2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P) | 978-1-6654-6512-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/EuroSP57164.2023.00066

Comprehensively Analyzing the Impact of Cyberattacks on Power Grids

Lennart Bader
Fraunhofer FKIE &
RWTH Aachen University
lennart.bader@fkie.fraunhofer.de

Martin Serror
Fraunhofer FKIE
martin.serror@fkie.fraunhofer.de

Olav Lamberts
Fraunhofer FKIE &
RWTH Aachen University
olav.lamberts@rwth-aachen.de

Ömer Sen
Fraunhofer FIT &
RWTH Aachen University
o.sen@iaew.rwth-aachen.de

Dennis van der Velde
Fraunhofer FIT
dennis.van.der.velde@fit.fraunhofer.de

Immanuel Hacker
Fraunhofer FIT &
RWTH Aachen University
i.hacker@iaew.rwth-aachen.de

Julian Filter
RWTH Aachen University
hallo@julianfilter.de

Elmar Padilla
Fraunhofer FKIE
elmar.padilla@fkie.fraunhofer.de

Martin Henze
RWTH Aachen University
& Fraunhofer FKIE
henze@cs.rwth-aachen.de

***Abstract*—The increasing digitalization of power grids and especially the shift towards IP-based communication drastically increase the susceptibility to cyberattacks, potentially leading to blackouts and physical damage. Understanding the involved risks, the interplay of communication and physical assets, and the effects of cyberattacks are paramount for the uninterrupted operation of this critical infrastructure. However, as the impact of cyberattacks cannot be researched in real-world power grids, current efforts tend to focus on analyzing isolated aspects at small scales, often covering only either physical or communication assets. To fill this gap, we present WATTSON, a comprehensive research environment that facilitates reproducing, implementing, and analyzing cyberattacks against power grids and, in particular, their impact on both communication and physical processes. We validate WATTSON's accuracy against a physical testbed and show its scalability to realistic power grid sizes. We then perform authentic cyberattacks, such as Industroyer, within the environment and study their impact on the power grid's energy and communication side. Besides known vulnerabilities, our results reveal the ripple effects of susceptible communication on complex cyber-physical processes and thus lay the foundation for effective countermeasures.**

1. Introduction

Over the past decades, the communication infrastructure of power grids has shifted from serial communication to Internet-compatible IP-based communication [66]. While this shift provides many benefits, such as enhanced flexibility, adaptability, and scalability [67], it also drastically increases the susceptibility to cyberattacks [50]. Past incidents, such as the cyberattacks on the Ukrainian power grid [112], highlight the severe real-world implications. However, besides rather abstract information on these incidents [112], little is known about the precise technical susceptibility of power grids and their communication infrastructure w.r.t. cyberattacks. For example, it is unknown which technically feasible attack on the communication infrastructure causes the most drastic impact on the power side and how severe this impact is. Consequently, academia and industry struggle to identify and prioritize the parts of the power grid communication infrastructure needing countermeasures to thwart attacks [40].

Thus, it is essential to evaluate and understand the impact of cyberattacks on power grids. Unlike traditional (office or server) networks, the power grid as a cyber-physical system also demands to include the physical, i.e., energy, side in this analysis [4]. However, evaluating cybersecurity in *existing* power grids is typically not feasible due to high availability requirements and the risk of permanently damaging critical physical assets. Thus, various streams of research study the impact of cyberattacks on *physical* grid operation using testbeds and simulations, e.g., [28], [60], [110]. From a different angle, related work concerning the security of the *communication* infrastructure of power grids ranges from game-theoretic analyses [11] over discrete-event simulations (DESs) [2] to network emulations [26]. Still, related work typically does not capture both the communication *and* the energy side in sufficient detail and scale required to thoroughly assess the impact of real cyberattacks on power grids.

To alleviate this situation, other domains report huge success with sophisticated simulation environments, accurately reflecting both the communication and physical side, e.g., for water distribution [74] and treatment [4], or the Internet of Things [92]. While these approaches show the feasibility of capturing the combination of communication and the corresponding physical side in sufficient detail, they work in significantly smaller scenarios than typically found in power grids. Respective power grid approaches often focus on aspects other than cybersecurity, most prominently grid operation (e.g., [27]). In turn, when considering cyberattacks, they abstract from actually used communication protocols (e.g., [79]) or study only small-scale scenarios, such as individual substations (e.g., [99]). Thus, an environment for the large-scale evaluation of cyberattacks against power grids covering both the com-

DOI 10.1109/EuroSP57164.2023.00066

munication and energy side has been identified as an important open research field [33], [72].

This paper addresses the challenge of accurately studying the impact of cyberattacks on power distribution grids of realistic size (e.g., 500 electrical assets and 150 communicating parties [70]). Therefore, we propose WATTSON, an open-source[1] research environment that combines a state-of-the-art power flow solver with a sophisticated network emulator to flexibly and accurately model the energy and communication side of real-world power grids.

We use WATTSON to deploy real network applications (e.g., attack tools) within a simulated power grid to comprehensively study the cyber *and* physical impact of attacks. Our study shows that such cyberattacks can have devastating impacts, particularly when attackers obfuscate immediate effects by manipulating measurements. Above all, they highlight that complex cyber-physical systems inseparably unite communication and physical processes and thus require a thorough view to analyze attacks.

Contributions. In summary, our contributions are:

- We classify research on cyberattacks against power grids, finding that current work tends to focus on the energy *or* communication side, leaving the actual susceptibility of power grids open. To understand the root cause of this issue, we analyze the state of research environments for power grids and conclude that a comprehensive environment for studying the combined impact of cyberattacks on the energy *and* communication side of power grids at realistic scales is missing (§2).
- To comprehensively analyze the cybersecurity of power grids, we present WATTSON[1], an open-source research environment for studying cyberattacks on power grids at realistic scales, allowing us to analyze their impacts on the power *and* networking side through co-simulation. We validate its accuracy by replicating an attack on a physical testbed and show its scalability to real-world grid sizes using benchmarking and reference grids (§3).
- We study the impact of authentic, real-world cyberattacks within a reference grid [70] (475 energy and 239 communication assets) within WATTSON. Among others, we cover the infamous Industroyer [21] as an actually conducted attack and sophisticated false data injection. We discussed and validated the practical relevance of these attacks with domain experts from a national cybersecurity agency and multiple grid operators. Our results show the susceptibility of power grids against sophisticated cyberattacks and that only the combined consideration of energy *and* communication allows extensive analyses of their impact, thus laying the foundation to secure them properly (§4).

Availability Statement. To spur further research on evaluating the impact of cyberattacks against power grids at realistic scales, WATTSON is available under an open-source license[1]. Furthermore, we provide the network traces and the physical grid states for our conducted attacks[1].

2. Cybersecurity in Power Grids

To lay the foundation for our work, we provide a brief background on power grids introducing their structure and operation (§2.1). We then classify possible cyberattacks against power grids (§2.2) based on a literature survey to identify current shortcomings and derive requirements for comprehensive cybersecurity research in power grids (§2.3). Finally, we analyze how related work covers this cross-domain research area and fulfills the derived requirements for evaluating cybersecurity in power grids (§2.4).

2.1. Background on Power Grids

The main task of a power grid is to transmit electric power from generators to consumers [75]. Therefore, it uses alternating currents with different voltage levels, ranging from low to extra-high voltage, following a hierarchic structure for efficiency and minimizing losses. While *transmission networks* supply larger industrial consumers and typically operate at higher voltage levels, the subordinated *distribution networks* supply smaller industrial and residential consumers and usually operate on lower voltage levels. Electrical *substations* connect these different networks and *transform* between the voltages as needed.

A major challenge in power grid operation is maintaining the balance between power demand and supply, as large quantities of electrical power cannot be stored efficiently. Deviations from the nominal frequency (e.g., 50 Hz in Europe; 60 Hz in the US) indicate an imbalance of demand and supply. Furthermore, the capacities of *transmission lines* and other physical assets limit the power flow, where overloading induces physical damage. *Protective relays* typically prevent such damage, whereas falsely tripping such relays may lead to (partial) blackouts.

Power grids usually rely on SCADA systems to monitor and control the correct operation and adapt to the growing number of digital assets, e.g., enabling *state estimations (SEs)* and remote tripping of relays [100]. Therefore, distributed Remote Terminal Units (RTUs) logically connect to a centralized Master Terminal Unit (MTU) for reporting measurements and receiving commands. Such systems thus depend on a communication infrastructure and suited protocols. Due to the increasing number of renewable power sources, they exhibit growing importance.

However, most protocols used to manage power grids, such as IEC 60870-5-104 (in Europe) or DNP3 (in the US), were not designed with security in mind [50]. Ongoing digitalization and insecure, hard-to-replace protocols make power grids vulnerable to various cyberattacks [42], [52]. While Information and Communications Technology (ICT) security mainly relies on dedicated networks with access restrictions, these networks provide insufficient internal security. Thus, they are especially vulnerable once attackers gain initial access, e.g., using credential theft, spear phishing, or physical intrusion into substations [36].

The cyberattacks on the Ukrainian power grid in 2015, 2016, and 2022 [21], [52], [69] demonstrated the vulnerabilities of power grids. In particular, they spur intensive research on understanding individual incidents with devastating consequences for the power grid and ultimately improving countermeasures. While this paper mainly focuses on the structure and operation of current and future European (alternating current (AC)) power grids, its contributions generally also apply to other power grids, such as North American ones.

---

[1] https://github.com/fkie-cad/wattson and https://wattson.it

TABLE 1. CLASSIFICATION OF POWER GRID ATTACKS INTO PHYSICAL, SYNTACTIC, AND SEMANTIC AND HOW CURRENT RESEARCH EVALUATES ATTACK IMPACT W.R.T. ICT AND POWER GRID.

	Attack Type	ICT	Power Grid
Phys.	Device Disconnect		[38], [95]
	Demand Manipulation		[39], [94] [93], [108]
Syn.	Denial-of-Service	[2], [11], [68] [115], **[96]**	**[96]**, [1], [29] [32], [58], [116]
	Replay	[53], [64], [114]	[41], [116], [119]
Sem.		[81]	[1], [41], [103]
	False Data Injection	[11], [45], [47] [46], [53], [107]	[1], [18], [58], [79] [15], [49], [63], [118] [23], [43], [56], [87]

2.2. Classification and Analysis of Attacks

For structuring and understanding the vast range of cyber threats against power grids, we propose a classification according to the required attackers' knowledge for successfully performing the attack. This approach classifies attacks based on their complexity and already supports the consideration of adequate countermeasures. As a result, we distinguish between *physical*, *syntactic*, and *semantic* attacks, which we introduce in the following while also discussing their assumptions and requirements.

Physical Attacks. Physical attacks comprise all incidents where attackers disturb the power grid by physically interrupting, tampering with, or destroying grid equipment, including any ICT components. Such attacks range from simple *device disconnects* (e.g., [95]) to more complex *demand manipulation*, where attackers simultaneously start up many high-wattage devices (e.g., [39]). As a prerequisite, attackers need (physical) access to the targeted devices, while limited knowledge about the power grid's ICT and processes typically suffices to execute such attacks.

Syntactic Attacks. In syntactic attacks, attackers use crafted, intercepted, or duplicated messages to interfere with the power grid's operation *through* the ICT. Although these messages are syntactically correct, attackers typically do not need profound process knowledge since they do not modify the messages' payload. Such attacks often manifest as *Denial-of-Service (DoS)* attacks, targeting the availability of (parts of) a power grid (e.g., [2]). Alternatively, attackers may *replay* messages to provoke unwanted or harmful behavior (e.g., [119]). Typically, attackers require (remote) access to the ICT to perform such attacks successfully.

Semantic Attacks. Semantic attacks have the same requirements as syntactic attacks, with the addition that attackers require some knowledge about the process, configuration, and topology to target power grid operations specifically. Then, attackers can unsuspiciously counterfeit measurements or commands within ICT messages, which may have devastating effects. Depending on the targeted process, it may even suffice to *replay* specific messages at the right time (e.g., [81]). However, the most prominent attack is *false data injection*, where attackers tamper with measurements to mislead grid operation (e.g., [87]).

To understand to which extent current research covers these attack types, we performed a survey of 36 scientific publications which analyze the impact of attacks against power grids. We follow the discussed attack classification, i.e., *physical*, *syntactic*, or *semantic* attacks, for comparing and structuring related work in Table 1. For each publication, we identify the considered impact of the investigated attacks, i.e., *ICT*, *power grid*, or even both.

