack weight matrix $w(k)$ is a matrix that all its column vectors are the same. Thus its Fourier coefficients are all zero except the zero frequency element and mathematically is $\mathcal{F}(w(k)) = [\cdots, \mathbf{0}, \mathbf{0}, \mathbf{W}_0, \mathbf{0}, \mathbf{0}, \cdots]$. Recall that the harmonic transfer function H_T is a block diagonal matrix where the central element, which corresponds to the zero frequency element, is the stable transfer function H. Thus we have:

$$\begin{aligned}\mathbf{y}_{est}(k) &= \mathcal{F}^{-1}(H_T\mathcal{F}(\mathbf{u}(k))) + \mathcal{F}^{-1}(H_T\mathcal{F}(w(k))) \\ &= \mathcal{F}^{-1}(H_T\mathcal{F}(\mathbf{u}(k))) + \mathcal{F}^{-1}(HW_0) \\ &= \mathcal{F}^{-1}(H_T\mathcal{F}(\mathbf{u}(k))) + Hw(k)\end{aligned} \tag{18}$$

Note that the first term is the normal estimation and the second term is the consequence of attack signal. Thus the consequence of attack signal on the observation $\tilde{\mathbf{Y}}(k)$ is the same with the estimated observation vector $\mathbf{y}_{est}(k)$. In this way, we prove that the numerator in the residual calculated by CLD would not be increased due to our attack. Furthermore, the increased $I_{gd,q}$ ensures that the denominator would not get smaller. Therefore, the detection criterion $R_{I_{gd,q}}(k)$ would not increase during our attack.

As for the VLD, since the voltage control loop ensures that the U_{dc} is close to U_{dc}^*, as long as the attack vector change slowly enough, the amplitude of the transient process would be omittable and would not be detected by the VLD. The same reason also applies to the MLSTM-based detector. Since it relies on the abrupt change in the observation, gradually changing the attack vector ensures that the process is smooth and decent, thus eliminating the possibility of being detected by the MLSTM-based detector.

We want to highlight that the key to designing attack matrices lies in the linear combination of column vectors of the stable transfer function $Hw(k)$. Only in this condition can we connect the effect of adding attack vector $\mathbf{z}_{att}(k)$ to the observation and the effect of adding attack vector $\mathbf{u}_{att}(k)$ to the control input and prove that our DIA could bypass the CLD.

V. Simulation and Experiment Results

In this section, we demonstrate the stealthiness of the designed DIA and its impact on point of common coupling (PCC) voltages through HIL experiments. We conduct the HIL experiments in a real-time Typhoon HIL testbed [23]. As shown in Fig. 3, the IEEE 34-node distribution grid connecting a power electronic converter-enabled PV farm is built in HIL 602+ emulator, and the PC is connected to the emulator to achieve supervisory control and data acquisition.

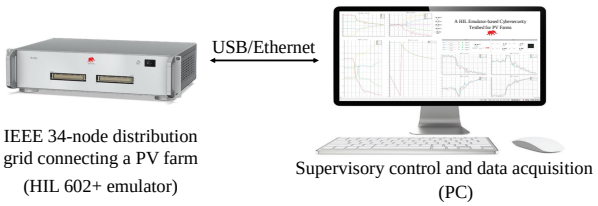


Fig. 3: Overview of the HIL-based grid-tied PV farm.

A. Datasets

To obtain the HSS-TF-based and MLSTM-based detectors, we design several types of normal cases by defining different environmental conditions. The irradiation on the PV panel varies in the range of 800, 900 and 1000w/m^2. The temperature varies within the range of 15, 25 and 35°C. Besides the conditions with constant irradiance and temperature, our experiments also consider gradual irradiance change and temperature change. For each scenario, 10s data are captured, and the sampling frequency is chosen as 1.3 kHz.

We conduct stealthy attacks under 900w/m^2 irradiation and 25°C. The attack lasts for 40s. In the first 20s, the attack weight vector changes from $[0.003, 0.0003, 0, 0]^T$ to $[0.024, 0.0024, 0, 0]^T$ with eaqual change every 0.25 seconds. For the last 20 seconds, the attack weight vector would change in the reverse direction and become $[0.003, 0.003, 0, 0]$ eventually.

