SAND2019-4555C

Hybrid Intrusion Detection System Design for Distributed Energy Resource Systems

A. Chavez, C. Lai, N. Jacobs, S. Hossain-McKenzie, C. B. Jones, J. Johnson, A. Summers
Sandia National Laboratories

***Abstract*—The integration of communication-enabled grid-support functions in distributed energy resources (DER) and other smart grid features will increase the U.S. power grid's exposure to cyber-physical attacks. Unwanted changes in DER system data and control signals can damage electrical infrastructure and lead to outages. To protect against these threats, intrusion detection systems (IDSs) can be deployed, but their implementation presents a unique set of challenges in industrial control systems (ICSs). New approaches need to be developed that not only sense cyber anomalies, but also detect undesired physical system behaviors. For DER systems, a combination of cybersecurity data and power system and control information should be collected by the IDS to provide insight into the nature of an anomalous event. This allows joint forensic analysis to be conducted to reveal any relationships between the observed cyber and physical events. In this paper, we propose a hybrid IDS approach that monitors and evaluates both physical and cyber network data in DER systems, and present a series of scenarios to demonstrate how our approach enables the cyber-physical IDS to achieve more robust identification and mitigation of malicious events on the DER system.**

***Keywords*–intrusion detection systems, distributed energy resources, cyber attacks, cyber-physical data**

I. Introduction

Inter-operable distributed energy resource (DER) grid-support functions enable high penetrations of renewable energy resources that would otherwise not be feasible. These commanded and configurable autonomous functions have been shown to:

- Improve voltage regulation on distribution circuits [1],
- expand distribution hosting capacity [2],
- provide wide-area damping [3],
- perform frequency control [4],
- and ancillary services [5].

Adding these inter-operable control functions has effectively added power generators to the Internet of Things (IoT), and raised a number of concerns regarding cyber-physical attacks. In theory, controlling an aggregation of DER devices could have grave impact on power system reliability, stability, and safety [6]. Fault detection models are able to flag malicious events that impact the grid, but are unable detect cyber actors that have gained unauthorized access to the network and are unable to detect cyber attacks early enough to thwart malicious actions. Therefore, more robust defense mechanisms are needed in the form of intelligent Intrusion Detection Systems (IDSs) [7]. Well-designed IDS systems provide the ability to detect malicious and abnormal events as quickly as possible, and to highlight relevant information to enable grid operators to respond appropriately to these events.

IDSs are responsible for detecting threats by monitoring one or more data streams. Intrusion prevention systems (IPSs) expand this capability by taking immediate action to contain the detected threat. The implementation of IDSs/IPSs for cyber-physical DER operations pose numerous challenges, since traditional IDSs focus only on cyber (i.e., network traffic) data, which includes communications between routers, switches, and endpoints [8]. These systems commonly use either signature-based or behavioral metrics to detect malicious network activities. Signature-based approaches monitor data and flag activity when known malware signatures are observed. In many cases, if a signature match is found, the IPS applies a predetermined rule–e.g., blocking traffic, quarantining data, etc. Some signature-based systems include the snort IDS [9], the Zeek (formerly Bro) IDS [10], and the Suricata IDS [11]. Behavioral approaches focus on recognizing or classifying anomalous patterns in network data compared to a baseline [12]. Behavioral-based [13] approaches can be trained on pre-existing data automatically or manually by an operator, and are often implemented using statistical machine learning algorithms.

While these IDS/IPS systems work reasonably well in IT environments, monitoring cyber data alone may not be enough to detect certain OT threats. There are major advantages to using a hybrid approach that incorporates both network and power system information. One advantage of the hybrid approach is that it is more challenging for an adversary to manipulate or spoof both cyber and physical data streams to execute an attack without detection.

To more effectively detect and respond to cyber-physical attacks in DER systems, we propose a hybrid IDS approach that integrates physical data with cyber network data, allowing us to not only capture the intricacies of power system models, but can also correlate physical measurements to specific events on the communications and control network.

II. Background

Current DER IDS research has explored a wide range of behavioral techniques for anomaly detection, with increasing interest in the detection of malicious attacks. For example,

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525; This material is based upon work supported by the U.S. Department of Energys Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Number 34234.

in [14], a combination of supervised, unsupervised, and ensemble algorithms are tested against several photovoltaic (PV) system attacks, as summarized below.

