Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies

Ioannis Zografopoulos, *Student Member, IEEE*, Juan Ospina, *Member, IEEE*, XiaoRui Liu, *Student Member, IEEE*, and Charalambos Konstantinou, *Senior Member, IEEE*

***Abstract*—Cyber-physical systems (CPS) are interconnected architectures that employ analog, digital, and communication resources for their interaction with the physical environment. CPS are the backbone of enterprise, industrial, and critical infrastructure. Thus, their vital importance, makes them prominent targets for malicious attacks aiming to disrupt their operations. Attacks targeting cyber-physical energy systems (CPES), given their mission-critical nature, can have disastrous consequences. The security of CPES can be enhanced leveraging testbed capabilities to replicate power systems operation, discover vulnerabilities, develop security countermeasures, and evaluate grid operation under fault-induced or maliciously constructed scenarios. In this paper, we provide a comprehensive overview of the CPS security landscape with emphasis on CPES. Specifically, we demonstrate a threat modeling methodology to accurately represent the CPS elements, their interdependencies, as well as the possible attack entry points and system vulnerabilities. Leveraging the threat model formulation, we present a CPS framework designed to delineate the hardware, software, and modeling resources required to simulate the CPS and construct high-fidelity models which can be used to evaluate the system's performance under adverse scenarios. The system performance is assessed using scenario-specific metrics, while risk assessment enables the system vulnerability prioritization factoring the impact on the system operation. The overarching framework for modeling, simulating, assessing, and mitigating attacks in a CPS is illustrated using four representative attack scenarios targeting CPES. The key objective of this paper is to demonstrate a step-by-step process that can be used to enact in-depth cybersecurity analyses, thus leading to more resilient and secure CPS.**

***Index Terms*—Cyber-physical systems, security, threat modeling, power grid, simulation, risk assessment, testbeds.**

I. Introduction

A. Background and Motivation

Over the past years, electric power systems (EPS) have diverged from a unidirectional generation and transmission model towards a more distributed architecture that supports traditional generation sources as well as distributed energy resources (DERs) in the form of distributed generation (DG), such as PV and wind, and distributed storage (DS) sources, such as battery energy storage systems (BESS) and thermal energy storage systems (TESS). The transformation of EPS to cyber-physical energy systems (CPES) is primarily enabled due to the introduction of information and communication technologies (ICT), automated control systems, remote sensing, and embedded industrial internet-of-things (IIoT) devices. According to the National Institute of Standards and Technology (NIST) [1], cyber-physical systems (CPS) refer to architectures that incorporate digital, analog, and physical components. The interaction of these components is determined by the dynamics of the system and the rules which orchestrate its operation. *CPES are energy-focused engineered systems that are transforming the way traditional EPS operate by seamlessly integrating physical entities with human, digital, and networking components designed to operate through integrated physics and computational logic.* As such, CPES contribute significantly towards the EPS modernization allowing for better planning, more flexible control, cyber-secure operations, system-wide optimization, transactive energy systems (TES), improvements in power quality, enhanced system reliability, resiliency, interoperability, and cleaner energy generation.

The security of CPS presents significant challenges in controlling and maintaining secure access to critical system resources and services (e.g., for CPES: generation reserves, frequency stability controls, power line protection, etc.), as well as ensuring the confidentiality, accessibility, and integrity of the information exchanged (e.g., control signals of supervisory control and data acquisition – SCADA systems). CPS, being large-scale complex systems of systems, employ numerous computing components such as remote terminal units (RTUs), programmable logic controllers (PLCs), and intelligent electronic devices (IEDs) that are often designed without security in mind. Typically, the hardware, software, and communication interfaces of these devices are developed utilizing commercial off-the-shelf components [2]. Thus, vulnerabilities within such components can be ported to the CPS environments creating potential entry points for malicious adversaries[1] aiming to disrupt CPS operations. An indicative incident of malicious behavior targeting CPS operation was reported in March 2019. Attackers targeted the United States (U.S.) grid infrastructure and performed a denial-of-service (DoS) attack through the exploitation of a known CPES vulnerability, namely a web interface firewall vulnerability [3], [4]. The attack resulted in a loss of communication between the utility's generation assets and the energy management system [5], causing brief interruptions in the utility's service. The number of cyber-attacks where adversaries exploit known and existing vulnerabilities to compromise CPS is increasing. This fact is validated by security reports stating that "*99% of the vulnerabilities exploited in 2020 are known to security professionals, while zero-day vulnerabilities only account for the 0.4% of vulnerabilities exposed during the past decade*"

[1]Throughout the paper, we use the terms *adversary*, *threat actor*, and *attacker* interchangeably.

[6].

The importance of CPS, and CPES in particular, for economic prosperity and public health at the national, state, and local level, can motivate attackers to compromise such systems in order to obtain financial or political gains. Hence, the evaluation of the CPES robustness and resilience against attacks in realistic scenarios is of paramount importance. At the same time, the quantification of cybersecurity risks is becoming more complex and challenging as EPS – also referred to as the 'largest interconnected machine on Earth' [7] – integrate numerous cyber-components at all levels and scales. In the past, the simulation of specific abnormal scenarios (e.g., faults, overvoltage conditions, frequency fluctuations, etc.) was sufficient to provide insights into EPS operations. However, current advances towards intelligent and interconnected CPES require more accurate models and representations capable of capturing the dynamic behavior of these interoperable systems. The enhancement of CPES security and reliability requires constant probing for potential weaknesses [8]. Security studies need to reflect the nature of the CPES infrastructure in actual testing environments that support the interfacing of actual hardware devices designed to operate in the 'real' system. In this context, hardware-in-the-loop (HIL) testbeds are effective in providing testing capabilities for evaluating the synergistic relationship between physical and virtual components in a controlled environment. Security-oriented HIL testbeds are invaluable in performing cybersecurity and risk analyses, identifying system vulnerabilities in various layers (hardware, firmware, software, protocol, process), implementing intrusion detection and prevention algorithms, and assessing the efficiency of mitigation techniques without inducing excessive economic burdens or safety hazards [2], [9].

The primary motivation of this paper is to develop a framework, which bridges theoretical and simulation-based security case studies and evaluates CPS system behavior leveraging testbed environments, leading to more secure CPES architectures. In order for testbeds to reliably capture the characteristics of the cyber-physical environment, testing and experimental case studies need to be described and modeled considering both the cyber and physical domains. The case studies require detailed descriptions of the resources and metrics that will be utilized for evaluating the CPES performance, reliability, and resilience. In addition, the testing setup must also capture the threat modeling characteristics of the adversary and the attack methodology. In terms of a potential adversary, the threat modeling characteristics are adversarial knowledge, resources, access to the system, and specificity. As for the attack methodology, the threat modeling characteristics include the attack frequency, reproducibility, discoverability, target level, attacked asset, attack techniques, and premise. Doing so, in a holistic and step-by-step approach, allows researchers and stakeholders to thoroughly examine and uncover security risks existing in the CPES under evaluation.

B. Research Contribution and Overview

The underlying goal of this manuscript is to provide a complete and detailed presentation of CPS security research studies by demonstrating a modular framework for assessing CPS security in the context of CPES. To this end, the paper describes all the required components for evaluating the behavior and performance of CPES under diverse and adverse operational scenarios. The framework showcases the modeling techniques used to represent the cyber and physical domains of the system, considers the resources used to model the CPES, and presents essential evaluation metrics for each corresponding case study. The contributions of this work *focusing on CPES security* can be summarized as follows:

- A literature review is provided that presents the research efforts in the area of CPS and CPES security, describes cyber-physical testbeds developed by prominent research centers and laboratories around the world, and illustrates current threat and risk modeling approaches widely used in the industry.
- A threat modeling methodology is proposed, comprised of two major parts, the adversary model and the attack model, allowing for an inclusive evaluation of malicious attack strategies.
- Leveraging our threat modeling approach, a risk assessment process is provided that takes into account risks related to the effectiveness of an attack, the targeted system component, and the criticality of the cyber-physical process being compromised.
- A framework is described that elucidates the crucial components and resources needed to accurately characterize CPS, making it essential for evaluating numerous studies (e.g., cyber, control, etc.). It is important to note that the proposed CPS framework can be used to characterize CPS in other sectors such as healthcare and transportation, but in this work, it is evaluated specifically for CPES.
- Four illustrative CPES attack case studies are presented, demonstrating the practicality of the CPS framework. For each case study, we provide the corresponding background and mathematical formulation, threat model, attack setup, and risk assessment. We also describe how each stage of the CPS analysis framework is applied to thoroughly model the specific characteristics of each case study.

A schematic overview of this paper is illustrated in Fig. 1. Section II presents the current state of CPES testbed research, a literature review of CPES security studies, and preliminary information for threat analysis and risk assessment of CPS. Section III delineates our comprehensive threat modeling and risk assessment methodology. Section IV provides the description of the proposed CPS framework with details on the modeling, resources, and performance metrics. In Section V, we discuss the background information and mathematical formulation for attack cases targeting CPES, and present such simulated test case scenarios accompanied by their experimental results implemented using the developed CPS framework. Finally, Section VI concludes this work.

II. Cyber-Physical Energy Systems (CPES): Testbeds, Studies, and Security Analysis

This section provides an overview of different CPES testbeds developed by various research centers, and presents their research objectives alongside the equipment used to realize them. We define different classes of CPES security studies

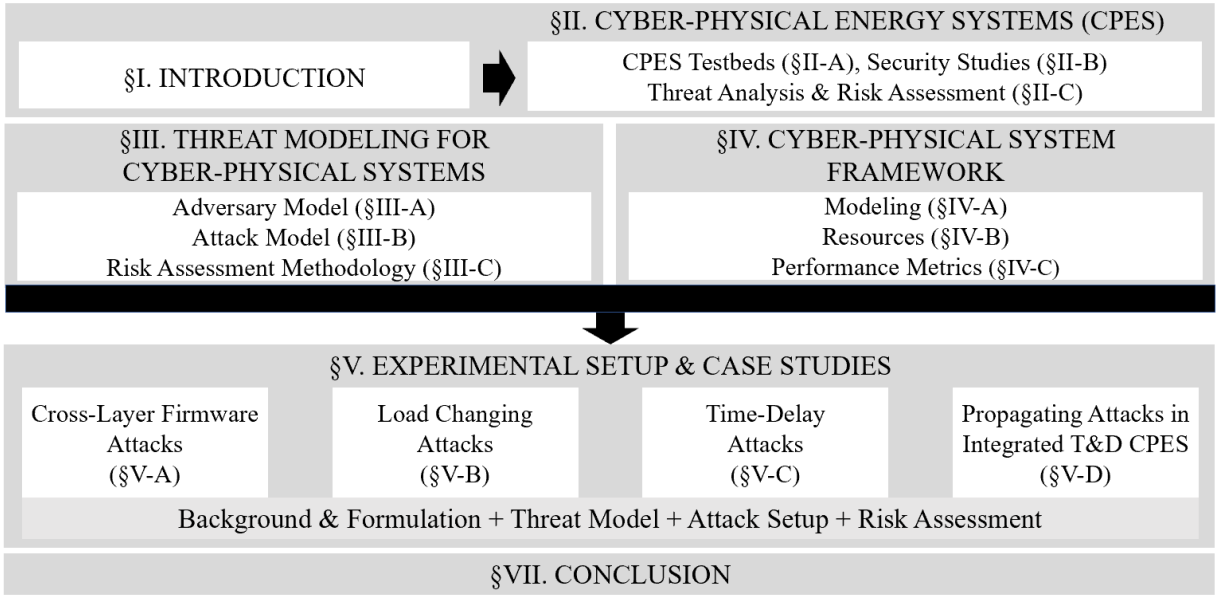


Fig. 1: Roadmap of the paper.

from the literature, and discuss prominent examples from such categories. Furthermore, we describe threat modeling and risk assessment methodologies and discuss how they can support security studies by defining, preventing, and mitigating threats.

A. CPES Testbeds

Throughout the years, EPS were designed and simulated following unidirectional structures in which power is generated at large bulk power generation facilities and then delivered through different stages of transmission and radial distribution systems to consumers. Minimum efforts were exerted to facilitate the integration of renewable energy sources (RES) and DERs [10]. However, the increasing penetration of RES and DERs along with the grid modernization efforts through ICT, increases the complexity of EPS [11]. On the one hand, RES and DERs can be used to meet consumer demands providing reliable, economic, and environmentally friendlier energy. On the other hand, attackers can exploit the fact that these resources are not centrally controlled (i.e., controlled directly by utilities) and stealthily plant their attacks on vulnerable system assets [12], [13]. The complex nature of modern EPS introduces a variety of potential entry points for attacks due to the fact that these systems depend on ICT for the communication between system assets [14]. Although the exigency for secure and resilient EPS is evident, our limited experience with dealing and coordinating such sophisticated architectures exacerbates the situation. We lack mechanisms to detect and mitigate the impact of unexpected adverse events on power system operation. The design of power system monitoring, control, and estimation algorithms, which are inherently secure, regardless of relying on CPES interconnected nature, relies heavily on the existence of representative frameworks where current and future security features and methodologies can be developed and evaluated.

CPES testbeds can provide an ideal environment where thorough system evaluations can be performed without any impact on the actual power system. The use of testbeds helps de-risk certain procedures before migration to the actual system, and avoid any potential adverse impact they could inflict. Such procedures include the testing and impact evaluation of new EPS equipment (e.g., integration of PV parks, EV charging stations, etc.), new control strategies (e.g., power dispatch prioritization between DER, RES, or other power generation resources), and mitigation methodologies for unexpected events (e.g., faults, equipment failures, cyber-attacks, etc.). The main structural components of such cyber-physical testbeds are depicted in Fig. 2. Below, we provide a list of the possible security-related tasks that can be performed on CPES testbeds:

- Train users and stakeholders in an simulated/emulated CPES environment.
- Validate interoperable systems' performance holistically, i.e., from the lowest level of operation (e.g., sensor, actuators, process, etc.) to the upper levels including communication between assets, distributed control, and monitoring applications.
- Develop and validate cyber-physical metrics and examine system security.
- Test novel security mechanisms such as intrusion detection and prevention systems (IDS/IPS), authentication protocols, and encryption algorithms.
- Evaluate the impact of attacks on the cyber and physical domains of the EPS.
- Examine the effectiveness of mitigation strategies against adverse cyber-physical events.

The importance of cybersecurity research for CPS and critical CPES infrastructures, has led many universities and U.S. national laboratories to develop in-house testbeds. A variety of testbeds have been designed and implemented based on the application field and the research objectives. In Table I, we provide a summary of some of the existing

TABLE I: Cyber-physical testbed architectures, accuracy, repeatability, cost characteristics, and example testbeds with their simulation resources.

Testbed Architecture	Accuracy	Repeatability	Cost	Example Testbed	Resources*	Year	References
Hardware Assisted	High	Low	High	INL	PS: *RTDS* NS: not applicable	2015	[15]–[21]
				NREL	PS: *RTDS*, *Opal-RT* NS: not applicable	2015	[22]–[24]
Software Assisted	Medium	High	Low	Texas A&M	PS: *RTDS* NS: *OPNET*	2014	[25], [26]
				TU Dortmund	PS: *Opal-RT* NS: *OPNET*	2014	[27], [28]
Hybrid: Physical Hardware & Simulated Cyber	Medium	Medium	Medium	FSU-CAPS	PS: *RTDS*, *Opal-RT* NS: *OPNET*, *EXataCPS*	2013	[29]–[32]
				PNNL	PS: *Opal-RT* NS:*OPNET*, *ns-2*, *ns-3*	2017	[33], [34]
Hybrid: Simulated Hardware & Physical Cyber	Medium	Medium	Low	HELICS	PS: *RTDS*, *Opal-RT* NS: *OPNET*, *OMNeT*	2017	[35]–[39], [40]

*PS, and NS stand for the corresponding CPS tetbeds' power and network simulators.

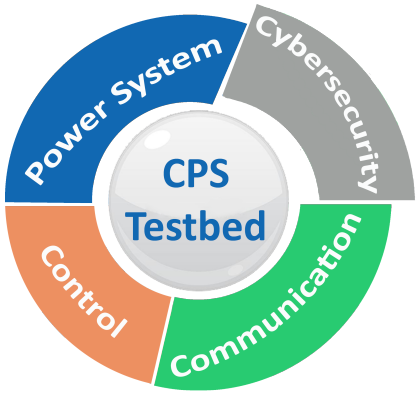


Fig. 2: Cyber-physical testbed components for EPS research.

real-time simulation CPS testbeds along with their inherent resources (i.e., simulation capabilities). We also categorize the cyber-physical testbeds based on their architecture, cost, and accuracy characteristics. Additionally, we present an in-depth overview of the differences between hardware and software-assisted testbeds.

Hardware-assisted testbeds are designed to explicitly study CPS while mostly incorporating a number of actual physical components encountered in the field. For instance, CPES hardware-assisted testbeds integrate physical equipment such as generators, relays, switchgear, energy storage systems – ESS, PV panels, wind turbines, etc. By replicating the behavior of the actual system with a considerable amount of physical equipment, these testbeds provide stakeholders the ability to: *i)* make decisions not only based on theoretical analyses but on practical studies leveraging the use of hardware resources, *ii)* evaluate the CPS behavior under abnormal operational scenarios without inhibiting the operation of the real system, and *iii)* preemptively assess cyber-attack or fault mitigation and control strategies before the corresponding hardware is deployed to the field, and thus, de-risking this cost-prohibitive and unpredictable process. Hardware-assisted testbeds, however, suffer from three major disadvantages: *i)* they are not cost-effective since they require the testbed components to match the actual equipment deployed in the field, *ii)* once the equipment and testbed configuration are setup in-place, any modification or expansion of the system architecture can be either time-consuming or practically and economically infeasible, and *iii)* scalability issues of representing large-scale EPS due to the requirement of procuring more assets (e.g., generators, inverters, etc.).

A typical example of a hardware-assisted research laboratory that leverages actual operational equipment to perform CPES security research is the Idaho National Laboratory (INL) of the U.S. Department of Energy (DOE) [15]. INL's Power and Energy Real-Time laboratory [16], [17], alongside their nuclear laboratory [18], [19], and micrgogrid (MG) testbed [20], [21], allow the simulation of realistic scenarios supported by actual hardware equipment and data generation routines. The real-time simulation capabilities of INL's testbeds allow researchers to create sophisticated scenarios involving power hardware devices that are interfaced with real-time simulation environments via HIL methodologies such as power hardware-in-the-loop (PHIL) and controller hardware-in-the-loop (CHIL) [17]. HIL allows controllers (CHIL) and parts of EPS (PHIL) to be extensively tested before their final integration to the main grid [41]. The National Renewable Energy Laboratory (NREL) of DOE also includes hardware-assisted testbeds [22]. NREL's Flatirons campus specializes in designing, analyzing, and providing accurate simulation models for wind turbines, hydropower, and hydrokinetic generation plants [23]. Their unique facilities drive the improvement of their high-fidelity simulation models, which are cross-

referenced to real assets, providing invaluable tools for power engineers performing system analyses incorporating off-shore, or distributed hydro and wind generation [24]. The actual power system assets of wind turbines and hydro-plants, as well as their simulation models, can be leveraged to investigate the potential impact of component failures or cyber-attack incidents with minimum cost, and most importantly, without compromising the actual EPS operation.

Hardware-assisted CPES testbeds do not exclusively utilize physical equipment. In most cases, the conducted research is supported by simulation software enabling the analysis of more complex systems. Since an actual duplicate of an operational CPS in the lab is typically infeasible, in the past years, a high number of software-based CPS testbeds have been developed following, the notion of digital-twin systems [42], [43]. The main difference between software-assisted testbeds and their hardware-assisted counterparts is that they do not possess any actual field equipment, thus limiting their testing scenarios. Moreover, software-assisted testbeds can be further segmented into sub-categories based on the simulation platform utilized for the system analysis. Some of them utilize widely available software simulators, e.g., *Matlab/Simulink*, *PowerWorld*, *PSSE*, etc., while others rely on real-time simulators (RTS) such as *Opal-RT*, *RTDS*, *Typhoon*, and *Speedgoat*. The main advantage of software-based CPS testbeds, compared to hardware-based testbeds, is the increased flexibility in designing, modifying, and scaling the systems under test. Also, their cost can be significantly lower for simulating large-scale CPES. However, the validity of the software-based simulated results relies heavily on the fidelity of the models (for emulation, virtualization, etc.) used to represent the corresponding real systems under investigation.

Examples of testbed environments with extensive CPS simulation capabilities include the ones at Texas A&M and TU Dortmund. At the Texas A&M CPS testbed, despite the lack of actual EPS equipment, CPES technologies such as smart grid controllers and RES can be virtualized and evaluated using software-based implementations. The testbed also includes RTS systems (*RTDS*) and supports the modeling of communications of CPES components via network simulators (*OPNET*). Furthermore, it allows researchers to evaluate how communication-enabled devices expand the threat surface [25].

The rapid penetration of ICT technologies in CPS is driving the design and development of large-scale software defined network (SDN) testbeds [44]. In such SDN-type testbeds, researchers can evaluate novel network technologies, communication protocols, custom data routing algorithms, etc. An example of such environment is the SDN4SmartGrids CPS testbed at TU Dortmund, where both SDNs and power system RTS are employed for experimentation with ICT-based smart grid applications [27]. In particular, the TU Dortmund's testbed is comprised of a RTS (*Opal-RT*) responsible for simulating the power system components. The infrastructure emulating the network topology and communication between the simulated grid assets (e.g., electric vehicles – EVs, ESS, etc.), management systems, and telemetry units (e.g, phasor measurement units – PMUs, advanced metering infrastructure – AMI, etc.) is implemented using the SDN and the *OPNET* network simulator [28].

In order to bridge the gap between the hardware and software-assisted CPS testbed methodologies, hybrid testbeds are considered as an effective alternative. As their name implies, hybrid approaches trade-off the utilization of the physical components, that can be found at the transmission and distribution (T&D) level of CPES, with the utilization of simulators and software suites designed to accurately represent the behavior of real energy systems. Hybrid testbeds enable diverse security investigations that can focus on the physical-system (e.g., programmable controllers, IEDs, grid assets, etc.), the cyber-system (i.e., SCADA communications, telemetry and remote control of assets, monitoring and measurement components, etc.), or any combination of the two. The main advantage of such testbeds is that they provide re-configurable platforms that can scale up, using simulation, to realistic systems' sizes, while also retaining the ability to investigate, with high granularity, the individual security and control properties present in physical devices. Consequently, hybrid CPS testbeds can evaluate holistically the impact of cyber-attacks on CPES, without any of the limitations encountered in hardware- or software- assisted testbeds.

A prime example of a hybrid CPES testbed framework is HELICS [35], [36]. The HELICS infrastructure enables the integration of different RTS operating at different time-steps as well as the interconnection of T&D system components. By timely simulating (depending on the temporal constraints) complex T&D architectures, cybersecurity assessments, including real-time impact analysis and risk mitigation strategies, can be conducted providing meaningful insights regarding the behavior of CPES [37]–[39]. The Pacific Northwest National Laboratory (PNNL) also features a hybrid testbed leveraging the aforementioned advantages. The testbed facilitates a variety of cybersecurity studies [33], and provides an effective framework for system vulnerability assessments, interactive simulations of CPES environments, threat scenario analyses, and risk mitigation strategy evaluations.

The facilities of the Center for Advanced Power Systems (CAPS) of Florida State University (FSU) also include a hybrid testbed setup. The testbed supports the use of RTS, based on the *RTDS* and *Opal-RT* platforms, power system simulation software such as *OpenDSS*, *PSCAD/EMTDC*, *Matlab/Simulink*, *RT-Lab*, *RSCAD/RTDSphysical*, and EPS components including generators, inverters, and flexible AC transmission systems (FACTS) [29]. The center's infrastructure can be segregated into two main subsystems able to perform both real-time and HIL simulations. The first subsystem is composed of 15 *RTDS*-enabled racks, each consisting of around 26-30 parallel processors. The subsystem can support real-time simulations comprised of more than $1,000$ electrical nodes (e.g., measurement points) and $5,000$ control units at time-steps in the range of 50μs. It should be noted that for time critical implementations, such as power electronics converters, the time-step of real-time simulation can be further reduced in the vicinity of 1μs. Fiber-optic networks facilitate the interconnection between the RTS and the physical EPS equipment. Namely, the physical equipment of the testbed includes a 4.16 kV distribution system, a 7.5 MVA on-site service transformer,

TABLE II: CPES security study categories and research examples.

Security Study Category	Example Literature	Reference
Attacks exploiting CPES vulnerabilities	Low-budget GPS spoofing attacks	[45]
	Load redistribution attacks	[46]
	Coordinated DoS attacks	[47]
	Data integrity attacks on state estimation	[48]
	FDIAs with limited resources	[49]
	Hall sensor spoofing attack	[50]
Evaluation of attack impact on CPES	Graph-theoretic quantitative security assessment	[51]
	Unsupervised learning-based evaluation	[52]
	Simulation-based impact evaluation	[53]
	Deep reinforcement learning analysis	[54]
	Stochastic modeling assessment	[55]
	Quantitative attack impact evaluation on substations	[56]
	Markov-process based reliability analysis	[57]
Attack detection algorithms in CPES	Unsupervised learning-based FDIA detection	[58]
	Wavelet transformation-based FDIAs in AC state estimation	[59]
	Historical data-assisted FDIA detection	[60]
	Distributed collaborative FDIA detector	[61]
	Machine-learning assisted FDIA detection	[62]
	Unsupervised learning anomaly detector	[63]
	Sensor- and process noise-based attack detection	[64]
	Autoencoder-based anomaly detection	[65]
Attack mitigations and defenses in CPES	Semi-supervised method for malware	[66]
	Markov-process based reliability analysis	[57]
	Data-driven and compressive sensing resilient state estimator	[67]
	Battery-based hardware security authentication	[68]
	Battery-power assisted risk mitigation	[69]
	Robust control-based defense mechanism	[70]
	Control flow integrity validation method	[71]
	Hardware security-based communication protocol extension	[72]

a 5 MW variable-voltage variable-frequency converter, a 5 MW dynamometer, and a 1.5 MVA experimental bus at 480 V_{ac} [30]. The second subsystem includes three *Opal-RT*-enabled racks, supported by multiple processor units along with *Xilinx field programmable gate array (FPGA)* computation units. The FPGA hardware accelerators perform the simulation of high-frequency power electronic converters with stringent timing constraints (i.e., in the ns range), while the rest of the EPS is simulated using μs time-steps. Both subsystems have support for multiple industrial protocols utilized for the communications between the physical or simulated EPS assets. Advanced control schemes and experimentation with communication network components are also supported via HIL simulations [31], [73]. Additionally, the impact of unexpected failures or cyber-attacks targeted at these components can be examined in a controlled environment where minimum risk exists [32].

B. CPES Security Studies

During the past decade, significant effort has been exerted into CPES security studies with the objective of enhancing CPES resiliency and alleviating cybersecurity vulnerabilities. For instance, a comprehensive work reviewing cybersecurity vulnerabilities and solutions for smart grid deployments is presented in [74]. Security solution evaluation, system threat classification, and future cybersecurity research directions are also considered. The authors in [75], investigate cyber-attacks on IoT-enabled grid deployments. They discuss how advancements in IoT technologies can drive the power grid modernization process, but at the same time increase the system's threat surface given its interconnected topology encompassing millions of IoT nodes. Researchers in [76] examine the security of modern power systems from the viewpoint of interconnection with microgrids. Emphasis is given on the cybersecurity and reliability challenges arising in these architectures. Essential approaches (e.g., testbed-assisted security studies) are discussed to enhance the security of future power systems. In addition, [77] provides a complete overview of the cyber-threats encountered on the infrastructure, network protocols, and application levels of power systems. Furthermore, attacks targeting the data availability, integrity, and confidentiality of microgrids are discussed in [78].