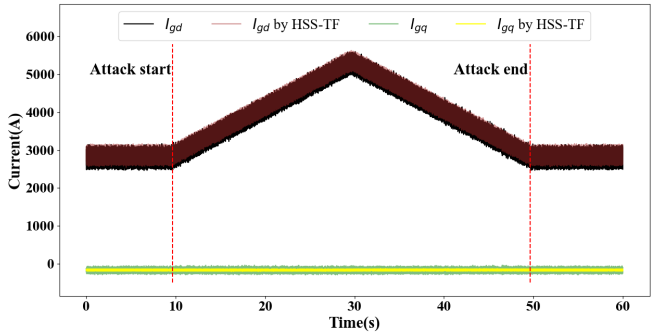
B. The Attack Stealthiness

Under the designed stealthy attacks, the detection results of the HSS-TF-based detector are shown in Fig. 4 and 5. Under the designed stealthy attack, the detection results of the HSS-TF-based detector are shown in Fig. 4 and 5. From Fig. 4(a), we can see that the detector's prediction I_{gd} by HSS-TF and I_{gq} by HSS-TF approximately equals to the real value during the attack period. In Fig. 4(b), note that the residuals of the I_{gd} and I_{gq} do not have obvious change from normal time and the residuals are both lower than the detection threshold. From Fig. 5(a), the oscillation of signal U_{dc} during attack does not have obvious change from normal time and would not be detected. The VLD detection residuals are always within the detection thresholds as shown in Fig. 4(b).

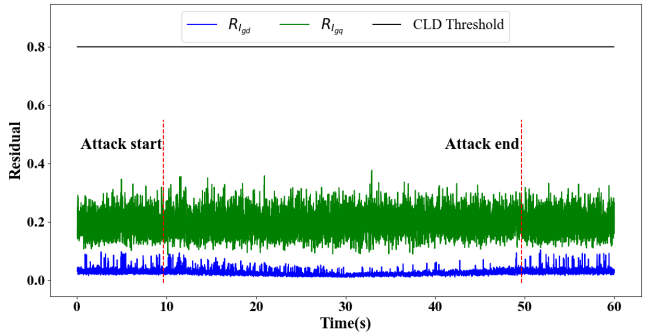
The detection results of the MLSTM-based detector under the stealthy attacks are shown in Table I. With different detection windows, the MLSTM-based detector all shows high precision but low recall. Precision is the fraction of correctly detected attacked instances among all the instances that labeled as attacked by the model, recall is the fraction of correctly detected attacked instances among all the attacked instances, and F1 score is a weighted average of precision and recall. The detection results indicate that the stealthy attacks are almost undetectable by the detector.

C. The Attack Impact Performance

We show the changes of the amplitude of PCC node voltage during attack in Fig. 6. Note that the attack leads to the

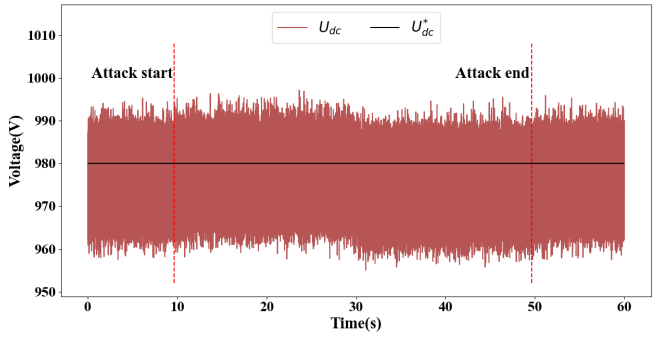


(a)

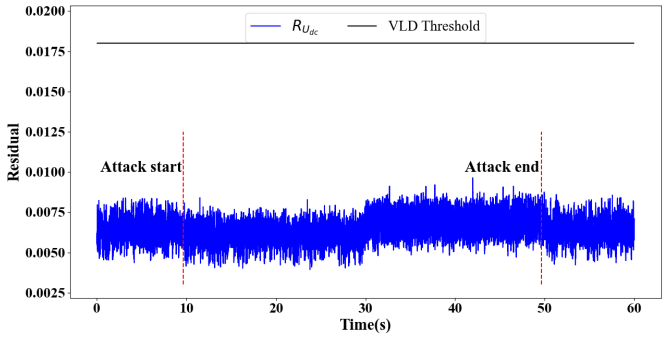


(b)

Fig. 4: (a) compromised measurement versus estimation of $i_{gd,q}$ under steathly attacks; (b) residual of CLD under stealthy attacks.



(a)



(b)

Fig. 5: (a) U_{dc} under stealthy attacks; (b) residual of VLD under steathly attacks.

Window Length	Recall	Precision	F1
50	0.41%	96.83%	0.0082
100	0.39%	95.77%	0.0078
200	0.44%	96.17%	0.0087

TABLE I: The detection results of MLSTM-based Detector.

increase of PCC nodes' voltages. Besides, at the maximum amplitude of the attack weight vector, that is about 30 second in the figure, the voltages of the PCC load would increase by approximately 1V that exceeds the normal oscillation range of PCC load voltages before attack.

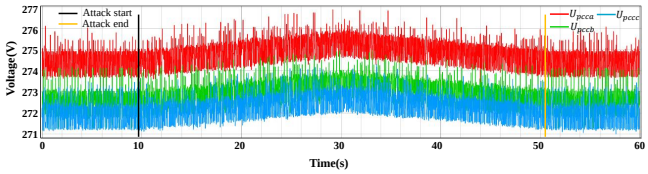


Fig. 6: The PCC voltage amplitudes under stealthy attack.

