IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 April 2026, accepted 26 May 2026, date of publication 1 June 2026, date of current version 4 June 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3699153

RESEARCH ARTICLE

Cyber–Physical Power System Dataset for Cyber Security of Digital Substations

NICOLA CIBIN[1], (Graduate Student Member, IEEE), NADINE KABBARA[2], ALFAN PRESEKAL[1], (Member, IEEE), IOANNIS SEMERTZIS[1], (Graduate Student Member, IEEE), VETRIVEL SUBRAMANIAM RAJKUMAR[1], (Graduate Student Member, IEEE), HIMANSHU GOYEL[1], (Graduate Student Member, IEEE), PETER PALENSKY[1], (Senior Member, IEEE), AND ALEXANDRU ŞTEFANOV[1], (Member, IEEE)

[1]Department of Electrical Sustainable Energy (ESE), Delft University of Technology, 2628 CD Delft, The Netherlands
[2]Utrecht University, 3584 CS Utrecht, The Netherlands

Corresponding author: Alexandru Ştefanov (a.i.stefanov@tudelft.nl)

This work was supported in part by the EU Horizon 2020 Marie Skłodowska Curie InnoCyPES Project under Grant 956433, in part by the EU Horizon Europe COCOON Project under Grant 101120221, in part by the EU Horizon Europe ESTELAR Project under Grant 101192574, and in part by the Dutch Research Council's RESCUE Project under Grant NWO ESI.2019.006.

ABSTRACT Cyber attacks targeting Intelligent Electronic Devices (IEDs) in digital substations can disrupt power system operation, causing equipment damage, instability, cascading failures, and even a blackout. Cyber–Physical Power System (CPPS) datasets are critically needed to develop novel methods for the detection and prevention of cyber attacks on digital substations. In this paper, a novel CPPS dataset is proposed for cyber security of digital substations, including real-time power system measurements, i.e., electromagnetic transient three-phase voltages and currents, communication network traffic, and virtual IED resource metrics. Various scenarios are simulated on an IEC 61850-compliant testbed consisting of Real-Time Digital Simulator (RTDS) and physical and virtual IEDs in hardware-in-the-loop configuration. The dataset contains different operating conditions and cyber attack scenarios, i.e., normal operation, single-phase-to-ground fault, network reconnaissance, resource exhaustion, and IEC 61850 Generic Object-Oriented Substation Event (GOOSE) and Sampled Values (SV) injection attacks. This work aims to provide the research community with a comprehensive and high-fidelity dataset to be used for the design and testing of novel methodologies to increase the cyber security of power grids.

INDEX TERMS Cyber–physical power system, cyber attacks, cyber security, dataset, digital substations, hardware-in-the-loop, IEC 61850.

I. INTRODUCTION

Electrical power systems are the backbone of modern society, as they supply electricity at national and continental levels for domestic consumption, economic activities, and critical industrial applications. As power systems are complex and highly interconnected, advanced digitalization is needed to accelerate the energy transition. The physical power systems are dependent on Operational Technology (OT) communication networks for real-time monitoring, protection, and control of physical facilities, e.g., power plants, transmission and distribution lines, and substations. Additionally, Information Technology (IT) solutions are increasingly applied to the energy utilities, interacting with their OT counterparts. The advanced integration of the physical power grid with IT-OT communication networks forms a complex and interdependent Cyber–Physical Power System (CPPS). However, such cyber–physical system interconnections require careful considerations regarding infrastructure cyber security. Indeed, with the increasing digitalization, cyber actors could target the IT-OT communication networks with a direct impact on power system operation. Such cyber attacks have

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Yuan Chen.

already occurred worldwide, with the most notable being BlackEnergy3 [1], CrashOverride [2], and IndustroyerV2 [3] in Ukraine in 2015, 2016, and 2022, respectively. Thus, cyber security and resilience of CPPSs are emerging issues, endangering the stability and secure operation of power systems.

One of the core components of the power grid is the electrical substation. The substations are essential for transforming voltage levels, reconfiguring the power grid topology, protecting the power grid from faults, and managing the power flows. The evolution from conventional to digital substations represents a significant shift in how power grids are designed, operated, and maintained, driven by advancements in digital technologies, automation, and data communication. In fact, analog equipment and communications over copper wiring are being replaced by digital devices and communications protocols like the ones defined in the IEC 61850 standard [4]. This transformation enables fast real-time data exchange, improved automation, and remote monitoring and control, resulting in higher efficiency, reliability, and security of electricity supply. However, this shift also paves the way for novel cyber threats due to the increased integration between OT and IT communication infrastructures [5]. Also, while the IEC 61850 standard is essential for the digitalization of power grid operations, it significantly falls short in incorporating basic network security measures such as message authentication and encryption [6]. Given the direct physical impact that a successful cyber attack on digital substations can have on power system stability, it is of utmost importance to design and develop novel methodologies for the detection and prevention of cyber attacks.

To conduct research in the field of digital substation cyber security, and in broader terms CPPS cyber security, it is crucial to have available comprehensive datasets including data acquired from both the cyber and physical system layers. However, due to the criticality of power grids and the required confidentiality of their architectures and operation, real-world data is often inaccessible to third-party entities and researchers. Thus, CPPS datasets are mostly generated in tabletop simulation environments, with strong limitations in terms of setup capabilities and scope of study. Furthermore, CPPS data can be generated through Hardware-in-the-Loop (HIL) experiments, involving industrial-grade equipment and communication networking devices, or with the deployment of simulated or emulated environments. The aforementioned approaches can also influence the timing of the data collection both in terms of time resolution and data synchronization. When the physical power system measurements are collected asynchronously to the communication network traffic, the quality of the acquired dataset degrades and its intrinsic usability for the design and evaluation of novel methodologies becomes limited. Finally, considering cyber security studies, the number of cyber attack scenarios considered and their impact on the physical systems are limited to the ones

TABLE 1. Comparison between existing datasets available in the literature and the one proposed in this work.

Dataset	Use case	Infrastructure	Physical	Cyber
Radoglou-Grammatikis et al. [10]	General ICS	Simulated industrial devices deployed on a physical communication network	Not available	DNP3
Matoušek et al. [11]	General ICS	Both virtual and real ICS	Not available	IEC 104, IEC 61850 MMS
Sahu et al. [14]	Transmission System	Emulated communication network and IEDs	Not available	DNP3
Huang et al. [15]	Transmission System	Emulated communication network and IEDs	Not available	DNP3
Zemanek et al. [16]	Substation	HIL configuration with physical IEDs, physical communication network	Not available	IEC 61850 GOOSE
Eynawi et al. [17]	Substation	Physical merging unit, IED, and Remote Terminal Unit (RTU). Current and voltage measurements generated with Omicron relay	Constant current and voltage waveforms	IEC 61850 GOOSE
Biswas et al. [18]	Substation	Emulated communication network and IEDs	RMS offline simulation	IEC 61850 GOOSE
Boakye-Boateng et al. [12]	Substation	Emulated communication network and IEDs	Not available	Modbus
Mlot et al. [13]	Substation	Physical IEDs and communication network	Not available	IEC 104, IEC 61850 GOOSE and SV, PTP, NTP
Quincozes et al. [19]	Substation	Synthetic network traffic	EMT simulation	IEC 61850 GOOSE and SV
This work	**Substation**	**HIL configuration with physical and virtualized IEDs, physical communication network**	**EMT simulation**	**IEC 61850 GOOSE and SV**

that can be comprehensively captured by the experimental setup.

As indicated above, due to the complexity and size of the CPPS, many different specialized datasets can be generated, depending on the scope of the study. In this work, the focus is on cyber attacks on digital substations and their physical impact on power system stability. The two main reasons why this work focuses on this research area are: (1) digital substations are fundamental operational nodes of the overall power system, as they control the power grid voltage levels, power flows, and topology; (2) digital substations are important targets for cyber attackers, as showcased in already occurred, real-world cyber attacks on the power grid in Ukraine in 2015, 2016 and 2022 [1], [2], [3].