Our survey reveals that current research primarily focuses on single attack types, with a few exceptions considering two or three different attacks [1], [53], [58]. However, none covers *all* proposed attack categories. Importantly, they exclusively focus on the ICT *or* the energy side, abstracting from the respective other domain, although successful attacks inevitably concern both domains. As an exception, [96] considers both the ICT and the energy side. However, with its focus on a DoS attack, it merely covers semantic attacks. As its mathematical modeling restricts the approach's potential for analyzing different attack types and relies on assumption-based attack effects, it is inappropriate for studying, e.g., semantic attacks and unforeseeable inter-domain side effects. Consequently, we argue that a meaningful analysis of cyberattacks on power grids requires a more comprehensive evaluation methodology, adequately reflecting the complex cyber-physical dependencies of the ICT and energy side, while also covering the entire range of possible attacks. In the following, we thus derive the requirements for such a methodology before continuing our related work analysis.

2.3. Requirements for Research Environments

For adequately studying sophisticated cyberattacks on power grids, an appropriate research environment needs to fulfill specific requirements. We derive a total of four requirements based on the results of our literature survey in §2.2 and explain them in the following:

Accuracy. Both the ICT and power grid components behave and interact according to their respective real components. This includes *real* communication behavior and the cross-domain interplay between ICT and power grid components during normal operations and cyberattacks. As some types of communication in real power grid networks affect the state and behavior of the grid component and, e.g., physical defects of these components can impact the communication behavior, this *cross-domain accuracy* is critical to assess the impact of cyberattacks.

Scalability. Research environments provide scalability w.r.t. the realizable scenarios. This ranges from small laboratory settings to realistic power grids with high complexity, i.e., with significantly more components, for comprehensively evaluating the impact of cyberattacks.

Flexibility. ICT networks, power grid topologies, and cyberattacks are flexible in evaluating various configurations and scenarios. Further, communication protocols, attack tools, defensive measures, and power grid processes are exchangeable, enabling extensive research opportunities.

Cybersecurity. Finally, a suitable research environment explicitly considers cybersecurity analyses in its design. Through the inclusion of real-world attack tools along with respective evaluation capabilities, studying realistic cyber threats becomes feasible.

Although individually achieving these design properties is already challenging, a particular challenge arises when fulfilling all of them as required. We thus continue

TABLE 2. CLASSIFICATION OF CO-SIMULATION APPROACHES REGARDING THEIR COMMUNICATION MODEL (CONTINUOUS OR DISCRETE) AND THEIR POWER MODEL (STEADY OR TRANSIENT STATE) ALONG WITH THEIR FULFILLMENT OF REQUIRED PROPERTIES.

Com. Model	**Power Model**	**Approaches**	**Accuracy**		**Scalability**		**Flexibility**		**Cybersecurity**		**Open Source**
			Com.	**Power**	**Com.**	**Power**	**Com.**	**Power**	**Com.**	**Power**	
Discrete	**Steady**	[7], [78]	□	▣	■	■	⬓	▣	□	□	✓
		[18], [16]	□	⬓*	■	■*	⬓	▣	□	▣	
		[68]	□	⬓*	■	■*	⬓	▣	▣	□	
		[9], [24], [27], [65], [71]	□	⬓*	■	■*	⬓	▣	□	□	✓
	Transient	[3], [8], [25], [37], [54], [57], [90], [105]	□	?	■	▣*	⬓	▣	□	□	
		[14], [31], [44], [76], [77]	□	?	■	■*	⬓	▣	□	□	✓
		[58], [59], [79]	□	⬓*	■	■	⬓	▣	□	■	
Continuous	**Steady**	[29], [30]	■	▣	⬓	▣*	▣	▣*	□*	■	✓
		[55]	■	▣	▣	■	▣*	▣*	□	□	✓
	Transient	[1]	▣*	■	⬓*	▣*	▣*	▣*	▣	■	
Continuous	**Steady**	**WATTSON**	■	▣	▣	■	■	■	■	■	✓

Requirement not □, marginally ⬓, mostly ▣, or thoroughly ■ fulfilled ∗ – Not evaluated by authors / uncertain ? – Unknown

with our related work discussion, analyzing to which extent current research environments for power grids cover the requirements with the intended depth.

2.4. Related Work on Research Environments

Evaluating cyberattacks and countermeasures is typically not possible in actually deployed power grids due to the high availability requirements of such systems and the risk of causing physical defects [98]. Although physical testbeds, i.e., deployments of real power grid components for test and development purposes, are optimal regarding the provided accuracy, they are limited in scalability and flexibility [13]. Consequently, research largely depends on modeling power grid components to analyze and evaluate cybersecurity. In turn, analytical models of power grids, such as [62], [91], offer a scalable and flexible way to analyze the impact of cyberattacks while lacking the required accuracy due to their abstract models. Hence, a better-suited approach is to simulate or emulate power grid components, enabling a comprehensive representation of the power grid. Here, a class of approaches exclusively focuses on simulating either the ICT, such as [61], [85], [99], or the power grid, such as [84], [102]. However, such approaches are extremely limited for cybersecurity research, as discussed in §2.2.

Thus, numerous approaches aim to *co-simulate* the ICT and the power grid. Such approaches follow general ICT and power simulation paradigms: A communication network simulation either uses a *discrete event* or a *continuous* approach. While discrete-event simulations (DESs) allow non-real-time simulations and offer high scalability, they cannot precisely represent actual network traffic, involved protocol stacks, and timing-related side effects due to the chosen level of abstraction [117]. As a result, they have limited accuracy and flexibility w.r.t. the evaluation of cyberattacks. Here, continuous approaches are favorable while having potential drawbacks regarding performance.

Similarly, for power grid simulations, a distinction is made between *steady state* and *transient state*. While the former only considers the final, i.e., steady, state after the system's parameters change, transient state simulation also models the process, i.e., the transient behavior, between these steady states. Transient state simulation thus offers more insights at the cost of higher complexity for model parameters and calculations [6]. In Table 2, we classify different co-simulation approaches based on these paradigms and summarize to which extent they fulfill the requirements derived in §2.3. We further indicate their (desirable) open-source availability, i.e., that they do not rely on proprietary soft- or hardware.

Most approaches rely on discrete-event simulation (DES), offering scalable means for analyzing large networks. However, discrete-event simulation requires precise prior knowledge about communication behavior and potential side effects [109]. This is particularly challenging for cybersecurity research, where attacks occasionally attempt to undermine such assumptions and where the communication behavior is often part of the research subject. Further, discrete models for all involved tools and applications are required, which is difficult when including real programs, e.g., malware. Hence, we assess that DES lacks the required accuracy and flexibility for conducting sophisticated cybersecurity research in power grids, i.e., covering the entire scope of possible attacks (cf. §2.2).

On the power grid side, a transient model offers a more precise representation of the grid's behavior while considering time-based effects. Such transient effects are particularly interesting when modeling safety equipment that protects power grid components from excessive currents or voltages. However, transient simulations require precise knowledge about the components' characteristics while being computationally expensive, thus often requiring dedicated hardware. Furthermore, the gained detail level can be neglected in many cases [6]. Therefore, steady-state simulations are similarly well-suited for evaluating the power grid's state and behavior while being computationally less expensive.

Explicitly focusing on approaches that rely on continuous communication network simulation (as favorable for

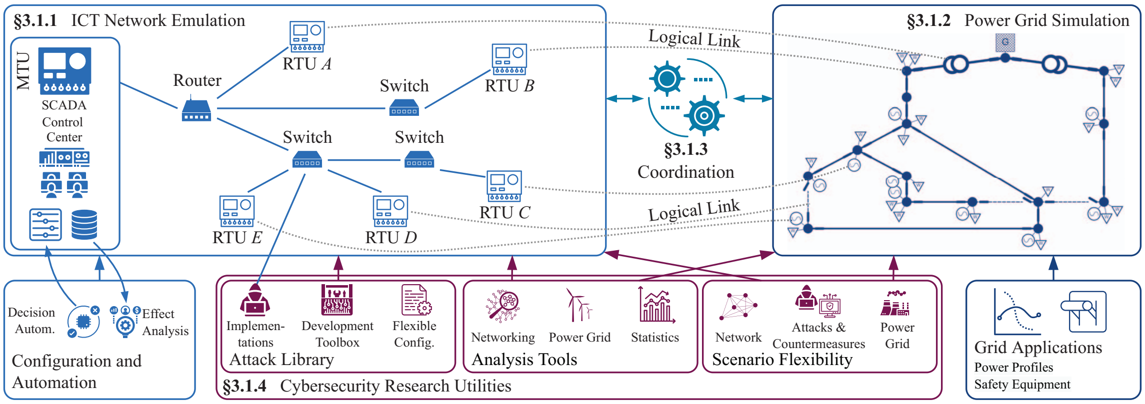

Figure 1. Wattson's architecture comprises three main components: an ICT network emulation, a power grid simulation, and a coordinator, complemented by several cybersecurity research utilities. The ICT network emulation deploys the network topology, emulates switches, links, routers, and hosts, and further supports the attachment of different applications, e.g., attackers or IDSs. RTUs are linked to the power grid via the coordinator and actively control and monitor changes within the power grid simulation, which can be configured with, e.g., power profiles and safety equipment.

analyzing the impact of cyberattacks), we find that Albarakati et al. [1] propose a transient approach where the co-simulator is based on OpenStack and OPAL-RT. However, the lack of validation of their approach's overall accuracy and scalability and the involvement of specialized and costly hardware exacerbates an independent evaluation of these critical requirements. Moreover, the proprietary power simulator and the involvement of hardware-based ICT components limit the flexibility of their approach regarding the range of possible scenarios and cyberattacks.

In turn, DSSnet [29], [30] follows a continuous, steady-state approach integrating Mininet and OpenDSS, which, however, suffers from poor scalability due to significant coordination overheads. Although DSSnet includes a continuous communication model, the authors confine themselves to simulating only the potential effects of cyberattacks instead of realistically conducting such attacks, which limits the assessability of the achieved flexibility and aptitude for cybersecurity research. Similarly, the co-simulator proposed by Li et al. [55] uses CORE (continuous) and GridLabD (steady state) to model ICT and power grid, respectively. While their distributed approach offers scalability, the authors neither validate its accuracy nor its suitability for cybersecurity evaluation.

To summarize, multiple co-simulation approaches address the challenge of enabling cross-domain research within power grids, although most do not specifically focus on cybersecurity research. Thus, design decisions specific to particular use cases, e.g., using a discrete communication model for scalability, impede their applicability for sophisticated cybersecurity research, as they cannot, amongst others, accurately implement real-world attacks. Further, the observed focus on either the communication or the power side (cf. §2.2) similarly holds for most research environments. Those few approaches that potentially offer the required cross-domain accuracy involve costly (proprietary) hardware, offer limited scenario flexibility, or exhibit indeterminate scalability. Consequently, a suitable research environment adhering to all requirements of §2.3 is still needed to facilitate comprehensive analyses of cyberattacks on power grids.

3. Wattson – A Foundation for Analyzing the Impact of Cyberattacks on Power Grids

Despite the sophisticated approaches for evaluating the impact of cyberattacks on power grids offered by related work, our analysis underlines the lack of cybersecurity-focused environments that accurately and flexibly cover the ICT and energy side while maintaining scalability to realistic grid sizes. Hence, we introduce Wattson, a cybersecurity research environment for power grids striving to achieve these properties (cf. §2.3) by combining a detailed ICT network emulation with a flexible power grid simulation and integrated cybersecurity research tools, covering attack frameworks, evaluation tools, and the possibility to deploy realistic countermeasures. Wattson combines valuable concepts from related work, while its cybersecurity-focused architecture makes it unique compared to these approaches. In the following, we present Wattson's design (§3.1) before evaluating its accuracy, scalability, flexibility, and cybersecurity features (§3.2).

3.1. Architecture Design

We depict Wattson's architecture and components in Fig. 1, consisting of a realistic *ICT network emulation*, an accurate *power grid simulation*, a dedicated *coordinator* linking and synchronizing the two, and *cybersecurity research utilities*. Hence, Wattson enables analyzing the interaction and interplay of the communication network (i.e., hosts, routers, switches, links) and the power grid (e.g., transformers, lines, loads, or generators). Besides a manual configuration of the ICT network, Wattson supports automatically deriving grid configurations through a dedicated modeling approach [48]. Wattson is primarily implemented in Python and we refer to our documentation website[2] for technical details. In the following, we detail the design of Wattson's components.

