Quantitative Risk Assessment of Cyber Attacks on Cyber-Physical Systems using Attack Graphs

Semertzis, Ioannis ; Subramaniam Rajkumar, Vetrivel; Stefanov, Alexandru; Fransen, Frank; Palensky, Peter

DOI
10.1109/MSCPES55116.2022.9770140

Publication date
2022

Document Version
Final published version

Published in
2022 10th Workshop on Modelling and Simulation of Cyber-Physical Energy Systems (MSCPES)

Quantitative Risk Assessment of Cyber Attacks on Cyber-Physical Systems using Attack Graphs

Ioannis Semertzis*, Vetrivel Subramaniam Rajkumar*, Alexandru Ştefanov*, Frank Fransen†, Peter Palensky*

*Department of Electrical Sustainable Energy
Delft University of Technology
Delft, the Netherlands
I.Semertzis@tudelft.nl

†Department of Cyber Security & Robustness
Netherlands Organization for Applied Scientific Research
Groningen, the Netherlands

2022 10th Workshop on Modelling and Simulation of Cyber-Physical Energy Systems (MSCPES) | 978-1-6654-6865-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/MSCPES55116.2022.9770140

Abstract — **Over the past decade, the number of cyber attack incidents targeting critical infrastructures such as the electrical power system has increased. To assess the risk of cyber attacks on the cyber-physical system, a holistic approach is needed that considers both system layers. However, the existing risk assessment methods are either qualitative in nature or employ probabilistic models to study the impact on only one system layer. Hence, in this work, we propose a quantitative risk assessment method for cyber-physical systems based on probabilistic and deterministic techniques. The former uses attack graphs to evaluate the attack likelihood, while the latter analyzes the potential cyber-physical impact. This is achieved through a dynamic cyber-physical power system model, i.e., digital twin, able to simulate power system cascading failures caused by cyber attacks. Additionally, we propose a domain-specific language to describe the assets of digital substations and thereby model the attack graphs. Using the proposed method, combined risk metrics are calculated that consider the likelihood and impact of cyber threat scenarios. The risk assessment is conducted using the IEEE 39-bus system, consisting of 27 user-defined digital substations. These substations serve as the backbone of the examined cyber system layer and as entry-points for the attackers. Results indicate that cyber attacks on specific substations can cause major cascading failures or even a blackout. Thereby, the proposed method identifies the most critical substations and assets that must be cyber secured.**

Keywords — attack graphs, cyber-physical systems, digital twin, cyber attacks, risk assessment

I. Introduction

Power systems rely on Operational Technology (OT) networks for real-time monitoring and control of the physical infrastructure. OT layers are coupled with the power grid forming an interdependent, complex Cyber-Physical System (CPS). Furthermore, the OT networks are integrated with Information Technology (IT) networks for non-operational functions. A direct consequence of this convergence is that traditionally segmented and air-gapped OT systems are interconnected with IT systems, which raises cyber security concerns [1].

Cyber attacks on power systems may cause severe disruptions, leading to power outages. The cyber attacks on the power grid in Ukraine in 2015 and 2016, show how malicious actors can disrupt the power system operation by gaining unauthorized access into the OT network through the corporate IT network [2], [3]. CPS is susceptible to cyber threats, as the existing OT equipment may have limited cyber security controls. Traditionally, unlike IT networks, the OT systems were not designed with cyber security considerations. Cyber security controls may conflict with the availability and real-time requirements of the OT systems.

The emergent threat of cyber attacks on power systems has prompted research into developing accurate CPS models and methods for impact analysis and risk assessment. Existing literature highlights that a holistic approach is needed to capture the complex interdependencies between the cyber and physical systems [4]. Many risk assessment methods reported in the literature include probabilistic analysis through Markov-chains [5] and Bayesian networks [6]. Monte-Carlo simulations are used to study the most critical attack scenarios [7]. Impact analysis is conducted based on qualitative factors such as equipment damage, employee health and safety, etc. [8], [9]. Attack graphs are typically based on generic probabilistic models. They are used to identify possible attack paths by examining interdependencies between the identified communication network vulnerabilities [10], [11], [12]. Therefore, existing risk assessment methods are either qualitative in nature or employ probabilistic models to study the impact on only one CPS layer. However, an accurate risk assessment of cyber attacks on critical infrastructures requires CPS domain-specific attack graphs, consideration of attackers' behavior, and impact analysis through quantitative criteria. Therefore, the contributions of this paper are summarized as follows:

1. We propose a quantitative risk assessment method for cyber-physical systems based on probabilistic and deterministic techniques. The former uses attack graphs to evaluate the attack likelihood through the Time-to-Compromise (TTC) and Mean-Time to Detect (MTTD) metrics. The latter quantifies the potential cyber-physical impact by computing impact indices and power system restoration factors. This is achieved through a dynamic cyber-physical power system model, i.e., digital twin, including various coordinated protection schemes for lines and generators and under frequency and under voltage load shedding. The digital twin computes system dynamics and simulates power system cascading failures caused by cyber attacks.
2. We propose a domain-specific language to describe the OT assets of digital substations and model the attack graphs. The generated attack graphs are used to calculate the overall TTC and quantify the likelihood of the examined attack scenarios.