A. RELATED DATASETS AND CONTRIBUTIONS

In the literature, the majority of datasets focus on general IT cyber security [7], [8], [9], whereas only a few tackle the cyber security issues of Industrial Control Systems (ICS) and OT for electrical power grids. Moreover, most of these datasets are limited to the acquisition of communication network traffic and lack data from the physical system [10], [11], [12], [13], even when the physical layer was simulated during the experiments [14], [15], [16], [17]. Up to date, the most complete cyber–physical dataset for digital substation cyber security is presented in Biswas et al. [18]. However, this study is limited to the IEC 61850 Generic Object-Oriented Substation Event (GOOSE) protocol, while the Sampled Values (SV) protocol is not considered. The deployed Intelligent Electronic Devices (IEDs) are simulated in software, and real hardware IEDs are not included. The physical system data is generated offline through Root Mean Square (RMS) simulations, thus limiting the analysis to electromechanical transients and ignoring the Electromagnetic Transients (EMT). In Quincozes et al. [19], a framework for the generation of an IEC 61850 intrusion detection dataset is presented. Even though the substation's physical system was simulated in real-time through an EMT simulation in PSCAD, no real communication network was deployed. In fact, the communication network traffic was synthetically generated by the proposed framework.

A comparison between the relevant datasets available in the literature and the one proposed in this study is provided in Table 1. Based on this comparison, and to the best of authors' knowledge, the dataset proposed in this paper is the first publicly available IEC 61850 digital substation dataset that combines: (i) EMT simulations of the power system, (ii) a HIL configuration including both IEC 61850-compliant physical and virtualized IEDs, and (iii) cyber and physical measurements collected during the actual execution of multiple representative cyber attacks types. This combination is not present in previous work, which lacks EMT-level dynamics and simultaneous deployment of physical and virtualized IEDs, or which do not capture real communication traffic from a deployed communication network.

The proposed CPPS dataset is designed to support research problems that cannot be adequately addressed with existing datasets, such as: (i) developing and benchmarking intrusion detection and prevention systems that jointly exploit cyber traffic and physical measurements including fast transients and protection actions; (ii) evaluating the cyber–physical impact of different attack classes on protection schemes and system stability. By providing synchronized cyber and physical data at EMT resolution within an IEC 61850 HIL setup, the dataset offers a level of fidelity that existing RMS-based or synthetic-traffic datasets cannot replace.

The remainder of the paper is organized as follows. In Section II, digital substation architecture and devices are introduced; moreover, common cyber attacks exploiting the GOOSE and SV protocols and their impact on the power system are briefly introduced. In Section III, the testbed used to acquire the dataset, simulated scenarios, and steps required to pre-process the data are described. In Sections IV and V, the acquired dataset is presented and then analyzed to provide technical information and insight valuable for further studies. Finally, Section VI concludes the paper.

II. DIGITAL SUBSTATION: ARCHITECTURE, CYBER ATTACKS, AND PHYSICAL IMPACT

Substation digitalization requires robust communication standards and protocols to ensure seamless interoperability between substation network devices. Commonly used protocols for substation digitization include MODBUS,

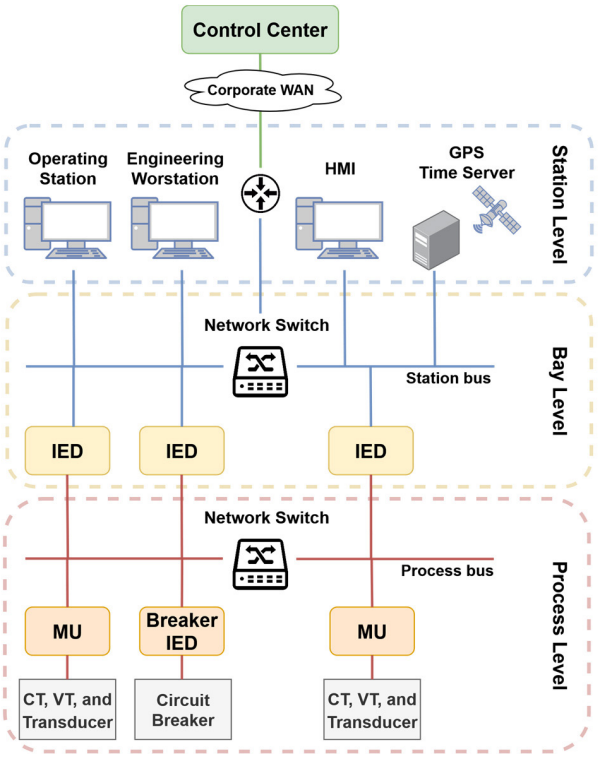

FIGURE 1. High-level representation of a digital substation communication network and deployed devices.

IEC 61850, IEC 60870-5-104, and DNP3. In recent times, IEC 61850 has become the preferred standard due to its comprehensive nature and ability to overcome most of the constraints and challenges encountered by other similar standards. Implementing this standard enables fast and reliable communication between IEDs in substations, enhancing the speed and accuracy of protection and control systems. As illustrated in Figure 1, IEC 61850 divides the substation into three hierarchical levels: the process, bay, and station levels. Communication within these levels is facilitated by the process bus and station bus. The digital process bus reduces the secondary wiring complexity for Current Transformers (CTs) and Voltage Transformers (VTs). It also streamlines the integration of substation automation features such as power quality monitoring and disturbance recording. IEDs and Merging Units (MUs) are the key operational devices driving the digitalization of substations. IEDs can interface with circuit breakers and can be configured for various protection functions. At the bay level, IEDs are typically configured with multiple protection functionalities, such as overcurrent, differential, and distance protection among others. Meanwhile, process-level IEDs directly interface with circuit breakers to perform control operations such as circuit breaker tripping, opening, or closing. Moreover, to further increase power system scalability, flexibility, and resilience, and reduce deployment and maintenance costs, the industry is moving towards the concept of virtualized Protection, Automation, and Control (vPAC). This concept consists of implementing centralized protection and control functions in virtualized environments deployed within virtualized IEDs (vIED) [20].

IEC 61850 defines the communication protocols for interactions between devices deployed in electrical substations. Key protocols include the Manufacturing Message Specification (MMS), GOOSE, and SV. The GOOSE protocol is used to transmit circuit breaker status and control commands, whereas the SV protocol publishes physical power system measurements, such as currents and voltages acquired from instrument Current and Voltage Transformers (CTs and VTs). However, due to the high frequency and stringent real-time processing requirements, message authentication and encryption are not implemented in GOOSE and SV [6]. Indeed, these signals must operate within 3 milliseconds, leading to latency being prioritized over cyber security.

The lack of basic security measures makes these protocols vulnerable to cyber attacks, as demonstrated in various previous experimental studies [6], [21]. Among these vulnerabilities, GOOSE messages, which control the tripping of circuit breakers, are particularly exposed to spoofing attacks. In such an attack, an attacker can pretend to be a legitimate sender and manipulate message fields to send false trip commands to circuit breaker relays, potentially causing malicious circuit breaker tripping [22]. Similarly, SV is vulnerable to cyber attacks such as denial of service, masquerading, and replay attacks. These attacks exhaust the IED resources by flooding the network with a large number of SV frames [23], [24] or exploit the protocol fields such as Application Protocol Data Units (APDU) [6] or sample count (smpCnt) [25] to report false measurements to IEDs. Power system protection schemes, such as distance protection, rely on voltage and current measurements from CTs and VTs. Cyber attacks that alter SV-reported measurements can compromise IEDs' functionality by blocking fault detection and delaying protection activation. On the other side, GOOSE spoofing attacks in a digital substation can cause malicious tripping of circuit breakers. Such delayed operation or malicious tripping can lead to the operation of backup protection zones and protection schemes in other substations, power system oscillations, generator desynchronization, system instability, and power system cascading failures. These disruptions may culminate in large-scale blackouts, further highlighting the criticality of vulnerabilities in protection mechanisms to cyber threats [22].

III. EXPERIMENTAL SETUP

The CPPS dataset is acquired with the deployment of a state-of-the-art testbed simulating in real-time power system measurements, i.e., currents and voltages, using the IEEE 5-bus test system, and emulating the communication network traffic of a digital substation. Given the reliance on Real-Time Digital Simulator (RTDS) to run the EMT simulation with a HIL configuration, and the deployment of real industrial-grade hardware, i.e., physical IEDs, the generated dataset provides the utmost levels of fidelity to the operation of cyber–physical systems. Moreover, the inclusion of virtualized IEDs [20] deployed in a general-purpose Virtual Machine (VM) allows for the analysis and comparison of the performance and resiliency of novel digital substation architectures. In the next subsections, the testbed, simulated scenarios, and dataset acquisition and processing phases are described in detail.