3.1.1. ICT Network Emulation. To facilitate *flexible* network topologies and the deployment of universal network-based applications, Wattson uses network emulation to

---

[2] https://wattson.it

model the ICT network of power grids. This approach enables the use of realistic communication protocols, e.g., IEC 60870-5-104, network monitoring tools such as Wireshark, and network-based attackers. We utilize Containernet [82], a fork of the network emulator Mininet [51], to create an ICT network comprising switches, routers, links, and process- and Docker-based hosts. The emulator uses Linux network namespaces to create several virtual hosts on the same physical host along with virtual switches based on Open vSwitch [83] and virtual links. Thereby, the precise configuration of devices and the underlying network, including link properties, e.g., jitter, bandwidth and delay, and realistic communication using the Linux networking stack down to Layer 2 become possible.

We extend Containernet to realize horizontal *scalability* of the network emulation by partitioning the network into several segments that can be distributed onto several physical hosts and connected via physical links. For emulating the ICT network model, we include implementations of RTUs and the corresponding MTU using the IEC 60870-5-104 protocol, and a graphical Virtual Control Center (VCC). RTUs serve as Intelligent Electronic Devices (IEDs) attached to one or multiple grid components. They transmit monitoring information, e.g., voltage measurements, to the MTU at the control center, which, in turn, can issue control commands to monitor and actively manage the grid. All components realistically communicate over real networking protocols (Ethernet, IPv4, TCP, IEC 104). Thus, WATTSON supports the evaluation of realistic ICT attacks such as DoS, network reconnaissance, Address Resolution Protocol (ARP) spoofing, and machine-in-the-middle (MitM) attacks.

3.1.2. Power Grid Simulation. WATTSON uses a steady-state simulation for the power grid, achieving *scalability* and meeting the real-time requirements induced by network emulation. Here, we utilize pandapower [101], a power flow solver supporting symmetric AC single- and three-phase systems. Consequently, changes in the configuration of the power grid, e.g., power adjustments, consumption changes, or topology changes resulting from opening or closing circuit breakers, trigger a simulation step that outputs the grid's power flow once a steady state is reached. Moreover, WATTSON supports adding Gaussian noise to power values of loads and generators before and after the power flow computation to model slight control deviations and measurement inaccuracies. Further, it provides load profiles representing the realistic behavior of the power demand over time.

While the steady-state approach allows the power simulation to keep up with the real-time network emulation, WATTSON cannot directly include any transient compensation processes. Still, we assess that the scalability and real-time capabilities of this approach offsets this imperfection and that the steady-state simulation suffices for comprehensively researching realistic cyberattacks.

3.1.3. Simulation Coordination. *Accurately* representing cyberattacks on power grids requires the consideration of the interplay between networking devices and physical grid components. Thus, we have to link the network emulation with the power grid simulation and coordinate their respective interactions. WATTSON includes a dedicated coordinator, interfacing with each host within the emulated network for interacting with the power grid simulation. In particular, each host is connected to a dedicated *management network* that solely transmits coordination traffic. By communicating with the coordinator, hosts can receive the grid's current state and update its configuration. As soon as the coordinator receives a request to change the grid configuration, it triggers the power simulation.

The functional linking of networking hosts to power grid components (cf. Fig. 1) is realized by a respective configuration of hosts, either manually or using an automated modeling approach [48]. For each monitoring and control information, the configuration defines the responsible RTU, the IEC 104 information object address (IOA), and information for interacting with the coordinator and the respective grid element. Through its close coupling with the power simulator, the coordinator further enables WATTSON to apply flexible logic applications to the power grid, e.g., load profiles, self-adjusting inverters, models for safety devices, or defects between simulation steps.

3.1.4. Cybersecurity Research Utilities. We design WATTSON as a thorough cybersecurity research environment, going beyond the scope of a pure co-simulator. For every component, we focus on including cybersecurity-related tools and functionality as well as their flexible extendability for future research directions. Particularly, WATTSON integrates three design aspects directly contributing to its cybersecurity focus.

Comprehensive Attack Library. WATTSON comes with an integrated cyberattack framework. Ranging from various DoS attack variants over re-implementations of existing malware, e.g., the Industroyer, to a sophisticated library for transparent semantic, e.g., false data injection, attacks, WATTSON provides a complex toolbox for conducting and evaluating the impact of such attacks in different scenarios. Due to ethical and risk considerations, we refrain from publishing this attack library.

Integrated Analysis Tools. WATTSON includes analysis tools per design to provide insight into the process and results of attacks and potential countermeasures. Based on configurable capture and export tools, e.g., power grid exports and targeted packet captures, WATTSON provides a library for analyzing these artifacts.

Scenario Flexibility. Finally, WATTSON targets to enable *flexible* research opportunities. Hence, we implement extensible scenarios, where changes and extensions to the power grid and the communication network's topologies and behaviors, including attack and countermeasure deployments, can be straightforwardly realized.

To summarize, WATTSON leverages two specialized tools for power flow computation and network emulation, coupled by a dedicated coordinator and enriched with advanced cybersecurity research tools. It thus enables the deployment and execution of real networking protocols and applications within the ICT network that interact with the power grid. We now analyze and discuss WATTSON's fulfillment of the demanded requirements (cf. §2.3).

3.2. Fulfillment of Requirements

For WATTSON to be usable and applicable to study the impact of cyberattacks on power grids, we need to ensure it fulfills the derived requirements, i.e., *accuracy*, *scalability*, *flexibility*, and suitability for *cybersecurity* research. In the following, we thus validate Wattson's accuracy by replicating an attack on a physical testbed. Then, we show that Wattson scales to realistic grid sizes based on benchmarking scenarios and reference grids. Finally, we discuss its flexibility and cybersecurity aspects.

3.2.1. Accuracy. To analyze Wattson's accuracy, we replicate a physical low voltage distribution grid testbed operated at RWTH Aachen University [88], [104] shown in Fig. 2 within Wattson and compare the behaviors of simulation and testbed. Hereby, we rely on the correct operation of both, Containernet [82] and pandapower [101], analyzing Wattson's combined, i.e., overall accuracy. The testbed comprises a medium/low voltage substation at 630 kVA, two photovoltaic (PV) inverters (Inv.), a battery inverter, three resistive loads at 20 kW maximum power consumption, and a corresponding ICT network. While all three inverters are controlled via IEC 104, the power consumption of the loads follows a specified time series power profile and is not controlled by the grid operator. During the experiment, the power output of the inverters is adjusted to match the respective demand. All inverters maintain a minimum power output and apply a predefined power factor of 0.95. The ICT network comprises a single MTU, three RTUs, three switches, and an attack host attached to the switch at the RTU of the *PV Inv. B*.

At the beginning of the experiment, the attackers perform an ARP spoofing attack against said RTU and the MTU to redirect the traffic between these hosts. At 123 s and 173 s into the experiment, the attackers inject a control command setting the power output of the *PV Inv. B* to its minimum value. There is a non-neglectable delay between the reception of the command and the visible realization in the measurements for all three inverters in the laboratory setup. We observe a random delay in the physical RTUs, resulting in a slight uncertainty regarding the actual delay due to measurement intervals of ≈2 s. Thus, we determine the earliest and latest command realization time based on the power measurements and model this uncertainty in the simulation by applying a random delay between the minimum and maximum real delay. We conduct 20 simulation runs of the physical testbed scenario, each using a different genuine random seed [106]. We plot the laboratory and simulation measurements for active and reactive power and network communication in Fig. 3. For the simulation, the plot indicates the minimum and maximum measurement and highlights the area of uncertainty.

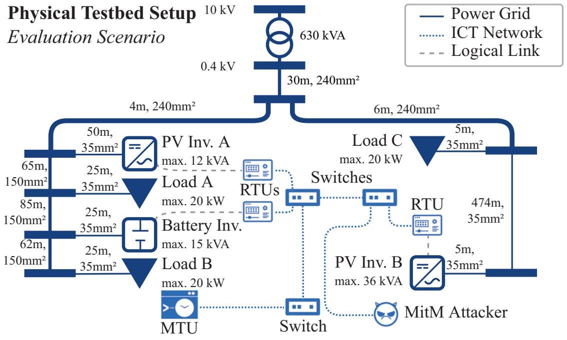

Figure 2. The physical testbed used to validate Wattson's correctness contains a single transformer, two inverters, one battery inverter, and three resistive loads. Each inverter can be controlled by a single MTU via a dedicated RTU from the ICT network. An attack host is attached to one of the switches to perform an ARP-spoof-based MitM attack.

As shown in Fig. 3(a), the simulation corresponds to the laboratory for both active and reactive power. We observe that the simulation falls within the uncertainty area over multiple runs during power level switching. Besides these time ranges, the simulation's active and reactive power measurements, voltages, and currents accurately match those from the laboratory. Notably, these consistencies are not limited to normal operations but continue during the ICT-driven attack. For the power level adjustments of the *PV Inv. B* induced by maliciously injected control commands, we observe dips of ≈4 kW in the laboratory measurements, which are not present in the simulation. These effects result either from overreactions of the inverter or minor measurement inaccuracies and are not observable with a steady-state simulator. Apart

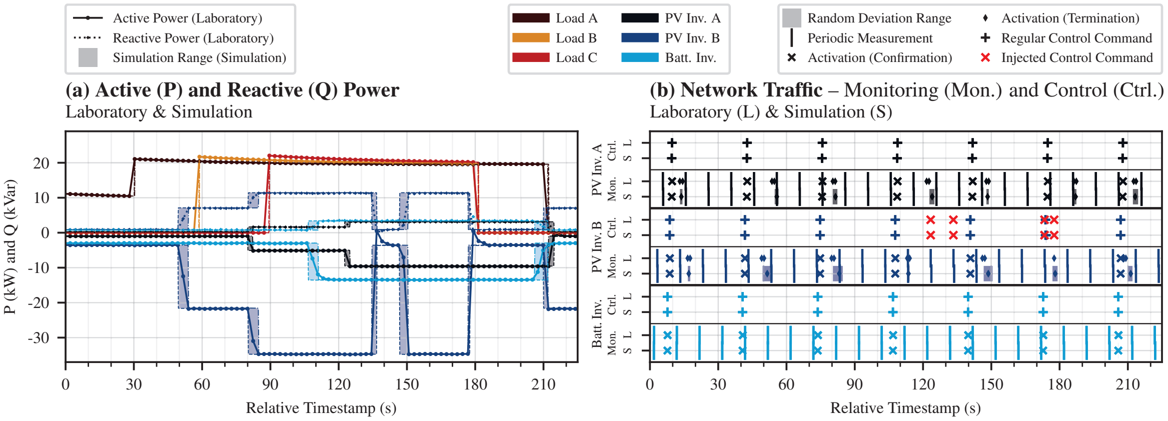

Figure 3. To validate Wattson's accuracy, we compare its behavior to a physical testbed for power grid and ICT networking under normal and attack conditions. For the power grid, the simulation precisely represents the testbed behavior, as course and absolute values of power measurements correspond to real-world measurements. Similarly, Wattson matches the communication patterns and contents of the testbed network and its devices.

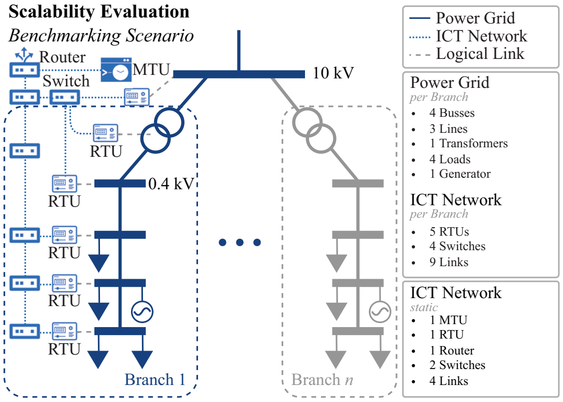

Figure 4. To evaluate WATTSON's scalability, we create linearly scalable power grids with *n* branches, each containing a single transformer, multiple buses, multiple assets, and corresponding ICT network components, resulting in a linearly growing number of hosts, switches, and links.

from these minor deviations, the behavior of the simulated power grid fully corresponds to the laboratory setup.

