

Risk-based Selection of Mitigation Strategies for Cybersecurity of  
Electric Power Systems

May 23, 2019

## **Abstract**

Electric power systems extensively rely on cyber physical systems to control physical components through cyber-based commands. Thus, the vulnerability to cyber threats requires an efficient allocation of resources to mitigate the risk of attacks. Common practices guide the selection of mitigation actions by prioritizing the cyber threat scenarios through a qualitative assessment. These practices can result in sub-optimal allocations of resources to protect the system. To overcome these drawbacks, we quantify the risk of cyber threats to the system through a comprehensive analysis of the system vulnerabilities. This analysis relies on Bayesian networks, which provide a solid framework for probabilistic risk assessment by representing cyber threat scenarios as combinations of cascading events. In addition, we develop an optimization model to determine the non-dominated mitigation strategies to protect the system from cyber threats. Specifically, the minimization of the risk of cyber threats supports the selection of mitigation actions, considering budget and technical constraints. The optimization model provides additional insight into risk management at different budget levels.

**Keywords:** Cyber physical systems, Cybersecurity, Electric power grids, Risk management, Multi-objective optimization.

# 1 Introduction

Cyber physical systems are physical systems in which operations are integrated, monitored and controlled through multi-core processors [1]. Such systems are increasingly employed in a wide range of industries, including electric power industry. Despite the substantial benefits to our society, the rapid proliferation of cyber physical systems also provides potential attackers with new opportunities to disrupt critical infrastructures [2].

Costly impacts can result from such attacks, for instance a cyber attack in 2015 caused the power outage of 225000 customers in Ukraine that lasted up to six hours. In that occasion, the operators at the three operations centers were unable to regain remote control of more than 50 substations affected by the incident. After the loss of over 130MW of load, the operators restored power by sending technicians to the substations and manually controlling the power system [3]. Besides critical infrastructures, cyber threats may affect all kind of institutions with potentially severe and costly impacts worldwide. For instance, the Petya and WannaCry cyber-attacks hit thousands of companies across the globe in 2017 [4]. Other relevant cases include the Stuxnet attack in 2010 to target an uranium enrichment centrifuge in Iran [5] and the attack on a German steel mill in 2014 to take over the plant control systems [6]. In recent years, cyber attacks have increased dramatically in terms of quantity, diversity and sophistication with significant economic losses [7].

These episodes prove the need for an effective deployment of security measures to mitigate the risk of cyber threats. Poolsappasit et al. [8] develop a mitigation strategy based on the likelihood of cyber attacks. A genetic algorithm supports the selection of a subset of mitigation actions by minimizing the cost of deployment and the expected damage to the system. Shameli-Sendi et al. [9] propose a dynamic framework for selecting optimal countermeasures to mitigate attacks. The selection is based on minimizing the cost of deployment and the impact on users and services. However, these optimization models do not consider the multiple impacts deriving from the cyber threat scenarios. Instead, mitigation strategies are selected on the cost and performance of individual actions. The resulting resource allocation could be sub-optimal for the cyber physical systems due to the lack of modeling the multiple impacts of cyber attacks [10]. Thus, the efficient allocation of resources to secure cyber physical systems involves challenges that we address in this paper.

Specifically, this paper fits into the first two functions of the National Institute of Standards and Technology (NIST) cybersecurity framework [11] in *detecting* system vulnerabilities and *protecting* the system from cybersecurity incidents. The NIST cybersecurity framework sets broadly accepted guidelines to improve the security of cyber physical systems. In this framework, we propose a methodology for the risk assessment of cyber threats based on a comprehensive analysis of the system vulnerabilities. This

methodology relies on Bayesian networks that model a probabilistic representation of combinations of events, possibly leading to severe outcomes. This model responds to the need for intuitive and computationally efficient methods for risk analysis, combining expert judgment and statistical analyses for the quantitative assessment of risks [12].

The proposed methodology leads to select the optimal portfolios of mitigation actions, based on the minimization of the risk of multiple impacts of cyber attacks. In particular, this paper focuses on mitigation strategies for protecting electric power systems, yet the framework has broader applications on cyber physical systems. Recently, the Electric Power Research Institute (EPRI) analyzed the cybersecurity failure scenarios and impacts for the electric sector [13]. The report provides insights on cybersecurity risks and potential mitigation actions to support risk assessment and resource allocation. Among applications on electric power systems, Ciapessoni et al. [14] propose a methodology to assess the security of such systems by analyzing the vulnerabilities to natural and human threats. On the other hand, Shelar and Amin [15] formulate a game theoretic framework to assess the security of an electricity distribution network, based on which the defender optimizes the security strategy of the network nodes.

In this paper, Section 2 reviews the practice proposed by the EPRI to select appropriate mitigation actions for the electric power system. Specifically, we present a critical analysis of the ranking procedure of individual cyber threats, which could lead to inefficient or unfeasible allocations for the system. This problem is addressed in Section 3, which provides an alternative to the EPRI practice by evaluating portfolios of mitigation actions to protect the cyber physical system against multiple cyber threat scenarios. In addition, an optimization model supports the selection of the mitigation strategies that minimize the expected impacts of cyber attacks, based on financial and technical constraints. Section 4 illustrates the methodology by analyzing the cyber threat scenarios concerning the Advanced Metering Infrastructure (AMI) of an electric power system. Section 5 discusses the potential and limits of the proposed framework, suggesting possible ways to overcome some inconveniences. Finally, Section 6 concludes the paper and outlines extensions for future research.