In this section, we outline the main topics of existing literature in the area. More specifically, the literature work is classified using the following categories: *i)* studies investigating the exploitation of CPES vulnerabilities, *ii)* studies evaluating the impact of cyber-attacks on CPES, *iii)* studies proposing and assessing algorithms (e.g., anomaly detection,

IDS/IPS, etc.) for the detection of cyber-attacks, and *iv)* studies focusing on mitigation and defense mechanisms. In Table II we provide an overview of recent CPES security studies classified under the four aforementioned categories.

1) Attacks Exploiting CPES Vulnerabilities: CPES are advancing towards decentralized interconnected systems in order to support increasing power demand while minimizing transmission losses, leverage MG deployments and their functionalities (e.g., grid connected or autonomous operations), and incorporate DERs. In addition, to enhance CPES control, reliability, and security, digital ICT equipment such as advanced measuring and monitoring units are being employed in geographically dispersed locations of decentralized CPES. For example, PMUs provide time-synchronized (using GPS) granular measurements for EPS related states including voltage, current, and power magnitudes and phase angles. However, it has been demonstrated that adversaries can leverage open-source public resources to perform GPS spoofing attacks against PMUs [45]. By introducing small undetectable timing delays (in the μs range) in the measurement signals (within the IEEE standard limits for synchrophasors C37.118 [79]), the phase differences between actual and measured angles can be significantly altered exceeding allowed limits, tripping circuit breakers (CBs), sectionalizing parts of the EPS, and causing power outages (e.g., brownouts, blackouts) [80].

Moreover, in [46], researchers introduce a coordinated load redistribution attack affecting power dispatch mechanisms. By attacking generators or transmission lines while falsifying load demand and line power flows, system operators are misled into increasing load curtailment. Furthermore, in [47], the authors investigate two types of DoS attacks along with their impact on EPS. The first attack is assumed to be a stealthy false data injection attack (FDIA) performed to mask the impact of the attack from detection algorithms. The second, assumed as a non-stealthy attack, aims to maximize the damage on power system operation by targeting the most vulnerable transmission line, impeding power dispatch, and causing load shedding. In [48], the authors propose a hybrid data integrity and data availability attack. They demonstrate how control center measurements can be manipulated leading to undetectable FDIAs. In more detail, by modifying some measurements (i.e., integrity attack) while making some others unavailable to the state estimation algorithm (i.e., availability attack), FDIAs can bypass bad data detection algorithms.

Ubiquitous power electronics bring new challenges in CPES operation. Future CPES are expected to be inverter-dominated systems. As such, vulnerabilities in such components can lead to abnormal system operations. In [50], the authors investigate how stealthy non-invasive attacks on grid-tied inverters can compromise their nominal operation and impact grid operation. Specifically, by spoofing the inverter's hall sensor they demonstrate fluctuations on the output voltage, active and reactive power while also introducing low-frequency harmonics to the grid. Similarly, by exploiting a vulnerability in the authentication mechanism of General Electric Multilin protection and control devices, the authors in [81] show that remote or local attackers can obtain weakly encrypted user passwords, which could then be reversed allowing unauthorized access. Furthermore, the authors in [12], [13] show that by coordinating the power usage of multiple devices, power reserve limits of EPS can be exceeded causing tripping of lines and shedding of loads. A botnet of IoT (internet-of-things)-connected high-wattage loads, such as washing machines, air-conditioning units, dryers, etc., are coordinated over the network, causing unexpected power usage profiles and pushing the grid to instability limits. Such attacks demonstrate that there is no requirement of strong adversarial knowledge nor considerable attack resources [49].

2) Evaluation of Attack Impacts on CPES: Impact evaluation and analysis studies are considered essential for prioritizing and safeguarding critical components in CPES. Such analyses explore the consequences of malicious attacks and can serve to proactively prepare systems for their adverse implications. Impact evaluations can expose critical system components, assist in prioritizing and securing them, and aid in the development of contingency plans in case these vulnerable components get compromised. For instance, the authors in [51] propose assessment metrics designed to evaluate the resiliency of CPES against adversarial attacks. Different techniques from game theory, graph theory, and probabilistic modeling have been utilized to assess the capability of CPES when supporting critical (or unsheddable) loads after they have been compromised or the system has suffered unexpected disturbances. Other works focus on analyzing the impact of cyber-attacks in TES [52]. Here, the authors investigate the system operation under two types of attacks that are designed to maliciously affect either the bid prices or the bid quantities. In view of the fact that IEDs, AMI, and smart inverters are penetrating EPS at a rapid pace, the authors in [53] and [82] demonstrate the adverse grid consequences if such devices are compromised. Specifically, the simulated impact of malicious smart inverter firmware modifications in MGs is demonstrated in [53]. Attacks targeting SCADA-controlled switching devices or monitoring devices impeding situational awareness (in an integrated T&D system model) are evaluated in [82].

Furthermore, cybersecurity assessment methodologies investigating the impact of RES integration to the grid are also investigated in the literature. For instance, the authors in [54] leverage open source intelligence and contingency analysis methods to discover the most critical system paths. Such transition paths could be utilized by an adversary to maximize the impact of cyber-attacks, leading to disastrous consequences for the EPS. A different approach which considers intrusion and disruption process modeling is proposed in [55], in which a stochastic game theory-based CPES security evaluation model is developed. The authors in [56] propose a mathematical framework to estimate the probability and evaluate the impact of malicious attacks on substation automation systems. In [57], the reliability and security of CPES are analyzed through a communication failure assessment process. Overall, assessment methodologies of attack impacts on CPES are designed with the purpose of aiding CPES evaluation studies. Thus, they should be leveraged as part of a defense-in-depth (DiD) portfolio when assessing potential damages and devising CPES defense strategies.

3) Attack Detection Algorithms in CPES: The severity of the effects of cyber-attacks in CPES underlines the need for accurate and effective attack detection mechanisms which can improve the situational awareness of system operators. Hence, remediation actions can be issued to avoid system and equipment failures, as well as ensure human safety. A plethora of detection schemes have been proposed especially for FDIAs in CPES [58]–[61]. For instance, in [61] researchers develop a distributed host-based collaborative mechanism for detecting false data measurements in PMUs. Each PMU is assigned a host monitor to probe its status (i.e., normal operation or anomalous) by comparing it with predefined nominal values. Then, a majority voting algorithm is executed to decide if the acquired measurements are valid by comparing the status of the under-investigation PMU with the corresponding neighboring PMUs. Unsupervised learning-based anomaly detection methods have also been proposed for cyber-attack detection in CPES [62]. An example of such anomaly detection scheme is presented in [63], where authors identify suspicious sensor activity using recurrent neural networks (RNNs). Other researchers have also demonstrated how data integrity attacks (DIA) can be identified when sensor and process patterns deviate from a residual-based fingerprinted data [64]. Furthermore, given the extensive use of Fieldbus communication devices in CPES, methodologies have been designed to detect anomalous network traffic in a variety of Fieldbus protocols [65]. All of the reviewed detection mechanisms have the objective of notifying system operators once incongruous sensor or monitor behavior is detected in the CPES. As a result, malicious incidents can be effectively handled, minimizing their impact on CPES operations.

4) Attack Mitigations and Defenses in CPES: The deployment of defense and mitigation mechanisms is critical to enhance the overall CPES security and minimize the adverse impact of cyber-attack scenarios. For example, mitigation strategies can protect CPES against FDIAs which could potentially result in generator equipment damage [69]. Specifically, BESS could be leveraged to assist the generators and reduce the load curtailment inflicted by malicious attacks. A hybrid control-based approach to safeguard systems against cyber-attacks is presented in [70]. The hybrid controller switches to the most secure controller, from a subset of available controllers, given that some of these controllers might have been compromised by an adversary. In [66], a semi-supervised learning mechanism is utilized to study malware patterns and defend the system from unknown malware targeting the CPES infrastructure.

Apart from software-based mitigation techniques and defenses, hardware-oriented mechanisms have also been proposed. In [68], the authors propose the use of hardware security primitives leveraging the intrinsic variation of BESS lithium cells to enhance communication protocol security. The practicality of the approach is validated in a simulated testbed environment [72]. Furthermore, in [71], an instrumentation-based defense technique is presented employing a sub-optimal plan to secure CPES in real-time. Even though the discussed defense and mitigation mechanisms may not be applicable for all cyber-attack scenarios, research and development in this direction contribute towards understanding attackers' tactics and defending against them in order to secure CPES.

C. Threat Analysis and Risk Assessment

Precise modeling is essential in order to investigate complex CPES architectures, discover any potential vulnerabilities, and extensively test and evaluate security features. The intricacies of CPS typically consist of multiple interconnected layers bridging assets of varying importance for the system operation, and leveraging ICT and communication protocols. Different methods are being used to review CPS architectures and assess their cybersecurity. Among them, the DiD and the Purdue models are the most popular ones. The DiD strategy was initially employed in military applications [83]. It ensures resiliency, redundancy, and the existence of multiple defenses if a vulnerability is exploited, a critical security flaw is identified, or a failure or unintentional fault occurs. Enforcing the DiD multi-layered topology has two main advantages from a security perspective. First, it delays the attack progress in the system since each layer provides an isolated execution environment. Second, it allows system operators to deal with the attack independently on multiple layers, rather than having to rely on a single point-of-defense. Similarly, the Purdue model for industrial control system (ICS) network segmentation [84], part of the Purdue Enterprise Reference Architecture (PERA), incorporates the DiD concept by demonstrating the interconnections and dependencies between layers and components, allowing for the design of secure CPS [85]. In the following parts (II-C1 and II-C2), we provide the essential information and related work regarding threat modeling and risk assessment methodologies with emphasis on industrial CPS and critical infrastructures.

1) Threat modeling: The term '*threat modeling*' refers to the procedure by which potential vulnerabilities are discovered before they can become system threats. This process is crucial for the design of security defenses and mitigation strategies. It is evident that performing threat modeling for CPES is essential since their compromise can have disastrous consequences to the grid operation and the economic and social well-being. However, CPES consist of multiple layers and assets, hence, it can be challenging, due to extensive time, modeling efforts, resources, and cost, to exhaustively examine all the possible scenarios that could arise as system vulnerabilities. To overcome such issues, without compromising the system's reliability, multiple threat modeling approaches have been proposed aiming to prioritize vulnerabilities and assist the implementation of potent security mechanisms. These methodologies provide a holistic view of the system by highlighting the significant assets, commonly referred to as crown-jewels [86], and assessing threats based on their potential impact and ease of deployment on the system.

STRIDE[2] and DREAD[3] are well-established threat modeling frameworks for the security assessment of products and

[2]STRIDE is an acronym for Spoofing, Tampering, Repudiation, Information disclosure, Denial-of-service, and Elevation of privilege.

[3]DREAD is also an acronym, and stands for Damage, Reproducibility, Exploitability, Affected Users, and Discoverability.

services throughout their life-cycle [87], [88]. For instance, STRIDE uses data flow diagrams for the threat modeling process. The data flow diagrams map system threats to the corresponding vulnerable system components (STRIDE per element approach). Given the interdependent nature of CPES, an attacker can compromise the system operation by exploiting different component vulnerabilities. Therefore, to guarantee the overall system security, vulnerabilities need to be addressed both at the component level as well as within the component interrelations (visualized in the data flow diagrams) [89]. DREAD can be leveraged to evaluate and rank the severity of threats. DREAD analysis is comprised of the following six steps: asset identification, system architecture formation, application decomposition, threat identification, threat documentation, and threat impact rating. DREAD and STRIDE methodologies can also be used jointly for comprehensive cybersecurity assessments [90].

Apart from STRIDE and DREAD, other methodologies for security assessments have been proposed and utilized in the cybersecurity arena. For instance, OCTAVE[4] Allegro is an alternative approach used by organizations when performing mainly information technology (IT) security evaluations and strategic planning for cyber-threats [91]. However, recent works validate the applicability of OCTAVE Allegro for CPS security assessments, both for the enumeration of potential risks as well as the design of countermeasures to maintain nominal system operation [82], [92]. The main steps followed in OCTAVE security assessments include: the development of risk evaluation criteria according to operational constraints, critical asset identification, critical asset vulnerabilities and corresponding threats discovery, and threat impact assessment. STRIDE, DREAD, and OCTAVE are well established tools when performing threat modeling analyses and identifying vulnerabilities in the pre-attack context.

The investigation of adversary behavior post-compromise is also important. At this point, the adversary has already overcome the first line of defense and has access to system resources. Notably, there is extensive research on initial exploitation and use of perimeter defenses [93], [94]. However, there is a knowledge gap of the adversary process after initial access has been gained. To address the aforementioned pitfall and support threat modeling, risk analysis, and mitigation methodologies, pre- and post-compromise events, MITRE developed the ATT&CK for Enterprise framework [95].

MITRE ATT&CK is an open-source knowledge-base that includes common adversarial attack patterns (e.g., attacks, techniques, and tactics). The ATT&CK database is constantly being updated with recent attack incidents to enhance enterprise cybersecurity by exposing system vulnerabilities and warrant safer operational environments for businesses and organizations. The framework describes the tactics, techniques, and procedures (TTPs) that an adversary could follow in order to make decisions, expand access, and stealthily compromise an organization while residing inside the enterprise network [96], [97]. In January 2020, MITRE corporation, realizing that ICS is an essential part of critical CPS infrastructures and with the objective of addressing cybersecurity issues arising by the diverse and interconnected nature of CPS, launched the ATT&CK for ICS framework [98].

The ATT&CK for ICS framework is also a free community-supported threat knowledge-base that includes information about TTPs that adversaries utilize when targeting ICS within CPS. The framework assists in understanding the adversarial attack chain and enhance the security standpoint of ICS and related CPS assets. ATT&CK for ICS is based on MITRE's ATT&CK for Enterprise framework, i.e., it ports many of the gathered threat intelligence from enterprise networks to ICS since industrial networks often have similarities with enterprise networks. The heterogeneity of ICS, however, with a plethora of operating systems (OS), network devices, and communications protocols co-existing with a variety of field devices (e.g., PLCs, IEDs, PMUs, RTUs, etc.) led to significant revisions from the ATT&CK for Enterprise to the ATT&CK for ICS.

The ATT&CK for ICS framework is designed to support a multi-layer reference approach for adversarial behavior evaluations. The framework is segregated into four core components, making it applicable to a wide spectrum of industrial CPS. The first component category includes *i) assets* which consist of control servers, engineering workstations, field controllers, human-machine interface (HMI), among others. All these assets might not be apparent in every system. This is factored by the ATT&CK methodology which investigates attacks targeting the respective assets independently as well as their cooperation with other industrial assets. The second core part of ATT&CK for ICS is the abstraction to focus on the *ii) functional levels* of the Purdue architecture. Such levels describe the depth of infiltration that the adversary has achieved. The level ranges from *Level 0*, which corresponds to the physical devices (e.g., sensors and actuators) that orchestrate the industrial process, all the way to *Level 2*, which includes the supervisory control systems, the engineering workstations, and HMIs. These functional levels are depicted in Table III. The last two parts of the framework revolve around the adversarial *iii) tactics* and *iv) techniques*. The term '*tactics*' refers to the reason why an adversary performs an action, i.e., adversary objective such as disrupting an industrial process control routine. *Techniques* describe the activities that the adversary uses to achieve the attack goal, i.e., represent "how" an attacker accomplishes his/her objectives by taking an action, e.g., through modifying the PLC control logic.

2) Risk Assessment: The term "*risk assessment*" refers to the process of identifying potential risks and their corresponding impact to the system operation as well as determining strategies to mitigate, defer or, accept these risks based on their criticality [91]. Cyber-threat risk assessment is a critical operation that CPES and their ICS need to perform regularly. The introduction of new technologies into CPES (i.e., DERs, EVs, control devices, etc.) along with the interoperable nature of the supported ICT infrastructure increases the risks arising from both the cyber (e.g., measurement, control commands or communication integrity attacks) and the physical domain (e.g., sensor and actuator compromise).

[4]OCTAVE acronym is for Operationally Critical Threat, Asset, and Vulnerability Evaluation.

TABLE III: ICS functional levels, equipment categories, and their corresponding components [99].

Functional Level	Category	Components
Level 2	Supervisory Equipment	Supervisory control functions, site monitoring, local displays
Level 1	Control Equipment	Protection devices, local control devices
Level 0	Process Control	Sensors, actuators

Typically, risk assessment methodologies rely on probabilistic analyses that leverage Markov-chains [100], Petri-nets [101], Bayesian belief networks [102], or game theory to estimate the impact of adverse events on system operation [103], [104]. In [104], for example, researchers model both the attackers and the system's defenses as agents with different action sets and objectives. Due to the contradictory roles of such agents, the corresponding action payoff depends on the ability to compromise the system's assets or the ability to detect the malicious attack from the perspective of attackers or defenders, respectively. Other works have proposed worst-case scenario risk assessment analyses that employ exhaustive Monte Carlo simulations and focus on diverse operation areas of EPS (e.g., automatic generation control –AGC, T&D system operations, etc.). Then, the interdependence of such EPS areas with specific risk mitigation mechanisms is analyzed [105], [106]. For instance, the authors in [106], review the impact on buses and transmission lines under abnormal operations caused by cyber-attacks. They also investigate how adverse scenarios can be mitigated if robust protection system strategies, i.e. coordinated bus and transmission line trippings, are correspondingly put in-place. Although probabilistic risk analyses and worst-case scenario assessments can provide useful results under specific constraints (i.e., if only part of a system is examined), applying such methods to dynamically changing large-scale T&D integrated models can be a challenging task. The multitude of T&D assets expands the search space of exhaustive methods such as Monte Carlo-based risk analyses [107]. For each asset and every investigated potential attack, the risk analysis process needs to be re-examined and re-computed. The risk calculation overhead is also exacerbated due to the interconnected CPS architecture.

The aforementioned methods, apart from being computationally intensive, they can also potentially suffer from poor accuracy. The security risk assessment accuracy of these methods relies on the precise modeling of the CPES physical components (e.g., generators, transmission lines, substations, etc.), their topology, as well as their interconnections with the cyber components (e.g., ICT nodes supporting EPS functions) [108]–[110]. Failure to properly model CPES can mask interconnection dependencies between components and their layers (cyber or physical), and thus, perturb the risk score calculation process. The presented risk assessment approaches in this section are credible if security assessment is performed partially, i.e., they fail to capture comprehensively system risks as their focus is on specific parts of a CPES ignoring the impact propagation to the rest of the infrastructure. In this work, the threat and system representation is performed meticulously during the threat modeling process (Section III) and the CPS framework stages (Section IV), respectively. As a result, our approach determines in advance a detailed system model, overcoming the drawbacks encountered when performing segmented risk evaluations.

In our analysis, system-specific characteristics are formalized and *Risk* scores are calculated by combining the attack *Threat Probability* along with the CPES objective priorities (Section III-C). The proposed methodology expedites the risk assessment analysis of CPS (since the threat modeling, CPS framework analysis, and performance metrics determination have been performed previously), and thus, mitigation policies can be evaluated iteratively until the corresponding *Risk* goals are met. For example, if an EPS asset is compromised, there might be multiple defense mechanisms that could be enforced to mitigate the attack. However, the implementation of some of those mechanisms might result in significant impacts (e.g., uneconomic operation, partial grid disconnections, etc.) or affect other parts of the system due to its interdependent nature. The ability to evaluate, in real-time, the effectiveness of risk mitigation mechanisms provides significant benefits for CPS, aiming to balance security objectives and system performance.

III. Threat Modeling for Cyber-Physical Systems

The fundamental property of any adverse failure is an artifact of the semantics and capabilities of building CPS from a diverse, possibly infinite, set of ways. It is crucial to mitigate any adverse event in CPS, regardless of whether it is accidental or intentional. However, there are distinctions that need to be made between these two types. For example, there is a high probability that a natural adverse event (e.g., short circuit fault) can be detected by the process, considering a built-in fault detection scheme in the system. In contrast, an intentional fault (possibly caused by an attacker) could alter the results of the system in a congruous way, hence causing the event to go undetected. Traditionally, fault monitoring and detection approaches do not consider the implications that arise due to adversaries and their attack goals. Their aim is solely to recover from transient faults overlooking the actions which trigger this abnormal behavior. Without considering a threat model that includes malicious and motivated adversaries, as well as sophisticated attacks, defense detection schemes can be potentially evaded by attackers entirely, despite the redundancy already built into control processes. A fault can become an exploited vulnerability and the compromised component, if not sanitized properly, can pose a danger to the entire CPS.

The complex nature of CPS, and consequently CPES, urges the identification of attack vectors on both the cyber and the physical domains of the system. Adversaries are constantly improving, adapting, and modifying their attack patterns to evade security mechanisms. As a consequence,

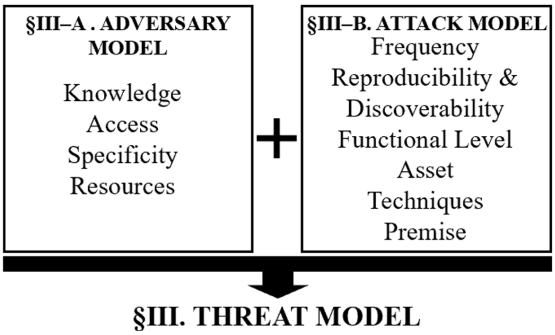


Fig. 3: Adversary model and attack model components comprising the comprehensive threat model architecture.

security researchers cannot passively await until an asset in the system is compromised to initiate remediation. To support the identification, anticipation, and mitigation of cyber-attacks in CPS, we develop a holistic threat model which incorporates the core components of MITRE's ATT&CK for ICS methodology while providing an additional dimension for security investigations. Specifically, the presented threat modeling approach extends MITRE's methods since:

- We incorporate an *adversary model* to allow for more granular and explicit threat modeling analyses.
- We rigorously define all aspects of potential cyber-attacks so that they can be implemented in CPS testbeds for security evaluations (e.g., evaluate defense mechanisms, mitigation strategy, detection schemes, etc.).
- We perform risk assessments considering the actual impact of cyber-attacks on the CPS and leveraging both the threat modeling and CPS framework resource mapping. Hence, every possible attacked CPS component is accounted towards the risk score calculation, aiding threat prioritization, and CPS security posture awareness.

In the developed threat modeling methodology, we evaluate threats and prioritize them based on the degradation that they can potentially inflict on the CPS. Our threat model consists of two major components, the *adversary model* and the *attack model*, as illustrated in Fig. 3. To understand the security implications of threats targeting CPS, the adversary model needs to capture specific information involving the adversary's capabilities, intentions, and objectives. In addition, it is essential to model attacks based on their specific methodology, targeted system component, and system impact, as well as define rules that enable multi-layer and severity attack analyses. The adversary and attack models compose the threat score index factored in the threat risk calculation process presented in Section III-C. For instance, the threat score of an attack performed by a stealthy and motivated adversary will be higher than the threat score of the same attack performed by an adversary with limited resources and oblivious knowledge about the system. Our versatile threat modeling approach can support various types of malicious events and enable end-users to adjust the desired level of threat model granularity.

A. Adversary Model

The capabilities of an attacker and the characteristics of the adversary model can be captured by factors such as resources, skills, knowledge of the system, access privileges, and opportunities (i.e., the means to carry out the attack and the number of failed attempts allowed) required to perform the attack. When it comes to system knowledge, distinction has to be made between *white-box* attacks, where an adversary has *complete* information, and *black-box* or *gray-box* attacks, where an adversary has *limited* information about the system [111]. In the gray-box threat model, the adversarial knowledge is limited to the target model, while in black-box attacks adversaries do not know the target model and can only query to generate adversarial samples [112]. We port such classification in the context of CPES, in which attackers may have *full, partial, or zero knowledge* of the system model and real-time power grid measurements. Existing work often assumes that adversaries have perfect knowledge of the system model, i.e., the information needed to create the measurement matrix (Jacobian) of the power system that depends on the network topology, the parameters of power lines, and the location of RTUs and PMUs [113]. However, in realistic attack scenarios, adversaries have limited knowledge of the system due to the dispersed, interconnected, and complex nature of the power grid, the restricted access to CPES control and monitoring functions, and errors in the data collection process [49], [114], [115].

Our adversary model takes into consideration the presented distinctions and defines a hierarchy of the available information to the attackers in order to characterize their knowledge capabilities. At the lowest level of the system knowledge hierarchy is an adversary that has no information about the system model. At the highest level is an adversary that knows the model characteristics, the algorithmic details, and all the grid measurements. In order for the adversary, however, to acquire system measurements or perform reconnaissance and monitoring, he/she should have – to some extent – access to the system. As such, our model delineates this access as the *accessibility level* an adversary needs to have in the target CPS.

In the MITRE ATT&CK for ICS framework, the term '*attacker tactics*' covers the attackers' access level with their corresponding intentions and objectives. In our modeling approach, however, the term is captured in two sub-categories, *access* and *specificity*, to allow for a more elaborate adversary classification. The access category defines the degree to which an attacker can interact with a system asset, while the adversarial specificity encapsulates the objectives and goals of the adversary who executes the attack on the CPS. Adversarial objectives are broadly lumped under *targeted* and *non-targeted*. In the case of targeted objectives, an adversary's intention is to execute an attack which can result in a specific target output (e.g., the miscalculation of EPS system states, due to topology modifications of wind integrated resources [116], or FDIAs [117]). On the other hand, non-targeted attacks can generically maximize malformed outputs of CPS algorithms (with respect to the ground truth) affecting the

operational reliability of the system. Finally, our adversary model also captures the adversary *resources*, differentiating between attackers with limited resources and attackers with a variety of intellectual and physical assets at their disposal.

Adversary Model Formulation: Our attacker model is decomposed into four dimensions: adversarial knowledge, resources, access, and specificity:

1) *Adversary Knowledge*

a) *Strong-knowledge adversary*: White-box attacks assume an adversary with *full knowledge* of the system model, parameters, and state vectors.

b) *Limited-knowledge adversary*: Gray-box attacks assume an adversary with *some knowledge* of the system's internals with partial understanding of the network and system model.

c) *Oblivious-knowledge adversary*: Black-box attacks assume an adversary with *zero knowledge* about the details of the system and can only estimate the system outputs using confidence scores. In such scenarios, the attacker does not have knowledge in regard to the system model.

2) *Adversary Access*

a) *Possession*: This type of attack requires the adversary to have physical access to the attacked component (e.g., IED, solar inverter, transformer, etc.) operating either in the digital or analog domain. The access could involve chassis intrusions (e.g., microprobing, memory flashing, circuit bending, etc.), or interface access to the device (e.g., side-channel analysis, power analysis, protocol decode, etc.).

b) *Non-possession*: In this type of attacks, the adversary cannot physically manipulate the asset under attack. Attacks can be performed leveraging proximity access (e.g., GPS spoofing, side-channel analysis), or by exploiting the network interfaces (e.g., replay attacks, rollback attacks, etc.).