The rest of this paper is organized as follows. Section II presents the cyber-physical system modelling. Section III describes the method for attack graph generation of digital

978-1-6654-6865-7/22/$31.00 ©2022 IEEE

substations. Section IV presents the quantitative risk assessment method. Section V presents the simulation results, while Section VI discusses the conclusions and future work.

II. Cyber-Physical System Modelling

The CPS model captures the dynamic behavior of an electrical power system. It is used to study the cascading failures due to cyber attacks. Both cyber-physical layers are modelled and co-simulated to develop a comprehensive and holistic CPS model. Thus, the interdependencies of the physical power system and OT networks are studied.

A. Power System Modelling

At the physical system layer, control schemes for generators are modelled to study the dynamical behavior of the power system, i.e., speed governors and Automatic Voltage Regulators (AVR). Multiple coordinated protection schemes are modelled for lines and generators that can disconnect power system elements during time domain simulations and lead to cascading failures. The interface protection for generators includes over/under voltage, over/under frequency, Rate of Change of Frequency (ROCOF), over flux, and out-of-step protection schemes. The settings are chosen based on national grid codes, as well as the IEEE C37.102 standard for generator protection [13]. Distance and overload protection are modelled for transmission lines. Load shedding schemes based on under frequency and under voltage conditions are considered. Time domain simulations are used to analyze power system stability and cascading effects.

B. Cyber System Modelling

The cyber system layer is represented in Fig. 1, which shows the connection between the control center and a substation from the station level down to the process level. The OTs comprise two types of packet-switching networks. The Local Area Network (LAN) of each digital substation and Wide Area Network (WAN) used for communication between the substations and control center. LAN consists of various OT devices, e.g., Intelligent Electronic Devices (IEDs), Merging Units (MUs), station control systems, network switches and routers. The LAN topology is based on vendor specifications [14]. The proposed WAN architecture is comprised of specific substations acting as data routing hubs between other substations and the control center. This is based on a decentralized design concept for WANs of CPS [15]. All measurement and control packets are communicated between substations and control center using TCP/IP.

III. Attack Graphs of Digital Substations

An attack graph is a representation of attackers' behavior on a specified network. It depicts the different attack paths an intruder can take to reach a target, by exploiting system vulnerabilities. In this paper, we utilize the Meta-Attack Language (MAL) to define OT domain-specific attack graphs [16]. MAL is a framework to create user-defined Domain-Specific Languages (DSL) and cyber threat models. In this research, we develop a probabilistic model for cyber security analysis that estimates the time-to-compromise for each attack step in the attack graph. TTC is defined as the number of days that an attacker needs to successfully conduct an attack step. The local TTC is linked to a single attack step, while the global TTC refers to the overall time needed to successfully compromise an OT target based on the entry point.

A. OT Asset Definitions and Attack Graph Generation

A new DSL, i.e., Substation-Lang, is developed to describe the OT assets, associations, and attack steps of digital substations and their interconnection to WAN. The attack graph model considers the substation equipment for monitoring, protection, and control. The attack graph is based on a simplified topological model of the OT network of digital substations, as specified in [17]. All OT assets and DSL elements are specified using the MAL syntax and connected as illustrated in Fig. 1: (i) *"WAN"* represents the wide area communication network between hub substations and control center, (ii) *"Gateway"* is the substation network router, (iii) *"OperatorConsole"* is the Human-Machine Interface (HMI), (iv) *"Controller"* represents the station control system, and (v) *"IED"* represents the bay level devices of the digital substation. The *"WAN"* and *"SubNetwork"* act as entry points for attackers, while the *"CircuitBreakers"*, *"LoadControl"*, and *"Generators"* DSL elements represent the final targets of attackers. The attack steps are based on the "MITRE ATT@CK for ICS" tactics [18]. The developed DSL enables attack scenario and cyber security studies.

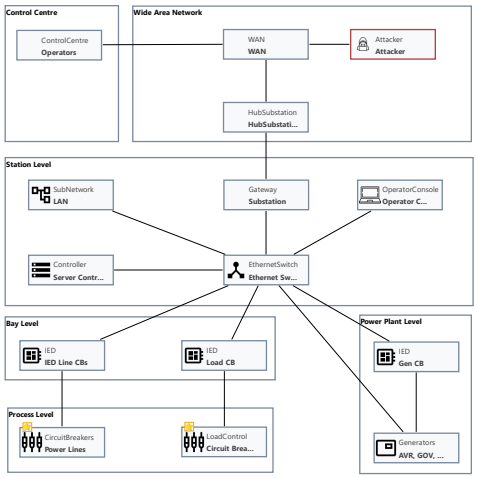


Fig. 1. Attack graph using an OT domain-specific language.