## 2 Analysis of the EPRI practice

Cybersecurity management calls for an extensive analysis of the system vulnerabilities, which leads to an efficient allocation of resources to protect the electric power system. In particular, the EPRI proposes the analysis of individual cyber threat scenarios based on attack graphs, multi-leveled diagrams describing threats on cyber physical systems and possible attacks to realize such threats [16]. Attack graphs are increasingly being applied to computer control systems, especially related to electric power systems, but they have also been used to analyze threats to physical systems [17]. Figure 1 illustrates the graphical

notation of two attack graphs, where a cyber threat scenario is represented through sequences of events (shown as diamonds) leading to the possible impacts of the cyber attack (shown as ellipses). The impacts of the cyber attack occur if a combination of events of the cyber threat scenario has proven to be successful, based on the binary representation of AND and OR gates (shown as solid and dashed lines, respectively). Attack graphs represents cyber threat scenarios, which are evaluated based on the *likelihood* of occurrence and *impact*. According to the EPRI analyses, the likelihood depends on 5 criteria whereas the impact depends on 15 criteria which are reported in Tables 3 and 4, respectively. These tables also report the EPRI scoring system for quantifying the likelihood and impact of cyber threat scenarios. Each score is an integer value in the range 0 – 9, thus the likelihood and impact are computed by summing the scores over the respective criteria. However, this scoring system can be questioned on the meaningfulness of the 0 – 9 scale. For instance, is a public accessible asset three times more accessible than a fenced asset with standard locks and nine times more accessible than a guarded/monitored asset? Furthermore, the additive model can be questioned on the sum of scores across different criteria. For instance, is “Public safety concern” comparable to “Long term economic damage”?

Mapping the likelihood and impact of all cyber threat scenarios in a risk matrix [18] makes it possible to rank the priority of individual cyber threats. This procedure clusters each cyber threat into *High, Medium* or *Low* likelihood and *High, Medium* or *Low* impact in order to prioritize the selection of mitigation actions. Specifically, cyber threats with *High* likelihood and *High* impact deserve the highest priority in the choice of mitigation actions, whereas priority decreases for cyber threats with lower likelihood and/or impact until the budget is depleted.

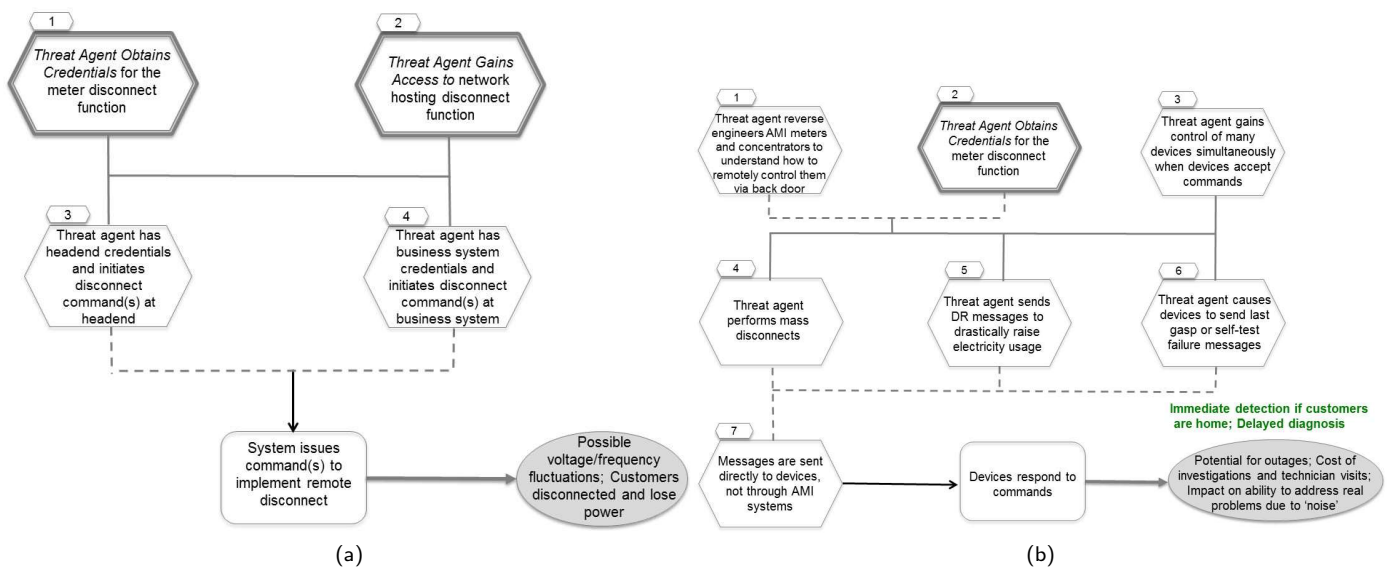


Figure 1: Attack graphs for (a) “Invalid disconnect messages to meters impact customers and utility” and (b) “Reverse engineering of AMI equipment allows unauthorized mass control” [13].