- **Disconnect attack**: An adversary gains control to a large number of PV inverters and issues a mass disconnect command; during heavy load conditions, this can cause line overloads, frequency/voltage violations, and system instabilities.
- **Power curtailment attack**: An adversary tampers the control algorithm parameters to reduce the allowable power output of the PV inverter. Detection of this type of attack can be difficult, as the system appears to be operating normally with partially reduced performance.
- **Volt-VAr attack**: An adversary manipulates inverter control to arbitrarily inject a different level of reactive power, affecting the voltage magnitude and phase angles in the grid.
- **Reverse power flow attack**: An adversary gains control of smart grid appliances and shuts them off or triggers circuit breakers to initiate a reduced demand response. This results in an increased reverse power flow under stressed conditions, which can cause line overloads and disrupt line voltage regulators that could lead to voltage collapse.

Although the algorithms perform reasonably well under a defined scenario, there is limited availability of real-world data for training the algorithms for deployment in an operational system. Some researchers have proposed the use of faster-than-real-time simulations of the power system to determine the impact of particular power system commands prior to executing the command [15]. However, the data collected from these simulations excludes network and host-based data that are commonly used to identify cyber attacks.

The predictability and regularity of communications and endpoint operations within energy delivery systems make IDSs especially effective. For example, Denial of Service (DoS) attacks using malformed packets, unauthorized reading of data, and unauthorized writing of data can be detected using the Snort IDS [16]. Additionally, system calls, the duration of software executing on an end device, the number bytes sent and received, and the processor/memory utilization are all useful features that can be extracted from end devices to detect anomalous behavior [17]. IDS mechanisms can perform well in these environments, operating strictly off of cyber data alone. However, spoofing physical data that signature-based and behavioral-based IDSs depend on can defeat such cyber-based detection mechanisms.

In DER systems, it is not sufficient to only detect cyber anomalies; it is critical to connect the detected cyber events with the effects in the physical power system. This is especially important for developing mitigation techniques and understanding the overall impact to the cyber-physical system. Furthermore, distinguishing malicious events from other sources of anomalies is particularly difficult.

Anomaly detection on the grid has largely focused on fault detection and location identification, which typically involves comparison between actual and predicted performance using power system models and physical sensor data [18]. While this approach is reasonably successful at producing warnings for imminent faults, it provides limited awareness of the underlying causes behind anomalies, which can result from systemic failures stemming from hardware or software, human error, or malicious intent. Consequently, little actionable insight is gained in terms of identifying the appropriate responses to ensure continued system availability.

III. Methodology

In this section, we will provide an overview of the hybrid IDS approach, impact of data availability, and representative, example attack scenarios.

A. Hybrid IDS Approach Overview

The hybrid IDS method performs anomaly detection analyses on both cyber and physical data to determine whether an attack had been conducted. Fig. 1 provides a general representation of this monitoring and analysis approach. The

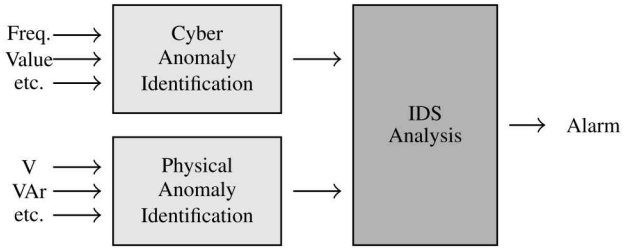

Fig. 1: The hybrid IDS methodology where parameters of interest are extracted from cyber and physical data streams to conduct the IDS analysis.

hybrid IDS captures data from DER communications and physical measurement equipment (e.g. advanced metering infrastructure, voltage/current sensors, phasor measurement units, DER devices), and extracts measurable features from the raw data (e.g. packet length, polling frequency, power factor, etc.). If malicious traffic is being sent from the DER device in a manner that is difficult to detect using network data, it can be detected via anomalies in the power data; if unexpected behavior occurs in the DER power system, the network data can be analyzed to determine whether the problem is caused by malicious commands on the network. However, a hybrid system must address additional complexities related to availability of data. Obtaining both network and physical data requires a higher throughput for the IDS system, and the physical data collected is often sparse, leading to challenges in real-time operation and response. Therefore, successful IDS operations are not only dependent on the accuracy and precision of the anomaly detection model, but also on meaningful data collection and feature extraction.