B. Calculation of Time-to-Compromise

The method to calculate the time-to-compromise of system components for quantitative risk estimation is introduced in [19]. It was applied for risk reduction of a Supervisory Control and Data Acquisition (SCADA) system, where vulnerable components are identified and patched. This method calculates a discrete TTC value for each attack step. In our work, we expand the method in [19] by calculating probability distributions to quantify the global time-to-compromise. The probability distributions capture a wide range of malicious actors targeting the CPS, whose skill levels may vary significantly. Attack steps are defined for each OT asset of the attack graph as presented in Table I. The National Vulnerability Database (NVD) [20] is used to identify the Common Vulnerabilities and Exposures (CVEs) of each OT asset. The known vulnerabilities of each asset are categorized based on the compromise type. The vulnerabilities identified in the substation OT equipment from major vendors are examined. Table I presents the number of CVEs categorized

per attack step. IEDs from multiple vendors are considered. Therefore, the number of known vulnerabilities varies per IED model. In this research, all IED models in a substation are assumed to be the same. No software patches are considered in the OT assets. Although security controls are not considered in this paper, firewalls and intrusion detection systems will be included in future work.

The proposed method to compute the local TTCs and probability distribution per attack step is presented in Fig. 2. The inputs are the number of known vulnerabilities V for each attack step of an OT asset, skill levels of attackers k given by a normal distribution, and the number of Monte-Carlo simulation samples S. By conducting one Monte-Carlo simulation, we calculate the local TTC of each sample s using the method in [19]. The local TTC per sample is computed using the attacker skill levels distribution. A histogram is generated based on the results of the Monte-Carlo simulations, i.e., all calculated local TTC values. The characteristics of the probability distribution per attack step are computed by performing a curve fitting on the histogram. The probability distribution characteristics serve as inputs for each attack step of an OT asset in the attack graph. The fitted probability distributions, based on the attack steps and number of CVEs, are shown in Table I.

TABLE I. OT ASSETS AND ATTACK STEPS OF THE ATTACK GRAPH.

Asset	Attack Step	No of CVEs	Fitted Distributions
Ethernet Switch	Discover Devices	1	$\mathbf{N}(9.2, 1.15^2)$
Operator Console	Command Line Interface	0	$\mathbf{N}(28, 1.91^2)$
Station Controller	Automated Collection	3	$\mathbf{\Gamma}(0.6, 0.1) + \mathbf{4}$
	Man in the Middle	4	$\mathbf{N}(4.1, 0.14^2)$
Gateway	Discover	2	$\mathbf{N}(5.3, 0.27^2)$
	Denial of Service	2	$\mathbf{N}(5.3, 0.27^2)$
	Connect	1	$\mathbf{N}(9.2, 1.15^2)$
IED	Denial of Service	1-8	Gamma/Normal
	Firmware Compromise	0-5	Normal

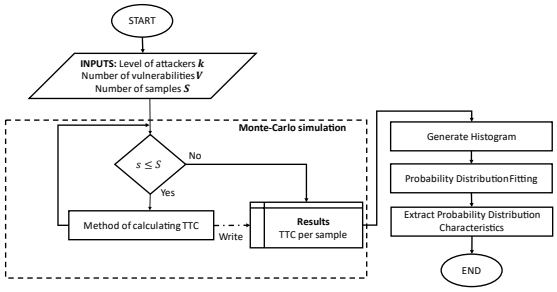


Fig. 2. Calculation of local TTC distribution parameters per attack step.

For the attack graph analysis, a new Monte-Carlo simulation is performed, considering all TTC probability distributions computed for all attack steps. For every sample of the attack graph analysis, different local TTCs are considered based on the calculated probability distributions. Therefore, all possible attack path combinations are generated. The global time-to-compromise of the target OT asset is calculated per simulation sample. The result of the attack graph analysis is a Cumulative Distribution Function (CDF). It accounts for the distribution of each individual TTC per Monte-Carlo simulation sample, based on Dijkstra's single-source shortest path algorithm [21]. The global time-to-compromise for an examined cyber attack scenario j is the average of all individual global TTCs. In the cyber attack scenarios with multiple targets, it is assumed that all attack paths are generated in parallel. Thus, the global TTC is calculated based on the maximum TTC of all targets.

IV. CYBER ATTACK RISK ASSESSMENT

The global TTC for a specific cyber attack scenario and a quantitative cyber-physical impact assessment are used to quantify the risk of a cyber attack on the cyber-physical system. The relation between risk, likelihood and impact on CPS is given in (1).

$$Risk(j) = Likelihood(j) \times \left(I_{Ph}(j) + I_{Cy}(j)\right) \times F_R(j) \quad (1)$$

where $Risk(j)$ is the calculated risk for attack scenario j, $Likelihood(j)$ describes the likelihood of success for the examined scenario, $I_{Ph}(j)$ is the quantitative impact on power system operation, $I_{Cy}(j)$ is the impact on the modelled communication network, and $F_R(j)$ is a factor proposed for quantifying the power system restoration efforts.