A. TESTBED

The dataset is generated by deploying a testbed consisting of RTDS [26], one physical IED implementing distance protection, and one virtualized IED providing instantaneous

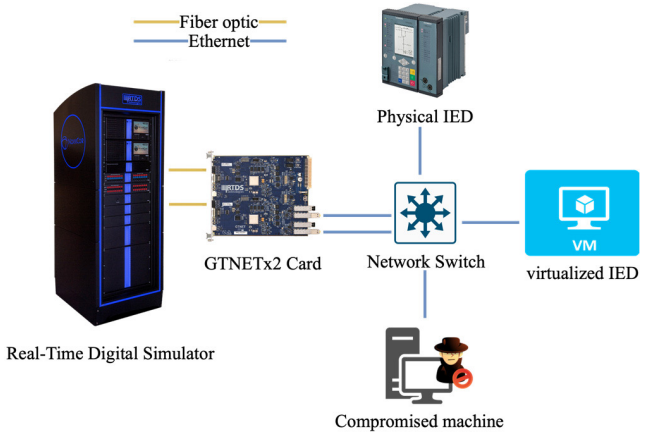

FIGURE 2. Testbed architecture.

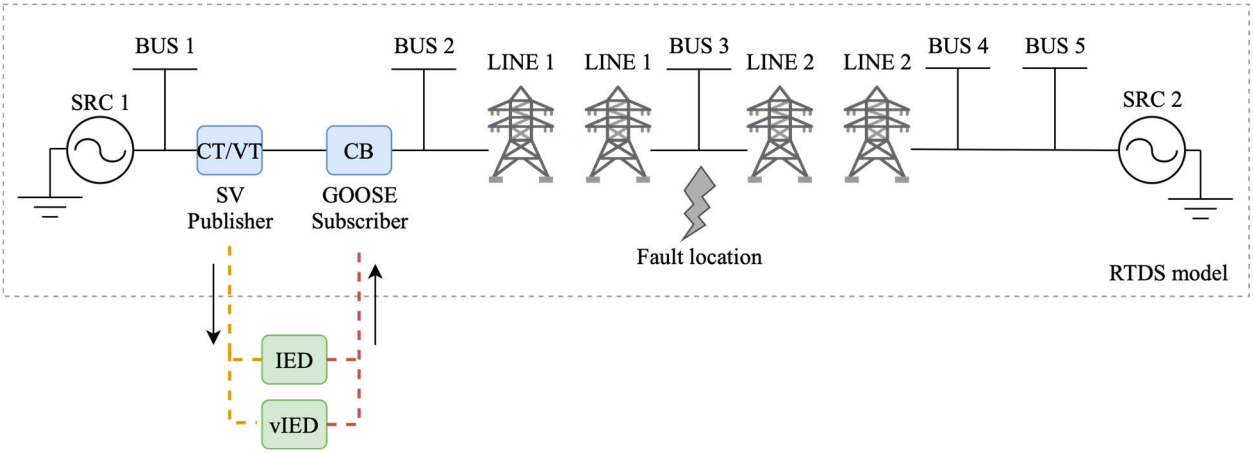

FIGURE 3. Power grid model simulated in RTDS and logic connection with physical and virtual IEDs 0.

overcurrent protection. The virtualized IED relies on an open-source C library [27] and is deployed on a VM equipped with a 4-core CPU @ 3.60 GHz, 16GB of RAM, and running Ubuntu 22.04 operating system. The two relays are coordinated such that the overcurrent element clears faults first, with the distance relay providing a time-delayed backup, achieving selectivity across the hybrid physical/virtual setup. The testbed architecture is represented in Figure 2. The physical and virtual IEDs are directly connected to a network switch. The RTDS communication network traffic is forwarded to the network switch through a GTNETx2 card [28]. In this IEC 61850 compliant configuration, RTDS is acting as IEC 61850 GOOSE subscriber and SV publisher. RTDS provides the 3-phase current and voltage measurements relative to one bus in the digital substation using IEC 61850 SV with a rate of 4800 samples per second. The two IEDs are subscribed to the SV stream published by RTDS, and as a response, they publish GOOSE messages to operate a circuit breaker simulated in the IEEE 5-bus test system in RTDS.

As previously mentioned, an EMT simulation is run in RTDS to simulate the IEEE 5-bus system's operation in real-time, with a time resolution of 50 microseconds. Contrary to RMS simulations which assume a dynamic operation and cannot model the transient behaviors, EMT simulations allow the capture of the instantaneous values of voltage and current measurements, making them ideal for the analysis of fast transient phenomena such as fault conditions and switching operations. The simulated power system one-line diagram created in RSCAD and loaded into RTDS is represented in Figure 3.

An additional workstation is connected to the network switch to conduct cyber attacks. In a real cyber attack scenario, the compromised device would correspond to a device, workstation, or server present in the digital substation, which was previously compromised by the intruder, and is being used to perform further actions in the digital substation communication network. The compromised device is used to launch the communication network reconnaissance, and GOOSE and SV injection attacks.

B. SIMULATED SCENARIOS

Different operating conditions and cyber attack scenarios are simulated to provide a comprehensive dataset containing various power systems dynamics, cyber kill chain steps, and intruder behaviours. Although these scenarios are implemented as isolated, single-stage events, they are designed to represent the fundamental building blocks of real-world attacks and can be combined to reconstruct more sophisticated, multi-stage cyber campaigns. In the following, each of the simulated scenarios is described in detail.

1) BASELINE

In the baseline scenario, a steady-state operating condition is simulated. The power grid operates without any changes in its loading, and as a result the instantaneous measurements obtained for currents and voltages have constant magnitudes and phase angles. The circuit breakers remain closed, while the protection IEDs do not operate.

2) FAULT

For the physical fault scenario, a single-phase-to-ground short-circuit is simulated. The type of simulated faults is limited to single-phase-to-ground short circuits as they represent the most common disturbance in power systems and are particularly relevant for studying protection schemes and the impact of cyber attacks. After the fault occurs, the short-circuit causes the magnitudes of the AC currents to increase while the voltage magnitudes sag. This change is detected by the protection equipment, namely the overcurrent and distance relays. In the simulated testbed, after the fault is detected, the tripping command from the IEDs is sent to the circuit breaker, opening all three phases. In the case without the reclosure mechanism, the circuit breaker will remain open. In the case with a reclosure mechanism implemented, at a predefined time after the first opening, the circuit breakers will automatically close. This option is utilized for short duration faults, in which the fault is cleared after a few

seconds. Thus, by reclosing the breakers, the affected line can be energized again. In case the fault is still occurring, the circuit breaker will open again, disconnecting permanently the affected power line.

3) RESOURCE EXHAUSTION

During the resource exhaustion attack, it is assumed that the attacker was able to compromise either the host machine, virtual machine, or container hosting the vIED. Then, this initial access is exploited by the intruder to exhaust the computational and memory resources of the vIED. To simulate this attack, the 'stress' [30], 'nice' [31], and 'renice' [32] Linux commands are used to exhaust the VM computing resources and to change the priority of the 'stress' and vIED processes. To allow the assessment of the attack impact under different conditions, three different priority levels are assigned to the 'stress' command process during the tests, i.e., 0, -15, and -19, with the smaller values referring to the highest priority levels. Furthermore, in the last case, the physical IED is completely disabled, and only the vIED is allowed to act on the occurring fault.

4) NETWORK RECONNAISSANCE

In this scenario, an active information-gathering process is simulated. The attacker exploits the already compromised machine to send probes to other devices in the network to gather information about open ports, exposed services, and devices and network configurations. This process is performed with the use of Nmap [33], a powerful open-source tool used for communication network discovery and cyber security auditing. To simulate different stealthiness levels, the communication network reconnaissance process is performed by setting Nmap to operate on four timing templates, i.e., 'polite', 'normal', 'aggressive', and 'insane' [34]. These timing templates govern how quickly Nmap performs a scan, handles responses, and decides when to resend packets. The following command is used to scan the physical and virtual IEDs simultaneously:

$$\text{nmap} - p1 - 1010, 4443, 8080, 55147 - \text{sV} \\ - T\{N\}\ 192.168.0.10, 12$$

where the option '-p' identifies the network ports to be scanned for each host, '-sV' enables the identification of the

TABLE 2. Description of the simulated scenarios.