VI. Conclusion and Future Works

In this paper, we proposed a generation scheme of stealthy DIAs that can bypass model-based [18] and data-driven [13] IDSs simultaneously in the scenario of PV farms. The model-based IDS is invalidated by tampering with the sensor measurements conforming to the physical dynamics, while the data-driven IDS is bypassed by changing the waveform data with an unobservable speed. Through HIL experiments, it is demonstrated that the designed stealthy DIA can obviously deviate PCC voltages from normal values without being perceived by data-driven or model-based IDSs.

To defend against the potential stealthy DIA, one possible direction is to utilize the idea of proactive detection, which aims to enhance the detection capability of IDSs by adding uncertainties to the adversary. Once the adversary cannot obtain accurate model knowledge of the PV system, then the designed stealthy DIA using outdated information is likely to be uncovered by IDSs. In addition to this, the information from multiple levels (device- and grid-levels) and domains (host, network, and physical domains) can be further used to improve the detection capability.

References

[1] U.S. Department of Energy, "Cybersecurity Considerations for Distributed Energy Resources on the U.S. Electric Grid," Tech. Rep., 2022. [Online]. Available: https://www.energy.gov/sites/default/files/2022-10/Cybersecurity%20Considerations%20for%20Distributed%20Energy%20Resources%20on%20the%20U.S.%20Electric%20Grid.pdf.

[2] North American Electric Reliability Corporation, "2020 Long-Term Reliability Assessment," Tech. Rep., 2020. [Online]. Available: https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/NERC_LTRA_2020.pdf.

[3] Ben Kellison, "The next five years will see massive distributed energy resource growth," 2021. [Online]. Available: https://www.woodmac.com/news/editorial/der-growth-united-states/.

[4] "Ieee standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*, pp. 1–138, 2018.

[5] Sandi National Laboratories, "Cyber Security Primer for DER Vendors, Aggregators, and Grid Operators," Tech. Rep., 2017. [Online]. Available: https://www.osti.gov/servlets/purl/1761987.

[6] ——, "Roadmap for Photovoltaic Cyber Security," Tech. Rep., 2017. [Online]. Available: https://www.osti.gov/servlets/purl/1782667.

[7] ——, "Recommendations for Trust and Encryption in DER Interoperability Standards," Tech. Rep., 2019. [Online]. Available: https://www.osti.gov/biblio/1761841.

[8] J. Ye, A. Giani, A. Elasser, S. K. Mazumder, C. Farnell, H. A. Mantooth, T. Kim, J. Liu, B. Chen, G.-S. Seo *et al.*, "A review of cyber–physical security for photovoltaic systems," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 10, no. 4, pp. 4879–4901, 2021.

[9] T. S. Ustun, "Cybersecurity vulnerabilities of smart inverters and their impacts on power system operation," in *2019 International Conference on Power Electronics, Control and Automation (ICPECA)*. IEEE, 2019, pp. 1–4.

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

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

[12] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *computers & security*, vol. 28, no. 1-2, pp. 18–28, 2009.

[13] 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 Transactions on Smart Grid*, vol. 13, no. 3, pp. 2369–2380, 2022.

[14] F. Li, Q. Li, J. Zhang, J. Kou, J. Ye, W. Song, and H. A. Mantooth, "Detection and diagnosis of data integrity attacks in solar farms based on multilayer long short-term memory network," *IEEE Transactions on Power Electronics*, vol. 36, no. 3, pp. 2495–2498, 2020.

[15] 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 Transactions on Smart Grid*, vol. 13, no. 2, pp. 1582–1597, 2021.

[16] A. Mustafa, B. Poudel, A. Bidram, and H. Modares, "Detection and mitigation of data manipulation attacks in ac microgrids," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2588–2603, 2019.

[17] O. A. Beg, L. V. Nguyen, T. T. Johnson, and A. Davoudi, "Signal temporal logic-based attack detection in dc microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3585–3595, 2018.

[18] J. Zhang, L. Guo, and J. Ye, "Cyber-attack detection for photovoltaic farms based on power-electronics-enabled harmonic state space modeling," *IEEE Transactions on Smart Grid*, 2021.

[19] 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 Transactions on Smart Grid*, vol. 14, no. 1, pp. 714–731, 2023.

[20] 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 Transactions on Smart Grid*, vol. 13, no. 5, pp. 3984–3996, 2022.

[21] A. J. Gallo, M. S. Turan, F. Boem, T. Parisini, and G. Ferrari-Trecate, "A distributed cyber-attack detection scheme with application to dc microgrids," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3800–3815, 2020.

[22] S. Sahoo, S. Mishra, J. C.-H. Peng, and T. Dragičević, "A stealth cyber-attack detection strategy for dc microgrids," *IEEE Transactions on Power Electronics*, vol. 34, no. 8, pp. 8162–8174, 2018.

[23] 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 *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2021, pp. 1–2.