A. Likelihood of a Cyber Attack Scenario

The likelihood of a successful cyber attack depends on the time to compromise the target OT assets. In addition to the global TTC, we also consider the Mean-Time to Detect (MTTD), which is a performance indicator defined by cyber security experts [22]. It is assumed that MTTD is a constant, describing the average time needed by cyber security teams to successfully detect an intrusion. Hence, the likelihood of an attack scenario j is formulated in (2).

$$Likelihood(j) = \frac{MTTD}{TTC_{avg}(j) + MTTD} \quad (2)$$

where $TTC_{avg}(j)$ is the average global TTC defined by the attack graph analysis. The likelihood function is in the range of [0,1]. If $TTC_{avg}(j) \ll MTTD$ the likelihood is close to 1. If an OT asset is cyber secure and it cannot be compromised by an attacker, then $Likelihood(j) \to 0$. Defining a method to calculate MTTD is beyond the scope of this paper, and it is assumed to be 14 days. This constant can vary depending on how an organization assess its detection capabilities, taking into account the complexity of the OT infrastructure.

B. Impact Assessment on Physical Power System

The states of the power system before and after the cyber attack are analyzed to compute the overall impact on the power system operation. The physical impact index is computed in (3).

$$I_{Ph}(j) = w_l \times I_L + w_f \times I_{Fr} + w_c \times I_C + w_v \times I_V \quad (3)$$

where w_l, w_f, w_c, and w_v are empirically chosen weighting factors for I_{Ph} to be in the range of [0, 100]. The functions I_L, I_{Fr}, I_V, and I_C assess the impact on the power system using a digital twin. The loss of load and voltage deviation indices are given in (4) and (5) [10].

$$I_L(j) = \sum_{i=1}^{N_{Loads}} \frac{\Delta P_{Load,i}(j)}{P_{init,i}(j)} \quad (4)$$

$$I_V(j) = \frac{1}{N_{Buses}} \sum_{i=1}^{N_{Buses}} \frac{|\Delta V_i(j)|}{\Delta V_{allowed}} \quad (5)$$

where $P_{init,i}$ is the power consumption of load i before the cyber attack, $\Delta P_{Load,i}$ is the loss of load, ΔV_i is the difference between the initial and final bus voltage magnitudes, and $\Delta V_{allowed}$ is the permissible bus voltage deviation.

Power system islanding due to cascading failures is considered in the impact analysis. An algorithm is used to identify the power system islands after the cyber attacks based on the measured generator frequencies.

ALGORITHM: Identification of power system islands

Inputs:

$f_{gen} = [f_{gen,i} \mid i = 1,2,...,N_{gen}]$: Measured frequency of generating units

ε : Measurement error

Outputs:

N_{subnet} : Number of detected islands

$n_{subnet} = [n_{subnet,k} \mid k = 1,2,...,N_{subnet}]$: Number of generators per island

$f_{subnet} = [f_{subnet,k} \mid k = 1,2,...,N_{subnet}]$: Operating frequency per island

Initialize $N_{subnet} = 0$, $n_{subnet} = \varnothing$, $f_{subnet} = \varnothing$

For $\forall f_{gen,i}$

  $N_{subnet} = N_{subnet} + 1$, $n_{subnet}[N_{subnet}] = 1$, $f_{subnet}[N_{subnet}] = f_{gen,i}$

  For $(\forall f_{gen,j} \mid j \neq i)$

    If $|f_{gen,i} - f_{gen,j}| \leq \varepsilon$

      $n_{subnet}[N_{subnet}] = n_{subnet} + 1$ **AND** delete $f_{gen,j}$ element

Return N_{subnet} , n_{subnet} , f_{subnet}

The frequency deviation index is computed in (6).

$$I_{Fr}(j) = \frac{1}{N_{Generators}} \left(\sum_{i=1}^{N_{Subnet}} \frac{|\Delta f_{subnet,i}(j)|}{\Delta f_{allowed}} \times n_{subnet,i} \right) \quad (6)$$

where $\Delta f_{subnet,i}$ is the frequency deviation of each detected island, $\Delta f_{allowed}$ is the permissible frequency deviation and $n_{subnet,i}$ is the number of generators that belong to the same power system island. The index for the number of disconnected power system components is calculated in (7).

$$I_C(j) = \frac{1}{N_{branch}} \left(\sum_{i=1}^{N_{Lines}} d_{Lines,i}(j) + \sum_{i=1}^{N_{Trafo}} d_{Trafo,i}(j) \right) \quad (7)$$

where $d_{lines,i}$ and $d_{Trafo,i}$ are binary status indicators of lines and transformers, respectively, i.e., the status is 1 if the component is disconnected, while N_{branch} is the total number of power system branches.