Scenario	Sub-scenario (# of runs)	Description
Baseline	Baseline (5)	Normal operating conditions
Fault	Fault without reclosure (5)	Single-phase to ground short circuit
	Fault with reclosure (5)	Single-phase to ground short circuit. Circuit breaker reclosure is implemented in the RTDS model
Network reconnaissance	Aggressive (1)	Nmap 'aggressive' timing template is used
	Insane (1)	Nmap 'insane' timing template is used
	Normal (1)	Nmap 'normal' timing template is used
	Polite (1)	Nmap 'polite' timing template is used
Resource exhaustion	Baseline priority A (5)	Resource exhaustion during normal conditions. 'Stress' process priority set to 0
	Fault priority A (8)	Resource exhaustion during a fault. 'Stress' process priority set to -15
	Fault priority B (10)	Resource exhaustion during a fault. 'Stress' process priority set to -19
	Only vIED (5)	As above, but the physical IED is disconnected from the network
GOOSE injection	IED spoofing baseline (5)	Malicious control signal to open the circuit breaker is sent during normal operating conditions while spoofing the physical IED identity
	IED spoofing during fault (5)	Malicious control signal to close the circuit breaker is sent during a fault while spoofing the physical IED identity
	IED spoofing during fault with reclosure (5)	As above, but the circuit breaker reclosure is implemented in the RTDS model
	vIED spoofing baseline (5)	Malicious control signal to open the circuit breaker is sent during normal operating conditions while spoofing vIED identity
	vIED spoofing during fault (5)	Malicious control signal to close the circuit breaker is sent during a fault while spoofing vIED identity
	vIED spoofing during fault with reclosure (5)	As above, but the circuit breaker reclosure mechanism is implemented in the RTDS model
SV injection	High current and voltage measurements (1)	High current and voltage measurements are maliciously injected into the SV frames
	Low current and voltage measurements (1)	Low current and voltage measurements are maliciously injected into the SV frames
	Replay fault (1)	Fault conditions measurements are maliciously reported in the injected SV frames
	Replay normal (1)	Normal conditions measurements are reported in the injected SV frames

versions of the services running on open ports, and '-T{N}' specifies the timing template to be used, with N parameter varying from 2 to 5 to refer to the 'polite', 'normal', 'aggressive', and 'insane' templates, respectively. To reduce the time required by the reconnaissance process to be completed, it has been limited to the IPs of the physical and virtual IEDs and to network ports that were likely to be open.

5) SV INJECTION ATTACK

In this scenario, the attacker injects malicious SV frames into the communication network to cause protection scheme malfunctions. Four different types of SV injection attacks are performed. In the first two cases, SV frames report extremely low or high current and voltage measurements, respectively. Then, in the last two cases, previously sniffed SV frames are reinjected into the communication network to perform a replay attack. In particular, in the second to last case, faulty conditions are replayed during normal operating conditions; conversely, in the last case, normal operating conditions are replayed during a fault situation.

For the cyber attacks to be successful, it is of fundamental importance to acquire a legitimate packet that was previously published by the legitimate SV publisher, in our case, RTDS. This allows the attacker to gain knowledge about the expected SV frame structure and to impersonate the legitimate SV publisher. A detailed description of how the attacks are performed is provided in the following.

The attack procedure can be divided into the following sub-steps:

1. *Sniffing*: by running Wireshark [35] on the attacker machine connected to the network switch, it is possible to acquire the SV frames being sent in multicast by the SV publisher, i.e., RTDS.
2. *Parsing*: after the original frames are sniffed, they are imported and parsed in Python with the usage of the Scapy module [36]. As previously mentioned, this step allows the attacker to gain knowledge about the expected SV frame structure and to impersonate the original SV frames publisher.
3. *Tampering*: depending on the performed attack type, the content of the frames is tampered by modifying the reported measurements or any other field available in the SV frame, e.g., the sample count field. How the reported measurements are verified is further described in Table 2. To allow a convenient and efficient injection of the generated malicious frames, the modified SV frames are exported and saved as a new PCAP Next Generation (.pcapng) format file which is then used in the last step.
4. *Injection*: the tampered SV frames are re-injected into the communication network by executing the tcpreplay [37] command and passing the just generated malicious.pcapng file as an argument. Other than the malicious.pcapng file, also the number of packets to be sent every second and the network interface to be used are specified with the usage of the '-p' and '-i' options, respectively. The usage of tcpreplay allows a straightforward injection of the malicious frames with a constant sending rate while spoofing the original SV publisher identity.

All attack steps are automated and implemented in a Python script which runs on the attacker's machine. It is crucial to note that, during the attack, RTDS SV publisher MAC address is spoofed, and the malicious SV frames have the same structure as the original one, thus the IED and vIED cannot distinguish between the original and the injected SV frames.

6) GOOSE INJECTION ATTACK

Similarly to the SV injection attack, in this attack scenario, it is assumed that the attacker can obtain the unencrypted GOOSE traffic by sniffing the messages multicasted throughout the communication network. Subsequently, the attacker modifies circuit breaker control parameters in the acquired GOOSE message using Scapy. This modification consists of changing the value of the control parameter from false to true [22]. In the experiment, one spoofed GOOSE message is injected every 100 ms, which leads to the malicious opening of the circuit breaker. The GOOSE injections are performed for both IED and vIED under various conditions, including normal, fault, and fault with reclosure. All the data from all tested scenarios are included in the dataset.

As for the SV injection attacks, the GOOSE injection attack procedure can be divided into the following sub-steps:

1. *Sniffing*: by running Wireshark [35] on the attacker machine connected to the network switch, it is possible to acquire the GOOSE messages being sent in multicast by the publisher.
2. *Parsing*: after the original frames are sniffed, they are imported and parsed in Python with the usage of the Scapy module [36]. As previously mentioned, this step allows the attacker to gain knowledge about the expected message structure and to impersonate the original GOOSE publisher.
3. *Tampering*: depending on the type of attack that is performed, the content of the frames is tampered with by modifying the sent control signals or any other field available in the GOOSE message. In particular, the status number (stNum) and sequence number (sqNum) fields are set to higher values with respect to the ones reported by the legitimate messages. This modification causes the GOOSE subscriber to discard legitimate messages and accept malicious ones. To allow a convenient injection of the generated malicious frames, the modified GOOSE frames are exported and saved as a new.pcapng file, which is then used in the last step.
4. *Injection*: the tampered frames are re-injected into the communication network by executing a Python script using the Scapy module.

C. DATA PROCESSING

To ensure the dataset is readily and conveniently usable by other researchers, a pre-processing phase was conducted. This phase involved the temporal alignment and filtering of both physical and cyber data to retain only data collected during simulation runs. Additionally, to prevent the disclosure of TU Delft and IEDs' vendor-sensitive information, the acquired network traffic was anonymized. The specific steps involved in this pre-processing phase are detailed in the following sections.

1) TEMPORAL ALIGNMENT

In our collected dataset, which includes both power system measurements and communication network packets, an inconsistency in time synchronization was identified. This discrepancy arises because the power system measurements from RTDS use a GPS clock for time reference, while the packets' timestamps are based on Coordinated Universal Time (UTC). Since there is a known time offset due to leap seconds between UTC and GPS time [38], this misalignment could impact the accuracy of our analysis. To address this, we applied a temporal alignment to ensure that all data points are consistently referenced to UTC time, enabling accurate comparisons and synchronization across the dataset. In the case of the physical measurements obtained through the RTDS, the time offset of 37 seconds was corrected in the collected dataset. Furthermore, as the timestamps for the communication network traffic and physical power system measurements need to be exact, a further data processing step is taken, in which the relative time from the start of the simulation is converted to a timestamp in ISO 8601 format.

2) FILTERING

The testbed's local area network was not completely isolated from the rest of the laboratory communication network. Therefore, additional communication network traffic unrelated to our study was acquired during the experiments. Since this extraneous data does not contribute to the objectives of our digital substation dataset, it was excluded from our network traffic acquisitions. To ensure that the dataset remains focused and relevant, the.pcapng files were filtered. This filtering process specifically selects only the devices involved in the experiment and the relevant communication protocols, allowing us to isolate the data that directly pertains to the digital substation environment.

3) ANONYMIZATION

For the anonymization phase, Tracewrangler was used [39]. This software is a network capture file toolkit running on Windows that supports.pcap and.pcapng file formats and allows for easy and effective sanitization and anonymization of communication network data acquisitions. Two different anonymization tasks are performed for IEC 61850 and non-IEC 61850 communication traffic.

TABLE 3. IP and MAC addresses of deployed systems and multicast/broadcast addresses assigned to IEC 61850 protocols.

Host/Endpoint	MAC Address	IP Address
RTDS	00:50:c2:f1:f1:f1CS	Not available
vIED	f5:a0:8d:e8:8f:2d	192.168.0.10
Attacker machine	4c:5c:06:ef:5e:c9	192.168.0.11
IED	12:c5:b1:f1:f1:f2	192.168.0.12
Network Switch	1c:60:19:52:b9:08	192.168.0.1
SV Multicast	01:0c:cd:04:00:00	Not available
GOOSE Multicast	01:0c:cd:01:00:00	Not available
Broadcast	ff:ff:ff:ff:ff:ff	Not available

For what concerns IEC 61850 GOOSE and SV traffic, the frames are kept intact apart for the source and destination MAC addresses which are replaced with the MAC addresses reported in Table 3. Also for the non-IEC 61850 traffic, MAC and IP addresses are replaced with the addresses reported in Table 3. Moreover, all the packets payload that are not recognized by Tracewrangler are anonymized by replacing all the payload bytes with zeros. This way, the statistical information of the communication network traffic such as packet lengths and network throughput are not influenced, but any sensitive information that might be contained in the payload is not disclosed.