Despite the intuitive appeal and simplicity, risk matrices do not necessarily recommend effective risk management decisions, instead they may lead to incorrect risk prioritization [19,20]. Thus, the sequential choices of mitigation actions may result in a sub-optimal resource allocation because they are based on an incorrect prioritization of cyber threats. Furthermore, this procedure does not consider technical and budget constraints across different scenarios. In conclusion, the EPRI practice presents several inconsistencies in assessing the risk of cyber threats and supporting the selection of mitigation actions.

### 3 Bayesian framework

We propose a Bayesian framework, which provides an alternative to the EPRI practice for the risk assessment of cyber threats and the risk-based selection of mitigation strategies. In particular, the proposed risk assessment is based on a comprehensive analysis of multiple cyber threats that can affect the cyber physical system [21]. The framework also includes an optimization model for determining non-dominated mitigation strategies in order to protect the system from cyber threats. Specifically, an optimization algorithm computes the portfolios of mitigation actions that minimize the expected impacts of cyber attacks, considering budget and technical constraints.

#### 3.1 From attack graphs to Bayesian network

In contrast to the EPRI analysis of individual cyber threat scenarios, the Bayesian framework relies on a comprehensive analysis of multiple attack graphs [22]. Each attack graph represents a single cyber threat scenario, however some events could be equivalent among different attack graphs. For instance, in Figures 1a and 1b the event “Threat agent obtains credentials for the meter disconnect function” is equivalent among both attack graphs. For this reason, multiple attack graphs can be integrated into a directed acyclic graph by combining the corresponding events into single nodes. This integration leads to a comprehensive representation of cyber threat scenarios that overviews the alternative opportunities to attack the system [23].

This directed acyclic graph can be converted into a Bayesian network [24], a probabilistic graphical model that consists of:

- *chance nodes* (shown as circles) representing the random events of cyber threat scenarios;
- *value nodes* (shown as diamonds) representing the possible impacts of the cyber attacks;
- *arcs* (shown as directed edges) indicating causal dependencies between nodes.

Specifically, chance nodes are connected by arcs to represent combinations of events leading to the respective final impacts [25]. In this framework, the combinations of events indicate the stages of cyber

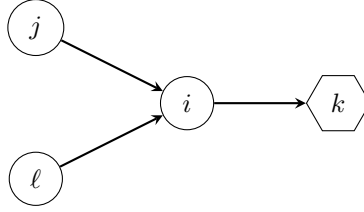


Figure 2: Example of a Bayesian network.

threat scenarios, whereas the final impacts indicate the possible outcomes of the cyber attacks. Arcs connects the nodes to represent the causal dependencies between the events of the attack graph. Figure 2 illustrates a Bayesian network, where each chance node represents a random event that encodes a finite set of discrete states, including a state of *No occurrence* of the event. Bayesian networks typically consider discrete states, nevertheless it is possible to include continuous variables under specific conditions [24]. Statistical analyses and expert judgment provide information to define the probability distributions of events that do not depend on any other chance node (nodes  $j$  and  $l$  in Figure 2). For events that show causal dependencies on other chance nodes through directed arcs (node  $i$  in Figure 2), the probabilistic representation is based on the state of the events they are depending on. Thus, it is necessary to define conditional probability tables for such nodes. Following the binary representation of attack graphs, conditional probability tables are derived from the information provided by AND and OR gates. Specifically, if the event  $i$  depends on the events  $j$  and  $l$  through AND(OR) gates, then the occurrence probability of the event  $i$  is 1 if the events  $j$  and(or)  $l$  occur as well, and 0 otherwise. For illustrative purposes, we consider the event “Threat agent performs mass disconnects” in Figure 1b as an example throughout the paper. Table 1 displays the conditional probability table of the event, based on the binary representation of *Occurrence* and *No occurrence* of the dependent events. The conditional probability table is derived from the information provided by AND and OR gates of the attack graph in Figure 1b.

Bayesian networks represent the events such that the occurrence probability is not necessarily limited to 0 and 1, but it is a real value in the set  $[0, 1]$ . This model leads to a more realistic representation of the stages of cyber threat scenarios, in contrast to the binary representation. Table 2 displays the conditional probability table of the event “Threat agent performs mass disconnects”, based on the multiple states of the events. This conditional probability table is not meant to represent any actual electric power system. According to the EPRI analyses, the occurrence probability of each event depends on (i) skill required, (ii) physical accessibility, (iii) logical accessibility and (iv) attack vector. In particular, the occurrence probability increases by enhancing the accessibility to equipment and information, while it decreases by requiring specialized knowledge and technical means to pursue the cyber threat.

The cascading events of the cyber threat scenarios finally lead to the possible impacts, assessed according to a set of criteria represented by the set  $K$  of value nodes [26]. The EPRI lists 14 possible impact criteria

Table 1: Conditional Probability Table Based on Binary States.

Threat agent reverse engineers AMI equipment	Threat agent obtains credentials	Threat agent gains control of devices	Threat agent performs mass disconnects	
			<i>Occurrence</i>	<i>No occurrence</i>
<i>Occurrence</i>	<i>Occurrence</i>	<i>Occurrence</i>	1	0
		<i>No occurrence</i>	0	1
	<i>No occurrence</i>	<i>Occurrence</i>	1	0
		<i>No occurrence</i>	0	1
<i>No occurrence</i>	<i>Occurrence</i>	<i>Occurrence</i>	1	0
		<i>No occurrence</i>	0	1
	<i>No occurrence</i>	<i>Occurrence</i>	0	1
		<i>No occurrence</i>	0	1

Table 2: Conditional Probability Table Based on Multiple States.