3) *Adversarial Specificity*

a) *Targeted* attacks occur in multi-class identification, control, and monitoring -based scenarios and misclassify CPS algorithms and operations to a specific malicious result category $x_j \in \mathcal{X}$ from all possible results $\mathcal{X}$. The adversary goal is to maximize the probability of the *targeted* class, i.e., maximize $\mathcal{P}(x_j)$.

b) *Non-targeted* attacks are similar to *targeted* attacks in terms of misclassification objective, however, the selection of category x_j is relaxed to any arbitrary output category except the correct one x_i.

4) *Adversarial Resources*

a) *Class-I* attackers that, despite their adversarial motivation, do not have the financial resources, equipment support, or access privileges, to successfully realize any attack without being detected.

b) *Class-II* attackers that can be funded individuals, organizations, or nation-state actors with large budgets and substantial access privileges, skills, and tools capable of realizing sophisticated attacks.

B. Attack Model

The second part of the proposed threat modeling method focuses on the specific characteristics of malicious attacks (e.g., frequency, reproducibility), the targeted CPS components, and the process aiming to achieve the system compromise. The attack model improves MITRE's taxonomy, which includes concepts such as the attack levels, assets and techniques, by incorporating other dimensions necessary for holistic security investigations. For instance, particular attention is drawn on aspects like the attack frequency, reproducibility/discoverability, along with the premise of the compromise. The aforementioned features enable comprehensive characterization of the attacks elucidating all their underlying elements, and as a result, they assist in performing threat and system impact evaluations for CPS environments.

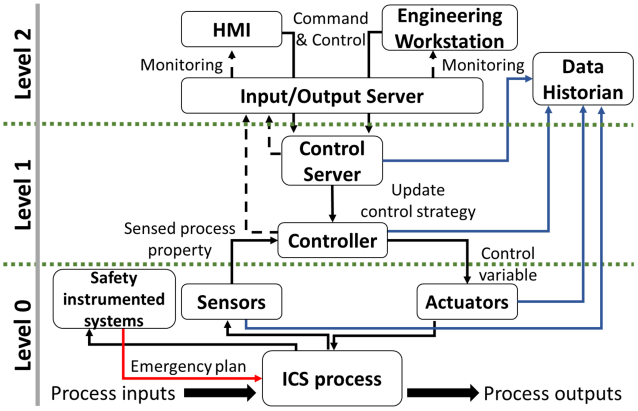


Fig. 4: CPS functional attack levels and assets overview.

The presented attack model accounts for the CPS structure and interconnections. Given that the same adversarial objective can be achieved following different attack paths, propagation scenarios with diverse attack entry points should be investigated. These attack paths can be initiated from process control devices such as sensors or actuators and propagate to supervisory and control equipment like HMIs. In particular, the attack model considers the attack *frequency*, i.e., the number of compromises required to achieve a particular adversarial objective, and the attack *reproducibility and discoverability*. The aspects of reproducibility and discoverability are crucial for CPS risk evaluations. This is attributed to the fact that even catastrophic attacks might not pose any actual danger for the CPS if materializing them is nearly impossible, or they can be easily discovered during their initial stages. The attack *functional level*, attacked *asset*, and attack *techniques* notions correspond to the definitions introduced in MITRE [98]. The only difference is that the selected attack techniques in our methodology represent some of the most common use cases encountered specifically in CPES. An overview of the CPS functional levels along with the corresponding attacked assets are illustrated in Fig. 4. Also, we consider the attack *premise* which indicates whether the attack is targeting the physical or cyber domain of a CPS, trailing the attack path origin and its expected impact. The formulated attack model with the additional aspects of attack frequency, attack reproducibility and discoverability, and attack premise allows to fine-tune each attack case study's model, and overall compose a well-defined CPS threat modeling approach.

Attack Model Formulation: Our attack model is decomposed into six dimensions: attack frequency, attack reproducibility, attack discoverability, attack level, attacked asset, attack techniques, and attack premise.

1) *Attack Frequency*

a) *Iterative attacks*: attacks that need multiple iterations to achieve the desired malicious output.

b) *Non-Iterative attacks*: attacks that only need to be realized once to achieve the desired malicious output.

2) *Attack Reproducibility and Discoverability*

a) *One-time attacks*: attacks that can only be realized once since they are detected after the first attempt.

b) *Multiple-times attacks*: attacks that can be reproduced multiple-times before they are identified and detected.

3) *Attack Functional Level*

a) *Level 0*: attacks that target CPS processes and their corresponding operational equipment (e.g., sensors, actuators, etc.).

b) *Level 1*: attacks that target the industrial control network (e.g., PLCs, system controllers, RTUs, etc.) and aim to stealthily manipulate functions that control CPS processes.

c) *Level 2*: attacks that target the SCADA, and monitoring devices (e.g., HMIs, engineering workstations, data historians, etc.) on the network level (i.e., LAN) overseeing CPS processes.

4) *Attacked Asset*

a) *Field controllers:* Such assets are low-level embedded devices (e.g., RTUs, PLCs, IEDs) that enable the control of CPS processes. They typically possess limited computation capabilities and they are in charge of coordinating industrial processes (e.g., generator governors, manufacturing process controllers, etc.).

b) *Control servers:* These devices cover the functionality of both programmable controllers (e.g., PLCs) as well as communication servers (e.g., SCADA master terminal units (MTU), distributed control servers, etc.). Thus, apart from interfacing with low-level CPS devices (e.g., sensors, actuators), they can also support software-based services in industrial environments.

c) *Safety instrumented systems (SIS):* These systems (e.g., protective relays, recloser controllers) are designed to perform automated remediation actions if an abnormal system behavior is detected (e.g., short-circuit, fault, etc.). The goal of protection systems is to keep the industrial CPS plant online, while avoiding hazard conditions.

d) *Engineering workstations:* These units are usually powerful and reliable computing configurations used for the monitoring and control of CPS, processes, and equipment. They are often accompanied by hardware components and software packages which enable the CPS supervision.

e) *Data historians:* Such elements are databases used to keep records and store process data. This information is stored in a time-series format that enables the examination, display, and statistical analysis of process control information.

f) *Human-machine interfaces (HMIs):* A graphical user interface that enables users to monitor system operations, diagnose malfunctioning system behavior, and initiate control and mitigation actions. HMIs can vary between vendors supporting different capabilities, graphical representations, and control interfaces (e.g., web-based, LAN-based, etc.). Additionally, different user groups can have access to different HMIs according to the systems they are monitoring and their clearance level for managing the CPS.

g) *Input/output (I/O) servers:* Such servers constitute the connecting link between system applications and the field devices which coordinate the ICS equipment under the control subsystems directions. I/O and data acquisition servers (DAS) operate as buffers since they can convert low-level control system data to packets, and forward them to the supervision locations (e.g., HMIs, engineering workstations). Additionally, they serve as intermediate translation units as they collect information from field devices (utilizing diverse communication technologies) and translate them to the predefined formats expected by system applications.

5) *Attack Techniques*

a) *Modify control logic:* In such attacks adversaries can cause the CPS to operate abnormally by modifying the code running on the system's control devices (e.g., PLC, RTU, IED). These system devices are orchestrating physical processes via actuators and other field equipment.

b) *Wireless compromise:* In these attack scenarios, adversaries can gain unauthorized remote access to the CPS network by exploiting: the vulnerabilities of devices with wireless connectivity, insecure wireless communication protocols, and/or network connections leaking sensitive information.

c) *Engineering workstation compromise:* In such attack setups adversaries, after granted access to a CPS engineering workstation, can cause system malfunctions via compromising CPS configurations controlled by engineering workstations, e.g., security systems, process controls, ICT infrastructure, etc.

d) *Denial-of-service (DoS):* Malicious adversaries performing DoS attacks can compromise a CPS asset by inhibiting its nominal functionality rendering it unresponsive. For instance, overflowing a device with artificial data, blocking its inbound or outbound communications, or even suspending/disrupting its operation can impact time-critical CPS.

e) *Man-in-the-middle (MitM):* During MitM attacks adversaries can maliciously intercept, modify, delay, block, and/or inject data streams exchanged between CPS asset communications. Depending on the adversary access level on the CPS networks, numerous attacks (e.g., modify or inject control commands, delay alarm messages, etc.) can be planted affecting CPS operations.

f) *Spoof reporting messages:* Adversaries performing this type of attack can broadcast malicious modified system messages. The attack goal is to either impact CPS operations by limiting the situational awareness (e.g., suppressing critical alarm messages), or misreport information (e.g., sensor measurements), thus, driving systems to unstable and potentially irreversible states.

g) *Module firmware:* In module firmware attack cases adversaries can upload maliciously modified code to embedded devices of CPS (e.g., PLCs, smart inverters, etc.). These actions can affect devices operation via modification of their

control objectives, and/or insertion of backdoor features (e.g., remote access, exploit system logs, etc.) allowing to stealthily manipulate CPS assets.

h) *Rootkits:* In this type of attacks adversaries employ rootkits, typically planted in the OS of devices, to disguise malicious software, services, files, network connections ports, etc. Rootkits provide attackers with user or even root-level privileges while hiding their presence from CPS defense mechanisms.

6) *Attack Premise*

a) *Attacks targeting the cyber domain*

i) *Communications and protocols:* refers to attacks targeting the in-transit CPS data, i.e., exchanged communications data including remote access credentials, measurements, system reports and warnings, etc., with the objective to get unauthorized access (data espionage) or insert malicious modifications (data alteration).

ii) *Asset control commands:* includes attacks targeting the CPS data integrity, i.e., mask counterfeit system data as genuine and unmodified, and trustworthiness (e.g., impersonate authorized user groups and issue, access, or modify control commands).

iii) *Data storage:* accounts for attacks targeting the accuracy and non-repudiation of CPS data (e.g., logs and history records of all the performed tasks such as asset setpoint modifications, user sign-ins and action histories, inbound/outbound connections and traffic, etc.).

b) *Attacks targeting the physical domain*

i) *Invasive:* attacks that require physical access to the CPS asset (e.g., PLC hardware including micro-controller, memory, integrated circuit – IC, etc.) in order to manipulate it (e.g., desoldering, depackaging) [118]. These attacks are time-consuming and require specialized equipment, however, they are difficult to detect.

ii) *Non-Invasive:* attacks that do not require any physical tampering of the ICs residing on the CPS assets, and performing them multiple times can be achieved with minimum effort. No traces are left after the attack is performed rendering them the most difficult type of attacks to detect. Common examples of non-invasive attacks include power analysis attacks, timing attacks, electromagnetic emission attacks, brute force attacks through physical means, hall sensor spoofing, etc. [50].

iii) *Semi-Invasive:* attacks that are a trade-off between invasive and non-invasive attacks, given that they are not as difficult to perform as invasive attacks and can be easily performed multiple times similar to non-invasive ones [119]. Common examples of semi-invasive attacks include fault injection, laser scanning, ultra-violet radiation, or control process tampering [120].

C. Risk Assessment Methodology

Risk assessment is a fundamental process in every cybersecurity analysis study. Its importance is further accentuated in the context of mission-critical CPS where operational disruptions can have disastrous impacts. Existing efforts often port IT risk assessment methodologies into operational technology (OT) security evaluations, and consequently, fail to holistically capture CPS constraints and objective [121]. Some key differences between IT and OT security revolve around the risks associated with loss of operation, asset availability, communication latency, architectural differences and contingency management strategies [122]–[124].

Qualitative assessments for cybersecurity risks require substantial system knowledge of the CPS structure and experience from the organizations and groups conducting the analysis [121], [125]. On the other hand, quantitative studies calculate exact risk scores aiding the prioritization and mitigation procedures [102], [126]. Other works employ simulation-assisted investigations in order to evaluate the corresponding impact of cyber-attacks [127]. Moreover, researchers have also considered dynamically adapting risk assessment models factoring the system and attack impact evolution for the risk score calculations [128]. Recent works have proposed combinations of different risk methods harnessing the advantages of more than one strategies and providing more realistic evaluations [129]–[131]. This mixed approach is motivated by the fact that in CPS we can have the same impact on system operation using different attack paths. Thus, although the system impact remains the same, the risk scores of these attacks would substantially differ. For example, such scenarios, i.e., following different attack procedures to achieve the same adversarial objective, would be difficult to capture using a qualitative-only risk assessment method.

In this paper, we utilize a hybrid risk assessment method bridging the advantages of both quantitative and qualitative methods. The hybrid approach adapts to dynamic system operation and adjusts risk scores based on the current system state. Specifically, we assess qualitatively the impact of attacks. To calculate the corresponding attack damage, however, we quantitatively prioritize CPS objectives. The threat probability is also assessed quantitatively to weigh the attack damage and model the risk. It is important to note that both the objective priority as well as the threat probabilities can change during the real-time system operation.

Such scenarios can be accommodated by our risk model. For instance, the loss of power at a residential area has the same outage impact, regardless if this is due to a natural disaster (e.g., hurricane, thunderstorm), a malicious attack, or EPS electrical faults (e.g., short-circuits). However, the threat probabilities and CPS objective priorities for the three aforementioned scenarios differ significantly. “People health and personnel safety” objective during a natural disaster has much higher priority compared to a power outage due to an EPS fault. The latter event, being not a life threatening situation, would have a higher “uninterrupted operation and service provision” priority.

Furthermore, the presented threat modeling methodology of Section III-A and III-B, which enables precise adversary and attack descriptions, serves as the backbone of our risk assessment method. Specifically, the definitive granularity of threat characterizations, not only exposes the vulnerable system assets but also can infer which CPS objective will be affected the most. The CPS objective is critical for the attack impact evaluations, while vulnerable assets demonstrate the feasibility of an attack. Thus, CPS attack risk scores can be calculated and their prioritization can be performed based on

TABLE IV: Example of attack damage calculation.

Objective	Priority	Attack Impact	Score
People health and personnel safety	4	`Low (1)`	**4**
Uninterrupted operation and service provision	3	`High (3)`	**9**
Organization financial profit	2	`High (3)`	**6**
Equipment damage and legal punishment	1	`Medium (2)`	**2**

the affected CPS objective. The threat *Risk* is defined as:

$$Risk = Threat\ Probability \times Damage \tag{1}$$

The *Threat Probability* portion of the risk formula accounts for the threat details, i.e., the adversary and attack models discussed previously, in addition to how likely it is for the investigated threat to materialize in the specific system context. The second component of Eq. (1) includes the *Damage*, which assesses the corresponding impact inflicted on the system. The *Damage* is defined as follows:

$$Damage = \sum_{k=1}^{n} Objective\ Priority \times Attack\ Impact \tag{2}$$

where the *Objective Priority* and the *Attack Impact* are used to address the consequences of the attack in the context of the specific CPS objectives. The *Attack Impact* is evaluated qualitatively using a number from 1 to 3, reflecting `Low`, `Medium`, or `High` impact, respectively. In addition, for every CPS, the objectives are ranked in order of importance. We utilize four ($n = 4$) main objective categories: *i)* people health and personnel safety, *ii)* uninterrupted operation and service provision, *iii)* organization financial profit, and *iv)* equipment damage and legal punishment. Numbers from 1 to 4 are used for the objective priorities; 1 indicates the least significant goal while 4 stands for the most critical objective.

In Table IV, we demonstrate a damage calculation example where we provide a subjective priority ranking as well as the attack impact values. Using Eq. (2), the total potential damage score can be calculated as $\sum(4 + 9 + 6 + 2) = 21$. Given a specific *Threat Probability* value, we can then assess the total *Risk* for the examined attack scenario. Overall, the presented application-aware risk assessment procedure is taking into consideration all the underlying components of sophisticated and multi-layer threats targeting complex CPS. In addition, it provides a universal method to assess attack risk, regardless of the particular CPS architecture or the corresponding operational objectives. Based on the assessment results, administrative authorities can prioritize which assets need immediate attention and which threats pose the highest risk to exploit the vulnerabilities of the in-scope CPS assets.

IV. Cyber-Physical System Framework: Modeling, Resources, and Metrics for CPES Studies

The framework depicted in Fig. 5 shows the different domains in which the proposed CPS framework is divided. The main objective of the framework is to provide a clear understanding of all the underlying concepts and components that make up the research around CPS. Specifically, the presented conceptual framework is intended to assist researchers towards identifying the models, resources, and metrics required to perform reliable CPES studies. Based on the study objectives, the framework can be used as a 'how-to guide' towards implementation of use cases and the development of CPES testbeds. This section, first, describes the *cyber-system* and *physical-system* layers that need to be considered in order to represent CPES. Then, we provide the description of the different factors that need to be taken into consideration when designing CPES studies, i.e., the modeling techniques, resources, and metrics.

Physical-System Layer: NIST's definition for CPS establishes that the physical-system layer of a CPS is composed of hardware and software components embedded into the system environment. These components have the capability of interacting with other physical-layer units through physical means, i.e., via sensors and actuators, or through the cyber-system layer using standard communication protocols. Some sectors where CPS can be extensively found are smart manufacturing [132], healthcare [133], robotics [134], transportation [135], and EPS [136]. In this paper, the developed framework focuses on the EPS sector, i.e., the models, resources, and metrics used in the physical-system layer are based on elements comprised in the generation, transmission, and distribution systems that make up CPES. Example components within the physical-system layer of CPES are PV panels, Li-ion batteries, wind energy systems, power converters, generators, voltage regulators, transformers, and T&D lines.

Cyber-System Layer: The cyber-system layer of a CPS is composed of the ICT structures deployed in the system. This encompasses communication and networking components such as hubs, modems, routers, switches, cables, connectors, databases, and wired and wireless network interface cards (NIC) [137], [138]. These components allow the interconnection of multiple computing devices using common communication protocols over digital links with the purpose of sharing, storing, and processing resources and data located across networking nodes. In this paper, our developed framework focuses on elements that make up communication networks in CPES, i.e., the models, resources, and metrics used in the cyber-system layer are related to components such as smart-meters, PMUs, EPS-related communication protocols (e.g., DNP3, IEC61850, IEEE 37.118, etc.), and other networking devices that support communication in EPS operations.

A. Modeling

Models able to represent systems by describing and explaining phenomena that cannot be experienced directly [139]. Such models are build from mathematical equations and/or data that are used to explain and predict the behavior and response of

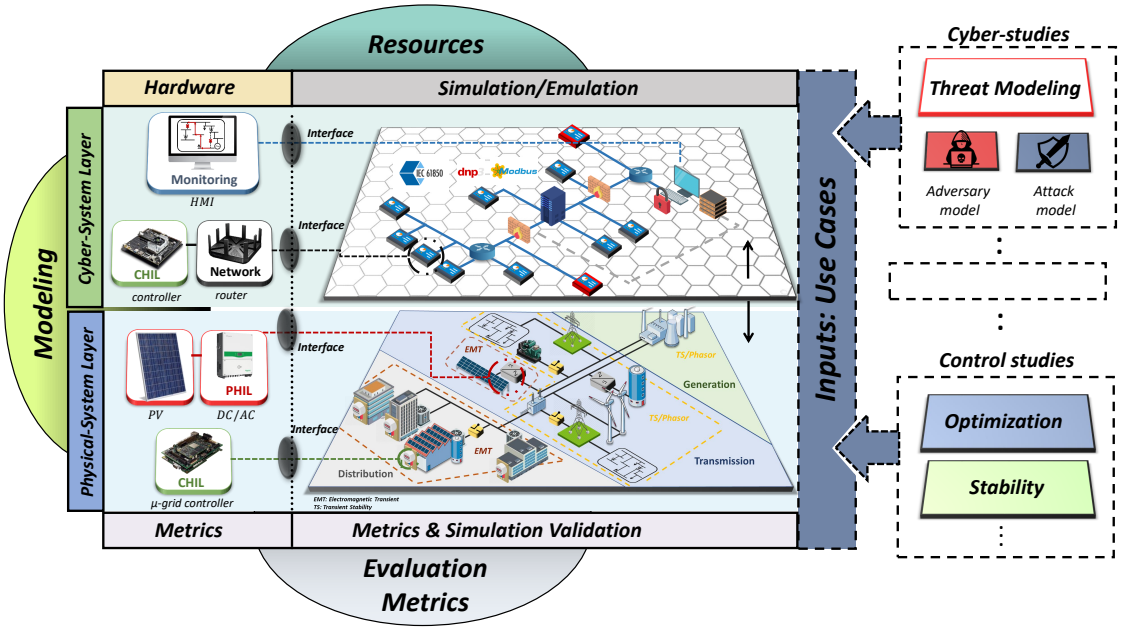


Fig. 5: Cyber-physical system (CPS) framework: the cyber-system and physical-system layers are presented with their respective factors, i.e., modeling, resources, and evaluation metrics, needed for conducting cyber-physical studies. Different use cases requirements can be adjusted to perform CPS related investigations.

complex systems. Specifically for CPES, researchers focus on creating models capable of replicating the behavior of the components that make up the cyber-system and physical-system layers of EPS, e.g., models for components such as PV systems, wind energy systems, ESS, transformers, transmission lines, distribution lines, smart meters, PMUs, routers, and switches, etc. In this part, we describe the different modeling techniques used to model both the cyber and physical layers of CPES.

1) Physical-System Layer: The design and modeling of the physical-system involve areas such as hardware design, hardware/component sizing, connection routing, and overall system testing. All components in this layer must be categorized based on their respective temporal and spatial requirements along with their intrinsic physical characteristics. In EPS, some of these characteristics and requirements are related to rated voltage, current, and power values, location of the generation and load resources, and physical characteristics of the lines (i.e., resistance, reactance, capacitance, and length). These features are utilized in developing models that represent the physical devices in the system. The objective is to capture and simulate system behavior so that a *digital twin* of the real system can be created. This 'virtualization' capability provides a significant advantage by allowing the analysis and study of different types of scenarios which can arise during the operation of the CPS. We can analyze and track physical processes, replicate potential harmful operating conditions or scenarios, and accelerate the testing of software and hardware components. More specifically, for EPS modeling, the current state-of-the-art simulation technology is based on electromagnetic transient (EMT) and transient stability (TS) simulation techniques [140]–[142].

a) *Electromagnetic transient (EMT)*: EMT simulation is a technique used to precisely reproduce the system response to fast dynamic events and system perturbations, that occur in the range of tens of microseconds or lower, caused by fast switching electromagnetic fields or loading events. Due to requirements, such as the unsymmetrical and instantaneous modeling of the signals and values that characterize the behavior of the system, nonlinear ordinary differential equations (ODE) are used to represent the system behavior in the EMT simulation environment. This detailed modeling provides improved accuracy, compared to TS-type simulations, when capturing the system behavior and response to fast transients. However, it requires high computational resources for simulating systems with large number of components. Typical applications where EMT studies are used include the simulation of power electronic devices, unbalanced distribution systems, and the impact evaluation of DER integration into modern power networks.

b) *Transient Stability (TS) Simulation*: TS simulation is a technique used to capture the slow dynamic events, i.e., events in the range of tens of milliseconds and higher, that occur in power systems. These events are related to the voltage stability, rotor angle stability, and frequency stability phenomena. In TS, the EPS is represented by nonlinear differential algebraic equations (DAE). These equations are used to solve the system states assuming that the fundamental power frequency (e.g., 50 or 60 Hz) is maintained throughout the system. Commonly, TS-type simulations are used for studies related to the analysis, planning, operation, and control of EPS elements with large time-steps, i.e., in the milliseconds range. Given that large time-steps and positive-sequence phasor-domain simulations are used in TS-type simulations, they allow users to simulate

large-scale T&D networks while requiring significantly less computational resources when compared to EMT-type simulations [140].

c) *Hybrid-Simulation (TS+EMT)*: Hybrid-simulation models make use of both EMT and TS simulation tools to leverage the benefits of two or more simulation environments, hence allowing even more comprehensive and accurate simulation studies. Some examples of this type of simulations are found in recent literature [143]–[145]. Integrated T&D co-simulations are a major field of study enabling the use of hybrid-simulation environments. Such environments can provide ways of simulating in detail, for example, power electronic converters interfaced with large-scale power networks. T&D co-simulation also provides an effective way of studying diverse impacts that anomalous events (e.g., unintentional faults or intentional malicious attacks) may have locally and globally in the overall physical-system layer of the CPES.

2) Cyber-System Layer: The design and modeling of the cyber-system layer involve communication network modeling, communication protocol implementation, design of information systems, and data storage processing. To model this layer, researchers must have a deep understanding of the communication infrastructure that needs to be replicated using the respective cyber-system layer models. Some of the characteristics that need to be taken into consideration for modeling the communication infrastructure are: *i)* the topology of the communication network, *ii)* physical characteristics (cable lengths, physical components, delays, etc.), and *iii)* Quality-of-Service (QoS), among others [138]. In a real-world CPS (e.g., cellular networks, military zones, or SCADA systems), multiple and diverse networking and computing components comprise the cyber layer. This greatly hinders the implementation of tests and studies designed to evaluate the operation and performance of the actual network, or to simply carry out any CPS-related investigation.

As discussed in Section II, carrying out evaluation type of studies in real systems can be dangerous for human safety, excessively costly, and may cause interruption or degradation of the network performance and the QoS (as perceived by the users). To address these issues, models can be used to simulate or emulate the behavior and performance of the cyber-system layer under different scenarios. In essence, simulation allows to replicate the behavior of cyber-system layer components, while emulation duplicates the behavior of these components and allows them to be used alongside real devices. The simulation and emulation of the cyber-system layer are fundamental tools towards understanding and studying topics related to complex network deployment, networking architectures, communication protocol features, and deployment of new services.

The simulation/emulation modeling process is often instantiated by identifying all the network components, commonly referred to as communication network entities. These entities, i.e., nodes and links configurations, constitute the network topology. Fig. 6 depicts a conceptual illustration of how the modeling process is performed in a communication network simulation. As seen in Fig. 6, in a network simulator/emulator architecture, a node is a key entity that represents any computing device connected to the overarching network. This

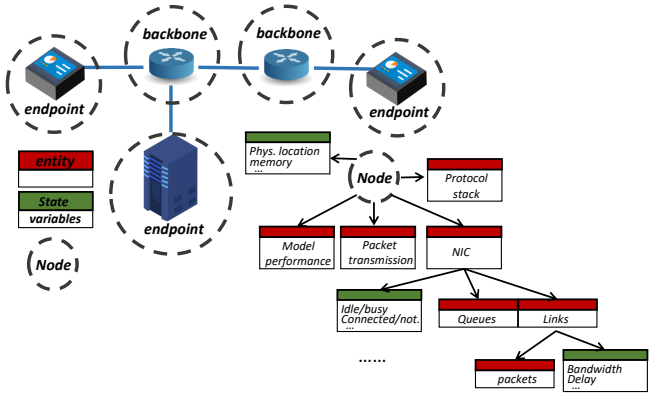


Fig. 6: Conceptual diagram of the modeling and simulation process of communication networks.

abstraction encapsulates all the possible representations of computing devices that may exist in a network setup. Some of these computing devices can refer to routers, switches, and hubs which embody the backbone of the network, while computers, RTUs, PLCs, meters, and servers constitute the endpoints of the network. A node is primarily characterized by its packet transmission entity attribute. In this packet transmission attribute, endpoints delineate the source or destination of the data packets while all backbone elements perform the forwarding tasks related to these packets. Other parameters, known as state variables, differentiate the behavior for each one of the modeled nodes. Some of these parameters are memory consumption, physical location, battery power, and CPU utilization. Additionally, other simulation entities, such as NIC, help to identify nodes in the network. These interfaces also have individual state variables that represent their state (i.e., idle or busy, and installed or not installed) while being in charge of transmitting, receiving, and processing the packets exchanged with other network nodes.