As can be noted from Table 3, the MAC address of each machine is replaced with a randomly generated MAC address, whereas the GOOSE and SV multicast MAC addresses and the broadcast one are kept unchanged. Regarding the IP addresses, they are mapped to a fictional subnetwork with an IP range in 192.168.0.0/16. All the IPs not belonging to the testbed, and thus not listed in Table 3, are replaced by randomly generated IPs.

IV. DATASET DESCRIPTION

All the data associated with this work is available in the associated repository [40]. For each simulated scenario, the dataset consists of 3 different files:

1. RTDS physical data.
2. vIED resource monitoring data.
3. vIED communication network traffic acquisition.

In the scenarios in which a communication network-based cyber attack was performed, the communication network traffic is acquired in the attacker's machine; further, ground-truth labels are provided in comma-separated values (.csv) format alongside the respective.pcapng files to distinguish legitimate and malicious packets. The data was monitored and acquired for 60 seconds for each of the simulated scenarios, except the reconnaissance scenario; in this latter case, the data acquisition lasted 200 seconds. The choice of increasing the monitoring time for the reconnaissance scenario was made to allow Nmap scans to complete before the end of the simulation. To further increase the dataset usability and result reproducibility, each scenario was simulated multiple times as reported in Table 2.

TABLE 4. Columns available in .csv data exported from RTDS.

Column name	Type	Description
Time	Float	Relative time since RTDS simulation start
I1_a	Float [kA]	Current phase A
I1_b	Float [kA]	Current phase B
I1_c	Float [kA]	Current phase C
FLT_input	Boolean	Fault state signal
CB1_state	Boolean	Circuit breaker state
ied_trip_signal	Boolean	IED circuit breaker trip signal
vied_trip_signal	Boolean	vIED circuit breaker trip signal
V1_a	Float [kV]	Voltage phase A
V1_b	Float [kV]	Voltage phase B
V1_c	Float [kV]	Voltage phase C
formatted_datetime	Datetime	Absolute time in ISO 8601 format

The physical power system data was exported from RTDS during EMT simulations of IEEE 5-bus test system. The exported data consists of .csv files, with the first row reporting the name of the column. Even though the simulation time step was set to 50 microseconds in RTDS, due to RTDS data exporting capabilities, the data was downsampled to a time resolution of 1.6 milliseconds by keeping only one sample out of thirty-two. The exported data consists of the three-phase current and voltage measurements on bus number 1, status of the controlled circuit breaker, and status of three control signals, i.e., fault, physical IED, and vIED control. Moreover, for each data record, the relative time from the start of the simulation and absolute timestamp in ISO 8601 format are provided. In Table 4, a detailed description of the data contained in the .csv file exported from RTDS is provided.

For what concerns the communication network data, the vIED resources usage information was obtained by running the 'top' command in the VM hosting the vIED, and passing the vIED Process IDentifier (PID) as command line argument [41]. The command provides detailed information about computational, memory, and network usage with a time resolution of 2 seconds, and the output is saved as .csv files. This lower resolution, compared to that of the physical measurements exported from RTDS, reflects the slower time-scale of vIED resource variations compared to the milliseconds-level power system dynamics. Even though this introduces a temporal resolution mismatch between vIED resource utilization traces and the high-fidelity physical data, the dataset still supports cyber–physical correlation by aligning all streams on a common time base and enabling analysis at coarser, event- or attack-phase-level aggregations.

Table 5 provides detailed information about the acquired data. It should be noted that the vIED was deployed as a single-core process, meaning that the maximum CPU usage that can be allocated to the process is 25%, which corresponds to fully consuming the computational resource of one CPU core out of four.

TABLE 5. Columns available in .csv file containing vied resource monitored metrics.

Column name	Type	Description
Timestamp	Datetime	Absolute time in ISO 8601 format
CPU 1 Usage	Percentage	CPU core 1 load
CPU 2 Usage	Percentage	CPU core 2 load
CPU 3 Usage	Percentage	CPU core 3 load
CPU 4 Usage	Percentage	CPU core 4 load
RAM Usage	Percentage	RAM load
Network Sent	Float [GiB]	Amount of data sent over the network
Network Received	Float [GiB]	Amount of data received over the network
PID	Integer	Process identifier
User	String	User running the monitored process
PR	Integer	Priority of the process
NI	Integer	The nice value of the process, which affects scheduling priority
VIRT	Float	Total amount of virtual memory used by the process
RES	Float	The amount of physical memory used by the process
S	String	Current state of the process
%CPU	Percentage	Total CPU usage consumed by the process
%MEM	Percentage	Total system RAM used by the process
TIME+	Datetime	Total CPU time the process has used since it started
COMMAND	String	The command or executable name of the process

Given that port mirroring was not available in the network switch used during dataset acquisition, to acquire all the relevant communication network traffic, Wireshark was run both in the vIED and in the attacker's machine. This acquisition configuration allows for the interception of all the packets relevant to the conducted experiments. Wireshark output consists of a list of captured network packets. Each packet contains various fields and layers associated with network protocols, from data link layer up to the application layer. Moreover, it provides detailed information about when the packets are received and if they are corrupted or not. The acquired packets are exported and stored in.pcapng format.

V. TECHNICAL VALIDATION

This section provides a technical validation of the dataset to ensure the consistency and usability of the acquired data for fault and cyber attack detection, prevention, and impact assessment. This is done first by discussing the impact of the single-phase to ground fault. Then, the network reconnaissance and GOOSE/SV injection attacks are analyzed to showcase the variations in communication network throughput and inter-frames arrival times during an ongoing cyber attack, respectively. Furthermore, we demonstrate how the

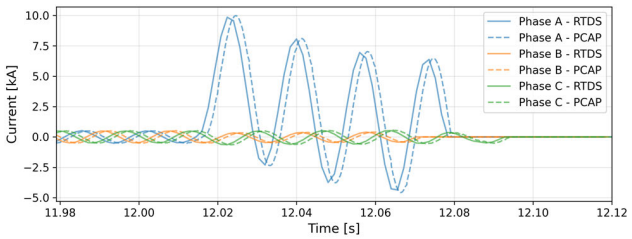

FIGURE 4. Validation of temporal alignment between physical data, i.e., three-phase currents exported from RTDS, and cyber data, i.e., three-phase currents reported in SV frames received by the vIED and monitored in Wireshark.

dataset is coherent with prior findings on the effects of SV injection attacks on physical IEDs.

A. PHYSICAL AND CYBER DATA TEMPORAL ALIGNMENT

As discussed in Section III-C, the dataset needed to be processed to ensure time consistency and alignment between cyber and physical data. To validate the correctness of the data processing, the timestamps of the physical measurements exported from RTDS were compared to those of the data reported in GOOSE and SV frames captured by monitoring the communication network traffic at the vIED. This comparison covered the three-phase current and voltage measurements and the trip signals exchanged between RTDS and vIED, with the three-phase current measurements presented in Figure 4 as a representative result. As can be appreciated, the three-phase current measurements obtained from the SV frames are received with a delay lower than 2 ms with respect to the data directly exported from RTDS. This delay is due to the time required by RTDS to process and transmit the data through the GTNETx2 card, by the frames to propagate through the communication network and reach the vIED, and by the vIED to process and timestamp the received frames. The delay experienced in our testbed is consistent with the one that would be experienced in a real digital substation communication network, in which MUs need to measure the currents and voltages, process and transmit them over the process bus, and IEDs need to receive and process them.

B. PHYSICAL IMPACT OF THE SINGLE-PHASE TO GROUND FAULT

To validate the protection functionality of the IEDs and consequently capture the associated GOOSE and SV communications, a single-phase to ground fault on phase A of line 1 is simulated for a duration of 100 ms.

As can be observed from Figure 5, following the fault, the current of phase A rises exponentially to almost 10 kA. Consequently, due to the high fault current, the vIED detects a fault, and a trip command is issued almost instantaneously. Following this, the circuit breaker opens after 50 ms to clear the fault as indicated by the orange line in Figure 4. As a result, the fault current is suppressed, and the fault is cleared at around 12.06 seconds simulation time. It can also be appreciated how the physical IED reacts to the fault too, but with a delay of around 20 ms with respect to the vIED. This delay is mainly due to the different protection schemes implemented in the two physical and virtual IEDs.