B. Cyber-Physical Data Features

In developing an IDS approach to detect intrusions or undesired behavior, one of the most important decisions to make is around what features to examine in the raw data.

There are significant differences in what a feature is and what it can tell you based on its source and the type of information contained. Here, we separate IDS features into

several categories, such as physical system features, network traffic features, and host based features. These can be further segregated as well, such as network traffic on a control network versus network traffic for a web-based management interface or portal.

While examining behavior for a hybrid IDS for DER and distribution control systems, the first set of features of interest are those around the physical performance and behavior of the system, which are collected from the power models and physical sensors operating in the system. These can include information such as:

- Physical System Features
  - ⋆ Current (AC/DC)
  - ⋆ Voltage (AC/DC) (sags, dips, spikes)
  - ⋆ Active, Apparent, & Reactive Power
  - ⋆ Frequency

As these variables are used in the various control schemes and algorithms applied to manage grid-support functions, being able to measure and determine whether the system is behaving as expected is critical for any hybrid IDS solution for DERs. Since the data is often temporally sparse, the IDS may incorporate ground truth data or power system models to provide expected values for the feature variables.

While physical features are sufficient for detecting faults, they are often insufficient for detecting nefarious activity. As control systems operate over a communications network, the network traffic may be analyzed to reveal behaviors outside the scope of the physical data. In the case of a malicious attack, altered data values and commands may produce a system response to impact grid performance without anomalies in the power system model. Moreover, the attacker may either spoof or block physical data from being passed to prevent normal fault detection mechanisms from being triggered. To effectively detect and respond against this type of threat, additional features that can be measured earlier in the attack cycle are required. This can include control network features, such as the following:

- Network Traffic Features (control signals)
  - ⋆ Frequency
  - ⋆ Setpoint values
  - ⋆ Destination & Source IP Address
  - ⋆ Destination and Source Port
  - ⋆ Sequence number
  - ⋆ TTL
  - ⋆ Checksum
  - ⋆ TCP flags
  - ⋆ Destination and Source MAC Address
  - ⋆ IP version
  - ⋆ Packet Length
  - ⋆ Throughput
  - ⋆ Latency

Grid-connected IoT devices often include a web portal for management and monitoring, and smart inverters and their management platforms may include such capabilities as well. These communications are usually segregated from the control network functionality. These portals create an additional attack surface that must be protected in the system. To detect malicious attacks in these devices, specific features should be parsed from the network data:

- Network Traffic Features (web management interface, optional)
  - ⋆ User-agent strings
  - ⋆ HTTP headers
  - ⋆ TCP headers
  - ⋆ Session ID
  - ⋆ Suspicious characters or data types
  - ⋆ Authentication logs

Not every indicator of system compromise can be discovered by examining network traffic alone. Advanced persistent threats on the system may operate in stealth, producing little or no network activity until triggered by an attacker. It is useful to examine differences in the endpoints of the system that are hosting the required control system functionality, in this case the smart inverters themselves. This can be done by measuring changes to elements of the inverter operating system or firmware, including files, network configuration, processing, and memory.

Additional memory usage or changes in timing patterns for processing may be indicative of additional processing being performed, potentially hidden from the user. Thus, the following features can be considered:

- Host-based features:
  - ⋆ File integrity
  - ⋆ Memory usage
  - ⋆ Processor usage
  - ⋆ Security logs

To achieve optimal detection of activities that may degrade DER security or performance, a variety of data sources must be used to collect relevant features. By applying a combination of signature-based and behavioral techniques to these features, the hybrid IDS can detect a wide range of scenarios covering the overall system attack surface. When training the behavioral models in the IDS, it is important to consider the priorities for detection and the quality of data that is available. Sensitivity analysis should be performed to identify the features that are best suited for detecting each type of anomaly. An IDS that incorporates response capabilities will need to prioritize features that indicate the earlier stages of an attack, such as network packets used for reconnaissance, and an IDS that seeks to prevent denial-of-service type attacks may want to prioritize relevant features such as throughput and memory usage. Moreover, these considerations need to be made with reliability in mind, as even the most precise model will fail to detect an event if the source data is not available over an extended period of time. For example, even if the detection model performs better on network data than on host data from a device, the host data may still have high importance if the network data is frequently unavailable.