Similarly, we observe matching traffic patterns between the emulated and the laboratory ICT network in Fig. 3(b), where we indicate timing variations between individual runs with faded areas. For all three RTUs, the periodically sent control commands are timed consistently. Analogously, the monitoring direction is equally present in every periodic measurement in the simulated network. Except for the variations related to the randomized PLC delays, the deviation areas exhibit a median of only 56 ms, indicating WATTSON's suitability for accurately modeling power grids and their corresponding ICT networks.

3.2.2. Scalability. We evaluate WATTSON's scalability to show that it can simulate realistic, large-scale power grids, covering architecture-induced overheads and corresponding delays. For our measurements, we use a single machine with two AMD EPYC 7551 processors, each offering 32 cores with 2 threads per core and 256 GiB of RAM. Besides widely-used realistic reference high/medium voltage grids (Cigre MV [97] and Simbench `1-MV-semiurb-0-sw` [70]), we design specific benchmarking scenarios to show the scalability of WATTSON. As shown in Fig. 4, we use power grid layouts with n branches ($n = 2, ..., 64$), all connected to a single medium voltage bus. Each branch contains four low-voltage buses with a linked RTU each, a transformer, four loads, a generator, and an RTU for the transformer, resulting in a linearly growing number of hosts, switches, and links. We express the combined number of assets within the ICT network and power grid as scenario size S.

We then investigate (i) the IEC 104 *communication latency* induced by network stacks and link delays; (ii) the *coordination overhead* from the delay between RTUs and coordinator; and (iii) the *power flow computation* time needed to simulate the grid state. Here, the *communication latency* per hop should closely adhere to the defined link delays of 2 ms plus a minor overhead induced by network stacks. The *coordination overhead* should be as low as possible and independent of scenario sizes. Finally, the *power flow computation* must not exceed delays induced by physical components in real-world power grids, ranging from tens of milliseconds up to 10 s [17].

To obtain reliable measurement results, we conduct 10 independent runs for each power grid layout (Cigre MV, Simbench `1-MV-semiurb-0-sw`, and our benchmarking grids with 2, 4, ..., 64 branches). During each run, the MTU performs 10 rounds of read and control commands, where all data points referring to bus voltages are read while all generators are set to active power outputs of 50 % or 100 %, alternating between rounds. Hence, the resulting traffic should scale with the size of the power grid while control commands trigger power flow computations.

We show our measurement results in Fig. 5, depicting the arithmetic mean over all runs with the 98 % mean confidence interval. First, we ascertain that power flow computations logarithmically scale with increasing grid size, exhibiting adequate scalability of the power simulation. Similarly, the coordination overhead is reasonably low, with a maximum mean coordination overhead of 1.5 ms. Hence, we assess that the simulation-induced overheads of power flow computations and coordination are suitable to match real-world deployments and thus enable conducting and evaluating cyberattacks at realistic scales.

Moreover, we observe strict adherence of the mean hop delay to the configured link delay of 2 ms, with a

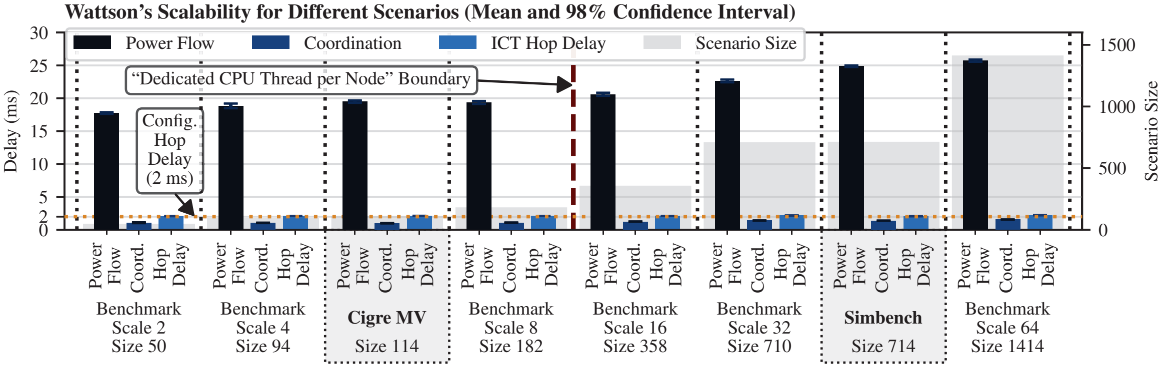

Figure 5. Despite increasing scenario sizes (growing number of assets in the power grid and nodes in the ICT network), our measurements indicate WATTSON's ability to scale to realistic grid size. We observe slightly increasing overheads for power flow computation, a reasonably low coordination overhead, and ICT network hop delays corresponding to the actual hop delay of 2 ms, enabling us to evaluate cyberattacks at realistic scales.

maximum mean of 2.2 ms and a maximum 98 % confidence interval width of 0.12 ms across all scenarios. Here, the delay slightly increases with growing scenario sizes due to scheduling effects when more nodes need to be simulated than CPU threads are available. These numbers are more than sufficient to evaluate cyberattacks at realistic scales. Still, for larger scenarios or less powerful hardware, WATTSON supports horizontal scalability by distributing the network over multiple systems.

3.2.3. Flexibility & Cybersecurity. To tap WATTSON's full potential, provided by its accuracy and scalability, we must ensure that its cybersecurity research capabilities offer extensive flexibility. Here, WATTSON addresses flexibility requirements on various levels. First, neither the ICT nor the power grid topology are fixed or restricted. They can be freely defined independently of each other, laying the foundation for flexible research scenarios. Second, their individual behavior as well as their coupling can be flexibly configured, e.g., the implementations of RTUs and any other host are freely exchangeable, and monitoring and control behavior can be adjusted as needed. In addition to these general flexibility aspects, WATTSON further offers flexibility regarding the research questions. Its network emulation approach allows for arbitrary attack implementations, ranging from custom-made attacks to deployment of real-world malware, whereby WATTSON already includes an exhaustive library for such attacks. Combined with capabilities for deploying potential countermeasures, e.g., IDSs or safety measures, WATTSON provides an extensive toolbox for various research scenarios, such that we assess its flexibility as suitable for cybersecurity research in power grids.

To summarize, our evaluation shows that WATTSON *accurately* replicates experiments in a physical testbed and can perform real-time simulations of reference and benchmarking scenarios with sufficiently low overhead, thus being able to *scale* to realistic power grid sizes. Further, it enables *flexible* topologies for the ICT network and power grid to conduct various customized cyberattacks and deploy arbitrary countermeasures. These combined properties underline WATTSON's suitability for sophisticated *cybersecurity* research in power grids.

4. Analyzing the Impact of Cyberattacks

Our evaluation shows that WATTSON fulfills the requirements for performing sophisticated cybersecurity research for power grids. Thus, we implemented several attacks using WATTSON, where we followed the attack classes from §2.2 and discussed and validated their practical relevance with domain experts from a national cybersecurity agency and multiple grid operators.

For all attacks, we rely on the realistic Simbench [70] `1-MV-semiurb-0-sw` medium voltage distribution grid, representing a suburban grid with two transformers connecting a 110 kV grid to a 20 kV grid, resulting in a total of 475 assets. The corresponding ICT network includes 119 RTUs, a total network node count of 239, resulting in a scenario size of 714. We define a link latency of 2 ms and a bandwidth of 100 Mbit/s.

Threat Model. We assume that attackers gained access to the ICT network either by breaking into an insufficiently secured facility, e.g., an unmanned substation, or through a remote attack, e.g., spear-phishing [50]. Thus, the attackers cannot fully control *where* they gain access to the network. However, they might use network reconnaissance to gather information about vulnerable devices and services to plan and execute the attack.

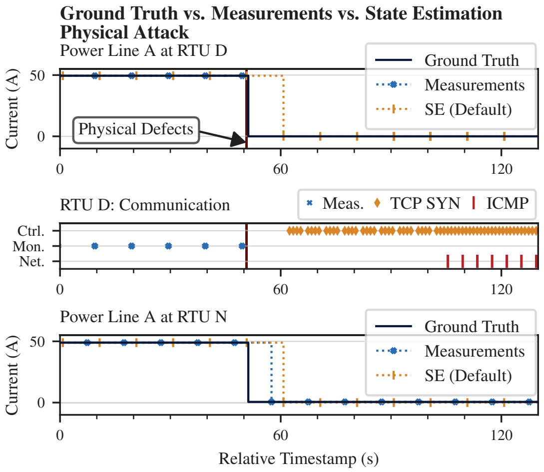

Figure 6. Physical attacks, such as disconnecting a substation (RTU *D*, transmission lines, and the bus bar), lead to a local blackout (top) and interrupted communication between the substation and the control center (middle). However, the adjacent RTU *N* measures the outage and informs the control center (bottom), leading to a correct SE.

We begin our analysis with physical attacks, i.e., disconnecting or destroying equipment (§4.1). Then, we conduct DoS and ARP flooding attacks to analyze the impact of syntactic attacks (§4.2). Finally, we study semantic attacks where the attackers manipulate communication to disrupt the safe operation of the power grid (§4.3).

4.1. Physical Attacks on Grid Equipment

The cross-domain representation of the power grid within WATTSON allows for studying the interplay of ICT and power grid components during an attack. As an example scenario, we implement a physical attack against a local substation, where the attackers physically disconnect or destroy ICT devices, i.e., the substation's RTU *D*, and local power grid equipment, i.e., transmission lines and the bus bar. Thus, the substation becomes inoperative for controls issued on-site and remotely. We plot the resulting effects on the power grid and communication in Fig. 6.

The attackers physically destroy the equipment at ≈50 s, disconnecting the entire substation from the grid and causing a locally constrained blackout. Consequently, the current over the power line connected to this substation immediately drops to 0 A. Furthermore, this attack interrupts the communication between the substation and the control center, mainly affecting the periodic measurements of the corresponding RTU *D*. Therefore, the MTU attempts to re-establish the connection after the TCP connection times out, indicated by the repeated TCP SYN packets in the communication plot. Since the RTU is disconnected from the network, the responsible router

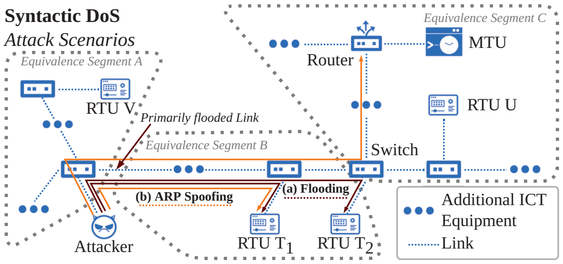

Figure 7. For both syntactic attacks, we attach an attack host via a 1 Gbit link to a switch in the network. During the flooding attack (a), attackers flood RTUs T_1 and T_2, mainly saturating the indicated link. During ARP spoofing (b), attackers poison the ARP caches of RTU T_1 and the router.

sends ICMP packets to the MTU, stating its unreachability. Since the measurements of an adjacent substation (RTU N) cover the impacted power line, the control center can localize the defects. As shown in Fig. 6, they are also covered by the state estimation.

Through conducting this attack with WATTSON, we can understand the locally limited impact of physical attacks on power grids. Furthermore, this simulation illustrates that operators have various means on the power grid and communication side to detect such attacks, as their immediate impact is comparable to a physical outage. However, attackers might also exploit their access to target a specific remote device, as presented in the following.

4.2. Syntactic Attacks on Grid Communication

Increasing the complexity of attacks, we continue with assessing the potential harm of syntactic attacks against the ICT network, aiming at a DoS condition: a TCP SYN flooding [19] attack against a network branch and ARP spoofing [86], [111] as a more precise attack measure.

4.2.1. TCP SYN Flooding. Depending on the networking equipment and the targeted RTUs, such an attack might saturate the network's throughput capabilities or exhaust the resources at the targeted host by creating thousands of TCP contexts. As a result, connections might get delayed, leading to higher packet loss and even collapsing TCP connections due to excessive delays. To comprehensively understand these effects, we divide the considered network into equivalence segments following the network configuration and topology, as shown in Fig. 7. We expect a comparable impact on the devices belonging to the same segment. With this approach, we can subsequently analyze the TCP SYN flooding impact on each segment's representative device to understand this impact. We use `hping3` with three processes for each target, flooding the target with randomized source addresses and a TCP payload of 1400 B. The attack is executed from a single host, attached to a switch via a 1 Gbit link.