C. NETWORK TRAFFIC THROUGHPUT INCREASE DURING NETWORK RECONNAISSANCE

When an attacker uses Nmap to perform network reconnaissance, specially crafted packets are sent to the target hosts and by analyzing the responses it enables to identify active systems, open ports, running services, and operating systems. This process inevitably increases the number of packets flowing in the communication network. For this reason, to balance network reconnaissance speed and stealthiness, Nmap provides predefined timing templates. In Figure 6, it can be appreciated how the number of packets received per second and the total number of received packets vary depending on the chosen timing template, i.e., ‘polite’, ‘normal’,

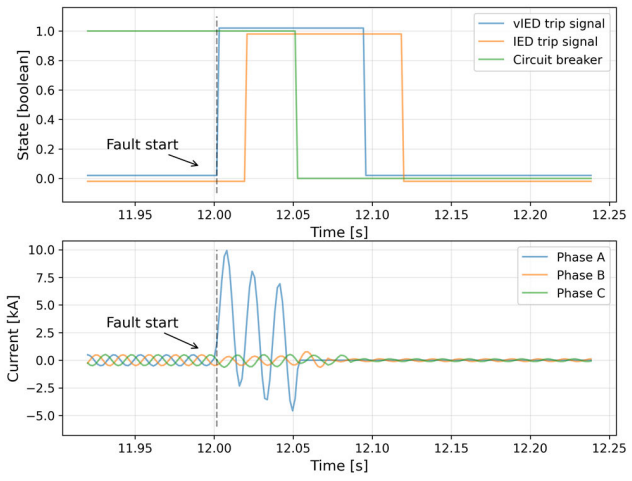

FIGURE 5. Successfully cleared single-phase to ground fault.

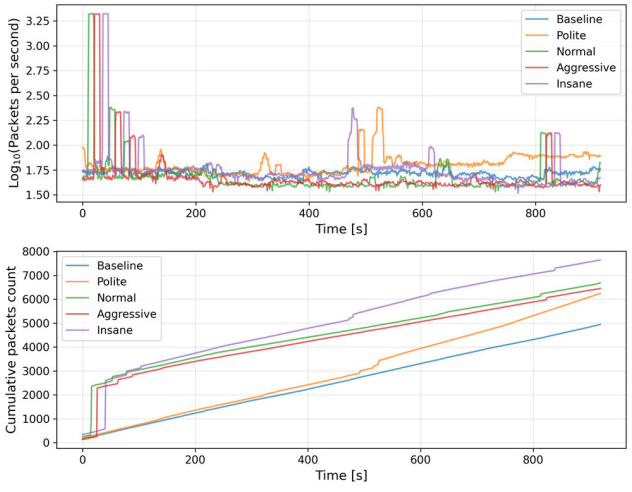

FIGURE 6. Number of packets received during network reconnaissance process when different intensities are used.

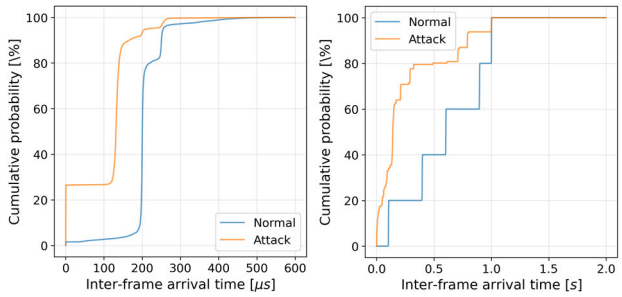

FIGURE 7. Cumulative probability of SV (left-side plot) and GOOSE (right-side plot) inter-frame arrival time during normal and injection attacks.

'aggressive', and 'insane'. To improve throughput visualization, SV frames are filtered out from the plotted network traffic acquisition given that SV protocol publishes frames with a constant rate of 4800 frames per second. As expected, it can be noted that with the three most intensive templates, the increase in communication network throughput is evident, thus providing a great informative edge for the detection of the cyber attack. On the contrary, the throughput difference between the baseline case, i.e., normal operations without an ongoing cyber attack, and the 'polite' timing template, is much more subtle but still noticeable. Also, an informative edge is still present given that by the end of the network reconnaissance process a similar total number of packets have been sent for each of the timing templates.

D. INTER-FRAME ARRIVAL TIME VARIATION DURING GOOSE AND SV INJECTION ATTACKS

As specified in the IEC 61850-9-2 standard, SV publishers are mandated to periodically send measurements at exactly pre-defined time intervals. This time interval depends on two factors, which are the measured signal frequency (f) and Samples Per Period (SPP) parameter. Specifically, two SPP values are specified in the standard, i.e., 80 and 256. Given these two factors, the number of SV frames published in each second can be computed by multiplying f by SPP. In the testbed used for generating the dataset presented in this paper, the system frequency is set to 60 Hz, whereas the SV publisher defined in RTDS is set to publish 80 SPP. This leads to the publication of 4800 SV frames per second, leading to an inter-frame arrival time of 208.33 microseconds. As represented in Figure 7, during normal conditions, almost 80% of the SV frames are received with an inter-frame arrival time slightly above 200 microseconds as mandated by the standard. Being the frame arrival time dependent on transmission, processing, queueing, and propagation delays, it is normal and expected that a fraction of the received frames doesn't match the expected frame arrival time. On the other hand, it can be noticed that during an SV injection attack, the inter-frame arrival time is considerably reduced. This is because the SV subscriber is receiving both the original and injected SV streams, leading to 9600 frames being received every second.

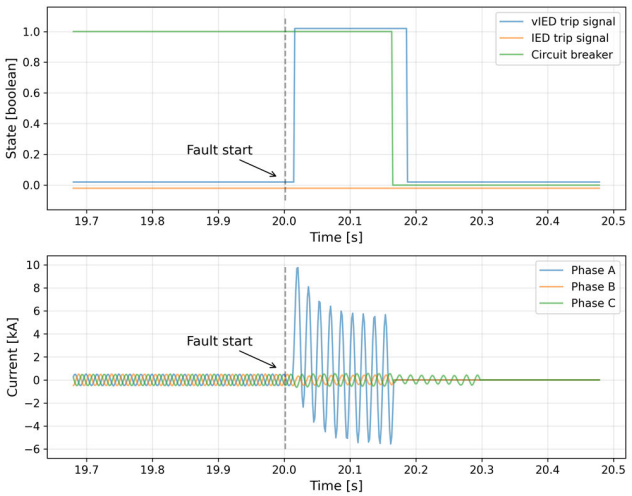

FIGURE 8. Response of IED and vIED during SV injection attack.

A similar behaviour can be observed in Figure 7, where GOOSE frames inter-arrival time is depicted. In the case of the GOOSE protocol, the standard does not mandate any periodicity on the publishing of GOOSE messages. Still, in the setup used during our experiment both the physical and virtual IEDs are publishing at least one GOOSE message per second. As for the SV injection case, when the GOOSE injection attack is ongoing, the inter-frame arrival time is reduced due to the presence of injected GOOSE messages flowing in the communication network and reaching the connected devices.

E. IMPACT OF SV INJECTION ON IEDS

As found in Rajkumar et al. [6], when an IED receives multiple SV streams identified by the same application identifier, i.e., APPID field in the SV frame, the IED ends up in a blocked state. This is because the IED does not know how to handle these conditions, and to prevent unintended commands from being sent it stops working. In the simulated scenario, the unexpected SV stream is the one injected by the attacker during the SV injection attack. This behaviour can be observed in Figure 8, where it can be noted that after the fault has occurred, the physical IED does not react and does not send any control signal to open the circuit breaker.

However, the vIED logic is implemented such that even if multiple SV streams using the same APPID are received, it keeps on operating and sending control signals to circuit breakers. Indeed, it can be observed that the vIED reacts to the fault and sends the control signal to open the circuit breaker and clears the fault. This further highlights the problem of different implementations and outcomes of the same standard, depending on vendor or library specifications [42].