Similar to the nodes, interfaces include other entities, such as queues and links, which represent realistic packet processing scenarios. Queues are modeled as buffers in the outgoing and incoming packet processes. Links are modeled as the connections between the two nodes communicating via the corresponding interfaces (i.e., communication medium). More specifically, links are modeled by defining communication parameters such as available bandwidth, propagation delays, jitter, and pre-defined packet loss rates. Furthermore, packets are modeled as entities which contain data exchanged between nodes in the network. For each node in the network, entities that represent the protocol stack must also be defined, while the packet sizes are determined by the corresponding communication protocol (e.g., TCP, UDP, etc.).

A protocol entity is responsible for managing the outgoing and incoming packets by adding and removing packet headers. Protocol modeling is also a key process. It covers the specific steps required to accurately emulate the behavior of the protocol stack. In this process, models are developed to capture elements and properties from the network access layer, internet layer, transport layer, and application layer. Finally, models for

performance evaluations, which do not represent real elements in the network, are also defined as additional entities that facilitate the implementation and evaluation of the network. Some representative examples of such entities are logging and helper utilities which can aid the network evaluation process [146].

B. Resources

The '*resources*' factor represents the different hardware and software systems that form, and can be used to model and simulate, the cyber- and physical-system layers of the CPES being studied. In this part, we make a distinction between the hardware and simulation/emulation resources that need to be considered for modeling the cyber- and physical-system layers using tools and techniques such as offline simulation, emulation, real-time simulation, and HIL.

1) Physical-System Layer: The simulation and hardware resources for the modeling and implementation of the CPES physical-system layer are presented below.

a) *Simulation*: A simulation provides a set of models or representations used to reproduce the behavior or operation of different processes of a particular system over time. Particularly for EPS, EMT- and TS-type simulation are the most prominent tools used to simulate the behavior of the different system components. These simulation classes can be further classified into two main categories: offline and real-time simulation [147].

i) *Offline simulation*: Offline simulation tools provide a simple and cost-effective way of conducting simulations on any generic computing device. These tools can execute models in slower or faster-than real-time speeds depending on the complexity of the model as well as the availability of computing resources. Figs. 7a and 7b show how the computation time of the system models, for both slower and faster-than real-time offline simulations, is not synchronized with the simulation clock, i.e., the real-time clock. Offline simulations allow the simulation of complex systems without considering real-time constraints, which for instance, enable researchers to simulate large periods of time, e.g., months or years, in a few minutes or seconds. Some tools and software which are available for performing this type of simulations include: *MATLAB/Simscape Electrical* (EMT & TS), *OpenDSS* (TS), *Gridlab-D* (TS), *eMegaSim* (EMT), *ePhasorSim* (TS), and *ETAP eMTP* (EMT).

ii) *Real-time simulation*: Real-time simulation tools provide the capability of generating results that are synchronized with a real-time clock. This allows physical devices to be interfaced with the simulated system via realistic data exchanges synchronized using a real-time clock. Fig. 7c demonstrates how for real-time simulation the computation time for the system is synchronized with the simulation clock. The computation time needed to solve all the states of the simulated system needs to be lower or exactly the same as the simulation clock, i.e., the real-time clock. Real-time simulation setups allow researchers to connect real devices using HIL techniques such as CHIL and PHIL. Some tools and software which are available for performing this type of simulations include: *eMegaSim* (EMT), *ePhasorSim* (TS), *HyperSim* (EMT), *RTDS* (EMT), and *Typhoon HIL* (EMT).

b) *Hardware*: Real-time HIL implementations allow the interconnection of external hardware devices to a real-time simulation through the appropriate I/O or networking interfaces. Two of these HIL techniques are CHIL and PHIL.

i) *Controller Hardware-in-the-Loop (CHIL)*: In CHIL, physical devices are in constant communication and interaction with a simulation running in the real-time environment. This interconnection includes sending control signals and receiving feedback signals through I/O and/or networking ports [148], [149]. As seen in the hardware section of Fig. 5, a physical device connected using a CHIL implementation can be interfaced directly: *i)* with the physical-system layer simulation using the appropriate interface, or *ii)* through the cyber-system layer using standard communication protocols and corresponding networking components.

ii) *Power Hardware-in-the-Loop (PHIL)*: In PHIL, a power hardware system such as a PV panel, inverter, or battery system is physically connected to the RTS through analog and digital I/O ports. A PHIL implementation needs the use of a power amplification unit that is responsible for the amplification and conversion of the digital voltage and current data signals – coming from the simulation environment – into analog voltage and current signals required by the connected actual/physical device. Interfacing algorithms are also essential to facilitate the interconnection between the software models and the physical-system [150].

2) Cyber-System Layer: The simulation/emulation and hardware resources related to the modeling and development of the cyber-system layer for the communication network are presented below.

a) *Simulation/Emulation*: As mentioned before, the main difference between simulation and emulation is that in a simulation, the models used are designed to replicate the behavior of the system while an emulation is designed to duplicate the behavior of the system. A more detailed description of the difference between simulation and emulation is given below in the context of the resources required to effectively replicate the cyber-system layer.

i) *Simulation*: In network simulations, theoretical and mathematical models are developed to create entirely virtual models of the corresponding networking components. Network simulation tools use discrete-event simulation approaches that generate sequences of discrete events which characterize the discrete cyberspace. The two critical components of such discrete-event driven simulators include the simulation time variable and a list of pending future events. The simulation time variable represents the current time at which the state of the system is known (in the simulation), while the list of pending future events contains all the state changes that have been scheduled to occur in the future, which guide the flow of the simulation. In a network simulation, external devices cannot be interfaced with virtual simulated devices, contrary to a network emulation, hence, the entire communication network needs to be simulated. Some of the available software tools that support this type of simulations are: *ns-2* [153], *ns-3*

TABLE V: Physical-system layer performance metrics. These metrics are divided according to the domain where they can be measured.

Name	Description	Domain
Rise time	Evaluates the time required for the output to rise from %x to %y of the steady-state response	Control
Percent overshoot	Evaluates the maximum peak value of the output minus the step value divided by the step value	Control
Settling time	Evaluates the time required for the output to reach and remain within a defined error band	Control
Steady-state error	Evaluates the difference between the input (command) and the output of the system	Control
Integrate absolute error (IAE)	Evaluates the absolute error of the system over time	Control
Voltage stability	Metrics that evaluate the voltage stability and regulation of the EPS according to defined limits	EPS
Frequency stability	Metrics that evaluate the frequency stability of the EPS according to defined limits	EPS
Optimization	Optimization metrics used to evaluate energy and power management functions, e.g., energy cost, efficiency	EPS
Power quality	Power quality metrics such as THD, transients, flickering, and voltage sags used to evaluate EPS operation [151]	EPS
Reliability	Reliability indices to evaluate EPS operation, e.g., SAIFI, SAIDI, ASAI, lost load % [151]	EPS
Command vs. Measured % error	Percentage error between signal command coming from controller and signal measured	Simulation (CHIL)
PHIL Accuracy	Metrics that evaluate the accuracy of the PHIL integration [152]	Simulation (PHIL)

[154], *SimPy* [155], and *EXata* [156], [146].

ii) *Emulation*: In network emulation, hardware and software solutions are designed to accurately replicate the behavior of networking components, exactly as if they were actual parts of an external network. Network emulation tools enable the configuration and manipulation of network parameters and constraints (e.g., packet loss, delays, jitter, etc.) to mimic the mirrored network. Some of the available software tools that support this type of network modeling are: the *Common Open Research Emulator (CORE)* [157], *NetEm* [158], and *EXata* [156]. Notably, some tools are capable of adapting network simulation models for emulation purposes by adding real-time synchronization mechanisms between the virtualized simulated environment and the real networking components [159], [160].

b) *Hardware (HIL)*: Similarly to the HIL implementations realized in the physical-system layer, HIL implementations of network components in the cyber-system layer can also be performed using the corresponding networking interfaces. Networking HIL provides emulation capabilities that allow the integration of real equipment into the emulated network through standard communication protocols. Commonly, a larger portion of the network or system is emulated and connected with external (real) devices. Such method provides high-fidelity responses – as expected from the actual device – while maintaining the scale of the emulation. Some software tools that support HIL with communication network models are *EXataCPS* [156], *ns-3* [154], and *CORE* through its *RJ45* utility [157].

C. Performance Metrics

A multitude of metrics exists to evaluate the performance of the modeled cyber and physical-system layers. The use of metrics allows the proper evaluation of the overall system alongside its corresponding subsystems. In essence, these metrics provide quantitative ways to measure and evaluate the performance of the system's operation at a particular time, both at the cyber and the physical-system layers.

1) Physical-System Layer: Some of the most commonly used metrics employed to evaluate the performance and operation of different functions that exist in the physical-system layer of CPES are presented in Table V and described below:

a) *Control systems*: Metrics related to control systems can be used to examine the performance of different control routines present at the physical-system layer. The evaluation can include the steady-state response of the system or other system performance indicators such as rise time, percent overshoot, settling time, steady-state error, and integrate absolute error.

b) *EPS resiliency, stability, and optimization*: Performance metrics can be defined in order to evaluate the performance of the system according to a predefined baseline behavior. For instance, in an EPS where the operation of a new MG controller is investigated, performance metrics related to voltage regulation, frequency regulation, energy cost, and power quality can be utilized. Similarly, especially for controllers, other performance metrics can be utilized to determine execution times, CPU utilization, and memory utilization, which are related or limited by the computing resources of the system.

c) *Simulation accuracy*: The simulation accuracy, either offline or real-time, can also be assessed based on different

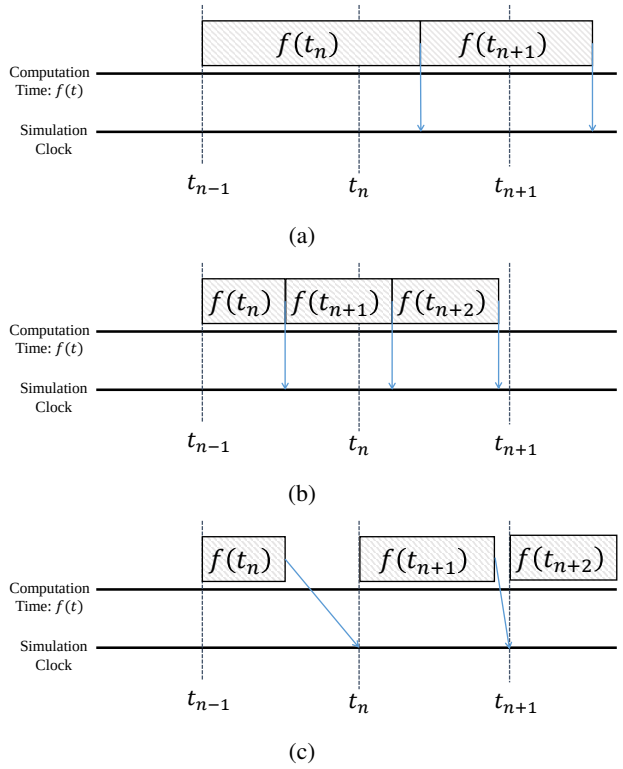


Fig. 7: Differences in the computation timing of offline simulation: (a) slower-than real-time, (b) faster-than real-time, and (c) real-time simulation.

performance metrics dependent on the stability and accuracy of the system response, respectively. The main objective of these metrics is to validate the response of different physical systems (being simulated), when compared to the actual response expected from the system under examination.

2) Cyber-System Layer: Different metrics can be utilized to evaluate the performance of the modeled cyber-system layer communication network. Here, we demonstrate, as a practical example, some of the most widely used metrics designed to evaluate the network performance at different layers of the open systems interconnection (OSI) model [146]. Table VI outlines some representative network performance metrics.

a) *Physical (L1) and Data Link Layers (L2)*: These layers describe how data should be generated and transmitted by network devices over the corresponding physical media.

b) *Network Layer (L3)*: This layer describes how data packets are transferred between a source and a destination node inside the network. It represents layer 3 of the OSI model. The main performance metrics described below are designed to evaluate two main routing functions: path selection, and network topology management. Path selection aims to determine the best path from source to destination, while network topology management defines how network entities are interconnected for data forwarding purposes.

c) *Transport (L4), Session (L5), Presentation (L6), and Application (L7) Layers*: These layers describe the shared communication protocols and interfacing methods used by the nodes in the network. In essence, these are the layers responsible for providing full end-user access to the communication network infrastructure.

It is important to note that many other network and physical performance metrics can be used to evaluate specific scenarios. The presented lists include a subsection of the available metrics discussed in literature. There are also application-specific metrics which can be defined according to each study requirements. Overall, researchers should carefully model their systems as well as select the corresponding resources and metrics to accurately represent the cyber- and physical-layer of the CPES under test. This will allow the integration of any external physical device, either through CHIL and/or PHIL, and ensure the holistic validation of the system's operation.

V. Experimental Setup & Case Studies

The case studies discussed in this section demonstrate how the presented threat modeling approach, the CPS framework, and risk assessment methodology can be utilized to perform detailed CPES studies. Table VII describes how each study can be formalized using our proposed threat modeling method. Following, the corresponding modeling layers, resources, and evaluation metrics are identified for each case study according to the conceptual CPS framework. Additionally, for each attack scenario, the specific background, and mathematical formulation are described and the corresponding threat model is provided based on Section III. The threat model describes the assumptions made for the adversary intentions and capabilities as well as the attack-specific details, demonstrating the practicality of our modeling approach for diverse attack scenarios. Furthermore, we demonstrate how the proposed risk assessment procedure can be applied to each case study and assist in prioritizing mitigation strategies. In our work, the objective priority for CPES is outlined in Table VIII. It should be noted that the order of objectives might change depending on the system's component being analyzed or the stakeholders' priorities. For instance, the impact of the "uninterrupted operation and service provision" objective could indicate less priority in the case of a compromised inverter serving as an ancillary power generation source in a residential deployment, in contrast to a T&D system-wide attack.

The attack cases presented in this section can be characterized as either *DIA* or *data availability attacks (DAA)*. Table IX provides the essential notation for the case studies. Each scenario follows a mathematical background as part of a CPS plant formulation:

$$x(k+1) = Gx(k) + Bu(k) \tag{3}$$

$$y(k) = Cx(k) + e(k) \tag{4}$$

where $x(k) \in \mathbb{R}^n$ represents the states of the system, $u(k) \in \mathbb{R}^l$ represents the control variables, and $y(k) \in \mathbb{R}^m$ represents the system measurements. $G \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times l}$, and $C \in \mathbb{R}^{m\times n}$ represent the system matrix, input matrix, and output matrix, respectively. The term $e \in \mathbb{R}^m$ represents measurement noise in the system's input measurements. As for the cyber part of the CPS, it can be generally expressed as:

$$u(k+1) = Hy(k) \tag{5}$$

TABLE VI: Cyber-system layer performance metrics. These metrics are divided according to the OSI model layer and connection where they can be measured.

Name	Description	Layer	Connection
Bit rate (R)	Number of bits conveyed per unit of time	*L1/L2*	Wired/Wireless
Bit-error rate (BER)	Ratio of the number of received bits altered during transmission to the total number of bits sent	*L1/L2*	Wired/Wireless
Packet-error rate (PER)	Ratio of the number of packets received incorrectly to the total number of packets received	*L1/L2*	Wired/Wireless
Nominal channel capacity (NCC)	Maximum number of bits that can be transmitted per unit of time	*L1/L2*	Wired/Wireless
Channel utilization (CU)	Ratio between NCC and the total number of bits received per transmission time	*L1/L2*	Wired/Wireless
Signal-to-noise ratio (SNR)	Ratio of the signal power to the background noise	*L1/L2*	Wired
Signal-to-interference-plus-noise ratio (SINR)	Similar to SNR but considers the interference power from other signals	*L1/L2*	Wireless
Spectral efficiency (SE)	Number of received bits per unit of time per unit of bandwidth and per unit areas (i.e., $\frac{b/s}{Hz \cdot m^2}$)	*L1/L2*	Wireless
Received signal strength indication (RSSI)	Signal strength measured at the receiver's antenna during packet reception	*L1/L2*	Wireless
Hop count	Minimum hop-count from source node to destination node	*L3*	Wired/Wireless
Round trip time (RTT)	Time that it takes for a signal to be sent and the acknowledgement received	*L3*	Wired/Wireless
Expected transmission time (ETT)	Time needed for a data packet to be correctly transmitted over a link	*L3*	Wireless
Betweenness	Number of shortest paths between any two nodes that pass through the evaluated node	*L3*	Wired/Wireless
Node degree	Number of nodes that depend on the evaluated node.	*L3*	Wired/Wireless
End-to-end delay	Time required to transmit a packet along the path between source and destination nodes	*L4/L5/L6/L7*	Wired/Wireless
Jitter	Packet delay variation	*L4/L5/L6/L7*	Wired/Wireless
Bandwidth	Overall bandwidth consumption in the network	*L4/L5/L6/L7*	Wired/Wireless
Link stress	Number of packet replicas traversing the same physical link	*L4/L5/L6/L7*	Wired/Wireless

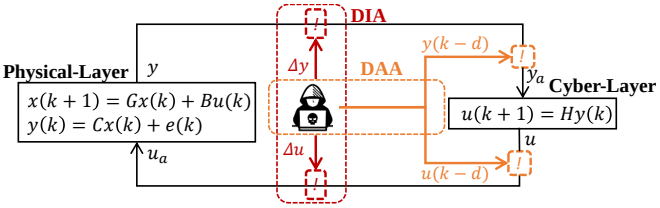


Fig. 8: Diagram of CPS plant under DIA and DAA scenarios.

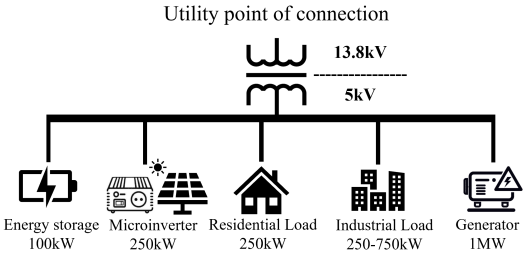


Fig. 9: Conceptual illustration of the MG system used for the cross-layer firmware case study.

where $H \in \mathbb{R}^{l \times m}$ represents the control matrix [161].

Fig. 8 depicts a diagram of the CPS mathematical formulation and the respective variables compromised by attackers during DIA and DAA scenarios. In the DIA case, either the measurements (y) or the control variables (u) can be compromised by attackers via modification or fabrication. On the other hand, in a DAA scenario, either the measurements (y) or controls (u) can be compromised by attackers via interruption, i.e., delaying their acquisition or utilization by the system.

A. Case Study 1: Cross-layer Firmware Attacks

Background & Formulation: Cross-layer firmware attacks refer to attacks targeting the firmware code of embedded devices (i.e., the device read-only resident code which includes microcode and macro-instruction level routines), aiming to generate and propagate impacts from the device layer to system and application layers, respectively. Typically, embedded devices in industrial CPS run on bare metal hardware without OS and directly boot monolithic single-purpose software. In

TABLE VII: Threat model of the attack case studies.

Threat Model	**Cross-layer Firmware Attacks**	**Load-changing Attacks**	**Time-delay Attacks**	**Propagating Attacks in Integrated T&D CPES**
Knowledge	Oblivious	Limited or Oblivious	Oblivious	Strong
Access	Possession	Non-possession	Non-possession	Possession
Specificity	Non-targeted	Targeted	Targeted	Targeted
Resources	Class I or II	Class II	Class I or II	Class II
Frequency	Iterative	Iterative	Iterative	Non-Iterative
Reproducibility	Multiple-time	Multiple-times	Multiple-time	One-time
Functional Level	L1	L1 or L2	L1	L2
Asset	PLC	PLC or HMI	Control Server	Engineering workstation
Technique	Modify control logic	Modify control logic or wireless compromise	Wireless compromise, MitM, Spoofing, DoS	Engineering workstation compromise
Premise	Physical: Invasive, or non-invasive Cyber: Asset control commands	Cyber: Communications and protocols or Asset control commands	Cyber: Communications and protocols	Cyber: Asset control commands

TABLE VIII: Objective Priority for CPES Risk Assessment.

Objective	**Priority**
People health and personnel safety	4
Uninterrupted operation and service provision	3
Equipment damage and legal punishment	2
Organization financial profit	1

TABLE IX: Symbols and notation for case studies formulation.

Domain	**Symbol**	**Definition**
CPS	k	timestep
	n	total number of states
	m	total number of measurements
	l	total number of controls
	x	system's states
	y	system's measurements
	y_a	attacked system's measurements
	u	control variables
	u_a	attacked control variables
	e	measurement noise
	G	system matrix
	B	input matrix
	C	output matrix
	H	control matrix
DIA	β	multiplicative attack factor
	W	additive White-noise attack factor
	$\mathrm{T}_{\mathrm{attack}}$	attack period (time)
	Δy	manipulated measurements used to create y_a
	Δu	manipulated controls used to create y_a
DAA	f_D	time-delay function
	d	discrete constant value or time function.
	s_r	compromised signal (could be either y or u)

such devices, tasks are executed on a single-threaded infinite loop. If the device firmware code execution is maliciously modified, adversaries could have total control over the embedded device. The effects of such attacks can have a cross-layer impact affecting multiple components and processes of the CPS. For example, in a CPES, by modifying the firmware controlling grid-tied inverters connected to BESS or EV chargers, an adversary could compromise system's measurements, thus causing frequency fluctuations, voltage sags, and system stability issues. Other scenarios could even cause wide-area outages, such as the Ukrainian power grid attack in 2015, in which attackers replaced the legitimate firmware of serial-to-Ethernet converters at substations causing them to become inoperable [14]. In general, cross-layer firmware attacks can be categorized as a DIA-type of attack since modifications at the firmware level could result in compromising the integrity of data at different CPS layers.

In this type of DIA, the adversary (though firmware modifications) can tamper with the input/sensed measurements (e.g., modify, scale, etc.), $y(k)$, and thus directly affect the inverter control strategy and variables, $u(k)$, driving the system into instability. This type of attack can be characterized as a *combined* DIA attack [162]–[164]. In more detail, the system's input measurements are modified using both an additive random/white noise component and an attack model in which nominal measurements are scaled (increased or decreased). These DIAs can be modeled as:

$$y_a(k) = \begin{cases} y(k) & \text{, when } k \notin \mathrm{T}_{\mathrm{attack}} \\ \beta y(k) + \mathrm{W} & \text{, when } k \in \mathrm{T}_{\mathrm{attack}} \end{cases} \tag{6}$$

where β represents the multiplicative attack term, W represents the additive random/white noise attack, $\mathrm{T}_{\mathrm{attack}}$ represents the period of time when the DIA is performed, and y_a represents the 'altered'/attacked input measurements. $\beta > 1$ represents increasing-type of attacks, and $\beta < 1$ decreasing attacks.

Following this combined-type DIA mathematical formulation, we demonstrate how the inverter operation can be compromised by spoofing its energy conversion module. The results of this compromise affect not only the inverter behavior, but also propagate and impact the MG operation as well.

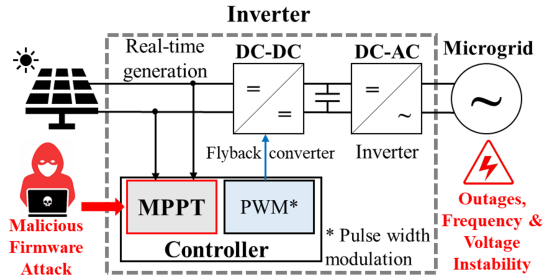


Fig. 10: Cross-layer firmware attack targeting the inverter's maximum power point tracking (MPPT) controller.

Threat Model: As presented in Section III, the threat modeling process for any attack can be characterized by the adversary model, and the attack model formulations. Specifically, in this cross-layer firmware attack case, we assume an *oblivious adversary* without full observability of the CPES, and who has direct physical access to the targeted hardware controller (i.e., *adversary access: possession*). Regarding adversarial specificity, the attack is presumed to be a *non-targeted attack*. The adversarial resources could range from the minimum, i.e., Class I, up to state-funded criminal organizations (Class II), in the worst-case scenario.

Furthermore, our case study assumes an attack that occurs *iteratively* and can be reproduced *multiple times*. The *targeted asset is a solar inverter controller*, so the attack level is defined as *Level 1*. Finally, the technique employed to compromise the system involves *control logic code modification*, and the attack premise can be categorized as either invasive or non-invasive (on the physical domain), or could target the inverter control (e.g., power conversion, power factor, active reactive injections, setpoints, etc.) using malicious commands (on the cyber domain).

Attack Setup & Evaluation: In this case study, a cross-layer firmware attack is modeled as a DIA that compromises physical components, more specifically a PV inverter, at the *physical-system layer* of the CPES. Both EMT and TS simulation modeling approaches are used to model a MG system comprised of a solar PV with its inverter, a Li-ion BESS, a diesel generator, and residential and industrial loads. The MG is connected to the main grid via a 13.8 kV/5 kV distribution substation transformer with a capacity of 250 MVA. The nameplate generation capacity for the diesel generator is set to 1 MW. The maximum generation capacity that the PV inverter can reach is 250 kW based on the provided solar irradiance profile. The BESS is capable of providing up to 100 kW and storing 100 kWh. The loads of the MG include aggregated residential loads with a constant power demand of 250 kW and a variable lumped industrial load whose power demand ranges between 250-750 kW. Fig. 9 shows a conceptual illustration of the described MG. The main software resource used to conduct the EMT and TS offline-simulations of the physical-system layer for this case study is *MATLAB/Simscape Electrical*.

In Fig. 10 we illustrate, the top-level architectural overview of an inverter, the core components comprising it, and the maximum power point tracking (MPPT) controller block that is the main target of this attack use case. Attackers can disrupt the nominal inverter operation by tampering with the firmware

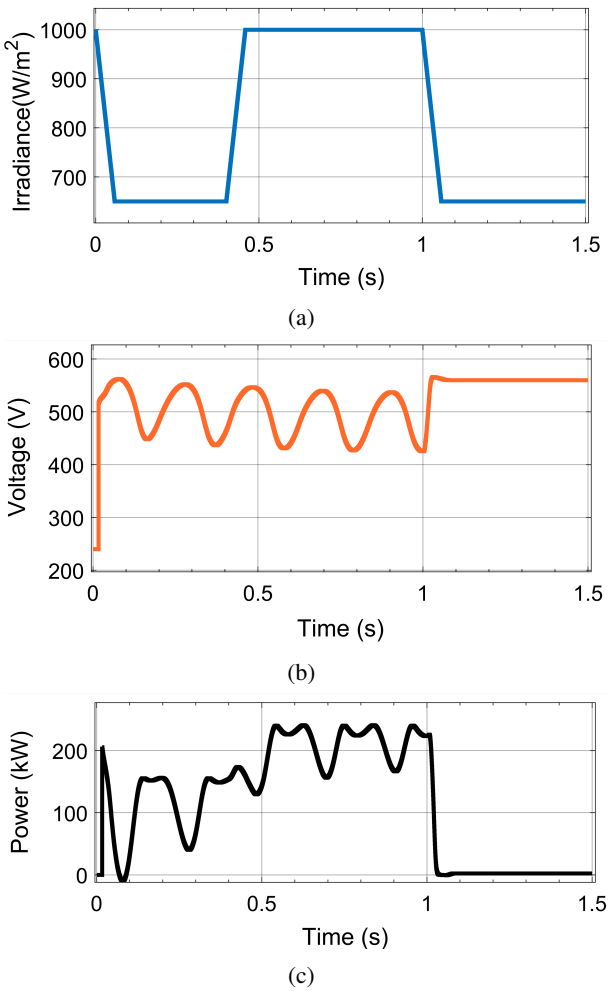


Fig. 11: Cross-layer firmware attack impact on the DC-DC converter stage: (a) solar irradiance profile, (b) impact on the DC voltage output, and (c) impact on the DC power output.

subroutines which control both the DC-DC boost and the DC-AC conversion stages. In particular, Figs. 11, 12 demonstrate the specifics of how the operation of an inverter can be affected by an adversary capable of compromising the firmware. For our use case, we employ a grid-tied solar inverter module provided by Texas Instruments [165]. The inverter leverages an *F2803x* series control card which is responsible for the management of the inverter's peripheral devices (e.g., sensing modules, analog-to-digital converters, transistor gate driver circuits, etc.) as well as the power conversion process (i.e., solar energy to electricity). By modifying the operation of the MPPT algorithm – within the firmware code of control card that the inverter utilizes to optimize the output power generated by the solar panels – the attacker is able to destabilize the operation of the converter.