C. Impact Assessment on Cyber System Layer

Latencies of data packets originating from the control center to OT devices at the substation bay level are used to assess the impact of cyber attacks on the communication network of CPS. Latency is increased by specific attacks affecting the OT network traffic, such as Denial-of-Service (DoS). The impact index for the cyber system layer is computed in (8).

$$I_{Cy}(j) = \sum_{i=1}^{N_{substations}} max \left(0, log \left(\frac{RTT_{avg,i}(j)}{t_{margin}} \right) \right) \quad (8)$$

where $RTT_{avg,i}$ is the average Round-Trip Time (RTT) of data packets for substation i and t_{margin} is the minimum acceptable latency, which is typically in the range of hundreds of milliseconds [23].

D. Power System Restoration Factor

Power system restoration is a multi-stage, complex optimization problem, where the restoration time depends on providing cranking power to non-black start units and gradual load pickup. A power system restoration factor is defined to quantify the effort required to restore the power system following cyber attacks. It considers the disconnection of generating units and generation capacity and type. The proposed restoration factor is given in (9).

$$F_R(j) = exp \left(\left(\sum_{i=1}^{N_{gen}} \frac{a_i(j) \times P_{nom,i}}{P_{tot}} \right) \times T \right) \quad (9)$$

where $P_{nom,i}$ is the nominal capacity of generator i, P_{tot} is the total installed capacity of the power system, a_i is the circuit breaker status of generator i, i.e., 1 if circuit breaker is open, and T is the power system restoration index. T is initialized by 0 and it is calculated based on the type of the disconnected generators.

$$T = \max_{0 \leq i \leq N_{gen}} (T, a_i(j) \times T_i) \quad (10)$$

where T_i is the restoration index of generator i. The restoration procedures for the disconnected generating units commence in parallel. The power system restoration index is given by the maximum generator restoration index. Generator units with black start capabilities, e.g., hydro power plants, have a restoration index of 0.5, while thermal power plants have an increased index of 0.8. The interconnectors with the neighboring power grid have a restoration index of 1 as the re-synchronization procedure can only be started after the power system restoration is completed.

V. Simulation Results

The cyber-physical system is co-simulated using DIgSILENT PowerFactory and Mininet, as shown in Fig. 3. IEEE 39-bus dynamic model is implemented at the physical layer. Multiple protection schemes for generators and lines are modelled and coordinated, i.e., over/under voltage, over/under frequency, ROCOF, over flux, out-of-step, distance, and overload protection. Under frequency and under voltage load shedding schemes are implemented for loads. The protection settings are defined based on national grid codes and IEEE C37.102 [13]. Time domain simulations are computed in real-time to analyze system dynamics and cascading failures initiated by cyber attacks. The cyber system model consists of substation LANs and WAN emulated in Mininet. Software-defined networking is used to emulate 27 substations consisting of network routers, Ethernet switches, HMIs, and hosts. The CPS network traffic is emulated in Mininet, while open-source tools, e.g., Wireshark and hping3, are used for packet monitoring and analysis. Power system measurements and control setpoints are exchanged between the physical and cyber system simulators in real-time using the Open Platform Communications Unified Architecture (OPC UA). The merging units and IEDs modelled in Mininet, communicate the measurements to the control center using TCP/IP, without implementing any specific power system communication standard. The attack graphs of digital substations and WAN are modelled in securiCAD using the defined OT domain-specific language. They are used for the calculation of global TTC per attack scenario. Simulation results from DIgSILENT PowerFactory, Mininet, and securiCAD are exported in CSV data format. The risk assessment is performed using Python 3.8. The assumptions considered in this study are: (i) system operators do not implement remedial actions to mitigate the impact of cyber attacks, (ii) skill level of attackers k is set at expert level in the normal distribution function, i.e., $\mathbf{N}(0.85, 0.04^2)$, and (iii) time needed by attackers to develop

additional tactics and techniques during the cyber attack is not considered.

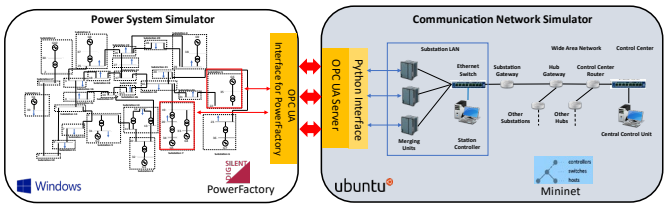


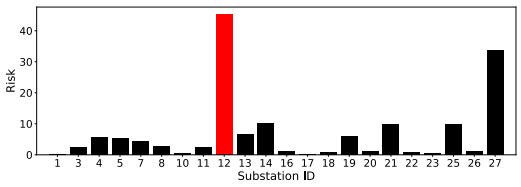
Fig. 3. Cyber-physical system co-simulation setup.