Fig. 8 shows the impact of this attack on the ability to reliably communicate measurements of the grid state for the different equivalence segments. Due to the attackers' position in the network, the IEC 104 communication of the two targeted RTUs T_1 and T_2 is not notably affected by the attack since the primarily flooded link is not on the path between either RTUs T_1 and T_2, and the MTU (equivalence segment B). For similar reasons, RTU U's communication is not impaired, as holds for all RTUs in equivalence segment C. However, since the communication of RTU V (and all other RTUs in equivalence segment A, 51 RTUs) entirely depends on the saturated link segment, it loses connectivity to the MTU during the attack, resulting in the absence of measurements and a loss of visibility for the control center.

In particular, the attack leads to a loss of periodic measurements from 51 of the 119 RTUs, further affecting the grid state estimation. As shown in Fig. 8, a loss of up to 42 % of measurements heavily impacts the state estimation accuracy, as both applied estimation modes exhibit significant inaccuracies during the attack. Here, assuming formerly reported measurements to remain valid when periodic transmissions are missing (*SE Default*) results in the estimation to stay close to these outdated measurements. In turn, dropping outdated measurements (*SE Decay*) occasionally achieves a more precise estimate of the actual grid state but also exhibits large deviations.

In summary, we could reveal the profound linkage between the ICT and the power grid, where even a simple flooding attack significantly impacts the visibility and controllability of the power grid. Moreover, the impact mainly depends on the attackers' location within the network and the timing of the attack, as the sole loss of visibility does not necessarily induce a critical grid state. However, in the presence of significant changes in power consumption, the lack of visibility might destabilize the grid, trigger wrong control action decisions or prevent their realization. Still, such attacks are rather evident to grid operators.

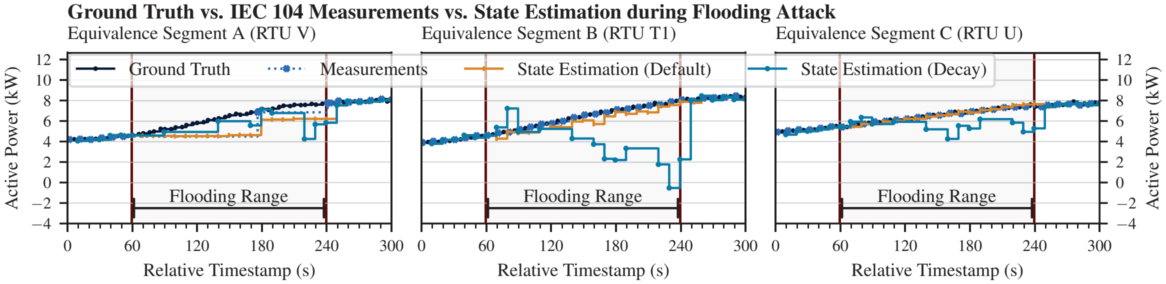

Figure 8. The flooding attack only affects the RTUs in equivalence segment A. During the attack, the majority of measurements of these RTUs do not reach the MTU, leading to a loss of visibility and divergence of real and observed states. Since 51 RTUs communicating over the flooded links are affected, only ≈58 % of the measurements reach the MTU. As a result, even the state estimations cannot fully compensate for the lack of information.

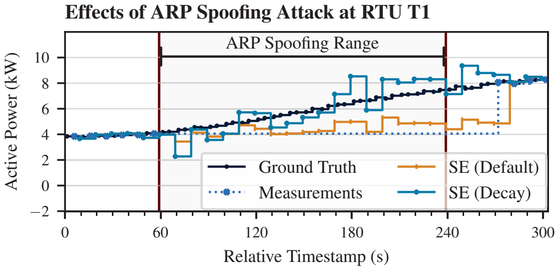

Figure 9. The ARP spoofing attack intercepts all traffic between the MTU and RTU T_1. As a result, the connection drops and no measurements reach the MTU. After the attack, it takes ≈30 s for the ARP cache to recover and re-establishing the connection. Although the attack only affects a single host, the state estimation accuracy for measurements of RTU T_1 decreases.

4.2.2. ARP Spoofing. The effectiveness of TCP SYN flooding strongly depends on the attackers' attachment point within the network. In contrast, ARP spoofing [86] is a more targeted attack strategy for DoS, potentially allowing attackers to intercept and interrupt IP-based communication within a subnet. By sending forged ARP replies, attackers can poison a victim's ARP cache and manipulate the mapping of IP addresses to MAC addresses. Essentially, the attackers announce themselves as the host with the IP addresses of their victims, causing corresponding packets to be forwarded to the attack host instead of the original destination. Consequently, they can eavesdrop on the communication and especially drop arriving packets.

For the spoofing attack, we use the same configuration as in the TCP SYN flooding attack scenario (cf. Fig. 7). We send spoofed ARP packets to the router and RTU T_1 using `arpspoof`. After poisoning the ARP caches, the traffic between the two hosts is redirected to the attack host, which drops all traffic and thus induces a DoS condition. Thus, the TCP connection between the MTU and RTU T_1 times out, and no more measurements from the RTU arrive at the control center. As the results in Fig. 9 indicate, even the loss of a single information source results in a decreased accuracy of the state estimation for the affected grid components, especially for the default SE behavior. After the ARP spoofing attack stops, the ARP cache takes another ≈30 s to recover from the poisoning, i.e., the attack effects last longer than the actual attack.

Compared to TCP SYN flooding attacks, ARP spoofing allows for more precise target selection since only the targeted RTU T_1 is affected by the attack. However, as the combined results of WATTSON show, even losing a single RTU affects the visibility within the grid and might even interfere with control commands. Still, such attacks can be easily detected, as missing measurements clearly indicate an anomaly (defects or attacks).

4.3. Comprehensive Semantic Attacks

While syntactic attacks (cf. §4.2) might achieve a loss of visibility and control within the power grid, they do not actively interfere with the grid operation. Further, they are easily detectable and preventable. Thus, we shift our focus from syntactic attacks to more sophisticated domain-specific semantic attacks. Here, attackers actively manipulate and inject commands and monitoring information by utilizing knowledge about used protocols, the grid's configuration, and the semantic meaning of network communication. In the following, we first analyze Industroyer [12], a real-world semantic attack that led to an actual blackout, before we study a stealthy false data injection attack.

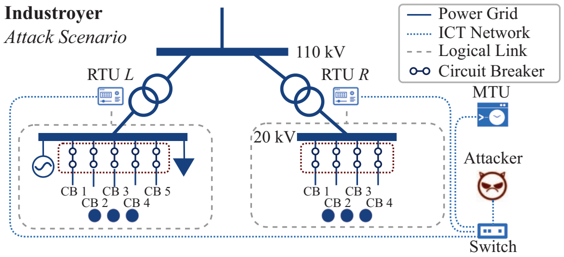

Figure 10. The Industroyer malware connects to two RTUs placed at the central substation of the Simbench reference grid. During the attack, the malware issues single-point commands to open all circuit breakers managed by these RTUs to disrupt the power supply of the grid.

4.3.1. The *Industroyer* Circuit Breaker Attack. The Industroyer family [12], which, among other things, targets IEC 104-based networks by sending control commands to RTUs, is a real-world example of such a semantic interference strategy. During the Ukrainian power grid attack in 2016 [69], the Industroyer malware successfully cut power for a fifth of the capital Kyiv. In the more recent Ukrainian power grid attacks in April 2022, its successor, the so-called *Industroyer2* [21], was (unsuccessfully) used by attackers aiming to again induce a blackout [21]. Industroyer attempts to connect to one or multiple pre-configured IP addresses as an IEC 104 client for issuing control commands. Thus, the attackers need prior insider knowledge on the devices (RTUs) w.r.t. their IP addresses and IEC 104 configuration, i.e., the malware configuration further includes the common address (COA) and associated IOAs for each configured RTU [12], [21]. By issuing control commands to open or close circuit breakers, Industroyer disconnects or connects the respective grid parts from the power supply.

We implement an attack reproducing the behavior of the Industroyer malware in WATTSON to analyze its potential impact on the power grid. As we depict in Fig. 10, we deploy an attack host configured to connect to two RTUs responsible for the central substation of the grid. Immediately after establishing these connections, the attackers repeatedly issue IEC 104 single-point commands for two minutes, causing the circuit breakers to open, effectively disconnecting most of the medium voltage grid from the high voltage grid. Closely following the implementation of the Industroyer malware, we use a dedicated thread per RTU, each issuing control commands in a 3 s interval [20].

In Fig. 11, we visualize the command injections, their effects on the circuit breakers, and the total power load. When the attack starts, the attackers issue control commands to open the pre-configured circuit breakers. After 12 s, a command for each circuit breaker has been issued, disconnecting more than 1.6 MW of load from the grid. Essentially, this creates a large-scale blackout with little

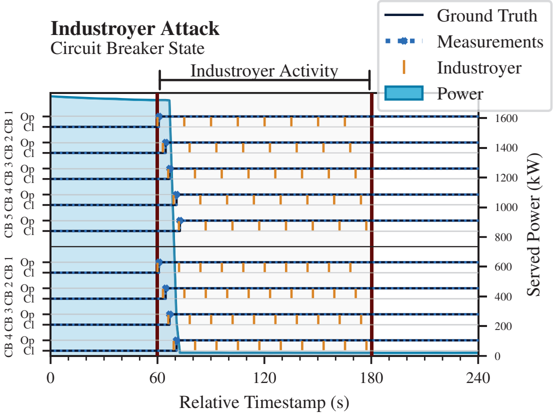

Figure 11. As soon as the Industroyer malware connects to both RTUs, it repeatedly issues control commands to open (Op) circuit breakers (CB). This ensures that even if the MTU issues commands to close (Cl) the circuit breakers, they are opened again. The attack results in more than 1.6 MW of load being disconnected from the grid, causing a blackout.

potential for effective countermeasures by the operator. Even after the attack stops, restoring the grid's full operation can take several hours to days, as equipment might be damaged, and the grid requires a gradual restart [80].

As Industroyer does not rely on ARP spoofing, preventing it requires further measures besides, e.g., static ARP tables. Even the detection of unauthorized devices would not suffice to detect and prevent this attack since, in the real-world attack, the malware was deployed on hosts controlled by the grid operator [21]. However, a stringent whitelist of allowed IEC 104 clients and/or message authentication can reduce the chance of a successful attack. We exemplary verified the impact of a whitelisting approach: Configuring RTUs to reject IEC 104 connections from hosts different from the MTU renders the current Industroyer attack unsuccessful (thus attackers, e.g., would additionally need to spoof the IP address of the MTU).

Our results obtained using WATTSON are twofold: First, they underline the significant security flaws of real-world power grids that enable such attacks. With only limited knowledge about the power grid and its network configuration, attackers are able to cause significant physical damage and induce a blackout. Second, they demonstrate that implementing countermeasures, e.g., whitelisting IEC 104 clients, effectively hamper attack success.

4.3.2. False Data Injection. Despite the severe consequences of the Industroyer attack, its effects are immediately visible to the grid operator. Therefore, we now assume sophisticated attackers with advanced knowledge, targeting to transparently manipulate measurements sent to the control center to obfuscate the attack's effects in a future-oriented scenario with numerous renewable power sources attached to the ICT network. Similar to the previous attacks, we deploy an attack host, as shown in Fig. 12, which establishes itself as an MitM between the MTU and multiple RTUs using ARP spoofing. Alternatively, attackers might also physically reconnect networking cables via the attack host. To evaluate the potential cascading effects of the attack, we deploy virtual safety devices at each circuit breaker, monitoring voltage, and current limits and disconnecting parts of the grid if these limits are exceeded.

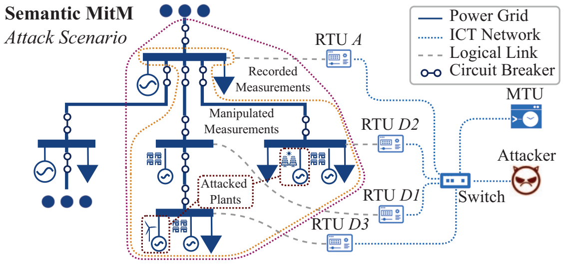

Figure 12. For the false injection, the attackers place themselves as an MitM between the MTU and four RTUs. After an initial recording phase, the attackers transparently inject commands for deactivating the power infeed of two renewable power plants. Subsequent measurement transmissions are manipulated to reflect the original grid state.