The next section will delve into how cyber-physical attacks may manifest in various cases and the indicators that may arise in DER systems.

C. Attack Scenarios

To develop and demonstrate this hybrid IDS approach, it is crucial to study the interplay between the features we can

observe in the communication network and the related power system effects. The various indicators of compromise or types of cyber-physical attacks in DER systems also needs to be connected to the observable information that can be used for intrusion detection.

The previous subsection discussed the various types of information that may be useful for detection and behaviors of interest, in this section we will connect these two aspects of the problem of intrusion detection using several hypothetical scenarios that mimic known cases of OT focused cyber attacks and display diversity in the types of information required for detection.

1) Scenario 1: False Data Injection (Control Settings): Data injection is a cyber attack type where incorrect data or commands are injected into an application. In control system networks this is a large concern as many common control protocols, such as Modbus and IEEE 1815 (DNP3), were not originally designed to incorporate security features that could prevent such an attack. Since then, Modbus has a version that adds TCP Security [19] and DNP3 secure authentication was added to the standard in 2012 [20] to reduce the threat–through these are rarely used in practice. In this scenario, this malicious data is either issued through a replay attack, man-in-the-middle attack, or some other technique to change the DER setpoints or falsify data sent back upstream to an aggregator, grid operator, or DER vendor. A variant of this scenario is false data injection of measurement data used for monitoring the distribution system, as seen in [21].

2) Scenario 2: Insider Threat: Insider threats are very difficult to detect because the threat is a valid, authorized user. Therefore, cyber features such as source IP, port number, and packet frequency indicators are not useful. Physical features may be more capable of detecting insider threat if the attacks impact the power system by changing the DER operations or performance. That is, utilizing knowledge of the physical system to block or thwart unallowable control settings can be used even if it appears said commands are coming from a valid source.

IV. Attack Scenario and Discussion

A. Experiment Setup

To better understand the implementation of hybrid IDS systems in a cyber-physical environment, an experiment was conducted on a distribution system simulation that contained three interoperable PV inverters. The feeder represented a distribution system located in Albuquerque with 440% PV penetration and was simulated using an Opal-RT 5600. The three utility-scale PV systems located on this feeder were 258 kW, 1 MW, and 10 MW and modeled using the EPRI PV Simulator. The PV simulator has the ability to be interfaced to a real-time power system simulation and includes DNP3 communication interfaces that allow power measurements (AC power, reactive power, AC voltage, frequency, etc.) to be captured and the power factor (PF) to be configured on the devices. An Advanced Distribution Management System (ADMS) software developer, Connected Energy, issued power factor settings to the devices based on a volt-var (VV) profile represented by the points: V = 92, 99, 101, 108% of nominal voltage and Q = 25, 0, 0, -25% of reactive power capacity of the DER device. The experimental setup is shown in Fig. 2. When configured correctly (like in this case), the VV function used the reactive power capabilities of DER devices to drive the power system toward nominal voltage. A 40-minute simulation of the power system was run for this VV curve. Then the ADMS company acted as an insider threat and reversed the sign on the reactive power (Q = -25, 0, 0, 25) to drive the power system away from nominal voltage. In the reversed case, the DER injected or absorbed reactive power to force the grid voltage away from nominal. It is noteworthy that the impact to the power system from a man-in-the-middle attack would have been the same.

B. Results

The power factor values for the 10 MW PV system for the normal and attacked scenarios are shown in Fig. 3. As shown in the figure, the DER device absorbs reactive power (negative PF) when the VV curve was programmed correctly. This kept the voltage at the Point of Common Coupling (PCC) of the PV system close to nominal. In the attacked case, the DER injected reactive power (positive PF) and the voltage increased significantly on Bus 12.

During the experiments, the DNP3 traffic to the DER devices was captured and the power system current and voltage were measured at the buses. In this scenario, the insider could potentially spoof the communications data so the power system data was used to create a simple classification mechanism. The voltage and reactive power injection of the PCC bus are plotted in Fig. 4. A simple alarm mechanism was devised by bounding the normal volt-var curve. As shown for the “good” and “bad” DER operations, the physical data measurements could easily determine when operations were abnormal and an alarm could be raised.