A. *Cyber Attacks with Different Switching Sequences*

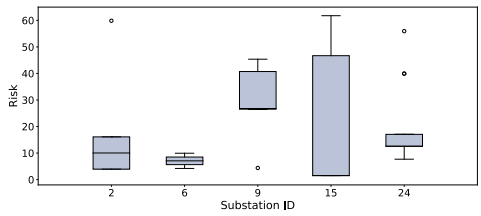
In this scenario, a cyber attack is conducted on each substation independently to disconnect all circuits from the power system. Circuit breakers can be maliciously opened, by compromising the data integrity of the IEDs [11]. The line circuit breakers of each substation are opened with a time interval of 2 seconds between the circuit breaker trip commands. All permutations of switching sequences are studied per substation. In total 248 dynamic simulations are performed. In this scenario, the attackers send a single malicious packet per line circuit breaker. Therefore, the cyber system layer is not affected by the cyber attack in terms of latency ($I_{Cy} = 0$). The highest risk indices calculated per switching sequence are presented in Table II.

TABLE II. CRITICAL SUBSTATIONS PER SWITCHING SEQUENCE.

Target Substation	Switching Sequence	TTC_{avg} (days)	*Likelihood*	I_{Ph}	F_R	*Risk*
15	2-1-3	15	0.48	92.5	1.39	61.7
2	3-1-2	16	0.47	91.6	1.39	59.8
24	4-3-5-1-2	20	0.41	60.6	2.25	55.9
12	1-2	26	0.35	93.3	1.39	45.4



(a)



(b)

Fig. 4. Risk assessment: (a) no variation of risk index and (b) with variation of risk index depending on the circuit breaker opening sequence.

The calculated risk indices for all substations are shown in Fig. 4. The substations for which the risk index is not affected by the switching sequence are given in Fig. 4 (a), while the substations with a varying risk index per switching sequence are shown in Fig. 4 (b). Substation 12 is considered critical as the impact remains high, i.e., 93.3, regardless the switching sequence with a constant risk index of 45.4. Substation 15 has a higher risk index of 61.7 for only two out of six switching sequence scenarios, while the other four scenarios have risk indices lower than 2.0. Therefore, substation 15 is considered less critical than substation 12 as the attackers must have knowledge of the precise attack switching sequence. On overall, out of the 248 cases, 8 are assessed as critical, i.e., $Risk(j) > 40.0$, while 19 cases of switching sequences resulted in major physical impact, i.e., $I_{Ph}(j) > 60.0$.

B. *Coordinated Cyber Attacks on Multiple Substations*

In the second scenario, coordinated cyber attacks are conducted on substations 5 and 7. The potential attack vectors and impact of coordinated cyber attacks on multiple locations of a power system are discussed in [24]. The attack graph and paths to reach the OT targets are presented in Fig. 5. The attacker's entry point is assumed to be in the LAN of substation 5. Attackers discover and compromise the gateway router of hub substation 7 by accessing the WAN via the substation 5 router. OT assets in substations 5 and 7 are compromised, i.e., operator console, Ethernet switch, station control system, and IEDs. The cyber attack is executed in three steps: (i) the voltage setpoint of generator G6 is manipulated at 5s, (ii) the circuit breaker of line 19-16 is opened at 10s, and (iii) a DoS attack is launched on the gateway router of the hub substation 7 at 10s. DoS affects the communication between the control center and all substations connected through the hub, impeding the monitoring and control of substations 5, 6, 7, 21, 24, and 25. The latencies are increased. The average round-trip time ranges between 230-530 ms, which is higher than the accepted limit of 100 ms.

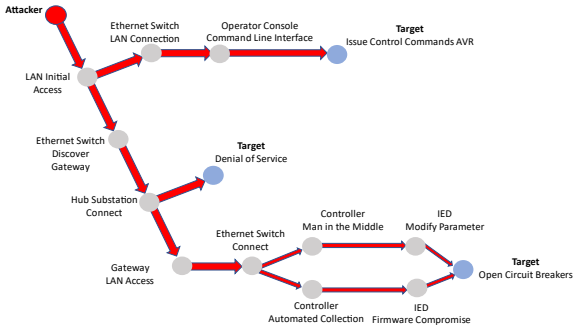


Fig. 5. Attack paths for cyber attack targeting two substations.

TABLE III. SEQUENCE OF CASCADING EVENTS FOR SCENARIO B.