As an MitM, the attack host records all measurements and control commands for 60 s. Afterward, it injects two control commands to shut down the power infeed of two renewable plants at RTUs $D2$ and $D3$, as indicated in Fig. 12, and suppresses the respective command confirmations (to hide the command injection from the control center). For each injection, manipulation, or suppression of packets, the attackers adjust the sequence and acknowledgment numbers for the TCP and IEC 104 connections. Hence, the original communication partners are entirely unaware of the intercepted communication. To conceal the physical impact of the injected commands, the attackers replace subsequently communicated measurements of the targeted RTUs with those matching the trend of the initial recording phase (*freezing*). Further, they suppress any control commands sent from the control center (while still acknowledging them), yielding the operator without control over the attacked RTUs.

As shown in Fig. 13, the active power infeed at the buses controlled and monitored by RTUs $D2$ and $D3$ drops close to 0 kW after the command injection at 150 s. However, the manipulated measurements still report an active power injection of 691 kW and 1.47 MW, respectively. Due to the inconsistencies between the measurements from the attacked RTUs and the remaining RTUs, the control center's state estimation miscalculates the grid state significantly, incorrectly suggesting an increased power drain of more than 2 MW at the bus at RTU A.

Besides hindering the localization of the actual event leading to these deviations and complicating the detection of the attack along with potential countermeasures by the grid operator, this attack might even provoke improper reactions by the grid operator, e.g., targeted load shedding at the bus controlled by RTU A. Depending on the attackers' technique for establishing themselves as an MitM, stopping the ongoing attack might require physical interaction of the grid operator and time-consuming reconnections of power grid assets. The combination of active interference with the grid operation by hiddenly injecting control commands and the subsequent false data injection for obfuscating the commands' existence and effects thus emerges as a particularly critical attack scenario.

To summarize, our evaluation using WATTSON illustrates the inseparable connection between ICT and physi-

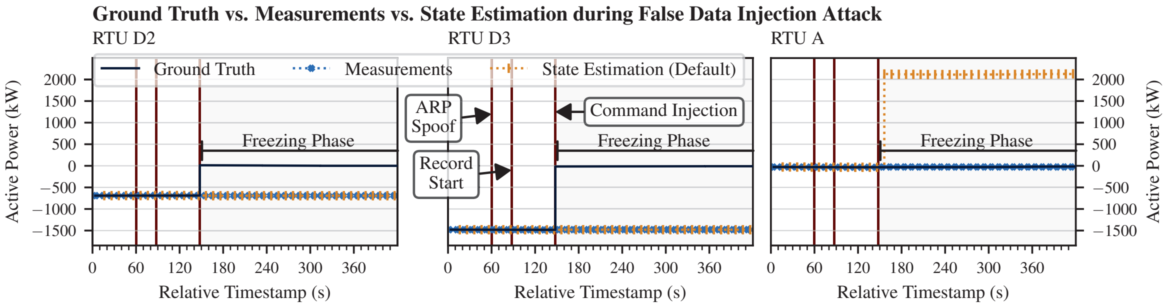

Figure 13. During the false data injection attack, the attackers establish themselves as an MitM and record monitoring information sent by multiple RTUs for 60 s. Afterwards, they inject two control commands to disable the power infeed of two renewable power plants controlled by RTUs $D2$ and $D3$. However, to obfuscate the effects of their attacks, the attackers replace subsequent measurements with values matching the grid state before the command injection. As a consequence, the actual grid state, reported measurements, and the state estimation (SE) results differ significantly.

cal processes in power grids. Further, we see that attackers can exploit knowledge of the power grid's structure and operation to amplify the immediate effects of attacks.

Overall, our analysis of the impact of various cyberattacks on power grids using WATTSON shows the importance of comprehensively considering effects on the communication and energy side of power grids to fully understand the attack capabilities, thus paving the way for adequately securing tomorrow's power grids.

5. Discussion and Future Research

Using WATTSON, we reproduced and analyzed different attacks from various attack classes (cf. §2.2) where the attackers increasingly require more knowledge about the power grid, correlating to the criticality of the attack's impact. In the following, we discuss the insights gained during our evaluation of the impact of cyberattacks (cf. §4) and identify further research potential.

As the least sophisticated attacks, *physical attacks*, e.g., destroying equipment, typically only have a locally limited impact on the power grid operation due to the redundant structure of such grids. Here, WATTSON enables an impact assessment of outages of different components, thus identifying which components require special protection. Moreover, it spurs improving the detection of outages, regardless of whether they result from a cyberattack or a purely technical failure, e.g., by identifying distinct artifacts of such outages w.r.t. to communication behavior.

In turn, *syntactic attacks* can impede the visibility of the grid operator. Moreover, we also understand that syntactic DoS attacks do not actively change the grid state, and the grid remains operational as long as no control commands are required. While power grids traditionally only occasionally required active control [73], increasing power demands and renewable power sources nowadays lead to a higher necessity for active control commands. Thus, such modern power grids are more affected by DoS attacks against the communication network. Nevertheless, albeit rarely performed in practice, such attacks are preventable and easily detectable, e.g., with static ARP tables, switch port authentication or deactivation, rate limiting, and IDSs. Here, WATTSON provides a safe environment to test such security measures before rolling them out.

Most importantly, we studied complex *semantic attacks*, illustrating how attackers can physically damage the power grid by (remotely) exploiting susceptible ICT. In particular, our results show that false data injection attacks involve a high risk for grid operators since they actively manipulate the grid's state while also obfuscating the attack's impact, leading to delayed reactions. In the extreme case, a sophisticated attack can cut off parts of the grid or permanently damage physical assets while simultaneously faking normal grid conditions to the grid operator. Besides statically freezing the current grid state, such an attack could also react to changes in other regions of the grid and influence the power flow, making it virtually impossible to detect the attack (both on the ICT and energy side) based on received measurements.

While we primarily designed and demonstrated WATTSON for analyzing the impact of cyberattacks, our discussion indicates enormous *further potential* for WATTSON to strengthen cybersecurity in power grids. As evident from our analysis and discussion, there is an inherent need to roll out measures to prevent (e.g., encryption and authentication of commands and measurements) and detect (e.g., intrusion detection systems) such cyberattacks.

However, power grid operators are typically reluctant to roll out any changes to their infrastructure as they fear negative impacts on grid operation and often feel overwhelmed with selecting, prioritizing, and adapting security measures. For example, power grid operators have been reluctant to introduce comprehensive authentication and encryption since intensive computations for long-living, resource-constrained devices contradict the power grid's stringent availability and latency requirements [22], [34], [35]. Likewise, when considering intrusion detection systems, operators face the challenge of reliably locating an attack's origin within geographically widespread power grid networks [5] without extensive sensor placements [10]. Here, WATTSON can serve as a safe environment to try out and evaluate such approaches for the prevention and detection of attacks, e.g., to ensure that security measures do not impact the grid's availability or to assist in identifying optimal sensor placement for detection approaches. Indeed, WATTSON has already been applied to evaluate an intrusion detection approach based on discrete-time Markov chains [113]. In ongoing work, WATTSON provides the foundation to study how network intrusion detection in power grids can be enhanced with automated facility monitoring [89].

Ethical Considerations. We acknowledge potential risks arising from our work, especially for power grid operators. However, we are convinced that the advantages of evaluating cyberattacks for understanding their impact and deriving suitable countermeasures outweigh the risks. We further adopt multiple measures to minimize such risks: First, we do not develop entirely new cyberattacks against power grids but focus on existing ones. Second, we refrain from publishing any implementation of the conducted attacks to prevent misuse of our research. This is in line with the judgment of domain experts from a national cybersecurity agency and multiple grid operators with whom we discussed the performed attacks.

6. Conclusion

This paper addresses the challenge of comprehensively researching cybersecurity within power grids as a cyber-physical system. Our related work analysis reveals that existing research environments do not sufficiently cover the requirements for such a comprehensive methodology. Therefore, we introduce WATTSON, an open-source research environment for studying sophisticated cyberattacks against power grids of realistic sizes and, primarily, their impact on communication and physical processes. We validated WATTSON's accuracy against a physical testbed, showed its scalability using benchmarking scenarios and reference power grids, and expounded its flexibility and suitability for cybersecurity research.

Accordingly, we used WATTSON to recreate and analyze different attacks of increasing complexity against a realistic reference distribution grid [70], including the sophisticated Industroyer attack, which was used to attack Ukrainian power grids [21]. Our analysis of the impact of different attacks reveals that with increasing technical knowledge of the grid, attackers can shift from a locally limited attack impact to an active disruption of the power grid's operation and even cause widespread blackouts.

While such attacks can already be devastating, obfuscating their effects by manipulating measurements and interfering with countermeasures further aggravates their impact. WATTSON provides the necessary means to researchers and grid operators for analyzing and comprehending the combined impact of such advanced cyberattacks on the communication and physical side of power grids. Moving forward in securing power grids based on this knowledge, WATTSON offers a safe and risk-free environment to test countermeasures, validate their effectiveness, and ensure the absence of unintended side effects on reliable power grid operation.

Acknowledgments

We thank Martin Unkel for his indispensable work on and support for the implementation of the C++ Python bindings[3] for the IEC 60870-5-104 protocol. These bindings represent a vital component used in WATTSON's RTU and MTU reference implementations.

---

[3] https://github.com/fraunhofer-fit-coop/104-connector-python

References

[1] Abdullah Albarakati, Bassam Moussa, Mourad Debbabi, Amr Youssef, Basile L Agba, and Marthe Kassouf. OpenStack-Based Evaluation Framework for Smart Grid Cyber Security. In *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids*, 2018.

[2] Anas AlMajali, Arun Viswanathan, and Clifford Neuman. Analyzing Resiliency of the Smart Grid Communication Architectures under Cyber Attack. In *Workshop on Cyber Security Experimentation and Test (CSET 12)*. USENIX Association, August 2012.

[3] Bhagya Amarasekara, Chathurika Ranaweera, Ampalavanapillai Nirmalathas, and Rob Evans. Co-simulation Platform for Smart Grid Applications. In *IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, 2015.

[4] Daniele Antonioli and Nils Ole Tippenhauer. MiniCPS: A Toolkit for Security Research on CPS Networks. In *Workshop on Cyber-Physical Systems-Security and/or PrivaCy*. ACM, 2015.

[5] Rebecca Gurley Bace and Peter Mell. Intrusion Detection Systems. *NIST Special Publication on Intrusion Detection Systems*, 2001.

[6] Stephan Balduin, Martin Tröschel, and Sebastian Lehnhoff. Towards domain-specific surrogate models for smart grid co-simulation. *Energy Informatics*, 2(1), 2019.

[7] Dhananjay Bhor, Kavinkadhirselvan Angappan, and Krishna M Sivalingam. Network and power-grid co-simulation framework for Smart Grid wide-area monitoring networks. *Journal of Network and Computer Applications*, 59, 2016.

[8] Desong Bian, Murat Kuzlu, Manisa Pipattanasomporn, Saifur Rahman, and Yiming Wu. Real-time Co-simulation Platform using OPAL-RT and OPNET for Analyzing Smart Grid Performance. In *IEEE Power & Energy Society General Meeting*, 2015.

[9] Gianni Celli, Paolo Atillio Pegoraro, Fabrizio Pilo, Giuditta Pisano, and Sara Sulis. DMS Cyber-Physical Simulation for Assessing the Impact of State Estimation and Communication Media in Smart Grid Operation. *IEEE Transactions on Power Systems*, 29(5), 2014.

[10] Hao Chen, John A Clark, Siraj A Shaikh, Howard Chivers, and Philip Nobles. Optimising IDS Sensor Placement. In *International Conference on Availability, Reliability and Security (ARES)*. IEEE, 2010.

[11] Pin-Yu Chen, Shin-Ming Cheng, and Kwang-Cheng Chen. Smart Attacks in Smart Grid Communication Networks. *IEEE Communications Magazine*, 50(8), 2012.

[12] Anton Cherepanov. WIN32/INDUSTROYER: A New Threat for Industrial Control Systems. Technical report, 2017.

[13] Mehmet Hazar Cintuglu, Osama A. Mohammed, Kemal Akkaya, and A. Selcuk Uluagac. A Survey on Smart Grid Cyber-Physical System Testbeds. *IEEE Communications Surveys & Tutorials*, 19(1), 2017.

[14] Selim Ciraci, Jeff Daily, Jason Fuller, Andrew Fisher, Laurentiu Marinovici, and Khushbu Agarwal. FNCS: A Framework for Power System and Communication Networks Co-Simulation. In *Proceedings of the Symposium on Theory of Modeling & Simulation - DEVS Integrative*, DEVS '14, 2014.