However, in the case where there was no voltage or current measurements at the PV PCC, PV voltage reads and PF writes from the ADMS could have been extracted from the DNP3 data. Since the EPRI inverter is configured per the DNP Application Note AN2013-001 information model, this data is relatively easily captured from the network traffic. Hybrid scenarios also exist. If only PF was issued and voltage was not measured from the DER, the voltage measurements from the power system would be necessary to conduct the alarm classification. It is more common that VV curves are programmed into DER devices, so if the equipment was programmed with the inverted “bad” VV curve, PF commands would not have been issued or captured. In that case, voltage and reactive power/power factors would need to have been collected from the DER or the power meter.

C. Comparison of Approaches and the Importance of Data Availability

As shown in Table I, depending on what data is available from the power system or from the DER network traffic, either the physical, cyber, or cyber-physical features are required to detect the malicious actions. For instance, in Case 2 only the cyber data is necessary to detect misprogramming of the DER device. Similarly for Cases 3-5 with the physical data.

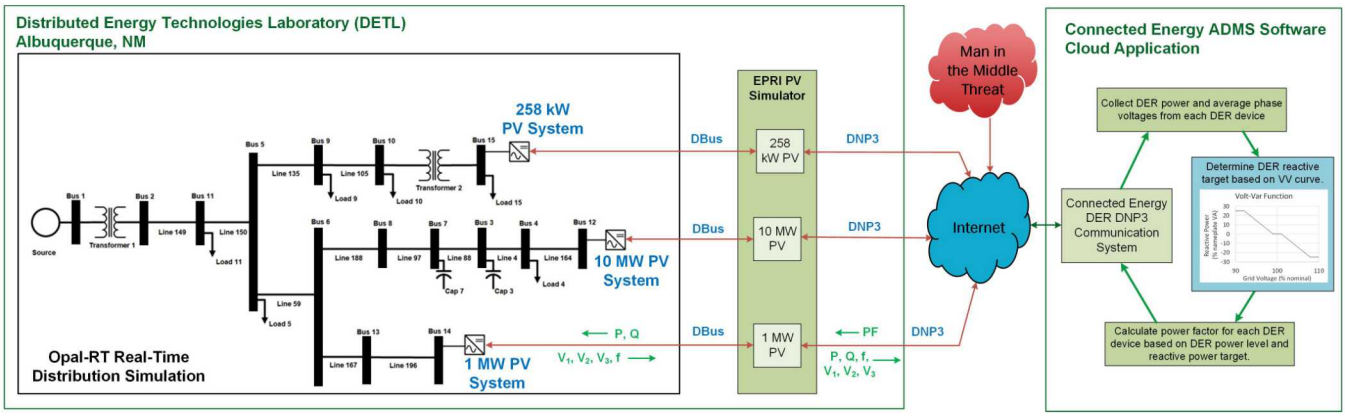

Fig. 2: Real-time power simulation connected to interoperable PV inverter simulators being issued PF commands.