Time (seconds)	Event
0	Start of simulation
5	Cyber attack on substation 5. Voltage setpoint of generator G6 is increased from 1.05 to 1.9 p.u.
10	Cyber attack on hub substation 7. Circuit breaker of line 19-16 is opened. DoS attack is launched on router
10.5	Generator G5 is tripped due to ROCOF protection (ROCOF settings: 2 Hz/s over 500ms)
11.278 -12.775	Multiple lines in vicinity of attack locations are disconnected by zone 3 of distance protection
12.858 -13.275	Generators G6 and G7 are disconnected due to over voltage and ROCOF protection (over voltage settings: 1.5 p.u. over 0.083s, i.e., 5 cycles)
14.044 -14.293	Under frequency load shedding is activated
14.295 -17.556	Additional lines are disconnected due to distance protection. Two islands are formed
18.116	Generator G9 is disconnected due to ROCOF protection
29.308 -29.549	Interconnector (G1) and generators G2 and G3 are disconnected due to over frequency protection. (Over frequency settings: 61.8 *Hz* over 5s)

These cyber attacks have a major impact on power system operation, causing cascading failures. The sequence of cascading events is presented in Table III. The protection schemes disconnect line 22-23, interconnector represented by

G1, and generators G3, G6, and G9 as shown in Fig. 6 (a) – (d). The likelihood of the cyber attack scenario is 32%, while the physical and cyber impact indices are computed as $I_{Ph} = 71.52$ and $I_{Cy} = 3.33$, respectively. The impact on the cyber layer is relatively small as only a limited area of the overall communication system is affected by the DoS attack. However, the overall risk is evaluated as high, i.e., $Risk = 56.28$, mainly due to the physical impact I_{Ph} and power system restoration factor of $F_R = 2.35$. The coordinated cyber attacks on the two substations result in cascading failures and a blackout, where 80% of the system generation capacity is lost.

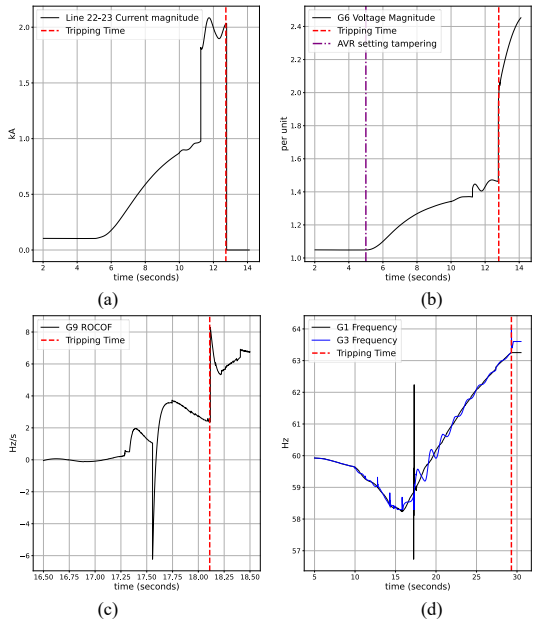


Fig. 6. Response of line and generator protection schemes: (a) distance protection, (b) over voltage, (c) ROCOF, and (d) over frequency.

VI. Conclusions and Future Work

In this paper, a quantitative risk assessment method for cyber attacks on cyber-physical systems is proposed using probabilistic and deterministic techniques. First, a holistic model of the cyber-physical system is developed. The CPS digital twin computes system dynamics and simulates power system cascading failures caused by cyber attacks. Second, OT domain-specific attack graphs are used to compute the time to compromise OT targets. The risk is assessed by calculating the likelihood of a cyber attack to succeed and impact on both cyber-physical system layers. The critical substations per cyber attack scenario are identified. The attack graphs analysis specifies the vulnerable OT systems that must be cyber secured. In future work, risk reduction techniques will be studied considering the implementation of cyber security controls and remedial actions of system operators. Furthermore, we foresee that digital twins can be deployed for system operations to assess in real-time the potential impact of cyber attacks based on the current power system state.

References

[1] B.W. Tuinema, J.L. Rueda Torres, A. Stefanov,F.M. Gonzalez, and M.A.M.M. van der Meijden, "Cyber-Physical System Modelling for Assessment and Enhancement of Power Grid Cyber Security, Resilience and Reliability," in *Probabilistic Reliability Analysis of Power Systems: A Student's Introduction*, pp. 237-270, Springer, 2020.

[2] R.M. Lee, M. J. Assante, and T. Conway, "Analysis of the cyber-attack on the Ukrainian power grid," *E-ISAC white paper*, SANS - Industrial Control Systems, 18 Mar. 2016.

[3] D.E. Whitehead, K. Owens, D. Gammel, and J. Smith, "Ukraine cyber-induced power outage: Analysis and practical mitigation strategies," in *Proc. 70th Annual Conference for Protective Relay Engineers (CPRE)*, Apr. 2017, pp. 1-8.

[4] R.V. Yohanandhan, R.M. Elavarasan, P. Manoharan, and L. Mihet-Popa, "Cyber-Physical Power System (CPPS): A Review on Modeling, Simulation, and Analysis with Cyber Security Applications," *IEEE Access*, vol. 8, pp. 151019-151064, Aug. 2020.