MPPT algorithms enable inverters to obtain high power conversion efficiencies. By constantly monitoring the solar PV outputs (i.e., PV generated voltage and current), MPPT algorithms regulate the converter's operating point achieving maximal power transfer. Given that the PV real-time generation measurements are critical for the MPPT operation, any perturbations of the sensed values can potentially com-

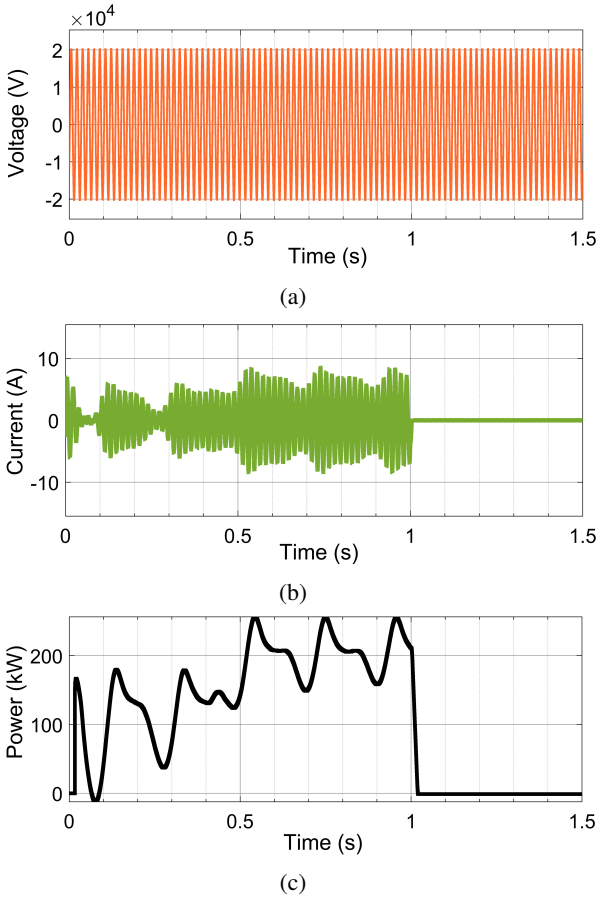


Fig. 12: Cross-layer firmware attack impact on the DC-AC converter stage (grid-tied inverter end): (a) impact on the AC voltage output, (b) impact on the AC current output, and (c) impact on the AC power output.

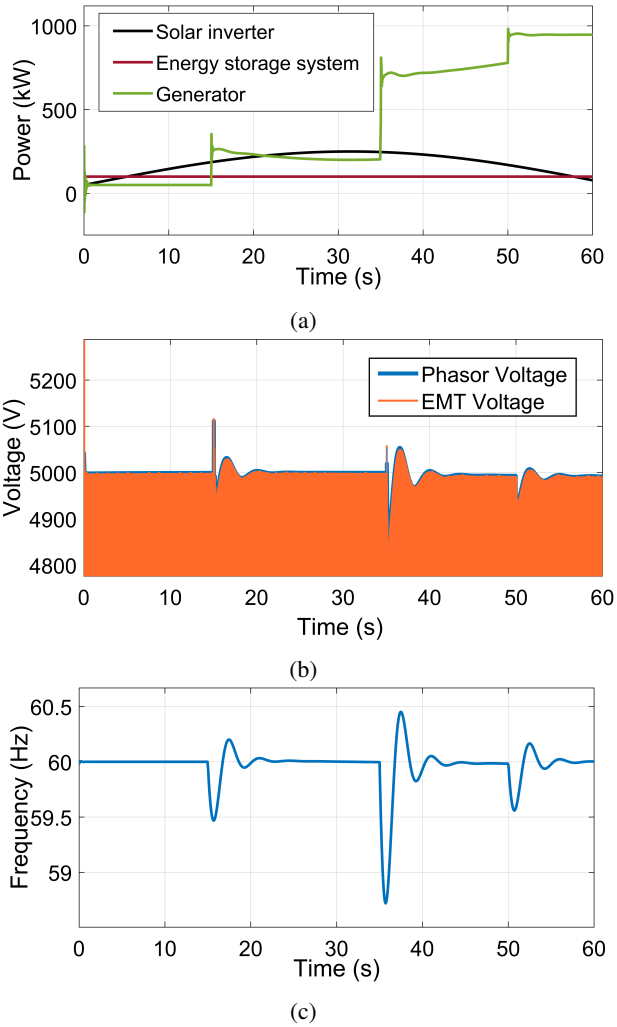


Fig. 13: Cross-layer firmware attack impact on the power grid: (a) renewable source generated power, (b) phasor and EMT magnitude of the MG voltage, and (c) MG operating frequency.

promise the inverter's nominal operation. For our case study, the modification of the inverter's firmware tampers with the inverter's MPPT function as well as the controls of the DC-DC and DC-AC converters. In the context of DIA attacks, the sensed inputs to the MPPT function, i.e., PV voltage and current, are maliciously modified. By tampering with the MPPT input measurements, down-scaling and introducing additive sinusoidal noise (combined-type DIA attack), we are able to generate the oscillatory behavior depicted in Fig. 11. This unstable behavior propagates through the inverter's power conversion process leading to anomalous behavior on the grid-tied inverter end, as seen in Fig. 12. The result of this compromise is the eventual disconnection of the inverter-enabled power resource ($t = 1sec$) in order to protect the rest of the MG devices and avoid operational disruptions.

The metrics used to evaluate the performance and behavior of the MG operation, based on the presented CPS framework, are the physical-system layer performance metrics related to *frequency stability* and *voltage stability*. Fig. 13 demonstrates the overall impact of malicious inverter operation on the MG, and how the grid's power, voltage and frequency are affected. In more detail, we notice that at $t = 35sec$ when a significant load increase in the MG occurs, the contribution of the anomalous inverter behavior significantly impacts the frequency causing potential stability issues. However, at $t = 15sec$ and $t = 50sec$, when the power generation of the inverter as well as its power contribution to the grid is much lower following the solar irradiance profile, the impact of the inverter's malicious behavior is reduced. Thus, from an adversarial perspective, targeting an inverter device during peak-hours when the solar generation is reaching its maximum can yield significant implication on the grid's operation. Fig. 11 shows the impact of the attack on the DC-side of the converter. It can be observed that both the DC voltage and current fluctuate, creating harmonic distortion at the output. Similarly, Fig. 12 demonstrates how the AC power generation is affected by the firmware modification attack. At $t = 1sec$ the oscillatory behavior causes an islanding scenario that disconnects the PV system from the rest of the MG. Fig. 14 shows the mapping of the presented case study with the CPS framework.

It is important to note that in the presented case study, it is assumed that the cross-layer firmware attack is performed by an adversary with the capability of compromising the

Layers	Modeling	Resources	Metrics
Cyber-System	-	-	-
Physical-System	EMT & TS : Offline	MATLAB/Simscape Electrical	Frequency & voltage stability

Fig. 14: Mapping of cross-layer firmware attack case study with CPS framework.

physical device, hence, modeling the cyber-system layer was not required. An extension of this study could involve the implementation of an over-the-air cross-layer firmware attack that compromises a device via the cyber-system layer. The implementation of such scenario would also require the modeling of the cyber-system layer, i.e., the communication network that serves as the medium and entry point for the attack.

Risk Assessment: Due to the inherent difficulty of getting simultaneous access to multiple devices in order to cause a severe impact on grid operation, the $Threat\ Probability$ for this type of attack is set to `Medium (2)`. For the resulting damage part of the $Risk$ formula, we use the priorities indicated in Table VIII, and set the "People health and personnel safety", "Uninterrupted operation and service provision", and "Equipment damage and legal punishment" attack impacts to `Low (1)`, while the "Organization financial profit" counterpart is set to `Medium (2)`. Thus, the comprehensive $Risk$ for the evaluated cross-layer firmware attack study is estimated to be $2 * \sum(4 + 3 + 2 + 2) = 22$.

B. Case Study 2: Load-changing attacks

Background & Formulation: In load-changing attacks, an adversary triggers an unexpected or sudden demand increase or decrease of IoT connected high-wattage appliances and DERs, with the objective of causing grid instabilities [12]. Although currently hypothetical, due to the low penetration rates of IoT-controllable high-wattage loads and DERs, load-changing attacks are projected to become a 'real' threat in the near future as the number of controlled DERs and loads is anticipated to grow exponentially [166]–[168]. Attackers able to install malware which could control DERs and load consumption, can therefore maliciously manipulate system operating conditions and affect the operation of CPES. One example of such an attack can entail an adversary capable of synchronously switching on and off high-wattage devices at unexpected times, causing power, voltage, and frequency instabilities, i.e., an Aurora-type attack at the load side [169]. This event could also potentially damage utility equipment or initiate cascading failures in distribution systems.

In terms of mathematical formulation, load-changing attacks can be framed as a DIA-type of attacks that maliciously modifies the control variables of loads in CPES, causing significant unexpected power variations that could, in turn, lead to circuit overflows or instabilities at certain vulnerable locations of the electric grid. This type of attack involves the malicious manipulation of high-wattage appliances and/or DERs that can significantly disturb the balance between power supply and demand. In order to perform this type of attack, we assume that the adversary accesses and controls multiple compromised elements through the cyber layer of the system, i.e., its communication network infrastructure, and then manipulates their control variables causing rapid fluctuations in the system's response. A load-changing attack differentiates from a 'measurements-altering' DIA in the sense that, instead of measurements being affected, the control variables are the one being directly manipulated by the adversary. Using the same CPS system described by Eqs. (3) – (5), the generalized DIA for the load-changing attack scenario is described by:

$$x_a(k+1) = Gx(k) + B\big(u(k) + \Delta u(k)\big) \tag{7}$$

$$y_a = C\big(x(k+1) + B\Delta u(k)\big) + e(k+1) \tag{8}$$

where x_a and y_a represent the states and measurements, respectively, 'altered' by the manipulation of the system's control variables Δu.

In order to map the above formulation to the load-changing attack case within CPES, the term u in Eq. (7) can be adapted to represent the controllable 'altered' load demand in the system as:

$$d_a(k) = d_i(k) + \Delta d(k) \tag{9}$$

where d represents the controllable load demand, d_i is the initial 'un-altered' load demand, Δd is the portion of the total load demand affected by the attack, and d_a represents the total load demand 'altered' by the load-changing attack. If the attackers simultaneously compromise more than one load in the system, Eq. (9) can be extended as:

$$D_T(k) = \sum_{l=1}^{m} d_{i,l}(k) + \sum_{j=1}^{n} d_{a,n}(k) + P_{loss} \tag{10}$$

where D_T represents the total demand in the system, m is the number of total 'unaltered' loads, n is the total number of loads compromised by adversaries, and P_{loss} is the total loss in the distribution network.

Based on the CPES requirement to balance load and generation in real-time in order to maintain frequency stability in the system [170], the summation of all generation output and all load demands and losses must be approximately equal:

$$D_T(k) \approx \sum_{g=1}^{N_g} P_g(k) \tag{11}$$

where N_g represents the number of g generators in the system. To understand the effect of sudden load changes in the frequency stability at each generator bus, we use the swing equations. The swing equations in Eq. (12) – (14) describe the relationship between the input mechanical power (P_m), output electrical power (P_e), and the rotational speed of the generator (ω) [171]. The term P_e is directly related to P_g, since it represents the generator power output plus electrical losses of the generating unit.

$$\frac{2H}{\omega_s}\frac{d^2\delta}{dt^2} = P_m - P_e \tag{12}$$

$$\frac{d\delta(t)}{dt} = \omega(t) - \omega_{\mathrm{s}} \tag{13}$$

$$\frac{2H}{\omega_s}\frac{d\omega(t)}{dt} = P_m - P_e \tag{14}$$

$$P_e = \frac{V_s V_r}{X} sin(\delta) \tag{15}$$

In these equations, H represents the constant normalized inertia, ω_s is the synchronous speed (i.e., 50 or 60 Hz), and δ is the power angle; the angle between the generator's internal voltage, i.e., the voltage at the generator bus V_s, and its terminal voltage, i.e., the voltage at receiving bus V_r. X is the reactance based on the classical model of a generator [172] . The relationship between the electrical frequency $\omega(t)$ with the power angle δ is shown in Eq. (13). Based on these relationships, any sudden change in load demand, caused by high-wattage loads turning on/off in the system, will affect P_e, and thus cause subsequent frequency fluctuations, as seen in Eq. (14).

Threat Model: In the load-changing attack case study, the adversary is assumed to be either oblivious, i.e., having no knowledge of the system topology, or with limited knowledge. Such limited information regarding the CPES could assist in optimally coordinate the attack and could be acquired, for example, via open-source intelligence techniques. The adversary can perform the attack remotely, thus, *non-possession* is presumed of the IoT devices controlling the high-wattage loads. Load-changing attacks are *targeted* attacks aiming to destabilize grid operation by causing blackouts, voltage sags, and/or frequency fluctuations. As a consequence, determined adversaries with significant resources at their disposal (*Class II* attackers) are required to successfully materialize such attacks.

As for the attack model of the load-changing scenario, the attack frequency component is considered *iterative* due to the fact that in order to cause a significant effect on the system, a single attack incident may not be sufficient. The reproducibility of such stealthy and indirect attacks is set to *multiple-times*. Furthermore, the attack functional level is at *Level 1 or 2*, per the assets that are vulnerable and enable this load-changing scenario (e.g., PLCs, controllers, HMIs, etc.). Last, the attack technique that the adversaries use can either include *control logic modifications* if PLCs are targeted, or *wireless compromise* if a wireless controller is affected. In both cases, the attack targets the *cyber domain*, and specifically, the integrity of the in-transit data issued from HMIs or SCADA MTUs (i.e., communications and protocols), or the control commands to PLCs.

Attack Setup & Evaluation: In order to demonstrate the effects of load-changing attacks on CPES, we simulate such attacks targeting multiple load buses in the IEEE-39 bus system. Three vulnerable load buses (bus 16, 23, and 29) are selected as the targets for the load-changing attacks [173], as shown in Fig. 15.

In this case study, it is important to examine the dynamic impact of frequency instabilities caused by load-changing attacks. Hence, to study these frequency instabilities, we model the physical-system layer using an EMT-approach with support for real-time simulation. At this layer, the generators are modeled as synchronous machines taking into consideration the dynamics of the stator, the field, and the damper windings. An excitation system is used for the system control and protection functions designed to handle any disturbances measured in the power system [174]. Loads are modeled as constant impedance, current, and power (ZIP) models [175]. The mapping of this load changing attack case study to the CPS framework is presented in Fig. 16. The main software resource used to conduct the EMT real-time simulations of the physical-system layer is *eMegaSim* (from *Opal-RT*). The metrics used to evaluate the performance and behavior of the IEEE-39 bus test system, based on the presented CPS framework, are the physical-system layer performance metrics related to *frequency stability* (Table V).

In order to evaluate the impact of the load-changing attack on the power grid frequency, we observe the frequency variations measured at the generators' connections to the grid. We develop four different scenarios of load-changing attacks in which the system is initialized with original load values from literature [176], and the system frequency is kept at a nominal value of 60 Hz. All load-changing attacks are triggered at $t = 4sec$ with a duration of $0.5sec$. Fig. 17a shows the effect of a 20% load demand increase at bus 29. Such sudden load demand increase causes the measured frequency to decrease to around 59.87 Hz on the nearby generator 9, while having a smaller impact on other generators. At $t = 4.5sec$, when the load demand increase is terminated, the frequency fluctuates and increases to around 60.11 Hz at bus 29. Fig. 17b shows the results of a simultaneous load-changing attack that causes a 20% load demand increase at buses 29 and 16. The main difference between this case compared to the first scenario is

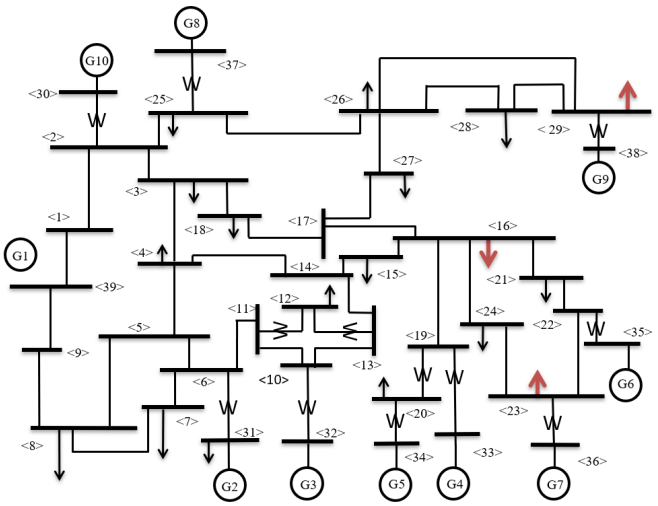


Fig. 15: Load-changing attack on IEEE-39 bus system. Load-changing attack targets are shown as red arrows in system topology.

Layers	Modeling	Resources	Metrics
Cyber-System	-	-	-
Physical-System	EMT : Real-time	eMegaSim	Frequency stability

Fig. 16: Mapping of load-changing attack case study leveraging the CPS framework.

the higher number of generators that are affected by the attack.

A load-changing attack with greater system impact is depicted in Fig.17c. In this scenario, an attack is simulated as a 50% load demand increase that affects simultaneously buses 29 and 16. Here, we observe that the frequency measured at multiple generators reaches around 59.85 Hz when the load-changing attacks are triggered at $t = 4sec$, and 60.23 Hz when the load demand reaches nominal values ($t = 4.5sec$). The final scenario is shown in 17d, where we implement an attack that suddenly increases the load demand by 50% at buses 29, 16, and 23. In this scenario, we observe how all generators in the system get heavily affected by the attack. The frequency measured at multiple generators reaches minimum and maximum values of 59.85 Hz and 60.23 Hz at the respective trigger and termination events of the attack. The most affected generators in this case study are generators 9 and 6.

In the aforementioned scenarios, the attacker is assumed to be able to alter the power consumption profiles of IoT-connected controllable loads, and therefore cause sudden load demand increase. The presented results demonstrate the feasibility and impact of load-changing attacks. The frequency fluctuations from such adverse events can lead in exceeding the nominal EPS frequency limits [177], [178], thus causing potential load-sheddings incidents or equipment failures [179]. As demonstrated in Fig. 18, EPS have in-built control and protect mechanisms to maintain the power system frequency within its nominal range. For example, the AGC mechanisms can adjust minute frequency deviations from its nominal value. However, if the EPS frequency deviates more than 0.036 Hz from the predefined grid frequency (i.e., 60Hz), the generator governor systems are employed to account for such frequency discrepancies and stabilize the system. On the other hand, during more severe incidents, such as overfrequency (at or above at 62.2Hz) or underfrequency (at or below 57.8Hz) events, switching equipment and relays will automatically trip to protect generator from such instantaneous and potentially catastrophic frequency fluctuations [180]. Furthermore, during underfrequency incidents load shedding is typically employed to bring the system frequency within acceptable operational limits (between 58.4Hz– 59.5Hz) [177]. Ancillary mechanism like the generator governors and AGC can then be utilized to bring the system back to its nominal frequency state. On the other hand, during severe events, where the frequency keeps decreasing even further, generators' CBs are tripped to protect the equipment from permanent damage.

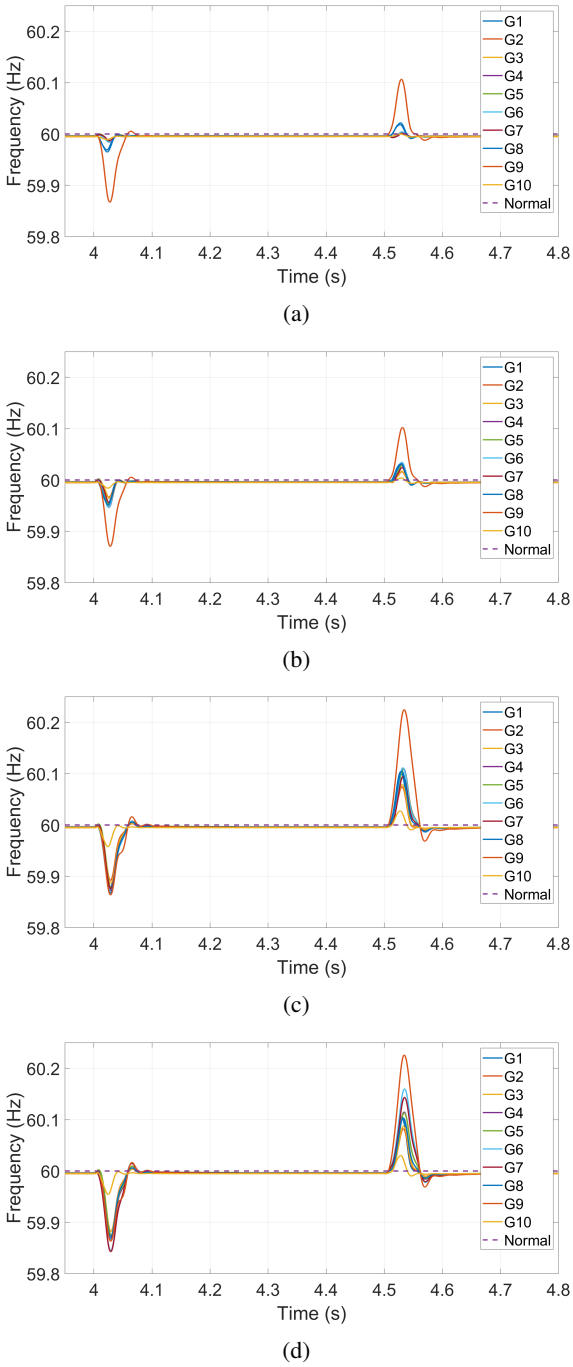


Fig. 17: Frequency variation impact on the power grid: (a) with 20% demand increased at bus 29, (b) with 20% demand increased at buses 29 and 16, (c) with 50 % load increase at buses 29 and 16, and (d) with 50 % load increase at buses 29, 16, and 23.

Risk assessment: Similar to the case study 1 of the cross-layer firmware attack, load-changing attacks require access to multiple devices to properly coordinate a successful attack. Thus, the $Threat\ Probability$ for this type of attack is set to `Medium (2)`. Following the same objective priorities depicted in Table VIII, we set the "People health and personnel safety", and "Equipment damage and legal punishment" attack impact to `Low(1)`. On the other hand, since potential protection mechanisms could be triggered in the event of a load-changing attack causing potential brownouts in order avoid cascading system effects [181], the "Uninterrupted operation, and service provision" as well as the "Organization financial profit" are set to `Medium (2)`. Consequently, the $Risk$ of the presented load-changing demand attacks are estimated to be $2 * \sum(4 + 2 + 6 + 2) = 28$.

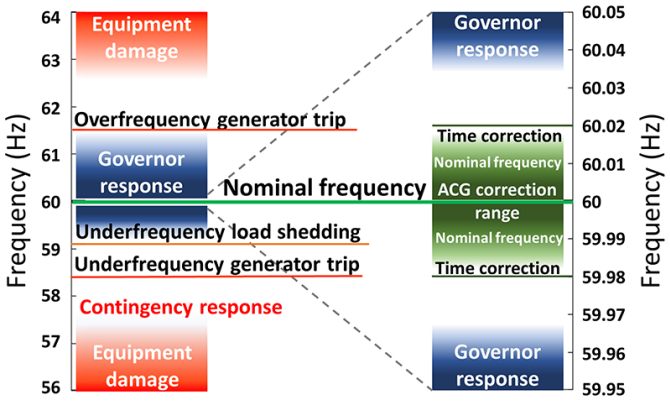


Fig. 18: Power system corrective mechanisms to maintain stability under different frequency deviations.

C. Case Study 3: Time-Delay Attacks

Background & Formulation: Time-delay attacks (TDA) are a type of DAAs where attackers aim to destabilize the operation of a compromised control system by delaying measurements and/or control commands of sensors and actuators. This type of attack does not require massive amount of attacker resources. For example, it can be implemented via network congestion caused by flooding the network with a huge amount of data, thus disrupting the nominal operation of the attacked system.

The mathematical formulation of TDAs is formulated as follows. Consider the CPS system described by Eqs. (3) – (5). If $\mathrm{T}_{\mathrm{attack}}$ is defined as the period of time when the TDA is performed, then the TDA can be structured as:

$$f_D\big(s_r(k)\big) = \begin{cases} s_r(k-d) & \text{, if } k \in \mathrm{T}_{\mathrm{attack}} \\ s_r(k) & \text{, otherwise} \end{cases} \tag{16}$$

where s_r represents the compromised signal (which can be either u, i.e., the control variable, or y, i.e., the measurements, in the CPS), f_D represents a time-delay function, and d represents either a discrete constant delay value or a time-varying delay function.

TDAs are considered a major threat to CPES due to their potential capability of disturbing the stability of islanded MGs, or even the overall power grid, by simply delaying measurements or control commands transmitted and received from sensing and control devices (e.g., smart meters, PMUs, etc.). Due to the importance of TDAs, existing literature aims to understand the complications such attacks could cause to CPES operations [32], [182], [183] . For instance, in [183], the authors present an analysis of different TDA concepts (e.g., TDA margins, boundaries, surfaces, etc.) regarding effective conditions for TDA disruptions against grid stability.

Threat model: In the TDA case study, we assume an *oblivious* adversary having essentially no knowledge of the system topology; such detailed information is not necessary to perform TDA events [184], [185]. Additionally, since this type of attack is performed by introducing substantial delays, mainly on the network level, *possession* of the targeted device is *not* required. Due to the objective of TDAs aiming to destabilize power grids by obstructing controls crucial for the system's assets operation, TDA can be seen as a *targeted* attack. Depending on the size and complexity of the compromised CPES, the adversaries might require fewer or an extensive array of skills and resources. Thus, adversaries resources for performing TDAs can be classified in either Class I or Class II type of attackers.

In order for TDAs to compromise CPES and severely impact their operation, TDAs should be performed *iteratively* and *multiple-times*. In addition, *Level 2* assets are commonly the ones being targeted by TDAs. As mentioned before, typically, TDAs occur on the *cyber domain*, i.e., *communications and protocols*, and target asset availability by tampering with control commands issued by *control server* devices. Consequently, *wireless compromise, MitM, spoofing, and DoS attacks* are the most prominent techniques adopted by adversaries to cause anomalous incidents and cascading failures based on TDAs.