F. IMPACT OF RESOURCE EXHAUSTION ATTACK ON VIED

During the execution of a resource exhaustion attack targeting the vIED, the attacker deprives the protection scheme process, running on the virtual machine, of the necessary

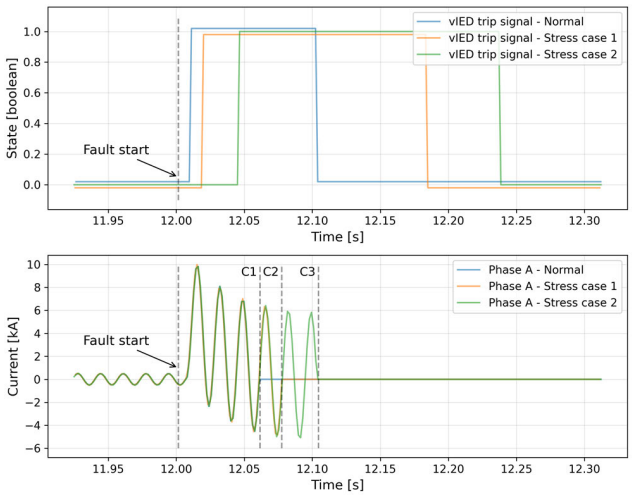

FIGURE 9. Delay in fault clearance due to a resource exhaustion attack targeting the vIED. The dotted lines labelled 'C1', 'C2', and 'C3' represent the fault clearance times under normal conditions, stress case 1, and stress case 2, respectively.

computational resources. This shortage can lead to a delay in issuing the control signal responsible for opening the circuit breaker and clearing the fault. As illustrated in Figure 9, when the vIED is subject to resource exhaustion, specifically in stress cases 1 and 2, the fault is cleared with a noticeable delay compared to normal conditions where no cyber attack is present. Under normal operation, the fault is cleared in approximately 60 ms; however, during a resource exhaustion attack, the fault clearing time can increase significantly, reaching up to 100 ms in the depicted scenario. Such delayed fault clearance can have a detrimental effect on the transient stability of the power system, potentially leading to oscillations or even loss of synchronism if the disturbance persists beyond the fault critical clearing time. As reported in El Hariri et al. [42], resource exhaustion attacks can result in even longer delays, and in extreme cases, may entirely prevent the vIED from reacting to the fault and issuing the required control command.

VI. CONCLUSION

This paper has presented a novel CPPS dataset specifically designed to support research on cyber security of digital substations. The dataset integrates real-time EMT measurements, network traffic, and virtual IED resource metrics, thereby providing a high-fidelity representation of both physical power system dynamics and cyber layer interactions. By leveraging an IEC 61850-compliant testbed with HIL integration of RTDS, physical IED, and virtual IED, diverse operating conditions and cyber attacks scenarios were systematically simulated. The resulting dataset covers normal operations, fault conditions, and multiple attack vectors, including network reconnaissance, resource exhaustion, and malicious GOOSE and SV injections. This comprehensive scope ensures that the dataset can serve as a benchmark for evaluating, training, and validating advanced cyber attack detection and prevention methodologies. By making this dataset available to the research community, we aim to facilitate the development of resilient and adaptive solutions for securing digital substations, ultimately contributing to the reliability and stability of future power grids. Future work will focus on extending the dataset to include additional attack types, fault scenarios, a larger-scale cyber–physical model, additional physical and virtual IEDs, and addressing real-world deployment considerations to further enhance its applicability.

REFERENCES

[1] D. U. Case, "Analysis of the cyber attack on the Ukrainian power grid," *Electr. Inf. Sharing Anal. Center*, vol. 388, nos. 1–29, p. 3, 2016.

[2] J. Slowik, "Crashoverride: Reassessing the 2016 Ukraine electric power event as a protection-focused attack," Dragos, Inc., Hanover, MD, USA, Whitepaper, Tech. Rep., 2019.

[3] P. Kozak, I. Klaban, and T. Šlajs, "Industroyer cyber-attacks on Ukraine's critical infrastructure," in *Proc. Int. Conf. Mil. Technol. (ICMT)*, May 2023, pp. 1–6.

[4] *Communication Networks and Systems for Power Utility Automation All Parts*, document IEC 61850, 2024. [Online]. Available: https://webstore.iec.ch/en/publication/6028

[5] D. Parsons, "The state of ICS/OT cybersecurity in 2022 and beyond," SANS Inst., North Bethesda, MD, USA, Survey Whitepaper, Tech. Rep., 2022.

[6] V. S. Rajkumar, M. Tealane, A. Stefanov, A. Presekal, and P. Palensky, "Cyber attacks on power system automation and protection and impact analysis," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur. (ISGT-Europe)*, Oct. 2020, pp. 247–254.

[7] D. Protic, "Review of KDD cup '99, NSL-KDD and Kyoto 2006+ datasets," *Vojnotehnicki glasnik/Military Tech. Courier*, vol. 66, no. 3, pp. 580–596, 2018.

[8] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *Proc. Mil. Commun. Inf. Syst. Conf.*, Nov. 2015, pp. 1–6.

[9] A. H. Lashkari, G. Kaur, and A. Rahali, "DIDarknet: A contemporary approach to detect and characterize the darknet traffic using deep image learning," in *Proc. 10th Int. Conf. Commun. Netw. Secur.*, Nov. 2020, pp. 1–13.

[10] P. Radoglou-Grammatikis, V. Kelli, T. Lagkas, V. Argyriou, and P. Sarigiannidis, "DNP3 intrusion detection dataset," IEEE Dataport, Piscataway, NJ, USA,, Tech. Rep., 2022, doi: 10.21227/s7h0-b081.

[11] P. Matoušek, O. Ryšavy, and P. Grofčik, "ICS dataset for smart grid anomaly detection," in *Proc. 10th Int. Conf. Commun. Netw. Secur.*, pp. 1–13.

[12] K. Boakye-Boateng, A. A. Ghorbani, and A. H. Lashkari, "Securing substations with trust, risk posture, and multi-agent systems: A comprehensive approach," in *Proc. 20th Annu. Int. Conf. Privacy, Secur. Trust (PST)*, Aug. 2023, pp. 1–12.

[13] E. D. Gutiérrez Mlot, J. Saldana, R. J. Rodríguez, I. Kotsiuba, and C. Gañán, "A dataset to train intrusion detection systems based on machine learning models for electrical substations," *Data Brief*, vol. 57, Dec. 2024, Art. no. 111153.

[14] A. Sahu, Z. Mao, P. Wlazlo, H. Huang, K. Davis, A. Goulart, and S. Zonouz, "Cyber-physical dataset for MiTM attacks in power systems," IEEE Dataport, Piscataway, NJ, USA, Tech. Rep., 2021, doi: 10.21227/e4dd-2163.

[15] H. Huang, "Cyber-physical dataset of hardware-in-the-loop cyber-physical power systems testbed under MiTM attacks," IEEE Dataport, Piscataway, NJ, USA, Tech. Rep., May 2022, doi: 10.21227/pn67-8408.

[16] S. Zemanek, I. Hacker, K. Wolsing, E. Wagner, M. Henze, and M. Serror, "PowerDuck: A GOOSE data set of cyberattacks in substations," in *Proc. 15th Workshop Cyber Secur. Experimentation Test*, Aug. 2022, pp. 49–53.