Threat agent reverse engineers AMI equipment	Threat agent obtains credentials	Threat agent gains control of devices	Threat agent performs mass disconnects [MW]			
			<i>No occurrence</i>	(0 50]	(50 100]	> 100
<i>Occurrence</i>	<i>Occurrence</i>	<i>None</i>	1	0	0	0
		<i>Few</i>	0.6	0.4	0	0
		<i>Moderate</i>	0.4	0.2	0.4	0
		<i>High</i>	0.3	0.1	0.2	0.4
	<i>No occurrence</i>	<i>None</i>	1	0	0	0
		<i>Few</i>	0.6	0.4	0	0
		<i>Moderate</i>	0.4	0.2	0.4	0
		<i>High</i>	0.3	0.1	0.2	0.4
<i>No occurrence</i>	<i>Occurrence</i>	<i>None</i>	1	0	0	0
		<i>Few</i>	0.6	0.4	0	0
		<i>Moderate</i>	0.4	0.2	0.4	0
		<i>High</i>	0.3	0.1	0.2	0.4
	<i>No occurrence</i>	<i>None</i>	1	0	0	0
		<i>Few</i>	1	0	0	0
		<i>Moderate</i>	1	0	0	0
		<i>High</i>	1	0	0	0

of cyber attacks on electric power systems, including financial, safety and service impacts. As a result, each value node of the Bayesian network represents a single impact criterion  $k$ , whose score depends on the state of events leading to that specific outcome. Ideally, the scores should be evaluated by a specific scale that reflects its unit of measure.

### 3.2 Probabilistic risk assessment

The probabilistic risk assessment of cyber threats is based on the computation of the expected impact for every impact criteria. Each chance node  $i$  represents a random event that encodes a finite set  $S_i$  of discrete states, including a state of *No occurrence* of the event. In particular, the occurrence probability of events that show causal dependencies relies on the occurrence probability of the events they depend on. For this reason, we define  $\Delta_i$  as the set of all possible combinations of states of the chance nodes



affecting the event  $i$ , such that

$$\Delta_i = \prod_{j|(j,i) \in E} S_j, \quad (1)$$

where  $E$  denotes the set of all arcs.

Let the random variable  $X_i$  represent the probability distribution of event  $i$  over the states  $s_i \in S_i$ . Then,  $\hat{\mathbf{X}}_i$  is a vector of random events on which  $X_i$  directly depends, meaning the vector of random variables  $X_j$  for all nodes  $j$  such that  $(j, i) \in E$ . For the d-separation property of Bayesian networks [24], the probability that the events affecting the event  $i$  meets a specific combination of states  $\delta_i \in \Delta_i$  is

$$P[\hat{\mathbf{X}}_i = \delta_i] = \prod_{s_j \in \delta_i} P[X_j = s_j] \quad \forall \delta_i \in \Delta_i. \quad (2)$$

Thus, the occurrence probability of the event  $i$  is computed by the *law of total probability* as the weighted average of the posterior probabilities across all  $\delta_i \in \Delta_i$ , such that

$$P[X_i = s_i] = \sum_{\delta_i \in \Delta_i} P[X_i = s_i | \hat{\mathbf{X}}_i = \delta_i] P[\hat{\mathbf{X}}_i = \delta_i]. \quad (3)$$

Because the occurrence probabilities are computed recursively, it is necessary to start the computation from the initial events throughout the dependent events of the cyber threat scenarios. The risk of cyber threats is then evaluated as the expected impact of the scenarios for each criterion  $k \in K$ , such that

$$E[V_k] = \sum_{\delta_k \in \Delta_k} P[\hat{\mathbf{X}}_k = \delta_k] V_k[\hat{\mathbf{X}}_k = \delta_k], \quad (4)$$

where  $V_k[\hat{\mathbf{X}}_k = \delta_k]$  is the score of the impact criterion  $k$  depending on the combination of states  $\delta_k$  of the events leading to that specific impact.

The expected impacts can be significantly reduced by deploying mitigation actions on the cyber physical system. Specifically, the mitigation actions affect the occurrence probability of one or multiple events in the cyber threat scenarios. In Bayesian networks, decision nodes (shown as squares) represent the choice of mitigation actions, as illustrated in Figure 3. Each arc directed from a decision node to a chance node

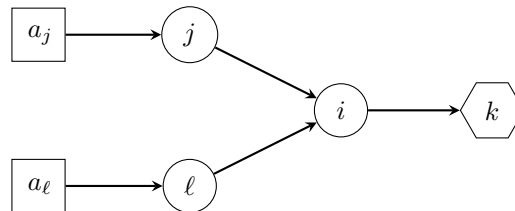


Figure 3: Example of a Bayesian network with decision nodes.

indicates that the deployment of the mitigation action affects the occurrence probability of the event represented by the chance node. Because this paper focuses on system design, the decision nodes do not depend on any event (no incoming arcs). Future research will focus on system control with decision nodes depending on other events.