Attack Setup & Evaluation: In this case study, we develop and simulate a TDA scenario in order to demonstrate its effect on a MG CPES. Specifically, in our study, a MG disconnects from the main grid by an intentional islanding command relayed from the MG controller at time $t = 10sec$. Due to the insufficient generation capacity in the system, the MG controller sends a load shedding command to a breaker that controls a controllable load. At this point, the adversary performs a TDA that will delay this load shedding command sent from the MG controller to one of the controllable loads, thus causing major disturbances at the *physical-system layer* of the CPES. The TDA occurs at the *cyber-system layer* of the CPES, so for this particular case study, models for the cyber-system layer and the physical-system layer are required to perform a *real-time co-simulation* of the respective layers.

The physical-system layer is modeled using an EMT-simulation approach with support for real-time simulation. At this layer, the MG is modeled as a test system composed by a conventional generator operated using a frequency control mechanism rated at 1 MW, a Li-ion BESS rated at 100 kW/100 kWh, two controllable loads rated at 300 kW (load #1) and 700 kW (load #2) , and a critical (non-sheddable) load rated at 200 kW. The main software resource used to conduct the EMT real-time simulations of the physical-system layer for this case study is *eMegaSim* (from *Opal-RT*). The cyber-system layer is modeled using a communication network emulation platform that supports co-simulation capabilities. Specifically, the software resource used to model the communication network that represents the cyber-system layer is *EXataCPS*.

Every MG component from the physical layer is mapped with a virtual communication node inside the network emulation platform. The backbone of the communication network is represented by a network router. The network router is responsible for sending control commands and receiving measurements from the MG components, i.e., BESS, loads, and generator, to the MG controller, respectively. The communication protocol used is the IEEE Std 1815, commonly known as DNP3. IEDs in the network are modeled as DNP3 outstations, and communicate with the MG controller which is modeled as a DNP3 master. The DNP3 master and outsta-

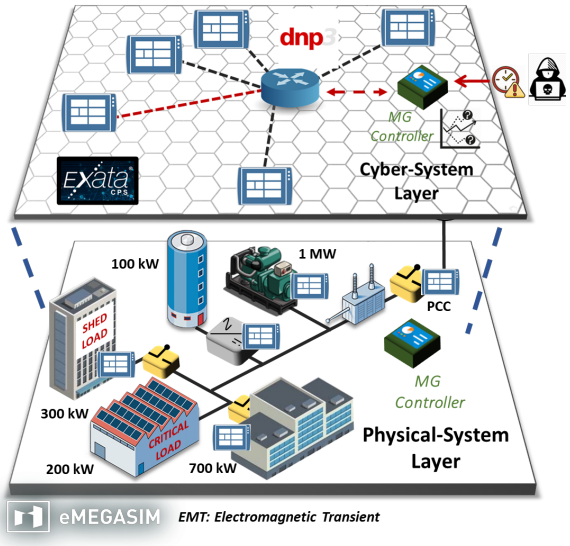
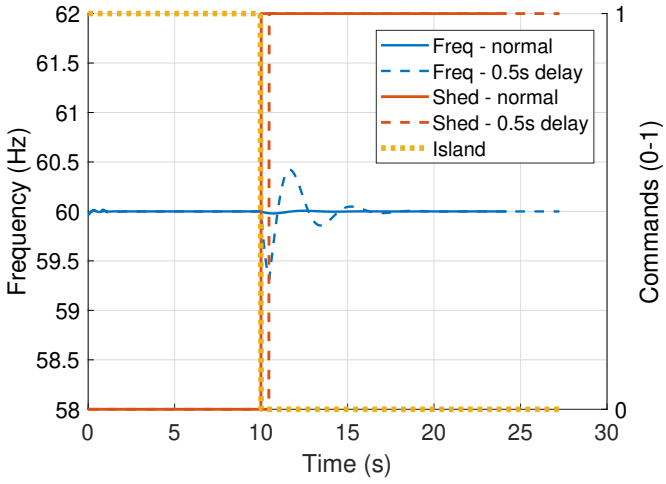


Fig. 19: Conceptual illustration of the real-time co-simulation MG system testbed used in the TDA case study.

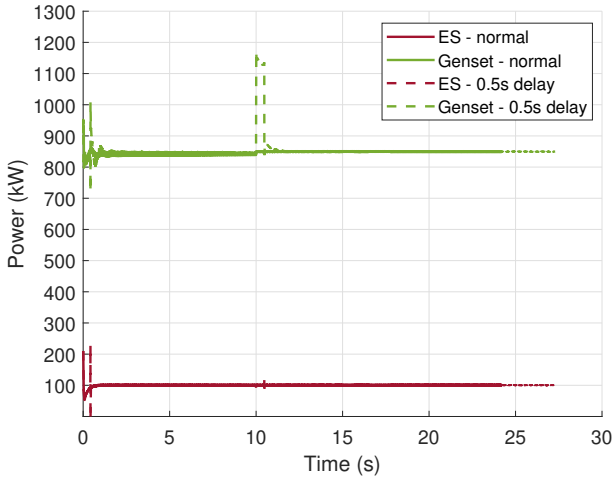
tion devices exchange data and control commands including, power generation, load consumption, breaker status, etc. The connections between the communication network nodes are modeled as wired 802.3 Ethernet connections with 100 Mbps bandwidth. Fig. 19 shows a conceptual illustration of the real-time co-simulation scenario designed to perform the described case study. The metrics used to evaluate the performance and behavior of the MG operation, based on the proposed CPS framework, are the physical-system layer performance metrics related to *frequency stability*, and the cyber-system layer performance metrics related to *average end-to-end delay* and *total number of packets delayed* by the TDA.

Based on the described setup, the impact of a malicious TDA in an islanded MG system is evaluated. An attacker compromises the communication link between the MG controller and the IED controlling the disconnection of the breaker at the controllable (sheddable) load #1 (300 kW). Three different attack test cases are evaluated by varying the time-delay duration of the TDA. These delays are $0.5sec$, $5sec$, and $15sec$ approximately. In the communication network, the attacks are modeled by modifying the exchanged packets with introducing a timing delay between the DNP3 master and the corresponding outstation.

The first attack scenario shows a $0.5sec$ TDA that blocks the load shedding command performed by the MG controller. Fig. 20a showcases the impact of the $0.5sec$ seconds TDA when compared to the normal operation of the MG system. In the graph, we observe how at $t = 10sec$ the breaker at the point of common coupling (PCC) is disconnected, i.e., breaker command goes from 1 to 0, in order to perform intentional islanding of the MG. Then, due to the insufficient generation capacity, the MG controller sheds controllable load #1 (shed command goes from 0 to 1). In the normal operation case, the shedding procedure is performed as soon as the MG islands, while in the TDA scenario the shedding procedure gets delayed



(a)

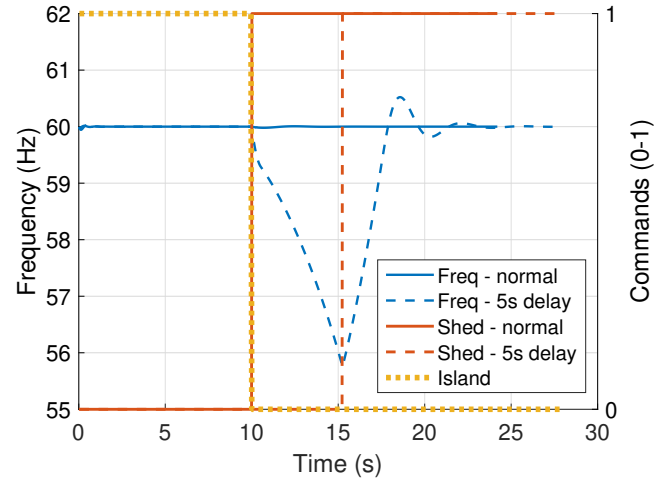


(b)

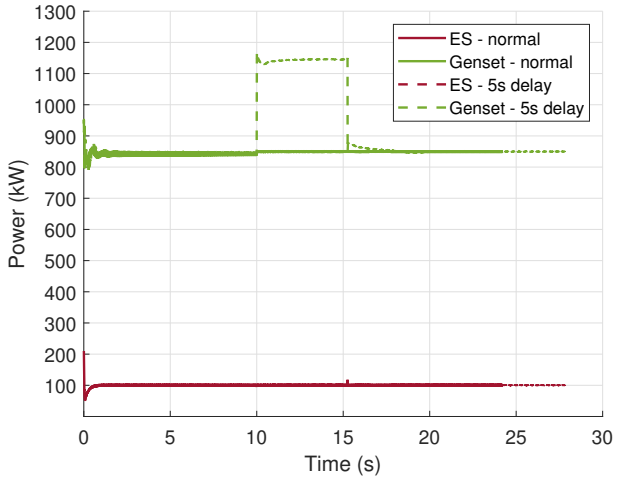
Fig. 20: Normal operation Vs. $0.5sec$ time-delay attack (TDA) scenario: (a) EPS frequency response during a $0.5sec$ time-delay attack (TDA) on the islanding command, and (b) generator power fluctuation during the $0.5sec$ time-delay attack (TDA).

by the amount of the time-delay attack. Notably, the maximum and minimum values of the MG frequency during the normal operation scenario are 60.02 and 59.71 Hz, respectively. On the other hand, the maximum and minimum values of the MG frequency during the $0.5sec$ TDA scenario are 60.42 and 59.32 Hz, indicating (see Fig. 18) that system operators would have to employ emergency corrective measures to maintain system stability. Fig. 20b depicts the output power of the generator set and the ESS during both scenarios.

Similarly, the second test scenario demonstrates a $5sec$ TDA that blocks the load shedding command performed by the MG controller. Fig. 21a presents the impact of the $5sec$ TDA when compared to the normal operation of the MG system. As seen, the impact in the operating frequency of the MG is greater than the first test scenario due to the sustained timing attack. The $5sec$ TDA causes a maximum and minimum MG frequency of 60.52 and 55.75 Hz, respectively. Granted

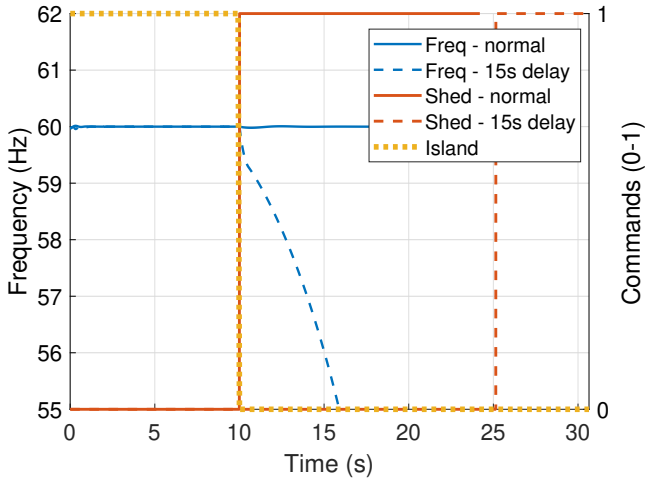


(a)

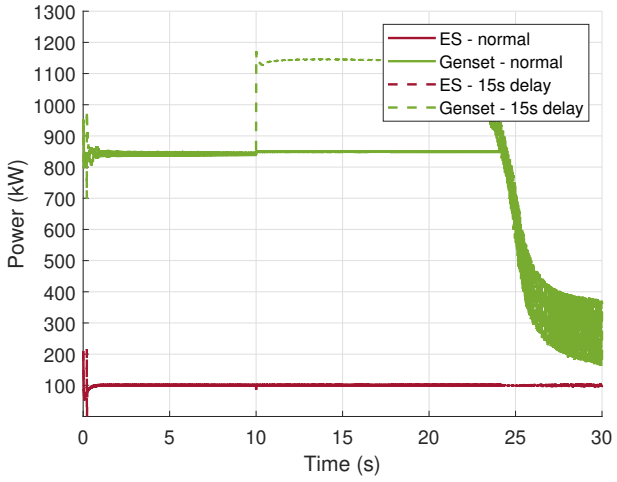


(b)

Fig. 21: Normal operation Vs. $5sec$ time-delay attack (TDA) scenario: (a) EPS frequency response during a $5sec$ time-delay attack (TDA) on the islanding command, and (b) generator power fluctuation during the $5sec$ time-delay attack (TDA).



(a)



(b)

Fig. 22: Normal operation Vs. $15sec$ time-delay attack (TDA) scenario: (a) EPS frequency response during a $15sec$ time-delay attack (TDA) on the islanding command, and (b) generator power fluctuation during the $15sec$ time-delay attack (TDA).

the substantial under-frequency incident, i.e., 55.75 Hz, load-curtailment along with generator tripping would have to be enforced to protect the EPS equipment and avoid the incident propagation leading to a generalized grid collapse (Fig. 18). As a result, this attack case demonstrates the potential of TDAs to greatly disrupt the operation of the system causing major equipment damages.

In the third test scenario, we perform a $15sec$ TDA that blocks the load shedding command performed by the MG controller. This case is analogous to a DoS attack, due to the long period of the TDA, which can greatly disrupt the operation of the MG's load shedding mechanism. As seen in Fig. 22a, this scenario demonstrates the worst-case scenario of a TDA to the CPES. The MG frequency decreases rapidly until it hits a minimum value of 15.31 Hz. Additionally, as depicted in Fig. 22b, the frequency-mode generator set is not capable of maintaining the stability of the system for such a prolonged period causing large oscillations in its power output. Notably, in realistic systems frequency violations should be averted before reaching such extreme values (e.g., 15.31 Hz). However, by leveraging the CPES framework we can perform worst-case scenario analyses, evaluate the system behavior under coordinated attacks (e.g., if an attacker disables automated grid safety mechanisms), and identify critical system components and contingencies without endangering the EPS operation.

In order to explore the behavior of the CPES at the cyber-system layer, we analyze two metrics that provide us important information regarding the response of the communication devices to the TDA. These two metrics are the *average end-to-end delay* at the communication network, and the *number of packets delayed* by the TDA. Fig. 23 shows the average end-to-end delay of all the network devices communicating using DNP3 at the cyber-system layer. Fig. 24 presents the total number of packets delayed due to the TDA that compromises the correct operation of the CPES according to two of the

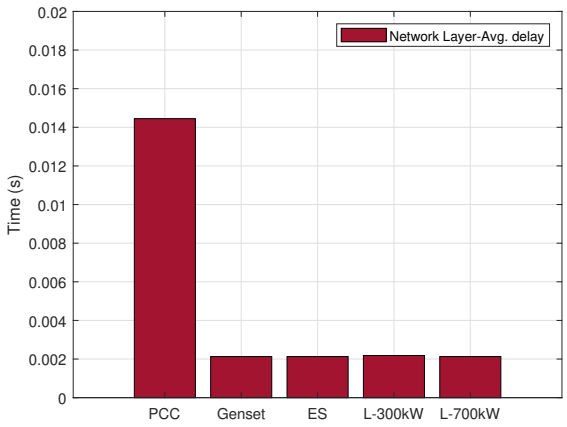


Fig. 23: Average end-to-end delay of all nodes in the communication network (cyber layer).

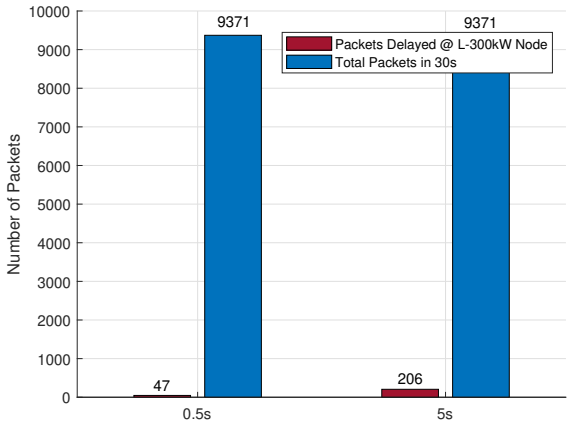


Fig. 24: Number of packets delayed by time-delay attack (TDA) at *delay* = $0.5sec$ and *delay* = $5sec$ Vs. *total number of packets in* $30sec$.

TDA scenarios ($0.5sec$ and $5sec$ TDA). As seen in Fig. 23, the average end-to-end delay of the communication network, operating under normal conditions, has a maximum value of $0.0144sec$. This value is related to the *master* DNP3 device located at the PCC that is communicating with all the DNP3 *outstations*. This is the average time that the MG controller takes to communicate the load shedding signal to the respective sheddable load under normal operating conditions. In contrast, the TDA compromises the system's operation by delaying the load shedding signal based on the scenarios presented previously. In order to get more details regarding the attack study, the total number of packets delayed by the TDA are measured and plotted in Fig. 24. Here, we observe a side-by-side comparison of the number of packets delayed in two of the presented test scenarios, i.e., $0.5sec$ and $5sec$ delay scenarios, and the total number of packets sent by the *master* and *outstation* devices in the $30sec$ real-time co-simulation. Fig. 25 shows the mapping of the presented case study with the CPS framework.

Layers	Modeling	Resources	Metrics
Cyber-System	Emulation	EXataCPS	- Avg. delay - # Packets delayed
Physical-System	EMT : Real-time	eMegaSim	Frequency stability

Fig. 25: Mapping of TDA case study with CPS framework.

Risk assessment: In this type of attack, an adversary does not require significant resources or capabilities to compromise the CPES, as long as the system has not been fortified with state-of-the-art defense mechanisms. This "low-bar" requirement of resources increases the probability of successfully performing such attack on a vulnerable CPES. As a result, the $Threat\ Probability$ for the TDA case study is set to `High (3)`. The impact on "People health and personnel safety" as well as the "Organization financial profit" are set to `Low (1)`. However, TDAs can potentially cause severe impact on the grid operation. For this reason the "Uninterrupted operation and service provision" is set to `Medium (2)`, and the "Equipment damage and legal punishment" objective priority is set to `High (3)`. The resulting $Risk$ for the TDA is estimated to be equal to $3 * \sum(4 + 1 + 6 + 6) = 51$.

D. Case Study 4: Propagating Attacks in Integrated Transmission and Distribution (T&D) CPES

Background & Formulation: As mentioned in Section IV, T&D integrated models for real-time simulation within CPES co-simulation testbeds can provide comprehensive and accurate simulation results able to capture the dynamic behavior of CPES. Specifically, integrated T&D model co-simulations can be used to holistically evaluate the impact of disruptions (e.g., malicious attacks, faults, etc.) in EPS, and exhibit how maloperations on the transmission system extend to the distribution and vice versa. Thus, in this case study, we present DIA-type of attacks as propagating processes, such as computer viruses do, evaluated in real-time integrated T&D simulation models.

EMT and TS power system simulations often model only the transmission or the distribution system of power grids. This is mostly due to the high computational power required to have a real-time simulation model of an entire EPS [143]. Aggregated distribution system sections are typically replaced by static or dynamic loads when simulating transmission system models [170]. Correspondingly, the transmission system's behavior is often abstracted using ideal voltage sources in distribution system modeling [186] . In addition, T&D models are usually simplified to a single-phase representation [187], [188]. Such modeling approach loses key information related to the behavior of highly unbalanced distribution systems. In reality, T&D systems are highly coupled [189], and in order to perform comprehensive and accurate security assessment and impact analysis studies in CPES, both T&D domains need to be accurately modeled and simulated in a coordinated fashion. This coordination involves a clock-synchronized loop in which, even if the two models are executed on different cores of a machine, they communicate in parallel to match boundary conditions (i.e., voltages, power values, etc.) at every simulation step, as seen in Fig. 26.

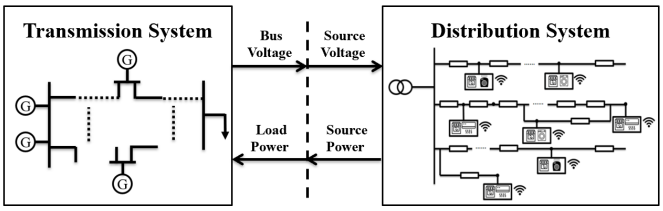


Fig. 26: Transmission and distribution (T&D) integrated simulation system setup.

There are different techniques that can be used to develop real-time integrated T&D models. Different platforms provide different solutions and methods that allow the parallel execution of different systems in real-time EMT environments. In general, the overall T&D system is separated into different groups (assigned to different cores of the machine) that are solved individually using a state-space approach. State-space equations and matrices are used to describe the system group dynamics, while the interaction between the groups is solved using a nodal admittance method [190]. In the state-space approach the physical-system is modeled as:

$$s' = A_q s + D_q v \tag{17}$$

$$o = E_q s + F_q v \tag{18}$$

where s is the state vector, v is the input vector, o is the output vector, and A, D, E, and F are the state-space matrices. The term q represents the size of the matrices.

In a typical EMT state-space implementation, such as the one available in *Matlab Simscape Power Systems*, every time a switch changes status (on/off), the entire state-space solutions are re-computed. Using such approach for real-time simulation ($<\approx 50\mu$s simulation time-step) of large interconnected T&D systems could be infeasible due to the required computational resources. With every single status change within the system model, the state-space outputs of the entire system would need to be re-computed. To address this computational issue, platforms such as *Opal-RT*, and its Advanced Real-Time Electro-Magnetic Solvers (*ARTEMiS*) package, use state-space nodal methods [191]. *ARTEMiS* implementations discretize, pre-compute, and store into cache memory, the state-space matrices for all the combinations of switch topologies which can occur. Then, using a nodal method, the common voltages, admittances, and currents of the system (i.e., shared values between groups) are solved as:

$$VY = I \tag{19}$$

where V, I, and Y are the respective common voltages, currents, and admittance matrices at the boundaries of the groups. In essence, the use of this approach improves the accuracy and computational execution time of the entire system's solutions. As a result, this is a feasible way for simulating a real-time integrated T&D system and evaluating the propagation impact of adverse disruptions, e.g., faults, attacks, etc.

Threat model: Integrated T&D models can be seen as complex structures. Depending on the T&D aspect targeted by an adversary and the type of the attack, the threat model may be adjusted to the specific details. For our use case, we assume an adversary with *strong knowledge* of the system's topology and its components. Additionally, in our setup the adversary aims to destabilize the integrated T&D system by maliciously controlling switching devices, i.e., the CBs, thus *possession* of the device is assumed. In the worst-case scenario analysis, the attackers could lead the CPES towards full system collapse, designating a *targeted attack* by *Class II* adversaries with abundant resources (e.g., nation-state funded groups).

In terms of the attack model formulation, the attack frequency is *non-iterative*, since compromising a critical system asset (crown-jewel) could impact the overall system. The reproduction of such types of attacks can be seen as impractical due to their high system impact. Thus, we model them as *one-time attacks*. The attack level is presumed to be *Level 2* since critical system components need to be compromised. Such assets for our case include engineering workstations since the attacker targets – in a DIA-type of event – the control and coordination between the T&D systems. The attack technique is correspondingly an *engineering workstation compromise*. Directly issuing malicious commands from an engineering workstation can also be a possible attack path, assuming a malicious insider scenario. However, in our case study, we assume a sophisticated and stealthy attack implemented on the *cyber domain* targeting the data integrity of the issued control commands from the engineering workstations (DIA). For instance, disruptions on the T&D can occur by falsifying the in-transit data exchanged between engineering workstations and CB control devices, triggering unexpected CB tripping and system sectionalizations.

Attack Setup: In this case study, an integrated real-time EMT T&D system is modeled in order to investigate different interactions of propagating attacks and disturbances between a transmission and an unbalanced distribution system. Specifically, we integrate a transmission system, modeled as the IEEE-9 bus system, with a distribution system, modeled as the IEEE-13 bus test system. In order to match the power generation and load consumption between the power grid benchmarks used, we scale some of the systems' parameters. For example, the active power and reactive power of the generators and the loads in the transmission system are reduced by an order of magnitude, while all the loads in the distribution system are increased by an order of magnitude. Additionally, as shown in Fig. 27, the load at bus 5 of the IEEE-9 bus transmission model is "replaced" by the IEEE-13 bus distribution system. Generator 1 (G1) is used as the slack bus. The EMT modeling and real-time simulation of this case study's physical-system layer are performed using *eMegaSim* of *Opal-RT*.

In order to evaluate the bi-directional impact of propagation attacks in integrated T&D models of CPES, we develop two attack scenarios in this case study. The first scenario assumes that the adversary has the capability of altering the EPS topology. This can be achieved by decoupling the T&D system at the PCC via a DIA attack on the EPS switch devices (i.e., the distribution feeder CBs). The second scenario demonstrates the impact on the distribution system when transmission system

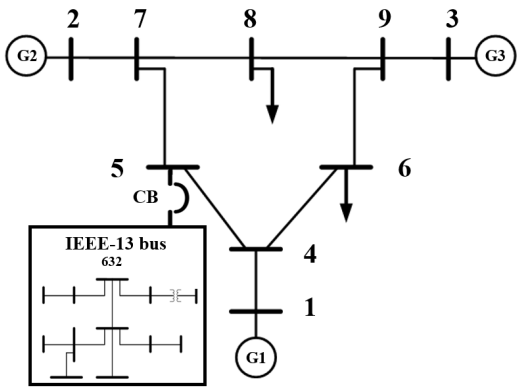


Fig. 27: The integrated T&D real-time simulation model: it includes the IEEE-9 bus system (transmission model) and IEEE-13 bus system (distribution model).

components are compromised. Following our CPS framework, the metrics used to evaluate the performance and behavior of the T&D system under the propagation attack scenarios, are the physical-system layer performance metrics related to *frequency stability* and *voltage stability*.

In the first propagation attack scenario, we assume that the adversary by tripping the CBs at the PCC between the T&D system can disturb the EPS frequency impacting its operation and potentially causing damages to field equipment (e.g., transformers, commercial and residential loads, etc.). Different attack paths can be pursued to compromise and decouple T&D systems. For instance, such adversarial objectives (i.e., T&D decoupling) can be achieved by *i)* intruding via the communication infrastructure and remotely manipulating the control tags issued by engineering workstations located at the system operation management facilities, *ii)* implementing DoS attacks on the targeted PLCs, disabling the CBs, *iii)* compromising the controller logic of IED-enabled switching equipment, or *iv)* penetrating the utility SCADA network and maliciously manipulating control settings (e.g., over/under-voltage or current limits) [192].