[17] A. Eynawi, A. Mumrez, G. Elbez, and V. Hagenmeyer, "Machine learning-based feature selection for intrusion detection systems in IEC 61850-based digital substations," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Sep. 2024, pp. 1–7.
[18] P. P. Biswas, H. C. Tan, Q. Zhu, Y. Li, D. Mashima, and B. Chen, "A synthesized dataset for cybersecurity study of IEC 61850 based substation," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Oct. 2019, pp. 1–7.
[19] S. E. Quincozes, C. Albuquerque, D. Passos, and D. Mossé, "ERENO: A framework for generating realistic IEC–61850 intrusion detection datasets for smart grids," *IEEE Trans. Dependable Secure Comput.*, vol. 21, no. 4, pp. 3851–3865, Jul. 2024.
[20] N. Kabbara, M. O. Nait Belaid, M. Gibescu, L. R. Camargo, J. Cantenot, T. Coste, V. Audebert, and H. Morais, "Towards software-defined protection, automation, and control in power systems: Concepts, state of the art, and future challenges," *Energies*, vol. 15, no. 24, p. 9362, Dec. 2022.
[21] P. Silveira, E. F. Silva, A. Galletta, and Y. Lopes, "Security analysis of digitized substations: A systematic review of GOOSE messages," *Internet Things*, vol. 22, Jul. 2023, Art. no. 100760.
[22] V. S. Rajkumar, A. Stefanov, A. Presekal, P. Palensky, and J. L. R. Torres, "Cyber attacks on power grids: Causes and propagation of cascading failures," *IEEE Access*, vol. 11, pp. 103154–103176, 2023.
[23] S. Ashraf, M. H. Shawon, H. M. Khalid, and S. M. Muyeen, "Denial-of-service attack on IEC 61850-based substation automation system: A crucial cyber threat towards smart substation pathways," *Sensors*, vol. 21, no. 19, p. 6415, Sep. 2021.
[24] M. El Hariri, E. Harmon, T. Youssef, M. Saleh, H. Habib, and O. Mohammed, "The IEC 61850 sampled measured values protocol: Analysis, threat identification, and feasibility of using NN forecasters to detect spoofed packets," *Energies*, vol. 12, no. 19, p. 3731, Sep. 2019.
[25] J. Hong, R. Karnati, C.-W. Ten, S. Lee, and S. Choi, "Implementation of secure sampled value (SeSV) messages in substation automation system," *IEEE Trans. Power Del.*, vol. 37, no. 1, pp. 405–414, Feb. 2022.
[26] RTDS Technologies. (2025). *Simulation Hardware*. [Online]. Available: https://www.rtds.com/technology/simulation-hardware
[27] M. Zillgith. *LIBIEC61850: Open-Source IEC 61850 Protocol Stack*, Accessed: Apr. 28, 2026. [Online]. Available: https://github.com/mz-automation/libiec61850
[28] RTDS Technologies. (2025). *GTNETx2: The RTDS Simulator's Network Interface Card*. [Online]. Available: https://knowledge.rtds.com/hc/en-us/articles/360034788593-GTNETx2-The-RTDS-Simulator-s-Network-Interface-Card
[29] N. Kabbara, A. Mwangi, A. Stefanov, and M. Gibescu, "A real-time implementation and testing of virtualized controllers for software-defined IEC 61850 digital substations," *IEEE Open J. Ind. Appl.*, vol. 5, pp. 300–310, 2024.
[30] (2025). *Stress(1)–Linux Manual Page*. [Online]. Available: https://linux.die.net/man/1/stress
[31] Free Software Foundation. (2025). *GNU Core Utilities: Nice*. [Online]. Available: https://www.gnu.org/software/coreutils/manual/html_node/nice-invocation.html
[32] Free Software Foundation. (2025). *GNU Core Utilities: Renice*. [Online]. Available: https://www.gnu.org/software/coreutils/manual/html_node/renice-invocation.html
[33] G. F. Lyon. (2025). *Nmap*. [Online]. Available: https://nmap.org
[34] G. F. Lyon. (2025). *Nmap Timing Templates*. [Online]. Available: https://nmap.org/book/man-performance.html
[35] Wireshark Foundation. (2025). *Wireshark Network Protocol Analyzer*. [Online]. Available: https://www.wireshark.org
[36] S. Bansal and N. Bansal, "Scapy—A Python tool for security testing," *J. Comput. Sci. Syst. Biol.*, vol. 8, no. 3, p. 140, 2015.
[37] A. Turnbull. (2025). *Tcpreplay*. [Online]. Available: https://tcpreplay.appneta.com
[38] W. Lewandowski and M. Marszalec, "A brief history of UTC leap second," *J. Telecommun. Inf. Technol.*, vol. 2023, no. 4, pp. 117–122, Apr. 2023.
[39] J. Richter. (2025). *TraceWrangler*. [Online]. Available: https://www.tracewrangler.com
[40] N. Cibin, N. Kabbara, A. Presekal, V. Rajkumar, H. Goyel, P. Palensky, and A. Stefanov, "Cyber-physical power system dataset for cyber security of digital substations," Zenodo, Geneva, Switzerland, Tech. Rep., May 2026, doi: 10.5281/zenodo.15371179.
[41] (2025). *Top(1)–Linux Manual Page*. [Online]. Available: https://man7.org/linux/man-pages/man1/top.1.html
[42] M. El Hariri, T. Youssef, and O. Mohammed, "On the implementation of the IEC 61850 standard: Will different manufacturer devices behave similarly under identical conditions?" *Electronics*, vol. 5, no. 4, p. 85, Dec. 2016.
[43] N. Kabbara, N. Cibin, H. Morais, A. Stefanov, and M. Gibescu, "A cybersecurity assessment for hybrid virtualized-physical digital substations," *Sustain. Energy, Grids Netw.*, vol. 43, Sep. 2025, Art. no. 101795.

NICOLA CIBIN (Graduate Student Member, IEEE) received the B.Sc. degree in information engineering and the M.Sc. degree in telecommunications engineering from the University of Padua, Italy, in 2019 and 2021, respectively. After working as a Research Assistant with Aalborg University, Denmark, from 2021 to 2023, he joined as a Ph.D. Student with the Cyber Resilient Power Grids (CRPG) Research Group, TU Delft, The Netherlands. His research interests include telecommunication networks, cyber security, intrusion detection and prevention systems, and smart grids.

NADINE KABBARA received the M.Sc. degree in control and electrical engineering from Paris-Saclay University, in 2021, and the Ph.D. degree from Utrecht University, in 2025. She was a Marie-Skłodowska Curie Doctoral Researcher with French Utility EDF, in their Research and Development Department as part of the InnoCyPES European Project on connected digital smart grids. Her research interests include digital substations, IEC 61850, virtualized IEDs, data modeling, and simulation frameworks.

ALFAN PRESEKAL (Member, IEEE) received the M.Sc. degree in secure software systems from the Department of Computing, Imperial College London, U.K., in 2016, and the Ph.D. degree in cyber-resilient power grids from Delft University of Technology, in 2025. He is an Assistant Professor of computer engineering with Universitas Indonesia. Previously, he was a Researcher with European Horizon projects HVDC-WISE and Cooperative Cyber Protection for Modern Power Grids (COCOON).

IOANNIS SEMERTZIS (Graduate Student Member, IEEE) received the Diploma degree in electrical and computer engineering from the Democritus University of Thrace, Greece, in 2019, and the M.Sc. degree in electrical power engineering from Delft University of Technology, Delft, The Netherlands, in 2021, where he is currently pursuing the Ph.D. degree with the Department of Electrical Sustainable Energy. His main research interests include cyber security, cyber–physical power systems, power system stability, and artificial intelligence for power system applications.

VETRIVEL SUBRAMANIAM RAJKUMAR (Graduate Student Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from Delft University of Technology, Delft, The Netherlands, in 2019 and 2025, respectively. He is currently a Policy Advisor with the Department of Systems Operations–Grid Security, Dutch TSO, TenneT. His research interests include cyber security and resilience for power grids.

HIMANSHU GOYEL (Graduate Student Member, IEEE) received the bachelor's degree in electrical engineering from the University of Mumbai and the Master of Science degree in electrical engineering from Indian Institute of Technology Madras, Chennai, India. He is currently pursuing the Ph.D. degree in cybersecurity for power systems with Technische Universiteit Delft, Delft, The Netherlands. His professional experience includes work as an Engineer with Grid-Sentry and Rakuten Mobile Inc., Tokyo. His research interests include power systems optimization, cybersecurity, digital substation, machine learning, and smart power grids.

PETER PALENSKY (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. and Habilitation degrees from Vienna University of Technology, Austria, in 1997, 2001, and 2015, respectively. He co-founded an Envidatec, a German startup on energy management and analytics, and joined the Lawrence Berkeley National Laboratory, Berkeley, CA, USA, as a Researcher, and the University of Pretoria, South Africa, in 2008. In 2009, he became appointed as the Head of the Business Unit on Sustainable Building Technologies, Austrian Institute of Technology (AIT), and later the first Principal Scientist of complex energy systems with AIT. In 2014, he was appointed as a Full Professor of intelligent electric power grids with TU Delft. He is active in international committees, such as ISO and CEN. His research interests include energy automation networks, smart grids, and modeling intelligent energy systems. He also serves as an IEEE IES AdCom Member-at-Large in various functions for IEEE. He is also the Editor-in-Chief of *IEEE Industrial Electronics Magazine*, and an associate editor of several other IEEE publications and regularly organizes IEEE conferences.

ALEXANDRU ŞTEFANOV (Member, IEEE) received the M.Sc. degree from the University Politehnica of Bucharest, Romania, in 2011, and the Ph.D. degree from University College Dublin, Ireland, in 2015. He is an Associate Professor in intelligent electrical power grids with the Department of Electrical Sustainable Energy, TU Delft, The Netherlands. He is the Director of the Control Room of the Future (CRoF) Technology Centre. He is leading the Cyber Resilient Power Grids (CRPG) Research Group. He holds the professional title of Chartered Engineer from Engineers Ireland. His research interests include cyber security of power grids, resilience of cyber–physical systems, and next generation grid operation.