Mitigation actions are numbered  $a \in \{1, 2, \dots, Ng\}$ , such that the binary variable  $z_a$  indicates the deployment of the mitigation action  $a$ . Specifically, the binary variable is  $z_a = 1$  for the deployment of the mitigation action  $a$  and  $z_a = 0$  otherwise. Thus, a portfolio is defined by the binary vector  $\mathbf{z}$  as a combination of binary variables  $z_a$  for all the possible mitigation actions. With no loss of generality, the vector  $\mathbf{z}$  lists binary variables such that

$$\mathbf{z} = [z_1, z_2, \dots, z_N]. \quad (5)$$

The deployment of mitigation actions reduces the occurrence probabilities of affected events. Bayesian networks compute probability updates of the cascading events throughout the cyber threat scenarios by the *law of total probability*, such that

$$P[X_i = s_i/\mathbf{z}] = \sum_{\delta_i \in \mathcal{I}_i} P[X_i = s_i/\hat{\mathbf{X}}_i = \delta_i] P[\hat{\mathbf{X}}_i = \delta_i/\mathbf{z}]. \quad (6)$$

Thus, the risk of cyber threats depends on the portfolio  $\mathbf{z}$  so that the expected impact of each criterion  $k \in K$  is

$$E[V_k](\mathbf{z}) = \sum_{\delta_k \in \mathcal{I}_k} P[\hat{\mathbf{X}}_k = \delta_k/\mathbf{z}] V_k[\hat{\mathbf{X}}_k = \delta_k]. \quad (7)$$

This framework aims to compute the risk of cyber threats for each impact criterion, making it possible to select mitigation strategies based on the minimization of the expected impacts.

### 3.3 Optimization model

The risk-based selection of mitigation strategies is performed through a multi-objective optimization model. Unlike the EPRI practice, the selection of mitigation actions is not based on the additive model of scores across different impact criteria. Instead, our optimization model determines the portfolios of mitigation actions that minimize the risk of cyber threats for every impact criteria. The selection is based on the analysis of expected impacts derived from the deployment of different mitigation strategies, so that the optimization model determines the portfolios that fulfill the Pareto condition

$$\mathbf{z} \succeq \mathbf{z}^{(j)} \iff \begin{cases} E[V_k](\mathbf{z}) \leq E[V_k](\mathbf{z}^{(j)}) & \text{for all } k \\ E[V_k](\mathbf{z}) < E[V_k](\mathbf{z}^{(j)}) & \text{for some } k \end{cases}. \quad (8)$$

This condition states that portfolio  $\mathbf{z}^*$  dominates  $\mathbf{z}$  if it reduces the risk of cyber threats for any impact criterion without increasing the risk for other impact criteria.

In addition to the Pareto condition, the optimal mitigation strategies need to fulfill budget and technical constraints. Budget constraints specify the financial feasibility of the deployment of a mitigation strategy. Each mitigation action  $a$  is associated to a cost  $c_a$ , thus the overall cost of portfolio  $\mathbf{z}$  must not exceed the budget  $B$  such that

$$\sum_a z_a c_a \leq B. \quad (9)$$

Technical constraints specify the properties of the system, such as mutually exclusive or mutually inclusive conditions of mitigation actions. For instance in Figure 3, the linear constraints

$$z_{a_j} + z_{a_\ell} \leq 1 \quad (10)$$

$$z_{a_j} - z_{a_\ell} = 0 \quad (11)$$

indicate that mitigation actions  $a_j$  and  $a_\ell$  cannot be deployed together or they must be deployed together, respectively.

Technical constraints also include risk acceptability limits that are represented by non-linear inequalities. In particular, specific regulatory conditions may apply to some events of the cyber threat scenarios. For such event  $i$ , the subset  $\tilde{S}_i \subseteq S_i$  includes the critical states whose occurrence probability must not exceed a risk acceptability threshold  $\epsilon_i$ , such that

$$\sum_{s_i \in \tilde{S}_i} \mathbb{P}[X_i = s_i | \mathbf{z}] \leq \epsilon_i. \quad (12)$$

Risk acceptability thresholds are usually provided by regulatory offices or internal company policies. Feasible portfolios belong to the set  $\mathbf{Z}_F$ , which includes all binary vector  $\mathbf{z}$  that fulfill linear and non-linear constraints. Then, the set of non-dominated solutions consists of the feasible portfolios that fulfill the Pareto condition for any other feasible portfolio, meaning that

$$\mathbf{Z}_{ND} = \{ \mathbf{z} \in \mathbf{Z}_F \mid \nexists \mathbf{z}' \in \mathbf{Z}_F \text{ such that } \mathbf{z}' \prec \mathbf{z} \}. \quad (13)$$

Generally, the set of non-dominated portfolios can include multiple alternative solutions, so the selection of a single mitigation strategy is not straightforward. For this reason, it is necessary to support the selection of the optimal mitigation strategy through additional analyses. A possible approach is the computation of the *core index* of each mitigation action. Analogously to Liesiö et al. [27], the core index

$CI(a)$  is defined as the fraction of non-dominated portfolios that include the mitigation action  $a$ , such that

$$CI(a) = \frac{|\{j \in \mathbf{Z}_{ND} \mid z_a = 1\}|}{|\mathbf{Z}_{ND}|}. \quad (14)$$

The analysis of the core indexes helps determine the mitigation actions that should be selected or rejected. If the core index of a mitigation action is 1, that measure is included in all non-dominated portfolios; on the other hand, if the core index is 0, that measure is not included in any non-dominated portfolio. Finally, mitigation actions whose core index is in the range  $(0, 1)$  require further analyses in order to be selected or rejected.