The results presented in Fig. 28a are measured at the generator buses, i.e., buses 1, 2, and 3 of Fig. 27. The frequency at the transmission side of the network rapidly increases when the CB is tripped at $t = 1.5sec$, and returns to its nominal values at around $t = 1.8sec$. The peak frequency value is around 60.23 Hz. Fig.28b shows the frequency response of the system when the attacker opens the CB between the T&D system at $t = 1.5sec$, and then closes it after 15 cycles (approximately $0.25sec$ later) which would avoid triggering any protection countermeasures during this intermittent frequency transient [169]. We observe how such attacks could stealthily destabilize the EPS just by tampering with the CB controls between different zones of the power grid. Here, two main fluctuations are observed following the CB tripping behavior, one between $t = 1.5sec$ and $t = 1.75sec$ when the CB is tripped open, and between $t = 1.75sec$ and $t = 2sec$ when the CB is closed. The last scenario assumes an attacker aiming to damage system components by asynchronously changing the status of the CB multiple times. Fig. 28c demonstrates the frequency

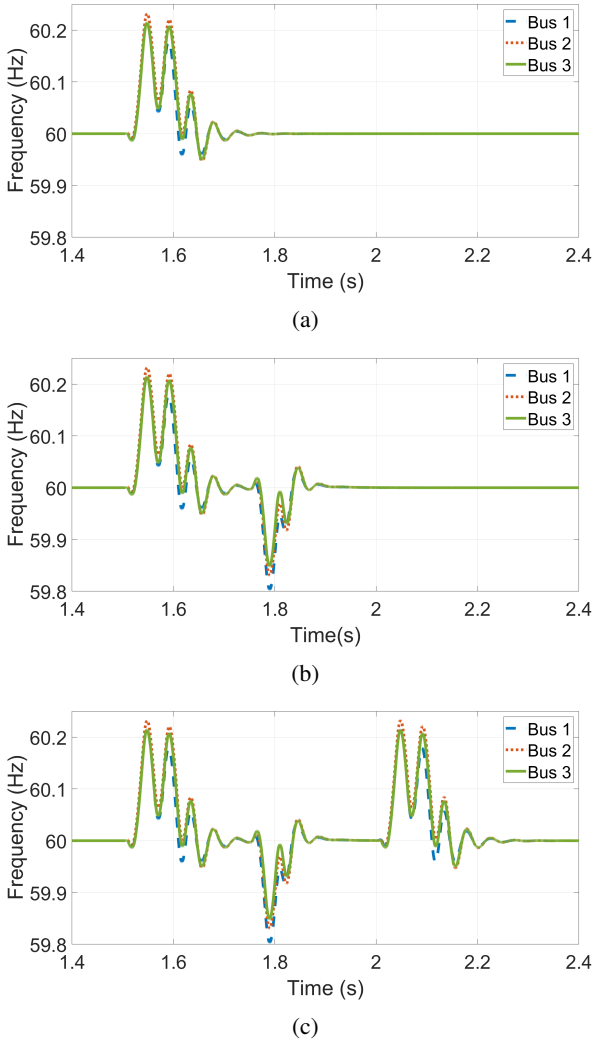
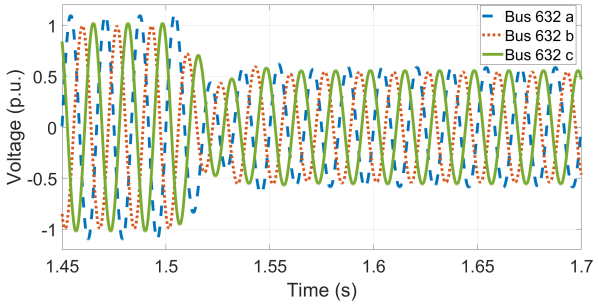


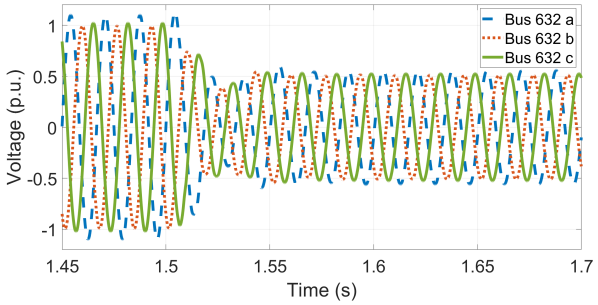
Fig. 28: Frequency response when the CB at bus 5 of the IEEE-9 bus system is: (a) opened at $t = 1.5sec$, (b) opened at $t = 1.5sec$ and then closed at $t = 1.75sec$, and (c) opened at $t = 1.5sec$, closed at $t = 1.75sec$, and opened again at $t = 2sec$.

fluctuations on the generator buses when the CB between the T&D systems is opened at $t = 1.5sec$, closed at $t = 1.75sec$, and opened again at $t = 2sec$. Notably, if safety mechanisms are not promptly enforced, the frequency instabilities occurring between $t = 1.5sec$ and $t = 2.4sec$ could affect frequency-sensitive grid components (i.e., consumer, commercial, and industrial loads), and impact grid equipment and control functions (e.g., generators, transformers, automated voltage control, etc.).

In the second scenario, it is assumed that an adversary has the capability of compromising components at the transmission side of the power grid. For example, such types of attacks have been experimentally evaluated and indicated that if they last around three minutes, they can cause permanent damage on generators [120]. In this use case, our aim is to evaluate how an attack on the transmission level can propagate on the distribution level and manifest as a voltage

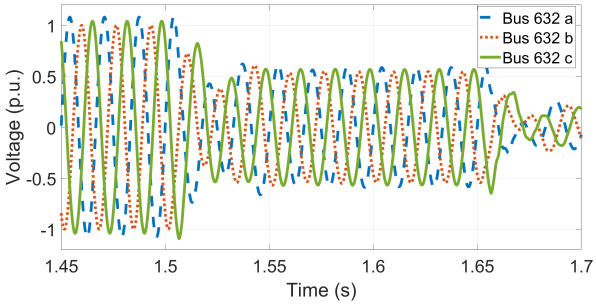


(a)

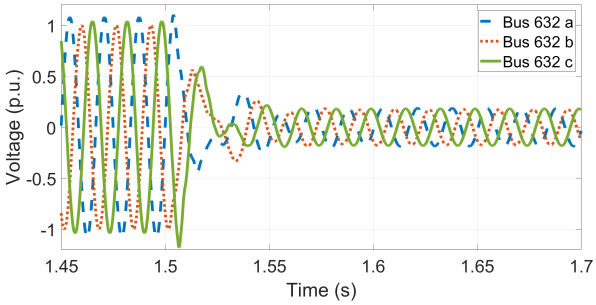


(b)

Fig. 29: Voltage response at bus 632 (distribution system) during N-1 transmission system contingencies when: a) generator G2 is disconnected at $t = 1.5sec$, and (b) generator G3 is disconnected at $t = 1.5sec$.



(a)



(b)

Fig. 30: Voltage response at bus 632 (distribution system) during transmission system contingency scenarios:a) N-1-1 contingency where generators G2 is disconnected at $t = 1.5sec$ and G3 at $t = 1.6sec$ consecutively, and (b) N-2 contingency where generators G2 and G3 are disconnected at $t = 1.5sec$ simultaneously.

variation. Contingency analysis is employed in the integrated simulation environment to study the effects of transmissions-side adverse events on the distribution system. In more detail, contingency analysis is a simulation-based system analysis tool used to assess the impact of various combinations of component failures occurring in transmission systems. The North American Electric Reliability Corporation (NERC) enforces a $N-1$ constraint for the U.S. power grid, which means that EPS transmission systems need to maintain nominal operation even if one component fails [193]. Such components include generators, transmission lines, transformers, etc. Depending on the security level of the EPS, a higher $N-k$ criterion may be required, where $k \geq 2$ represents two or more contingency events. For example, a nuclear plant may be required to satisfy a $N-2$ constraint, allowing the grid to withstand the simultaneous failure of two components.

For the purpose of our study, we assume that the attacker is able to compromise one or more components of the transmission system, causing under-voltage events at the distribution-side. When component failures occur, the system aims to maintain stability. However, the intermittent transmission system instability along with the potential inability to support power demand, results in voltage deviations which are also propagated to the distribution level. Four main sub-cases are designed in this second scenario to illustrate the corresponding voltage impact at bus 632 of the distribution system. The first two sub-cases consider an attacker that compromises one generator (G2 or G3) once at a time ($N-1$), where the rest sub-cases consider an attack on two generators (G2 and G3) consecutively ($N-1-1$), or simultaneously ($N-2$). In all sub-cases, we evaluate the voltage variation (depicted in per unit – p.u.) measured at bus 632.

As seen in Fig. 29a and Fig. 29b, the voltage measured at bus 632 of the distribution system drops from 1 p.u. to 0.5 p.u. at $t = 1.5sec$, i.e., when one of the generators (G2 or G3) is disconnected from the transmission system ($N-1$). Fig. 30a demonstrates the voltage variations of the $N-1-1$ contingency event in which G2 and G3 are disconnected at $t = 1.5sec$ and *t=1.6sec*, respectively. During this case, the bus voltage initially drops from 1 p.u. to 0.5 p.u. (G2 disconnection), and then to 0.2 p.u. when G3 is also disconnected. In the $N-2$ case, presented in Fig. 30b, the simultaneous disconnection of G2 and G3 from the system lowers the voltage significantly at $t = 1.5sec$. The voltage measured at bus 632 of the distribution system decreases to under 0.2 p.u. within $0.05sec$. Fig. 31 illustrates the mapping of the propagating attack case study in T&D systems with the CPS framework.

Risk Assessment: Compromising T&D systems requires determined adversaries possessing both strong knowledge of the system architecture as well as ample resources, since these can enhance the probability of materializing successful attacks. Thus, we set the *Threat Probability* to `High (3)` [194]–[197]. Attackers could perform stealthy and disastrous attacks by leveraging the knowledge of system topology, asset

Layers	Modeling	Resources	Metrics
Cyber-System	-	-	-
Physical-System	EMT : Real-time	eMegaSim	Frequency & voltage stability

Fig. 31: Mapping of T&D case study with CPS framework.

placement information, power demand profiles, etc. As a result, by targeting mission-critical system components during peak utilization periods (e.g., peak power demand time of the day), adversaries could maximize the corresponding attack impact [82]. A power system collapse (e.g., blackout), which could be the impact of the successful propagation of T&D attacks, can significantly affect "People health and personnel safety", "Uninterrupted operation and service provision", and "Equipment damage and legal punishment". Based on these assumptions, the $ResultingDamage$ is set to `High (3)`, while the "Organization financial profit" is set to `Low (1)`. The aggregated $Risk$ for this type of attack can be estimated to be $3 * \sum(12 + 9 + 6 + 1) = 84$.

E. Post-risk assessment discussion

The next step after the risk score calculation includes the risk prioritization. The risk identification, assessment and prioritization serve as preliminary steps and are critical for the decision making and formulation of mitigation plans. Specifically, in this work, we have considered four diverse attack use cases aimed at CPES while targeting different cyber or physical subsystems or components. In more detail, we discuss cross-layer firmware attacks with a calculated risk score equal to 22, load changing attacks with risk score equal to 28, TDAs with a risk score evaluated to 51, and finally propagating attacks targeting integrated T&D CPES with an 84 risk score. These risk scores provide a useful way to perform one-to-one comparisons between attacks even if their specifics are unknown.

Attacks with higher risk scores (e.g., the T&D propagation attacks) will induce higher impact on the system when compared to other attacks such as the cross-layer firmware attack with a less pronounced score. In Section III-C, we justify how the use of each case's risk score variations depend on the corresponding attack characteristics (e.g., threat probability, objective priorities, and potential impact on the CPES operation). As a result, attacks similar the one targeting the integrated T&D system, aim to affect almost every CPES operational objective. They are attractive from an adversarial perspective due to the maximization of the inflicted system disruption. Hence, such attacks will obtain high risk scores. The same cannot be argued for attacks which can be sustained even post compromise, targeting less critical CPES equipment.

The risk score-based ranking helps categorizing the attacks (and their corresponding risks) into pools [91]. For example, assuming that we have four pools, the most devastating attacks (i.e., with scores greater than a system-defined threshold) would be placed into pool 1, while less critical attacks – with smaller risk scores – would be allocated to pools 2 - 4 in a descending risk score fashion. For each of the pools, predefined strategies are designed to mediate potential attacks. Typically, attacks belonging to pool 1 should be mitigated at all costs since they can compromise the whole system (in our case the CPES). However, the mitigation of attacks belonging to lower ranked pools might either be *i)*deferred if they do not pose significant threats to system operation, *ii)* transferred to other parties instead of allocating system resources to resolve them, or *iii)* accepted, if the cost of mitigating them outweighs the impact that will be inflicted on the system. Thus, risk assessment does not only provide better awareness of system risks and an efficient way to perform risk comparisons, but it can also automate the process of handling risks and administering corrective measures.

VI. Conclusions

In this work, we provide a comprehensive analysis of CPS security, with particular emphasis on CPES applications. The first step in this process encompasses an extensive threat modeling procedure, where adversary and attack models are constructed. The adversary and attack models provide an in-depth understanding of attackers' motives and capabilities, in addition to the attack's details including potential entry points, attack techniques, and end goals. The next step in the analysis includes the presentation of a CPS framework, where the resources, metrics, and modeling techniques needed to effectively evaluate CPS, and more specifically CPES, are discussed in detail. This framework is designed with the objective of assisting researchers and stakeholders identify the models and resources required to perform high-fidelity and reliable CPS studies. Furthermore, we present a risk assessment methodology that leverages both the treat modeling as well as the CPS framework to characterize system risks.

In order to illustrate the suitability of the overall methodology and description of the CPES security landscape, we investigate four attack case studies. For each scenario, we provide a fundamental background alongside its mathematical formulation, and discuss the corresponding threat model and attack setups. The presented case studies are simulated under nominal and abnormal operating conditions to uncover their system-wide impacts. Risk assessment analysis is also performed as part of each case's security investigation. During the risk assessment stage, we calculate the relative risk scores indicating the severity of each compromise. The risk scores correspond to the discussed studies, the threat scenarios, and the targeted assets (e.g., microinverters, T&D system, time-delay, etc.). These scores can be utilized for the ranking and prioritization of possible disruptions, and the determination of proper risk mitigation strategies to address malicious attacks implications.

The holistic approach and studies presented in this paper provide guidelines for modeling CPS threats as well as designing, simulating, and evaluating detailed CPS models. The presented framework can promote rigorous security analysis of CPS. Our future work will extend this framework and advance its capabilities even further, allowing for:

- Secure and resilient CPES operation: In this work, we have stressed the importance for cyber-secure CPES as well as

that the integration of contemporary cyber features and new physical components can increase the attack surface. Emphasis though, should not only be placed on detecting attacks, limiting and mitigating them but also in designing fault-tolerant and resilient CPES. Having identified potential vulnerabilities present in the CPES and leveraging our framework, we will define resiliency methodologies and metrics to assess CPES posture. In more detail, the resiliency methodologies will serve as CPES design best practices promoting the design of robust systems with in-built redundancy mechanisms if adverse scenarios occur. On the other hand, the resiliency metrics will be ported to our current framework and have a twofold objective, *i)* they will indicate how effectively the system can handle adverse circumstances, and *ii)* they will serve as criteria for the categorizations of CPES based on their ability to withstand attacks.

- Autonomous CPES operation and simulation-aided risk assessments: CPES are becoming more sophisticated and support a plethora of automated processes (e.g., automated control mechanisms, PLCs, AGC, etc.). Such automated systems should be capable to make real-time decisions, especially for time-critical parts of CPES, and coordinate the dynamic system behavior. It is expected that CPES will become more complex and densely interconnected as they integrate more features (remote access and control, assets, communications protocols, etc.). During their autonomous operation, the system might encounter unexpected states (e.g., unintended faults during natural disasters, or malicious attacks) that might require specific handling. Thus, determining and evaluating their security should be facilitated in a dynamic, albeit abstract way. Following this approach guarantees that every unexpected scenario will be accounted for, as well as that adverse situations will be timely prevented. Digital twin system configurations can achieve these objectives and enable the design and real-time evaluation of risk mitigation strategies. As a result, a CPES testbed will be designed supporting fully-automated operation, and incident-response structures, where attacks can be promptly detected and optimally mitigated, eliminating any adverse consequence on the actual system.
- Dynamic reconfiguration and self-healing capabilities: Securing CPES should be viewed from two directions. The first direction includes the security measures and practices which should be employed to protect system operations and avert attackers. On the other hand, the second direction features the policies and strategies which should be pursued post-compromise or during dire circumstances. The first direction has been extensively discussed in this paper; we aim to account for the second direction in our future framework extensions. Specifically, utilizing our framework and system resources we will provide classes of crisis-handling plans promoting CPES self-healing capabilities. These classes will provide tailor-made strategies to overcome emergencies, depending on the current state of the CPES and under-investigation scenario characteristics. For example, during a transmission system contingency, the corresponding class would provide alternative ways to dispatch power overcoming this issue and potential predicaments. These dynamic re-configurations and self-healing CPES capabilities will stimulate the design of future secure and resilient systems and prove invaluable tools for system operators.

REFERENCES

[1] NIST, "Cyber-Physical Systems." [Online]. Available: https://www.nist.gov/el/cyber-physical-systems

[2] S. McLaughlin, C. Konstantinou, X. Wang, L. Davi, A.-R. Sadeghi, M. Maniatakos, and R. Karri, "The cybersecurity landscape in industrial control systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1039–1057, 2016.

[3] NIST, "CVE-2018-0296 Detail," https://nvd.nist.gov/vuln/detail/CVE-2018-0296, 2018.

[4] MITRE, "CVE-2018-0296," https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-0296, 2018.

[5] North American Electric Reliability Corporation, "NERC lesson learned, risks posed by firewall firmware vulnerabilities," 2019.

[6] Gartner, "Focus on the Biggest Security Threats, Not the Most Publicized," https://www.gartner.com/smarterwithgartner/focus-on-the-biggest-security-threats-not-the-most-publicized/, 2017.

[7] J. Weeks. U.s. electrical grid undergoes massive transition to connect to renewables. [Online]. Available: https://www.scientificamerican.com/article/what-is-the-smart-grid/

[8] D. Gritzalis, M. Theocharidou, and G. Stergiopoulos, "Critical infrastructure security and resilience," *January): 0–16*, 2019.

[9] A. Keliris, C. Konstantinou, N. G. Tsoutsos, R. Baiad, and M. Maniatakos, "Enabling multi-layer cyber-security assessment of industrial control systems through hardware-in-the-loop testbeds," in *2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*. IEEE, 2016, pp. 511–518.

[10] C. Konstantinou, "Towards a secure and resilient all-renewable energy grid for smart cities," *IEEE Consumer Electronics Magazine*, 2021.

[11] S. Muyeen and S. Rahman, *Communication, control and security challenges for the smart grid*. The Institution of Engineering and Technology, 2017.

[12] J. Ospina, X. Liu, C. Konstantinou, and Y. Dvorkin, "On the feasibility of load-changing attacks in power systems during the covid-19 pandemic," *IEEE Access*, vol. 9, pp. 2545–2563, 2021.

[13] A. Dabrowski, J. Ullrich, and E. R. Weippl, "Grid shock: Coordinated load-changing attacks on power grids: The non-smart power grid is vulnerable to cyber attacks as well," in *Proceedings of the 33rd Annual Computer Security Applications Conference*, ser. ACSAC 2017. New York, NY, USA: Association for Computing Machinery, 2017, p. 303–314. [Online]. Available: https://doi.org/10.1145/3134600.3134639

[14] C. Konstantinou and M. Maniatakos, "Security analysis of smart grid," pp. 451–487, 2017. [Online]. Available: https://digital-library.theiet.org/content/books/10.1049/pbpo095e_ch15

[15] "Idaho National Lab Grid Resilience Program." [Online]. Available: https://inl.gov/research-programs/grid-resilience/

[16] "Idaho National Lab Resilience Optimization Center." [Online]. Available: https://factsheets.inl.gov/FactSheets/INLResilienceOptimizationCenter.pdf

[17] "Idaho National Lab Infrastructure and Capabilities." [Online]. Available: https://factsheets.inl.gov/SitePages/AboutINLFactSheets-Internal.aspx

[18] "Idaho National Lab Nuclear Programs." [Online]. Available: https://factsheets.inl.gov/FactSheets/NuclearPrograms.pdf

[19] "Idaho National Lab Nuclear Laboratory." [Online]. Available: https://factsheets.inl.gov/FactSheets/NationalNuclearLaboratory_Overview.pdf

[20] "Idaho National Lab Energy Systems Laboratory." [Online]. Available: https://factsheets.inl.gov/FactSheets/EnergySystemsLaboratory.pdf

[21] "Idaho National Lab Energy, Environment Science and Technology." [Online]. Available: https://factsheets.inl.gov/FactSheets/EESandT.pdf

[22] "National Renewable Energy Laboratory (NREL)." [Online]. Available: https://www.nrel.gov/index.html

[23] "National Renewable Energy Laboratory Flatirons Campus." [Online]. Available: https://www.nrel.gov/flatirons-campus/

[24] "Increasing Power Expands Research Capabilities at NREL's Flatirons Campus." [Online]. Available: https://www.nrel.gov/news/program/2020/increasing-power-at-flatirons-campus.html

[25] B. Chen, K. L. Butler-Purry, A. Goulart, and D. Kundur, "Implementing a real-time cyber-physical system test bed in rtds and opnet," in *2014 North American Power Symposium (NAPS)*, 2014, pp. 1–6.

[26] C. Queiroz, A. Mahmood, and Z. Tari, "Scadasim—a framework for building scada simulations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 589–597, 2011.

[27] N. Dorsch, F. Kurtz, H. Georg, C. Hägerling, and C. Wietfeld, "Software-defined networking for smart grid communications: Applications, challenges and advantages," in *2014 IEEE international conference on smart grid communications (SmartGridComm)*. IEEE, 2014, pp. 422–427.

[28] H. Georg, S. Müller, N. Dorsch, C. Rehtanz, and C. Wietfeld, "Inspire: Integrated co-simulation of power and ict systems for real-time evaluation," 10 2013, pp. 576–581.

[29] M. J. Stanovich, I. Leonard, K. Sanjeev, M. Steurer, T. P. Roth, S. Jackson, and M. Bruce, "Development of a smart-grid cyber-physical systems testbed," in *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, 2013, pp. 1–6.

[30] CAPS–Florida State University, "Center for Advanced Power Systems Infrastructure," [Online]: https://www.caps.fsu.edu/media/1256/caps-flyer.pdf, 2020.

[31] C. Ogilvie, J. Ospina, C. Konstantinou, T. Vu, M. Stanovich, K. Schoder, and M. Steurer, "Modeling communication networks in a real-time simulation environment for evaluating controls of shipboard power systems," in *2020 IEEE CyberPELS (CyberPELS)*, 2020, pp. 1–7.

[32] J. Ospina, I. Zografopoulos, X. Liu, and C. Konstantinou, "Demo: Trustworthy cyberphysical energy systems: Time-delay attacks in a real-time co-simulation environment," in *Proceedings of the 2020 Joint Workshop on CPS & IoT Security and Privacy*, ser. CPSIOTSEC'20. New York, NY, USA: Association for Computing Machinery, 2020, p. 69. [Online]. Available: https://doi.org/10.1145/3411498.3422926

[33] S. Sridhar, A. Ashok, M. E. Mylrea, S. Pal, M. J. Rice, and S. N. G. Gourisetti, "A testbed environment for buildings-to-grid cyber resilience research and development," 9 2017.

[34] J. N. Haack, B. A. Akyol, N. D. Tenney, B. J. Carpenter, R. M. Pratt, and T. E. Carroll, "Volttron™: An agent platform for integrating electric vehicles and smart grid," 12 2013.

[35] HELICS. (2020) Hierarchical engine for large-scale infrastructure co-simulation (HELICS). [Online]. Available: https://gmlc-tdc.github.io/helics.org/

[36] ——. (2020) Tools with HELICS Support. [Online]. Available: https://docs.helics.org/en/latest/Tools_using_HELICS.html

[37] T. D. Le, A. Anwar, R. Beuran, and S. W. Loke, "Smart grid co-simulation tools: Review and cybersecurity case study," in *2019 7th International Conference on Smart Grid (icSmartGrid)*. IEEE, 2019, pp. 39–45.

[38] N. Duan, N. Yee, B. Salazar, J. Joo, E. Stewart, and E. Cortez, "Cybersecurity analysis of distribution grid operation with distributed energy resources via co-simulation," Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), Tech. Rep., 2019.

[39] B. S. Palmintier, "Helics for integrated transmission, distribution, communication, & control (tdc+ c) modeling," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2019.

[40] C. Siaterlis, B. Genge, and M. Hohenadel, "Epic: A testbed for scientifically rigorous cyber-physical security experimentation," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 2, pp. 319–330, 2013.

[41] M. O. Faruque and V. Dinavahi, "Hardware-in-the-loop simulation of power electronic systems using adaptive discretization," *IEEE transactions on industrial electronics*, vol. 57, no. 4, pp. 1146–1158, 2009.

[42] S. Boschert and R. Rosen, *Digital Twin—The Simulation Aspect*. Cham: Springer International Publishing, 2016, pp. 59–74. [Online]. Available: https://doi.org/10.1007/978-3-319-32156-1_5

[43] M. Zhou, J. Yan, and D. Feng, "Digital twin framework and its application to power grid online analysis," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 3, pp. 391–398, 2019.

[44] T. Huang, F. R. Yu, C. Zhang, J. Liu, J. Zhang, and Y. Liu, "A survey on large-scale software defined networking (sdn) testbeds: Approaches and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 891–917, 2016.

[45] A. Keliris, C. Konstantinou, M. Sazos, and M. Maniatakos, "Low-budget energy sector cyberattacks via open source exploitation," in *2018 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)*. IEEE, 2018, pp. 101–106.

[46] Y. Xiang, L. Wang, and N. Liu, "Coordinated attacks on electric power systems in a cyber-physical environment," *Electric Power Systems Research*, vol. 149, pp. 156–168, 2017.

[47] J. Tian, B. Wang, T. Li, F. Shang, and K. Cao, "Coordinated cyber-physical attacks considering dos attacks in power systems," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 11, pp. 4345–4358, 2020.

[48] H. Tu, Y. Xia, K. T. Chi, and X. Chen, "A hybrid cyber attack model for cyber-physical power systems," *IEEE Access*, vol. 8, pp. 114 876–114 883, 2020.

[49] K. Pan, A. Teixeira, M. Cvetkovic, and P. Palensky, "Data attacks on power system state estimation: Limited adversarial knowledge vs. limited attack resources," in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, 2017, pp. 4313–4318.

[50] A. Barua and M. A. Al Faruque, "Hall spoofing: A non-invasive dos attack on grid-tied solar inverter," in *29th {USENIX} Security Symposium ({USENIX} Security 20)*, 2020, pp. 1273–1290.

[51] V. Venkataramanan, A. Hahn, and A. Srivastava, "Cp-sam: Cyber-physical security assessment metric for monitoring microgrid resiliency," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1055–1065, 2019.

[52] Y. Zhang, V. Krishnan, J. Pi, K. Kaur, A. Srivastava, A. Hahn, and S. Suresh, "Cyber physical security analytics for transactive energy systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 931–941, 2019.

[53] A. P. Kuruvila, I. Zografopoulos, K. Basu, and C. Konstantinou, "Hardware-assisted detection of firmware attacks in inverter-based cyberphysical microgrids," *arXiv preprint arXiv:2009.07691*, 2020.

[54] X. Liu, J. Ospina, and C. Konstantinou, "Deep reinforcement learning for cybersecurity assessment of wind integrated power systems," *IEEE Access*, vol. 8, pp. 208 378–208 394, 2020.

[55] H. Orojloo and M. A. Azgomi, "A stochastic game model for evaluating the impacts of security attacks against cyber-physical systems," *Journal of Network and Systems Management*, vol. 26, no. 4, pp. 929–965, 2018.

[56] Y. Fan, J. Li, D. Zhang, J. Pi, J. Song, and G. Zhao, "Supporting sustainable maintenance of substations under cyber-threats: an evaluation method of cybersecurity risk for power cps," *Sustainability*, vol. 11, no. 4, p. 982, 2019.

[57] Y. Yang, S. Wang, M. Wen, and W. Xu, "Reliability modeling and evaluation of cyber-physical system (cps) considering communication failures," *Journal of the Franklin Institute*, 2018.

[58] L. Wei, D. Gao, and C. Luo, "False data injection attacks detection with deep belief networks in smart grid," in *2018 Chinese Automation Congress (CAC)*. IEEE, 2018, pp. 2621–2625.

[59] J. James, Y. Hou, and V. O. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, 2018.