[5] L. Pièrre-Cambacédès, and M. Bouissou, ''Attack and defense modeling with BDMP,'' in *Proc. Int. Conf. Math. Methods, Models, Archit. Comput. Netw. Secur. Cham*, Switzerland: Springer, 2010, pp. 86–101.

[6] K. Huang, C. Zhou, Y-C. Tian S. Yang, and Y. Qin, "Assessing the Physical Impact of Cyberattacks on Industrial Cyber-Physical Systems," *IEEE Trans. Ind. Electron.*, vol. 65, no. 10, pp. 8153-8162, Oct. 2018.

[7] W. Wang, A. Cammi, F. Di Maio, S. Lorenzi, and E. Zio, ''A Monte Carlo-based exploration framework for identifying components vulnerable to cyber threats in nuclear power plants,'' in *Rel. Eng. Syst. Saf.*, vol. 175, pp. 24–37, Jul. 2018.

[8] W. Wu, R. Kang, and Z. Li, "Risk assessment method for cybersecurity of cyber-physical systems based on inter-dependency of vulnerabilities," in *Proc. IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore, 2015, pp. 1618-1622.

[9] I. Zografopoulos, J. Ospina, X. Liu, and C. Konstantinou, "Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies," *IEEE Access*, vol. 9, pp. 29775-29818, Febr. 2021.

[10] A. Stefanov, C.C. Liu, M. Govindarasu, and S.S. Wu, "SCADA modeling for performance and vulnerability assessment of integrated cyber physical systems," *Int. Trans Electr Energy Sys*t, vol. 25, no. 3, pp. 498-519, Dec. 2013.

[11] Y. Zhang, L. Wang, Y. Xiang, and C.-W. Ten, "Power System Reliability Evaluation with SCADA Cybersecurity Considerations," *IEEE Trans. Smart Grid,* vol. 6, no. 4, pp. 1707-1721, Jul. 2015.

[12] A. Ghazo, "A2G2V: Automatic Attack Graph Generation and Visualization and Its Applications to Computer and SCADA Networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 10, pp. 3488-3498, 2020.

[13] *IEEE Guide for AC Generator Protection*, IEEE Standard C37.102-2006 (Revison of IEEE Std C37.102-1995), Nov. 2006, pp.1-177.

[14] ABB (2019). *Centralized Protection and Control-White Paper* [Online].Available: https://library.e.abb.com/public/6b20916a4d2e412daabb76fbada1268e/Centralized_Protection_and_Control_White_paper_2NGA000256_LRENA.pdf

[15] Y. Wang, P. Yemula, and A. Bose, "Decentralized Communication and Control Systems for Power System Operation," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 885-893, Mar. 2015.

[16] P. Johnson, R. Lagerström, and M. Ekstedt, "A Meta Language for Threat Modelling and Attack Simulations," in *Proc. 13th International Conference on Availability, Reliability and Security (ARES)*, New York, USA, 2018.

[17] Siemens AG (2017). *Power Engineering Guide-Edition 8.0* [Online]. Available: https://assets.new.siemens.com/siemens/assets/api/uuid:5bfb815b0db95760272f17c1329cc56c0c402686/peg8-final-160812.pdf

[18] O. Alexander, M. Belisle, and J. Steele (2020). *MITRE ATT&CK® for Industrial Control Systems: Design and Philosophy* [Online]. Available: https://collaborate.mitre.org/attackics/img_auth.php/3/37/ATT%26CK_for_ICS_-_Philosophy_Paper.pdf

[19] M. McQueen, W. F. Boyer, M. A. Flynn, and G. A. Beitel, "Time-to-Compromise Model for Cyber Risk Reduction Estimation," in *Proc. 1st Workshop Qual. Prot.*, Milan, Italy, Sep. 2005, pp. 49-64.

[20] NIST (2021). *National Vulnerability Database* [Online]. Available: https://nvd.nist.gov/vuln

[21] *securiCAD Professional.* (2014) Foreseeti Inc. Accessed: Sep. 28, 2021. [Online]. Available: https://docs.foreseeti.com/docs/simulation-logic

[22] S. I. Pérez, S. Moral-Rubio, and R. Criado, "A new approach to combine multiplex networks and time series attributes: Building intrusion detection systems (IDS) in cybersecurity," *Chaos, Solitons and Fractals*, vol. 150, Sep. 2021.

[23] M. Kuzlu, M. Pipattanasomporn, and S. Rahman, "Communication network requirements for major smart grid applications in HAN, NAN and WAN," *Comp. Netw.*, vol. 67, pp. 74-88, Jul. 2014.

[24] S. Liu, B. Chen, T. Zourntos, D. Kundur, and K. Butler-Purry, "A Coordinated Multi-Switch Attack for Cascading Failures in Smart Grid," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1183-1195, May 2014.