[15] Athanasios Dagoumas. Assessing the Impact of Cybersecurity Attacks on Power Systems. *Energies*, 12(4):725, 2019.

[16] Evandro de Souza, Omid Ardakanian, and Ioanis Nikolaidis. A Co-simulation Platform for Evaluating Cyber Security and Control Applications in the Smart Grid. In *IEEE International Conference on Communications (ICC)*, 2020.

[17] Dieter Dohnal. On-load Tap-changers for Power Transformers. Technical report, Maschinenfabrik Reinhausen GmbH, 2013.

[18] Nan Duan, Nathan Yee, Benjamin Salazar, Jhi-Young Joo, Emma Stewart, and Ed Cortez. Cybersecurity Analysis of Distribution Grid Operation with Distributed Energy Resources via Co-Simulation. In *IEEE Power & Energy Society General Meeting (PESGM)*, 2020.

[19] Wesley Eddy. TCP SYN Flooding Attacks and Common Mitigations. Technical report, 2007.

[20] Erik Hjelmvik. Industroyer2 IEC-104 Analysis. https://www.netresec.com/?page=Blog&month=2022-04&post=Industroyer2-IEC-104-Analysis, 2022. Accessed at 2022-07-11.

[21] ESET Research. Industroyer2: Industroyer reloaded. https://www.welivesecurity.com/2022/04/12/industroyer2-industroyer-reloaded/, 2022. Accessed at 2022-07-11.

[22] Ertem Esiner, Daisuke Mashima, Binbin Chen, Zbigniew Kalbarczyk, and David Nicol. F-Pro: a Fast and Flexible Provenance-Aware Message Authentication Scheme for Smart Grid. In *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2019.

[23] Shibin Gao, Jieyu Lei, Jian Shi, Xiaoguang Wei, Ming Dong, and Zhu Han. Assessment of Overloading Correlations Among Transmission Lines Under Load Redistribution Attacks. *IEEE Transactions on Smart Grid*, 2021.

[24] Michele Garau, Gianni Celli, Emilio Ghiani, Fabrizio Pilo, and Sergio Corti. Evaluation of Smart Grid Communication Technologies with a Co-Simulation Platform. *IEEE Wireless Communications*, 24(2), 2017.

[25] Hanno Georg, Sven Christian Müller, Christian Rehtanz, and Christian Wietfeld. Analyzing Cyber-Physical Energy Systems:The INSPIRE Cosimulation of Power and ICT Systems Using HLA. *IEEE Transactions on Industrial Informatics*, 10(4), 2014.

[26] Uttam Ghosh, Pushpita Chatterjee, and Sachin Shetty. A Security Framework for SDN-enabled Smart Power Grids. In *International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, 2017.

[27] Tim Godfrey, Sara Mullen, David W. Griffith, Nada Golmie, Roger C. Dugan, and Craig Rodine. Modeling Smart Grid Applications with Co-Simulation. In *International Conference on Smart Grid Communications*. IEEE, 2010.

[28] Adam Hahn, Ben Kregel, Manimaran Govindarasu, Justin Fitzpatrick, Rafi Adnan, Siddharth Sridhar, and Michael Higdon. Development of the PowerCyber SCADA Security Testbed. In *Workshop on Cyber Security and Information Intelligence Research*. ACM, 2010.

[29] Christopher Hannon, Jiaqi Yan, and Dong Jin. DSSnet: A Smart Grid Modeling Platform Combining Electrical Power Distribution System Simulation and Software Defined Networking Emulation. In *ACM SIGSIM-PADS*, 2016.

[30] Christopher Hannon, Jiaqi Yan, Dong Jin, Chen Chen, and Jianhui Wang. Combining simulation and emulation systems for smart grid planning and evaluation. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 28(4):1–23, 2018.

[31] Jacob Hansen, Thomas Edgar, Jeff Daily, and Di Wu. Evaluating Transactive Controls of Integrated Transmission and Distribution Systems using the Framework for Network Co-Simulation. In *2017 American Control Conference (ACC)*, 2017.

[32] Md Abul Hasnat and Mahshid Rahnamay-Naeini. A Data-Driven Dynamic State Estimation for Smart Grids under DoS Attack using State Correlations. In *North American Power Symposium (NAPS)*, 2019.

[33] Martin Henze, Lennart Bader, Julian Filter, Olav Lamberts, Simon Ofner, and Dennis van der Velde. Poster: Cybersecurity Research and Training for Power Distribution Grids – A Blueprint. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, November 2020.

[34] Martin Henze, Benedikt Wolters, Roman Matzutt, Torsten Zimmermann, and Klaus Wehrle. Distributed Configuration, Authorization and Management in the Cloud-based Internet of Things. In *2017 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2017.

[35] Jens Hiller, Martin Henze, Martin Serror, Eric Wagner, Jan Niklas Richter, and Klaus Wehrle. Secure Low Latency Communication for Constrained Industrial IoT Scenarios. In *Proceedings of the 43rd IEEE Conference on Local Computer Networks (LCN)*, 2018.

[36] Junho Hong, Chen-Ching Liu, and Manimaran Govindarasu. Integrated Anomaly Detection for Cyber Security of the Substations. *IEEE Transactions on Smart Grid*, 5(4), 2014.

[37] Kenneth Hopkinson, Xiaoru Wang, Renan Giovanini, James Thorp, Ken Birman, and Denis Coury. EPOCHS: A Platform for Agent-Based Electric Power and Communication Simulation Built From Commercial Off-the-Shelf Components. *IEEE Transactions on Power Systems*, 21(2), 2006.

[38] Md Jakir Hossain and Mahshid Rahnamy-Naeini. Line Failure Detection from PMU Data after a Joint Cyber-Physical Attack. In *IEEE Power & Energy Society General Meeting*, 2019.

[39] Bing Huang, Alvaro A. Cardenas, and Ross Baldick. Not Everything is Dark and Gloomy: Power Grid Protections Against IoT Demand Attacks. In *USENIX Security Symposium*. USENIX Association, August 2019.

[40] Xiaoge Huang, Zhijun Qin, and Hui Liu. A Survey on Power Grid Cyber Security: From Component-Wise Vulnerability Assessment to System-Wide Impact Analysis. *IEEE Access*, 6, 2018.

[41] Takashi Irita and Toru Namerikawa. Detection of Replay Attack on Smart Grid with Code Signal and Bargaining Game. In *American Control Conference (ACC)*, 2017.

[42] Mark James, Adam McGovern, Justin Somelofske, Claire Valentine-Fossum, and Kristen Zweifel. Improving the Cyber Security of the Electric Distribution Grid – Identifying Obstacles and Presenting Best Practices for Enhanced Grid Security. Institute for Energy and the Environment, Vermont Law School, 2019.

[43] Ming Jin, Javad Lavaei, and Karl Henrik Johansson. Power Grid AC-Based State Estimation: Vulnerability Analysis Against Cyber Attacks. *IEEE Transactions on Automatic Control*, 64(5), 2019.

[44] Brian M. Kelley, Philip Top, Steven G. Smith, Carol S. Woodward, and Liang Min. A Federated Simulation Toolkit for Electric Power Grid and Communication Network Co-simulation. In *2015 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, 2015.

[45] Jinsub Kim and Lang Tong. On Topology Attack of a Smart Grid: Undetectable Attacks and Countermeasures. *IEEE Journal on Selected Areas in Communications*, 31(7), 2013.

[46] Jinsub Kim, Lang Tong, and Robert J. Thomas. Data Framing Attack on State Estimation. *IEEE Journal on Selected Areas in Communications*, 32(7), 2014.

[47] Tùng T. Kim and H. Vincent Poor. Strategic Protection Against Data Injection Attacks on Power Grids. *IEEE Transactions on Smart Grid*, 2(2), 2011.

[48] Benedikt Klaer, Ömer Sen, Dennis van der Velde, Immanuel Hacker, Michael Andres, and Martin Henze. Graph-based Model of Smart Grid Architectures. In *IEEE International Conference on Smart Energy Systems and Technologies (SEST)*, 2020.

[49] Oliver Kosut, Liyan Jia, Robert J. Thomas, and Lang Tong. Malicious Data Attacks on the Smart Grid. *IEEE Transactions on Smart Grid*, 2(4), 2011.

[50] Tim Krause, Raphael Ernst, Benedikt Klaer, Immanuel Hacker, and Martin Henze. Cybersecurity in Power Grids: Challenges and Opportunities. *Sensors*, 21(18), 2021.

[51] Bob Lantz, Brandon Heller, and Nick McKeown. A Network in a Laptop: Rapid Prototyping for Software-Defined Networks. In *ACM HotNets*, 2010.

[52] Robert M. Lee, Michael J. Assante, and Tim Conway. Analysis of the Cyber Attack on the Ukrainian Power Grid. *SANS Industrial Control Systems*, 2016.

[53] Hongwei Li, Rongxing Lu, Liang Zhou, Bo Yang, and Xuemin Shen. An Efficient Merkle-Tree-Based Authentication Scheme for Smart Grid. *IEEE Systems Journal*, 8(2), 2014.

[54] Weilin Li, Antonello Monti, Min Luo, and Roger A. Dougal. VPNET: A Co-simulation Framework for Analyzing Communication Channel Effects on Power Systems. In *IEEE Electric Ship Technologies Symposium*, 2011.

[55] Xin Li, Qiuyuan Huang, and Dapeng Wu. Distributed Large-Scale Co-Simulation for IoT-Aided Smart Grid Control. *IEEE Access*, 5, 2017.

[56] Gaoqi Liang, Junhua Zhao, Fengji Luo, Steven R. Weller, and Zhao Yang Dong. A Review of False Data Injection Attacks Against Modern Power Systems. *IEEE Transactions on Smart Grid*, 8(4), 2017.

[57] Vincenzo Liberatore and Ahmad Al-Hammouri. Smart Grid Communication and Co-Simulation. In *IEEE 2011 EnergyTech*, 2011.

[58] Hua Lin, Yi Deng, Sandeep Shukla, James Thorp, and Lamine Mili. Cyber Security Impacts on All-PMU State Estimator - A Case Study on Co-Simulation Platform GECO. In *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, 2012.

[59] Hua Lin, Santhosh S. Veda, Sandeep S. Shukla, Lamine Mili, and James Thorp. GECO: Global Event-Driven Co-Simulation Framework for Interconnected Power System and Communication Network. *IEEE Transactions on Smart Grid*, 3(3), 2012.

[60] Ren Liu and Anurag Srivastava. Integrated Simulation to Analyze the Impact of Cyber-Attacks on the Power Grid. In *Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, 2015.

[61] Ren Liu, Ceeman Vellaithurai, Saugata S. Biswas, Thoshitha T Gamage, and Anurag K Srivastava. Analyzing the Cyber-Physical Impact of Cyber Events on the Power Grid. *IEEE Transactions on Smart Grid*, 6(5), 2015.

[62] Shan Liu, Salman Mashayekh, Deepa Kundur, Takis Zourntos, and Karen Butler-Purry. A Framework for Modeling Cyber-Physical Switching Attacks in Smart Grid. *IEEE Transactions on Emerging Topics in Computing*, 1(2), 2013.

[63] Yao Liu, Peng Ning, and Michael K. Reiter. False Data Injection Attacks against State Estimation in Electric Power Grids. *ACM Transactions on Information and System Security*, 14(1), June 2011.

[64] Yining Liu, Chi Cheng, Tianlong Gu, Tao Jiang, and Xiangming Li. A Lightweight Authenticated Communication Scheme for Smart Grid. *IEEE Sensors Journal*, 16(3), 2016.

[65] Martin Lévesque, Da Qian Xu, Géza Joós, and Martin Maier. Co-Simulation of PEV Coordination Schemes over a FiWi Smart Grid Communications Infrastructure. In *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, 2012.

[66] Kelvin Mai, Xi Qin, Neil Ortiz, Jason Molina, and Alvaro A. Cardenas. Uncharted Networks: A First Measurement Study of the Bulk Power System. In *Proceedings of the ACM Internet Measurement Conference*, 2020.

[67] Kwok-Hong Mak and B.L. Holland. Migrating electrical power network SCADA systems to TCP/IP and Ethernet networking. *Power Engineering Journal*, 16(6), 2002.

[68] Malaz Mallouhi, Youssif Al-Nashif, Don Cox, Tejaswini Chadaga, and Salim Hariri. A Testbed for Analyzing Security of SCADA Control Systems (TASSCS). In *Innovative Smart Grid Technologies (ISGT)*, 2011.