[60] D. Ye and T.-Y. Zhang, "Summation detector for false data-injection attack in cyber-physical systems," *IEEE transactions on cybernetics*, vol. 50, no. 6, pp. 2338–2345, 2019.

[61] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "Distributed host-based collaborative detection for false data injection attacks in smart grid cyber-physical system," *Journal of Parallel and Distributed Computing*, vol. 103, pp. 32–41, 2017.

[62] A. Sayghe, Y. Hu, I. Zografopoulos, X. Liu, R. G. Dutta, Y. Jin, and C. Konstantinou, "Survey of machine learning methods for detecting false data injection attacks in power systems," *IET Smart Grid*, vol. 3, pp. 581–595(14), October 2020. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-stg.2020.0015

[63] J. Goh, S. Adepu, M. Tan, and Z. S. Lee, "Anomaly detection in cyber physical systems using recurrent neural networks," in *2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE)*. IEEE, 2017, pp. 140–145.

[64] C. M. Ahmed, M. Ochoa, J. Zhou, A. P. Mathur, R. Qadeer, C. Murguia, and J. Ruths, "Noiseprint: Attack detection using sensor and process noise fingerprint in cyber physical systems," in *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, 2018, pp. 483–497.

[65] P. Schneider and K. Böttinger, "High-performance unsupervised anomaly detection for cyber-physical system networks," in *Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy*, 2018, pp. 1–12.

[66] S. Huda, S. Miah, M. M. Hassan, R. Islam, J. Yearwood, M. Alrubaian, and A. Almogren, "Defending unknown attacks on cyber-physical systems by semi-supervised approach and available unlabeled data," *Information Sciences*, vol. 379, pp. 211–228, 2017.

[67] O. M. Anubi and C. Konstantinou, "Enhanced resilient state estimation using data-driven auxiliary models," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 639–647, 2020.

[68] I. Zografopoulos and C. Konstantinou, "DERauth: A Battery-Based Authentication Scheme for Distributed Energy Resources," in *2020 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2020, pp. 560–567.

[69] T. R. B. Kushal, K. Lai, and M. S. Illindala, "Risk-based mitigation of load curtailment cyber attack using intelligent agents in a shipboard power system," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4741–4750, 2018.

[70] C. Kwon and I. Hwang, "Cyber attack mitigation for cyber–physical systems: hybrid system approach to controller design," *IET Control Theory & Applications*, vol. 10, no. 7, pp. 731–741, 2016.

[71] X. Hao, M. Lv, J. Zheng, Z. Zhang, and W. Yi, "Integrating cyber-attack defense techniques into real-time cyber-physical systems," in *2019 IEEE 37th International Conference on Computer Design (ICCD)*. IEEE, 2019, pp. 237–245.

[72] I. Zografopoulos, J. Ospina, and C. Konstantinou, "Special session: Harness the power of ders for secure communications in electric energy systems," in *2020 IEEE 38th International Conference on Computer Design (ICCD)*, 2020, pp. 49–52.

[73] J. Ospina, N. Gupta, A. Newaz, M. Harper, M. O. Faruque, E. G. Collins, R. Meeker, and G. Lofman, "Sampling-based model predictive control of PV-integrated energy storage system considering power generation forecast and real-time price," *IEEE Power and Energy Technology Systems Journal*, vol. 6, no. 4, pp. 195–207, 2019.

[74] M. Z. Gunduz and R. Das, "Cyber-security on smart grid: Threats and potential solutions," *Computer Networks*, vol. 169, p. 107094, 2020.

[75] K. Kimani, V. Oduol, and K. Langat, "Cyber security challenges for iot-based smart grid networks," *International Journal of Critical Infrastructure Protection*, vol. 25, pp. 36–49, 2019.

[76] B. Canaan, B. Colicchio, and D. Ould Abdeslam, "Microgrid cyber-security: Review and challenges toward resilience," *Applied Sciences*, vol. 10, no. 16, p. 5649, 2020.

[77] Z. El Mrabet, N. Kaabouch, H. El Ghazi, and H. El Ghazi, "Cyber-security in smart grid: Survey and challenges," *Computers & Electrical Engineering*, vol. 67, pp. 469–482, 2018.

[78] F. Nejabatkhah, Y. W. Li, H. Liang, and R. R. Ahrabi, "Cyber-security of smart microgrids: A survey," *Energies*, vol. 14, no. 1, p. 27, 2021.

[79] "Ieee standard for synchrophasor data transfer for power systems," *IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–53, 2011.

[80] C. Konstantinou, M. Sazos, A. S. Musleh, A. Keliris, A. Al-Durra, and M. Maniatakos, "Gps spoofing effect on phase angle monitoring and control in a real-time digital simulator-based hardware-in-the-loop environment," *IET Cyber-Physical Systems: Theory & Applications*, vol. 2, no. 4, pp. 180–187, 2017.

[81] A. Keliris, C. Konstantinou, and M. Maniatakos, "GE Multilin SR protective relays passcode vulnerability," *Black Hat USA*, 2017.

[82] Zografopoulos, I. and Konstantinou, C. and Tsoutsos, N. G. and Zhou, D. and Broadwater, R., "Security Assessment and Impact Analysis of Cyberattacks in Integrated T&D Power Systems," [Online]: https://dss-lab.github.io/pub/misc/PESGM2021.pdf, 2020.

[83] Department of Homeland Security, "Recommended Practice: Improving Industrial Control System Cybersecurity with Defense-in-Depth Strategies." [Online]. Available: https://www.us-cert.gov/sites/default/files/recommended_practices/NCCIC_ICS-CERT_Defense_in_Depth_2016_S508C.pdf

[84] T. J. Williams, "The purdue enterprise reference architecture," *Comput. Ind.*, vol. 24, no. 2–3, p. 141–158, Sep. 1994. [Online]. Available: https://doi.org/10.1016/0166-3615(94)90017-5

[85] Ackerman, Pascal, "Industrial Cybersecurity." [Online]. Available: https://subscription.packtpub.com/book/networking_and_servers/9781788395151

[86] MITRE Enterprise Engineering. Crown Jewels Analysis. [Online]. Available: https://www.mitre.org/publications/systems-engineering-guide/enterprise-engineering/systems-engineering-for-mission-assurance/crown-jewels-analysis

[87] Shostack, Adam, "Experiences Threat Modeling at Microsoft." [Online]. Available: https://adam.shostack.org/modsec08/Shostack-ModSec08-Experiences-Threat-Modeling-At-Microsoft.pdf

[88] J.D. Meier, Alex Mackman, Michael Dunner, Srinath Vasireddy, Ray Escamilla and Anandha Murukan, "Improving Web Application Security: Threats and Countermeasures." [Online]. Available: https://docs.microsoft.com/en-us/previous-versions/msp-n-p/ff648644(v=pandp.10)

[89] R. Khan, K. McLaughlin, D. Laverty, and S. Sezer, "Stride-based threat modeling for cyber-physical systems," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. IEEE, 2017, pp. 1–6.

[90] Mahmood, Haider. Application Threat Modeling using DREAD and STRIDE. [Online]. Available: https://haiderm.com/application-threat-modeling-using-dread-and-stride/

[91] Caralli, Richard A., Stevens, James F., Young, James F. and Wilson, William R., "Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process." [Online]. Available: https://resources.sei.cmu.edu/asset_files/TechnicalReport/2007_005_001_14885.pdf

[92] P. Radanliev, D. De Roure, J. R. Nurse, R. Nicolescu, M. Huth, S. Cannady, and R. M. Montalvo, "Integration of cyber security frameworks, models and approaches for building design principles for the internet-of-things in industry 4.0," 2018.

[93] K. Stouffer, J. Falco, and K. Scarfone, "Guide to industrial control systems (ics) security," *NIST special publication*, vol. 800, no. 82, pp. 16–16, 2011.

[94] S. D. Antón, D. Fraunholz, C. Lipps, F. Pohl, M. Zimmermann, and H. D. Schotten, "Two decades of scada exploitation: A brief history," in *2017 IEEE Conference on Application, Information and Network Security (AINS)*, 2017, pp. 98–104.

[95] "MITRE ATT&CK." [Online]. Available: https://attack.mitre.org/

[96] MITRE ATT&CK for Enterprise. MITRE ATT&CK Groups Overview. [Online]. Available: https://attack.mitre.org/groups/

[97] MITRE Cybersecurity. 7 Steps For an APT Detection Playbook using ATT&CK. [Online]. Available: https://www.mitre.org/capabilities/cybersecurity/overview/cybersecurity-blog/7-steps-for-an-apt-detection-playbook-using

[98] "MITRE ATT&CK for Industrial Control Systems." [Online]. Available: https://collaborate.mitre.org/attackics/index.php/Main_Page

[99] The MITRE Corporation. MITRE ATT&CK for ICS levels. [Online]. Available: https://collaborate.mitre.org/attackics/index.php/All_Levels

[100] L. Piètre-Cambacédès and M. Bouissou, "Attack and defense modeling with bdmp," in *International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security*. Springer, 2010, pp. 86–101.

[101] X. Li, C. Zhou, Y.-C. Tian, N. Xiong, and Y. Qin, "Asset-based dynamic impact assessment of cyberattacks for risk analysis in industrial control systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 608–618, 2017.

[102] K. Huang, C. Zhou, Y.-C. Tian, S. Yang, and Y. Qin, "Assessing the physical impact of cyberattacks on industrial cyber-physical systems," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 10, pp. 8153–8162, 2018.

[103] R. E. Bloomfield, P. Popov, K. Salako, V. Stankovic, and D. Wright, "Preliminary interdependency analysis: An approach to support critical-infrastructure risk-assessment," *Reliability Engineering & System Safety*, vol. 167, pp. 198 – 217, 2017, special Section: Applications of Probabilistic Graphical Models in Dependability, Diagnosis and Prognosis. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832017305963

[104] E. Bompard, C. Gao, R. Napoli, A. Russo, M. Masera, and A. Stefanini, "Risk assessment of malicious attacks against power systems," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 5, pp. 1074–1085, 2009.

[105] W. Li, *Risk assessment of power systems: models, methods, and applications*. John Wiley & Sons, 2014.

[106] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Li, "Power system risk assessment in cyber attacks considering the role of protection systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 572–580, 2017.

[107] W. Wang, A. Cammi, F. Di Maio, S. Lorenzi, and E. Zio, "A monte carlo-based exploration framework for identifying components vulnerable to cyber threats in nuclear power plants," *Reliability Engineering & System Safety*, vol. 175, pp. 24 – 37, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832017308621

[108] G. Giannopoulos, R. Filippini, and M. Schimmer, "Risk assessment methodologies for critical infrastructure protection. part i: A state of the art," *JRC Technical Notes*, 2012.

[109] M. Theocharidou and G. Giannopoulos, "Risk assessment methodologies for critical infrastructure protection. part ii: A new approach," *Scientific and Technical Research Reports*, 2015.

[110] C. Vellaithurai, A. Srivastava, S. Zonouz, and R. Berthier, "Cpindex: Cyber-physical vulnerability assessment for power-grid infrastructures," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 566–575, 2015.

[111] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, "Adversarial attacks and defences: A survey," *arXiv preprint arXiv:1810.00069*, 2018.

[112] K. Ren, T. Zheng, Z. Qin, and X. Liu, "Adversarial attacks and defenses in deep learning," *Engineering*, vol. 6, no. 3, pp. 346 – 360, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S209580991930503X

[113] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2016.

[114] J. Zhang, Zhigang Chu, L. Sankar, and O. Kosut, "False data injection attacks on power system state estimation with limited information," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5.

[115] 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)*, 2012, pp. 3153–3158.

[116] M. Mohammadpourfard, A. Sami, and Y. Weng, "Identification of false data injection attacks with considering the impact of wind generation and topology reconfigurations," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1349–1364, 2018.

[117] J. Zhao, G. Zhang, M. La Scala, Z. Y. Dong, C. Chen, and J. Wang, "Short-term state forecasting-aided method for detection of smart grid general false data injection attacks," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1580–1590, 2017.

[118] C. Konstantinou and M. Maniatakos, "A case study on implementing false data injection attacks against nonlinear state estimation," in *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*, ser. CPS-SPC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 81–92. [Online]. Available: https://doi.org/10.1145/2994487.2994491

[119] S. P. Skorobogatov and R. J. Anderson, "Optical fault induction attacks," in *Cryptographic Hardware and Embedded Systems - CHES 2002*, B. S. Kaliski, ç. K. Koç, and C. Paar, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 2–12.

[120] J. Meserve, "Staged cyber attack reveals vulnerability in power grid," [Online]: http://www.cnn.com/2007/US/09/26/power.at.risk/index.html, 2007.

[121] X. Lyu, Y. Ding, and S. Yang, "Safety and security risk assessment in cyber-physical systems," *IET Cyber-Physical Systems: Theory Applications*, vol. 4, no. 3, pp. 221–232, 2019.

[122] W. A. Conklin, "IT vs. OT Security: A Time to Consider a Change in CIA to Include Resilience," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 2016, pp. 2642–2647.

[123] R. Paes, D. C. Mazur, B. K. Venne, and J. Ostrzenski, "A guide to securing industrial control networks: Integrating it and ot systems," *IEEE Industry Applications Magazine*, vol. 26, no. 2, pp. 47–53, 2020.

[124] P. K. Garimella, "It-ot integration challenges in utilities," in *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, 2018, pp. 199–204.

[125] R. J. Rodríguez, "On qualitative analysis of fault trees using structurally persistent nets," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 2, pp. 282–293, 2015.

[126] A. A. Cárdenas, S. Amin, Z.-S. Lin, Y.-L. Huang, C.-Y. Huang, and S. Sastry, "Attacks against process control systems: risk assessment, detection, and response," in *Proceedings of the 6th ACM symposium on information, computer and communications security*, 2011, pp. 355–366.

[127] J. Shi, "Security risk assessment about enterprise networks on the base of simulated attacks," *Procedia Engineering*, vol. 24, pp. 272–277, 2011.

[128] W. Wu, R. Kang, and Z. Li, "Risk assessment method for cyber security of cyber physical systems," in *2015 First International Conference on Reliability Systems Engineering (ICRSE)*, 2015, pp. 1–5.

[129] Q. Zhang, C. Zhou, N. Xiong, Y. Qin, X. Li, and S. Huang, "Multimodel-based incident prediction and risk assessment in dynamic cybersecurity protection for industrial control systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 10, pp. 1429–1444, 2016.

[130] Y. Peng, T. Lu, J. Liu, Y. Gao, X. Guo, and F. Xie, "Cyber-physical system risk assessment," in *2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2013, pp. 442–447.

[131] E. Zio, "The future of risk assessment," *Reliability Engineering & System Safety*, vol. 177, pp. 176 – 190, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832017306543

[132] J. Wan, S. Tang, D. Li, M. Imran, C. Zhang, C. Liu, and Z. Pang, "Reconfigurable smart factory for drug packing in healthcare industry 4.0," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 507–516, 2019.

[133] Y. Zhang, M. Qiu, C. Tsai, M. M. Hassan, and A. Alamri, "Healthcps: Healthcare cyber-physical system assisted by cloud and big data," *IEEE Systems Journal*, vol. 11, no. 1, pp. 88–95, 2017.

[134] Y. Chang, "Architecture design for performing grasp-and-lift tasks in brain–machine-interface-based human-in-the-loop robotic system," *IET Cyber-Physical Systems: Theory Applications*, vol. 4, no. 3, pp. 198–203, 2019.

[135] J. K. Naufal, J. B. Camargo, L. F. Vismari, J. R. de Almeida, C. Molina, R. I. R. González, R. Inam, and E. Fersman, "A2cps: A vehicle-centric safety conceptual framework for autonomous transport systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 6, pp. 1925–1939, 2018.

[136] J. Johnson, J. Quiroz, R. Concepcion, F. Wilches-Bernal, and M. J. Reno, "Power system effects and mitigation recommendations for der cyberattacks," *IET Cyber-Physical Systems: Theory Applications*, vol. 4, no. 3, pp. 240–249, 2019.

[137] C. B. Vellaithurai, S. S. Biswas, R. Liu, and A. Srivastava, *Real Time Modeling and Simulation of Cyber-Power System*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 43–74. [Online]. Available: https://doi.org/10.1007/978-3-662-45928-7_3

[138] X. Fan, S. G. Aksoy, Q. Huang, J. P. Ogle, D. Wang, A. Tbaileh, and T. Fu, "Coordination of transmission, distribution and communication systems for prompt power system recovery after disasters:report – grid and communication interdependency review and characterization of typical communication systems," 3 2019.

[139] Science Learning Hub, "Scientific modeling," [Online]: https://www.sciencelearn.org.nz/resources/575-scientific-modelling, 2019.

[140] V. Jalili-Marandi, V. Dinavahi, K. Strunz, J. A. Martinez, and A. Ramirez, "Interfacing techniques for transient stability and electromagnetic transient programs IEEE task force on interfacing techniques for simulation tools," *IEEE Transactions on Power Delivery*, vol. 24, no. 4, pp. 2385–2395, 2009.

[141] P. Le-Huy, G. Sybille, P. Giroux, L. Loud, J. Huang, and I. Kamwa, "Real-time electromagnetic transient and transient stability co-simulation based on hybrid line modelling," *IET Generation, Transmission Distribution*, vol. 11, no. 12, pp. 2983–2990, 2017.

[142] D. Shu, X. Xie, V. Dinavahi, C. Zhang, X. Ye, and Q. Jiang, "Dynamic phasor based interface model for EMT and transient stability hybrid simulations," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3930–3939, 2018.

[143] R. Huang, R. Fan, J. Daily, A. Fisher, and J. Fuller, "Open-source framework for power system transmission and distribution dynamics co-simulation," *IET Generation, Transmission & Distribution*, vol. 11, no. 12, pp. 3152–3162, 2017.

[144] A. Hariri and M. O. Faruque, "A hybrid simulation tool for the study of pv integration impacts on distribution networks," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 648–657, 2017.

[145] Y. N. Velaga, G. Krishnamoorthy, A. Dubey, A. Chen, and P. K. Sen, "Advancements in co-simulation techniques in combined transmission and distribution systems analysis," *The Journal of Engineering*, vol. 2019, no. 12, pp. 8432–8438, 2019.

[146] M. S. Obaidat, F. Zarai, and P. Nicopolitidis, *Modeling and simulation of computer networks and systems: Methodologies and applications*. Morgan Kaufmann, 2015.

[147] J. Mahseredjian, V. Dinavahi, and J. A. Martinez, "Simulation tools for electromagnetic transients in power systems: Overview and challenges," *IEEE Transactions on Power Delivery*, vol. 24, no. 3, pp. 1657–1669, 2009.

[148] M. D. Omar Faruque, T. Strasser, G. Lauss, V. Jalili-Marandi, P. Forsyth, C. Dufour, V. Dinavahi, A. Monti, P. Kotsampopoulos, J. A. Martinez, K. Strunz, M. Saeedifard, Xiaoyu Wang, D. Shearer, and M. Paolone, "Real-time simulation technologies for power systems design, testing, and analysis," *IEEE Power and Energy Technology Systems Journal*, vol. 2, no. 2, pp. 63–73, 2015.

[149] A. Newaz, J. Ospina, and M. O. Faruque, "Controller hardware-in-the-loop validation of a graph search based energy management strategy for grid-connected distributed energy resources," *IEEE Transactions on Energy Conversion*, vol. 35, no. 1, pp. 520–528, 2019.

[150] G. F. Lauss, M. O. Faruque, K. Schoder, C. Dufour, A. Viehweider, and J. Langston, "Characteristics and design of power hardware-in-the-loop simulations for electrical power systems," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 1, pp. 406–417, 2016.

[151] T. Key and K. Forsten, "Security, quality, reliability and availability: Metrics definition: Progress report," *EPRI*, 2005.

[152] W. Ren, "Accuracy evaluation of power hardware-in-the-loop (PHIL) simulation," *Florida State University*, 2007.

[153] USC ISI. The network simulator - ns-2. [Online]. Available: https://www.isi.edu/nsnam/ns/

[154] ns-3 project. ns-3 network simulator. [Online]. Available: https://www.nsnam.org/

[155] MIT. Simpy - discrete event simulation for python. [Online]. Available: https://simpy.readthedocs.io/en/latest/

[156] SCALABLE Network Technologies. Scalable software solutions. [Online]. Available: https://www.scalable-networks.com/

[157] The CORE project. CORE: Common Open Research Emulator. [Online]. Available: https://github.com/coreemu/core

[158] Linux Man Pages. NetEm - Network Emulator. [Online]. Available: https://www.linux.org/docs/man8/tc-netem.html

[159] E. Weingärtner, F. Schmidt, H. Vom Lehn, T. Heer, and K. Wehrle, "Slicetime: A platform for scalable and accurate network emulation," in *Proceedings of the 8th USENIX conference on Networked systems design and implementation*, 2011, pp. 253–266.

[160] E. Weingärtner, F. Schmidt, T. Heer, and K. Wehrle, "Synchronized network emulation: matching prototypes with complex simulations," *ACM SIGMETRICS Performance Evaluation Review*, vol. 36, no. 2, pp. 58–63, 2008.

[161] Z. Zhang, Y. Wang, and L. Xie, "A novel data integrity attack detection algorithm based on improved grey relational analysis," *IEEE Access*, vol. 6, pp. 73 423–73 433, 2018.

[162] L. Ma, Z. Wang, Q. Han, and H. Lam, "Variance-constrained distributed filtering for time-varying systems with multiplicative noises and deception attacks over sensor networks," *IEEE Sensors Journal*, vol. 17, no. 7, pp. 2279–2288, 2017.

[163] M. S. Chong, M. Wakaiki, and J. P. Hespanha, "Observability of linear systems under adversarial attacks," in *2015 American Control Conference (ACC)*, 2015, pp. 2439–2444.

[164] G. Na and Y. Eun, "A multiplicative coordinated stealthy attack and its detection for cyber physical systems," in *2018 IEEE Conference on Control Technology and Applications (CCTA)*, 2018, pp. 1698–1703.

[165] Texas Instruments, "Grid-tied Solar Micro Inverter with MPPT," [Online]: https://www.ti.com/tool/TIDM-SOLARUINV, 2019.

[166] U.S. DOE, "Annual energy outlook 2019 with projections to 2050," [Online]: https://www.eia.gov/outlooks/aeo/pdf/aeo2019.pdf, 2019.

[167] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, D. Kallitsis, Michalis Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the mirai botnet," in *26th {USENIX} security symposium ({USENIX} Security 17)*, 2017, pp. 1093–1110.

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

[169] C. Konstantinou and M. Maniatakos, "Impact of firmware modification attacks on power systems field devices," in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2015, pp. 283–288.

[170] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.

[171] J. D. Glover, M. S. Sarma, and T. Overbye, *Power system analysis & design, SI version*. Cengage Learning, 2012.

[172] I. Boldea, *Synchronous generators*. CRC press, 2015.

[173] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic load altering attacks against power system stability: Attack models and protection schemes," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2862–2872, 2016.

[174] S. K. Singh, B. P. Padhy, S. Chakrabarti, S. Singh, A. Kolwalkar, and S. Kelapure, "Development of dynamic test cases in opal-rt real-time power system simulator," in *2014 Eighteenth National Power Systems Conference (NPSC)*. IEEE, 2014, pp. 1–6.

[175] A. Bokhari, A. Alkan, R. Dogan, M. Diaz-Aguiló, F. De Leon, D. Czarkowski, Z. Zabar, L. Birenbaum, A. Noel, and R. E. Uosef, "Experimental determination of the zip coefficients for modern residential, commercial, and industrial loads," *IEEE Transactions on Power Delivery*, vol. 29, no. 3, pp. 1372–1381, 2013.

[176] M. Pai, *Energy function analysis for power system stability*. Springer Science & Business Media, 2012.

[177] NERC, "2019 Frequency Response Annual Analysis," https://www.nerc.com/, 2019.

[178] NYISO, "Transmission and Dispatch Operations Manual," https://www.nyiso.com/documents/20142/2923301/trans_disp.pdf/9d91ad95-0281-2b17-5573-f054f7169551, 2020.

[179] T. Shekari, A. Gholami, F. Aminifar, and M. Sanaye-Pasand, "An adaptive wide-area load shedding scheme incorporating power system real-time limitations," *IEEE Systems Journal*, vol. 12, no. 1, pp. 759–767, 2016.

[180] NERC, "NERC Reliability Standard PRC NERC Reliability Standard PRC-024-1," https://www.nerc.com/, 2020.

[181] H. Seyedi and M. Sanaye-Pasand, "New centralised adaptive load-shedding algorithms to mitigate power system blackouts," *IET generation, transmission & distribution*, vol. 3, no. 1, pp. 99–114, 2009.

[182] A. Sargolzaei, K. K. Yen, and M. Abdelghani, "Time-delay switch attack on load frequency control in smart grid," *Advances in Communication Technology*, vol. 5, pp. 55–64, 2013.

[183] J. Wang and C. Peng, "Analysis of time delay attacks against power grid stability," in *Proceedings of the 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids*, 2017, pp. 67–72.

[184] A. Sargolzaei, K. K. Yen, and M. N. Abdelghani, "Preventing time-delay switch attack on load frequency control in distributed power systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1176–1185, 2016.

[185] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135 – 148, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109814004488

[186] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2012.

[187] M. Emmanuel and R. Rayudu, "The impact of single-phase grid-connected distributed photovoltaic systems on the distribution network using pq and pv models," *International Journal of Electrical Power & Energy Systems*, vol. 91, pp. 20–33, 2017.

[188] A. Pradhan, A. Routray, and S. M. Gudipalli, "Fault direction estimation in radial distribution system using phase change in sequence current," *IEEE Transactions on Power Delivery*, vol. 22, no. 4, pp. 2065–2071, 2007.

[189] A. Singhal and V. Ajjarapu, "Long-term voltage stability assessment of an integrated transmission distribution system," in *2017 North American Power Symposium (NAPS)*. IEEE, 2017, pp. 1–6.

[190] C. Dufour, H. Saad, J. Mahseredjian, and J. Bélanger, "Custom-coded models in the state space nodal solver of ARTEMiS," in *Proceeding of the 2013 International Conference on Power System Transients (IPST-2013)*, 2013.

[191] Opal-RT. ARTEMiS User Guide, v 6.1. [Online]. Available: https://www.opal-rt.com/

[192] S. Adepu, N. K. Kandasamy, J. Zhou, and A. Mathur, "Attacks on smart grid: Power supply interruption and malicious power generation," *International Journal of Information Security*, pp. 1–23, 2019.

[193] NERC, "Reliability Considerations from the Integration of Smart Grid," [Online]: https://www.nerc.com/files/SGTF_Report_Final_posted_v1.1.pdf, 2010.

[194] A. Nicholson, S. Webber, S. Dyer, T. Patel, and H. Janicke, "Scada security in the light of cyber-warfare," *Computers & Security*, vol. 31, no. 4, pp. 418 – 436, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167404812000429

[195] L. Maglaras, M. A. Ferrag, A. Derhab, M. Mukherjee, H. Janicke, and S. Rallis, "Threats, protection and attribution of cyber attacks on critical infrastructures," 2019.

[196] Y. Mo, T. H. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber–physical security of a smart grid infrastructure," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 195–209, 2012.

[197] C. Wilson, *Cyber Threats to Critical Information Infrastructure*. New York, NY: Springer New York, 2014, pp. 123–136. [Online]. Available: https://doi.org/10.1007/978-1-4939-0962-9_7