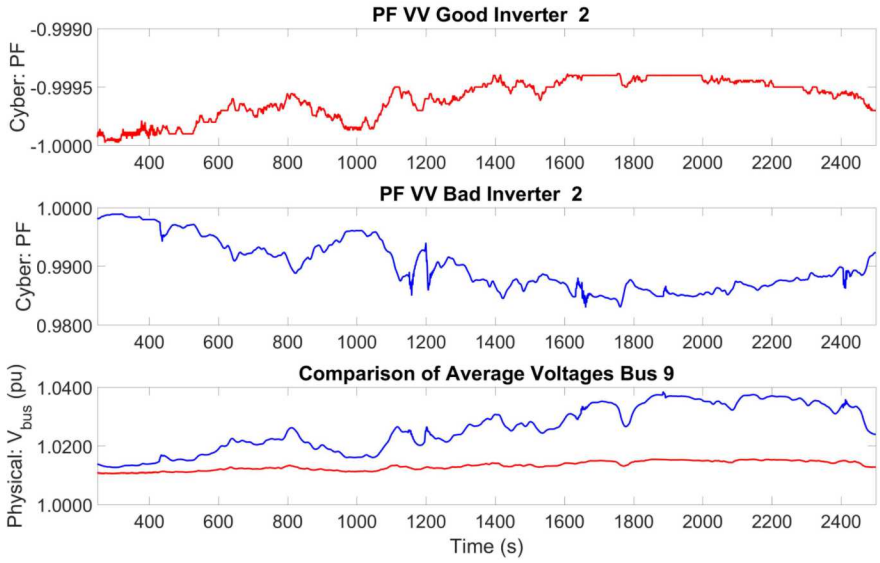

Fig. 3: Power factor values with normal and malicious commands and the associated voltage on Bus 9.

However, in Case 6, cyber and physical data are necessary to detect the cyber attack. It is worth noting that, collecting duplicative data (e.g., voltage measurements from the power system and also from the DER devices) is useful because this corroborates the integrity of the DER communications and power measurements. If there are discrepancies in multiplicative measurements, this would be a clear indication of a cyberattack or sensor faults. Collecting cyber-physical measurements provides higher confidence by marrying data streams. To fool an hybrid IDS would require extra steps to spoof or disable various data streams and collection mechanisms, which increases the difficulty of compromising the DER devices while remaining undetected. It is also worth mentioning, there comes a point where there is not enough data to conduct these assessments with cyber-physical data, such as in Case 7, in which there is no data about the DER PF or reactive power to determine if the device is behaving as intended.

V. Conclusions

As DER systems increase in penetration on the power system, the risk to the power system from malicious control of DER devices also increases. This necessitates new sophisticated defense mechanisms that incorporate cyber-physical data sets. A hybrid IDS approach is described in this paper that is suitable for DER systems which monitors and alerts using cyber and physical features. The need for this joint analysis is exemplified by an insider threat scenario; the results demonstrated the effectiveness of the hybrid approach and its broader applicability with limited data. It is recommended that the hybrid IDS approach is further developed using machine learning and IDS algorithms and its deployment and testing an a high-fidelity emulation environment.

References

[1] J. Seuss, M. Reno, R. Broderick, and R. Harley, "Evaluation of reactive power control capabilities of residential pv in an unbalanced distribution feeder," in *IEEE PVSC*, June 2014, pp. 2094–2099.

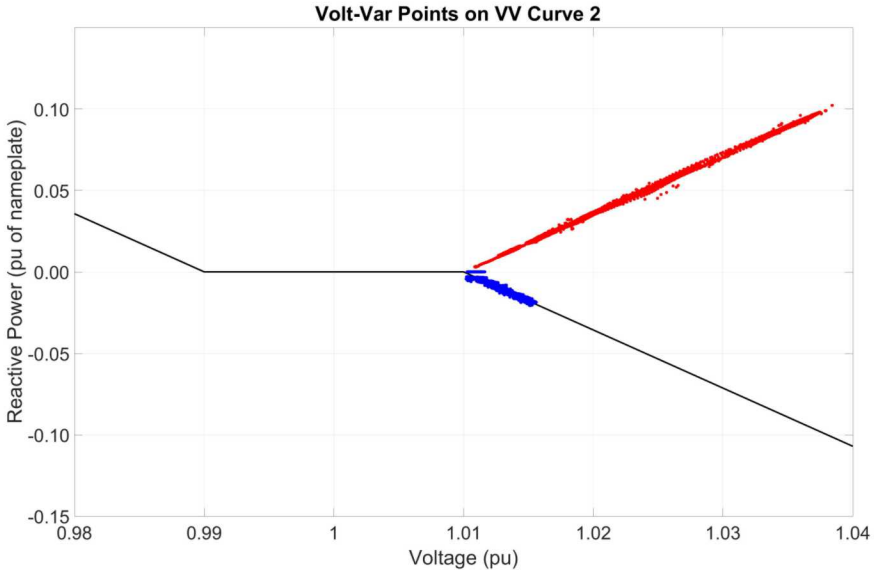

Fig. 4: Deviation from expected Volt-VAr curve due to insider threat.