An implicit enumeration algorithm computes the set of non-dominated portfolios that minimize the risk of cyber threats over the impact criteria. The algorithm is an adaptation of Liesiö [28] and has been proposed by Mancuso et al. [29] for multi-objective optimization. This optimization algorithm is computationally efficient but it may be time consuming for a large amount of mitigation actions (over 40). In this case, evolutionary algorithms are a possible alternative to approximate non-dominated solutions for a lower computational time [30].

## 4 Case study

We illustrate the potential of the Bayesian framework by optimizing the selection of mitigation strategies for the Advanced Metering Infrastructure (AMI) of an electric power system. AMI systems have raised many security concerns since they connect traditionally self-contained power system operations with unreliable customer sites that are widely dispersed. The deployment of AMI systems is introducing millions of components to the electric grid that support two-way communication for next-generation grid applications. Although these systems can increase operational efficiency and enable new capabilities such as demand-response, they also increase the attack opportunity for potential adversaries. For this reason, electric power companies must address these new cybersecurity risks as part of their risk management strategy.

Information about AMI systems is provided by the National Electric Sector Cybersecurity Organization Resource (NESCOR), a program funded by the U.S. Department of Energy to protect electric power systems from cybersecurity incidents, both malicious and non-malicious. The NESCOR document “Electric Sector Failure Scenarios and Impact Analyses” [16] provides short descriptions of approximately 125 cyber threat scenarios in seven domains of the electric sector: *Advanced Metering Infrastructure*, *Distributed Energy Resources*, *Wide Area Monitoring, Protection and Control*, *Electric Transportation*, *Demand Response* and *Distribution Grid Management*. Furthermore, the NESCOR document “Analysis of Selected

Electric Sector High Risk Failure Scenarios” [13] presents the analyses of a selection of these cyber threat scenarios. Specifically, each analysis includes an attack graph that details the logical dependencies of events leading to a successful cyber attack. In addition to the attack graph, several of the analyses also provide a detailed description of each scenario.

Based on the NESCOR analyses, we select 8 cyber threats with the highest priority for AMI systems, in particular: invalid disconnect messages to meters impact customers and utility; reverse engineering of AMI equipment allows unauthorized mass control; threat agent obtains credentials for system or function; threat agent uses social engineering; threat agent gains access to network; threat agent exfiltrates data; authorized employee brings malware into system or network; threat agent exploits firewall gap.