[69] Michael McFail, Jordan Hanna, and Daniel Rebori-Carretero. Detection Engineering in Industrial Control Systems. Ukraine 2016 Attack: Sandworm Team and Industroyer Case Study. Technical report, MITRE CORP MCLEAN VA, 2022.

[70] Steffen Meinecke, Džanan Sarajlić, Simon Ruben Drauz, Annika Klettke, Lars-Peter Lauven, Christian Rehtanz, Albert Moser, and Martin Braun. SimBench—A Benchmark Dataset of Electric Power Systems to Compare Innovative Solutions Based on Power Flow Analysis. *Energies*, 13(12), 2020.

[71] Kevin Mets, Tom Verschueren, Chris Develder, Tine L Vandoorn, and Lieven Vandevelde. Integrated Simulation of Power and Communication Networks for Smart Grid Applications. In *2011 IEEE 16th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. IEEE, 2011.

[72] Peter Mihal, Martin Schvarcbacher, Bruno Rossi, and Tomáš Pitner. Smart grids co-simulations: Survey & research directions. *Sustainable Computing: Informatics and Systems*, 35:100726, 2022.

[73] Antonello Monti and Ferdinanda Ponci. Power Grids of the Future: Why Smart Means Complex. In *Complexity in Engineering*. IEEE, 2010.

[74] Andres Murillo, Riccardo Taormina, Nils Tippenhauer, and Stefano Galelli. Co-Simulating Physical Processes and Network Data for High-Fidelity Cyber-Security Experiments. In *Industrial Control System Security Workshop (ICSS)*, 2020.

[75] Pedro H. J. Nardelli, Nicolas Rubido, Chengwei Wang, Murilo S. Baptista, Carlos Pomalaza-Raez, Paulo Cardieri, and Matti Latva-aho. Models for the modern power grid. *The European Physical Journal Special Topics*, 223(12), 2014.

[76] James Nutaro. Designing power system simulators for the smart grid: combining controls, communications, and electro-mechanical dynamics. In *2011 IEEE Power and Energy Society General Meeting*, 2011.

[77] James Nutaro, Phani Teja Kuruganti, Laurie Miller, Sara Mullen, and Mallikarjun Shankar. Integrated Hybrid-Simulation of Electric Power and Communications Systems. In *IEEE Power Engineering Society General Meeting*, 2007.

[78] Bryan Palmintier, Dheepak Krishnamurthy, Philip Top, Steve Smith, Jeff Daily, and Jason Fuller. Design of the HELICS High-Performance Transmission-Distribution-Communication-Market Co-Simulation Framework. In *2017 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, 2017.

[79] Kaikai Pan, André Teixeira, Claudio David López, and Peter Palensky. Co-simulation for Cyber Security Analysis: Data Attacks against Energy Management System. In *Int'l Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2017.

[80] Georgios Patsakis, Deepak Rajan, Ignacio Aravena, Jennifer Rios, and Shmuel Oren. Optimal Black Start Allocation for Power System Restoration. *IEEE Transactions on Power Systems*, 33(6), 2018.

[81] L. Pavithra and D. Rekha. Prevention of Replay Attack for Isolated Smart Grid. In *Next Generation Information Processing System*. Springer Singapore, 2021.

[82] Manuel Peuster, Holger Karl, and Steven van Rossem. MeDICINE: Rapid Prototyping of Production-Ready Network Services in Multi-PoP Environments. In *IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, November 2016.

[83] Ben Pfaff, Justin Pettit, Teemu Koponen, Ethan Jackson, Andy Zhou, Jarno Rajahalme, Jesse Gross, Alex Wang, Joe Stringer, Pravin Shelar, Keith Amidon, and Martin Casado. The Design and Implementation of Open vSwitch. In *12th USENIX symposium on networked systems design and implementation (NSDI 15)*, 2015.

[84] Shiva Poudel, Zhen Ni, and Naresh Malla. Real-time Cyber Physical System Testbed for Power System Security and Control. *International Journal of Electrical Power & Energy Systems*, 90, 2017.

[85] Panagiotis Radoglou-Grammatikis, Panagiotis Sarigiannidis, Ioannis Giannoulakis, Emmanouil Kafetzakis, and Emmanouil Panaousis. Attacking IEC-60870-5-104 SCADA Systems. In *IEEE World Congress on Services*, volume 2642, 2019.

[86] Vivek Ramachandran and Sukumar Nandi. Detecting ARP Spoofing: An Active Technique. In *International Conference on Information Systems Security*. Springer, 2005.

[87] Danda B. Rawat and Chandra Bajracharya. Detection of False Data Injection Attacks in Smart Grid Communication Systems. *IEEE Signal Processing Letters*, 22(10), 2015.

[88] Ömer Sen, Dennis van der Velde, Philipp Linnartz, Immanuel Hacker, Martin Henze, Michael Andres, and Andreas Ulbig. Investigating Man-in-the-Middle-based False Data Injection in a Smart Grid Laboratory Environment. In *Proceedings of IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, 2021.

[89] Martin Serror, Lennart Bader, Martin Henze, Arne Schwarze, and Kai Nürnberger. Poster: INSIDE - Enhancing Network Intrusion Detection in Power Grids with Automated Facility Monitoring. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2022.

[90] Chong Shum, Wing-Hong Lau, Tian Mao, Henry Shu-Hung Chung, Kim-Fung Tsang, Norman Chung-Fai Tse, and Loi Lei Lai. Co-Simulation of Distributed Smart Grid Software Using Direct-Execution Simulation. *IEEE Access*, 6, 2018.

[91] Rezoan A. Shuvro, Zhuoyao Wang, Pankaz Das, Mahshid R. Naeini, and Majeed M. Hayat. Modeling Impact of Communication Network Failures on Power Grid Reliability. In *North American Power Symposium (NAPS)*. IEEE, 2017.

[92] Shachar Siboni, Vinay Sachidananda, Yair Meidan, Michael Bohadana, Yael Mathov, Suhas Bhairav, Asaf Shabtai, and Yuval Elovici. Security Testbed for Internet-of-Things Devices. *IEEE Transactions on Reliability*, 68(1), 2019.

[93] Mikhail Simonov, Fabrizio Bertone, and Klodiana Goga. Detecting the Manipulation of Demand via IoT. In *International Conference on Event-Based Control, Communication, and Signal Processing (EBCCSP)*, 2019.

[94] Saleh Soltan, Prateek Mittal, and H. Vincent Poor. BlackIoT: IoT Botnet of High Wattage Devices Can Disrupt the Power Grid. In *USENIX Security Symposium*. USENIX Association, August 2018.

[95] Saleh Soltan and Gil Zussman. Power Grid State Estimation after a Cyber-Physical Attack under the AC Power Flow Model. In *IEEE Power & Energy Society General Meeting*, 2017.

[96] Pirathayini Srikantha and Deepa Kundur. Denial of Service Attacks and Mitigation for Stability in Cyber-Enabled Power Grid. In *IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2015.

[97] Kai Strunz, Ehsan Abbasi, Chad Abbey, Christophe Andrieu, F Gao, T Gaunt, A Gole, N Hatziargyriou, and R Iravani. Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources. *Cigre Task Force C*, 6(04-02), 2014.

[98] Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. Cyber security of a power grid: State-of-the-Art. *International Journal of Electrical Power & Energy Systems*, 99, 2018.

[99] Eniye Tebekaemi and Duminda Wijesekera. Designing An IEC 61850 Based Power Distribution Substation Simulation/Emulation Testbed for Cyber-Physical Security Studies. In *Int'l Conference on Cyber-Technologies and Cyber-Systems*, 2016.

[100] Mini S. Thomas and John Douglas McDonald. *Power System SCADA and Smart Grids*. CRC press, 2017.

[101] Leon Thurner, Alexander Scheidler, Florian Schäfer, Jan-Hendrik Menke, Julian Dollichon, Friederike Meier, Steffen Meinecke, and Martin Braun. Pandapower – An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems. *IEEE Transactions on Power Systems*, 33(6), 2018.

[102] Leon M. Tolbert, Fred Wang, Kevin Tomsovic, Kai Sun, Jingxin Wang, Yiwei Ma, and Yunting Liu. Reconfigurable Real-Time Power Grid Emulator for Systems With High Penetration of Renewables. *IEEE Open Access Journal of Power and Energy*, 7, 2020.

[103] Thien-Toan Tran, Oh-Soon Shin, and Jong-Ho Lee. Detection of Replay Attacks in Smart Grid Systems. In *Int'l Conference on Computing, Management and Telecommunications (ComManTel)*, 2013.

[104] Dennis van der Velde, Martin Henze, Philipp Kathmann, Erik Wassermann, Michael Andres, Detert Bracht, Raphael Ernst, George Hallak, Benedikt Klaer, Philipp Linnartz, Benjamin Meyer, Simon Ofner, Tobias Pletzer, and Richard Sethmann. Methods for Actors in the Electric Power System to Prevent, Detect and React to ICT Attacks and Failures. In *Proceedings of the 6th IEEE International Energy Conference (ENERGYCon)*, 2020.

[105] Eric MSP Veith, Jawad Kazmi, and Stephan Balduin. Large-Scale Co-Simulation of Power Grid and Communication Network Models with Software in the Loop, 2020.

[106] John Walker. HotBits: Genuine random numbers, generated by radioactive decay. https://www.fourmilab.ch/hotbits/.

[107] Jiankang Wang, Gonzalo Constante, Christian Moya, and Junho Hong. Semantic analysis framework for protecting the power grid against monitoring-control attacks. *IET Cyber-Physical Systems: Theory & Applications*, 5(1), 2020.

[108] Lizhi Wang, Lynn Pepin, Yan Li, Fei Miao, Amir Herzberg, and Peng Zhang. Securing Power Distribution Grid Against Power Botnet Attacks. In *IEEE Power & Energy Society General Meeting*, 2019.

[109] Klaus Wehrle, Mesut Günes, and James Gross, editors. *Modeling and Tools for Network Simulation*. Springer Science & Business Media, 2010.

[110] Alexandre Gustavo Wermann, Marcelo Cardoso Bortolozzo, Eduardo Germano da Silva, Alberto Schaeffer-Filho, Luciano Paschoal Gaspary, and Marinho Barcellos. ASTORIA: A Framework for Attack Simulation and Evaluation in Smart Grids. In *IEEE/IFIP Network Operations and Management Symposium*, 2016.

[111] Sean Whalen. An Introduction to ARP Spoofing, 2001.

[112] David E. Whitehead, Kevin Owens, Dennis Gammel, and Jess Smith. Ukraine Cyber-Induced Power Outage: Analysis and Practical Mitigation Strategies. In *Conference for Protective Relay Engineers (CPRE)*, 2017.

[113] Konrad Wolsing, Eric Wagner, Antoine Saillard, and Martin Henze. IPAL: Breaking up Silos of Protocol-dependent and Domain-specific Industrial Intrusion Detection Systems. In *Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses (RAID)*, 2022.

[114] Dapeng Wu and Chi Zhou. Fault-Tolerant and Scalable Key Management for Smart Grid. *IEEE Transactions on Smart Grid*, 2(2), 2011.

[115] Syed Ali Raza Zaidi and Mounir Ghogho. Stochastic Geometric Analysis of Black Hole Attack on Smart Grid Communication Networks. In *IEEE Int'l Conference on Smart Grid Communications*, 2012.

[116] Sven Zemanek, Immanuel Hacker, Konrad Wolsing, Eric Wagner, Martin Henze, and Martin Serror. PowerDuck: A GOOSE Data Set of Cyberattacks in Substations. In *Proceedings of the 15th Workshop on Cyber Security Experimentation and Test*, CSET '22. ACM, 2022.

[117] Hang Zhang, Bo Liu, and Hongyu Wu. Smart Grid Cyber-Physical Attack and Defense: A Review. *IEEE Access*, 9:29641–29659, 2021.

[118] Zhenghao Zhang, Shuping Gong, Aleksandar D. Dimitrovski, and Husheng Li. Time Synchronization Attack in Smart Grid: Impact and Analysis. *IEEE Transactions on Smart Grid*, 4(1), 2013.

[119] Junfeng Zhao, Jing Wang, and Lei Yin. Detection and Control against Replay Attacks in Smart Grid. In *Int'l Conference on Computational Intelligence and Security (CIS)*, 2016.