	Physical Data				Cyber Data			
Case	Current Phasor	Voltage Phasor	Reactive Power	Detect	PF Write	V Read	Detect	Cyber & Physical Detect
1	✓	✓	✓	✓	✓	✓	✓	✓
2					✓	✓	✓	✓
3	✓	✓	✓	✓				✓
4	✓	✓		✓	✓			✓
5	✓	✓		✓				✓
6			✓			✓		✓
7		✓				✓		

TABLE I: IDS attack detection with different non-falsified physical, cyber, and cyber-physical data sources.

[2] M. Rylander and J. Smith, "Stochastic analysis to determine feeder hosting capacity for distributed solar PV," EPRI Report 1026640, Tech. Rep., 12 2012.

[3] J. Neely, J. Johnson, R. Bryne, and R. T. Elliott, "Structured optimization for parameter selection of frequency-watt grid support functions for wide-area damping," *DER Journal*, vol. 11, no. 1, pp. 69–94, 2015.

[4] J. Neely, S. Gonzalez, J. Delhotal, J. Johnson, and M. Lave, "Evaluation of pv frequency-watt function for fast frequency reserves," in *IEEE Applied Power Electronics Conference (APEC), Long Beach, CA*, March 2016.

[5] J. Johnson, J. Neely, J. Delhotal, and M. Lave, "Photovoltaic frequency-watt curve design for frequency regulation and fast contingency reserves," *IEEE Journal of Photovoltaics*, vol. 6, no. 6, pp. 1611–1618, 2016.

[6] J. Johnson, J. Quiroz, R. Concepcion, F. W. Bernal, and M. Reno, "Power system effects and mitigation recommendations for der cyber attacks," *IET Cyber-Physical Systems: Theory & Applications*, 2019.

[7] J. Johnson, "Roadmap for photovoltaic cyber security," Sandia Technical Report, SAND2017-13262, Tech. Rep., 12 2017.

[8] B. Zhu and S. Sastry, "Scada-specific intrusion detection/prevention systems: a survey and taxonomy," in *Proceedings of the 1st workshop on secure control systems (SCS)*, vol. 11, 2010, p. 7.

[9] C. Valli, "Scada forensics with snort ids," 2009.

[10] H. Lin, A. Slagell, C. Di Martino, Z. Kalbarczyk, and R. K. Iyer, "Adapting bro into scada: building a specification-based intrusion detection system for the dnp3 protocol," in *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*. ACM, 2013, p. 5.

[11] R. Bray, D. Cid, and A. Hay, *OSSEC host-based intrusion detection guide*. Syngress, 2008.

[12] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[13] B. Zhu and S. Sastry, "Scada-specific intrusion detection/prevention systems: a survey and taxonomy," in *Proceedings of the 1st workshop on secure control systems (SCS)*, vol. 11, 2010, p. 7.

[14] D. M. Shila, K. G. Lore, T. Weiz, T. Lovetty, and Y. Cheng, "Catching anomalous distributed photovoltaics: An edge-based multi-modal anomaly detection," *CoRR*, vol. abs/1709.08830, 2017. [Online]. Available: http://arxiv.org/abs/1709.08830

[15] S. Meliopoulos, G. Cokkinides, R. Fan, L. Sun, and B. Cui, "Command authentication via faster than real time simulation," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2016, pp. 1–5.

[16] S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, and A. Valdes, "Using model-based intrusion detection for scada networks," in *Proceedings of the SCADA security scientific symposium*, vol. 46. Citeseer, 2007, pp. 1–12.

[17] A. R. Chavez, J. Hamlet, and W. Stout, "Artificial diversity and defense security (addsec) final report." Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), Tech. Rep., 2018.

[18] M. D. Benedetti, F. Leonardi, F. Messina, C. Santoro, and A. Vasilakos, "Anomaly detection and predictive maintenance for photovoltaic systems," *Neurocomputing*, vol. 310, pp. 59 – 68, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231218305484

[19] Modbus Organization, Inc., "MODBUS/TCP Security: Protocol Specification," MB-TCP-Security-v21_2018-07-24, Tech. Rep., 07 2018.

[20] IEEE Power and Energy Society, "IEEE Std. 1815: IEEE Standard for Electric Power Systems Communications - Distributed Network Protocol (DNP3)," Tech. Rep., 10 2012.

[21] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *2012 IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 3153–3158.