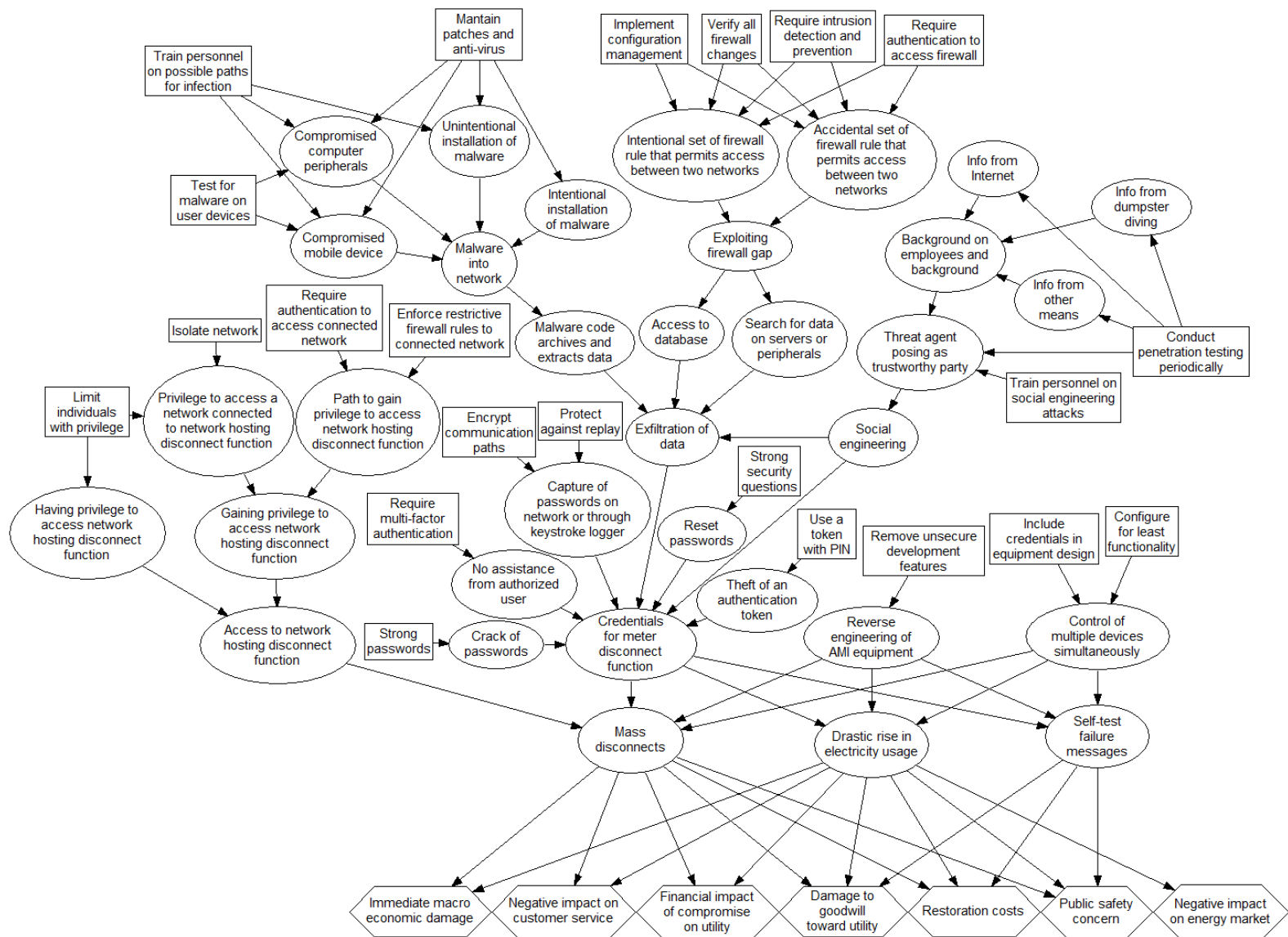


Figure 4: Bayesian network for selected cyber threat scenarios to the Advanced Metering Infrastructure of an electric power system.

These cyber threats potentially lead to “Threat agent performs mass disconnects”, “Threat agent sends demand-response messages to drastically raise electricity usage” and “Threat agent causes devices to send last gasps or self-test failure messages”, which indicate the possible outcomes of cyber attacks to the AMI systems. In Figure 4, the Bayesian network is based on the attack graphs of the 8 cyber threat scenarios to represent the alternative opportunities to attack the system. In particular, the circles represent the events of the cyber threat scenarios, the diamonds indicate the possible impacts of cyber attacks whereas the squares show mitigation actions that could be deployed for protecting the AMI system from cyber threats. Note that the event “Threat agent obtains credentials for the meter disconnect function” is equivalent among both cyber threat scenarios in Figures 1a and 1b. For this reason, this event has been represented by one chance node named “Credentials for meter disconnect function” in the Bayesian network. Reducing redundancies of equivalent events in multiple cyber threat scenarios facilitates the comprehensive analysis of cyber threats as a Bayesian network. In addition, the events “Threat agent has headend credentials and initiates disconnect(s) at headend” and “Threat agent has business system credentials and initiates disconnect(s) at business system” in Figure 1a are not considered in the Bayesian network because it is sufficient that the threat agent gains access to the network hosting the meter disconnect function and obtains the relative credentials to cause “Possible voltage/frequency fluctuations with disconnected customers”.

In the Bayesian network, probability distributions of the chance nodes have been set according to information provided by the NESCOR documents. For instance, the psychological manipulation (social engineering) of an employee may be expensive and it could lead to a public disclosure if the attempt fails, which summarizes to a low occurrence probability. However, such information is not sufficient to specif-

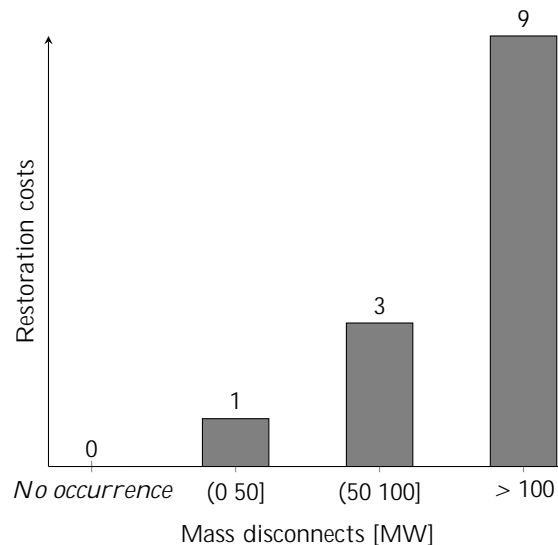


Figure 5: Illustrative impact scores for “Restoration costs”.

ically quantify the occurrence probability. For this reason, the occurrence probabilities of this example are not meant to be representative of any existing AMI system, but they are illustrative values to prove the viability of the Bayesian framework. The value nodes list the impact criteria of Table 4, in particular the ones affected by a possible cyber attack to the AMI systems. For illustrative purposes, impacts  $V_k$  have been quantified based on the scoring system of Table 4 due to the lack of detailed information in literature. For instance, the event “Threat agent performs mass disconnects” is quantified in different states of mass disconnects: *No occurrence*,  $(0\ 50]$ MW,  $(50\ 100]$ MW,  $> 100$ MW. Thus, each value node maps the impact score depending on the states of that event, as illustrated in Figure 5 for the impact criterion “Restoration costs”. Note that the impacts of cyber threats are not necessarily evaluated by every criteria of Table 4. For instance, the event “Threat agent performs mass disconnects” does not affect the impact criterion “Loss of privacy” for any state.

Assuming that event  $i$  is “Threat agent performs mass disconnects” and the value node represents the impact criterion “Restoration costs” in Figure 2, the expected impact of “Restoration costs” [ $k = RC$ ] is the weighted average of the impact scores for every state in Figure 5, such that

$$E[V_{RC}] = P[X_i = 50MW] V_{RC}[X_i = 50MW] + \dots + P[X_i = 100MW] V_{RC}[X_i = 100MW]. \quad (15)$$

The NESCOR documents also list possible mitigation actions that could be deployed to protect the AMI systems from cyber threats, specifying the events affected by each mitigation action. The deployment of a mitigation action affects the occurrence probability of the cyber threats according to the effect of the action. In particular, this case study accounts for 22 possible mitigation actions, which lead to  $2^{22}$

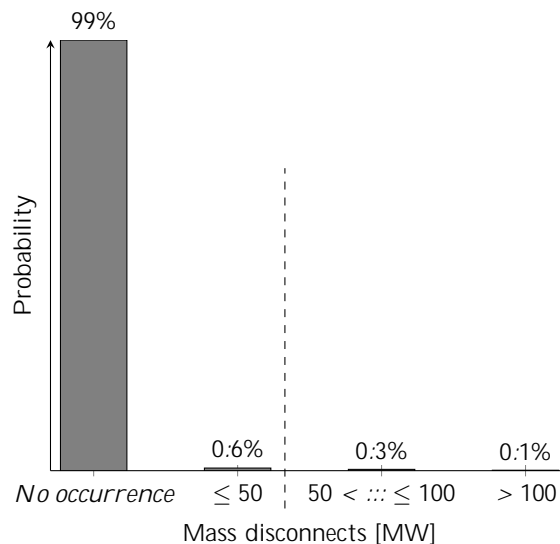


Figure 6: Illustrative probability distribution for mass disconnects.

mitigation strategies. Tables 5-10 list the 22 mitigation actions for the selected cyber threat scenarios, specifying the affected events based on the NESCOR analyses. The first column of the tables lists the index of the action in the portfolio  $\mathbf{z}$ , whereas the third column lists the cost of each mitigation action. The illustrative costs of mitigation actions aim to include a budget constraint to the optimization model. In addition, the optimization model includes a technical constraint on the risk acceptability of mass disconnects above 50MW. Figure 6 illustrates the probability distribution of the event “Threat agent performs mass disconnects” deriving from the deployment of a generic portfolio  $\mathbf{z}$ . Assuming that experts set the risk acceptability threshold to 0.5%, then the occurrence probability of the critical states must fulfill the constraint

$$\mathbb{P}[X_i > 50MW|\mathbf{z}] \leq 0.5\%. \quad (16)$$

The results of the multi-objective optimization show a decrease of the risk of every impact criteria by increasing the budget level. Figure 7 shows that larger budgets lead to more effective mitigation strategies to reduce the risk of every impact criteria. In this case study, the risk profiles of some impact criteria are overlapping because the impact scores are based on the same 0 – 9 scale that limit the quantification of the impacts. The analysis of the risk profiles supports the definition of the optimal budget by selecting the budget level above which the risk converges for every impact criteria, such as  $B = 400$  in this example. Computational time is around one hour on a regular laptop, however it depends on the constraints limiting the set of feasible portfolios. For instance, relaxing the budget constraint leads to higher computational time because the algorithm considers a larger set of feasible portfolios. In Figure 7, the risk profiles consider all the non-dominated portfolios selected by the optimization algorithm for each budget level.

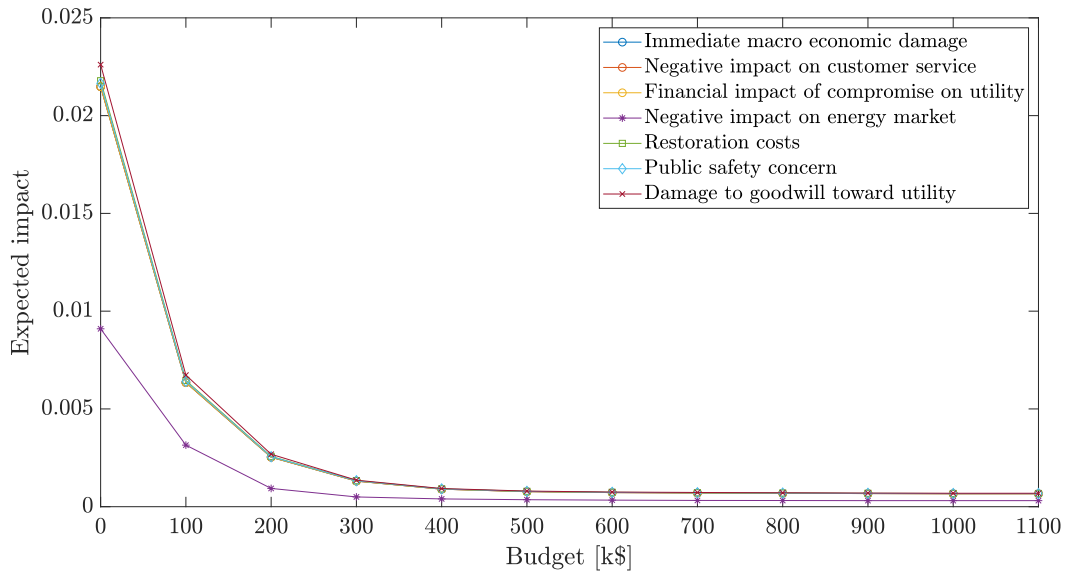


Figure 7: Expected impact of each impact criterion for different budget levels.



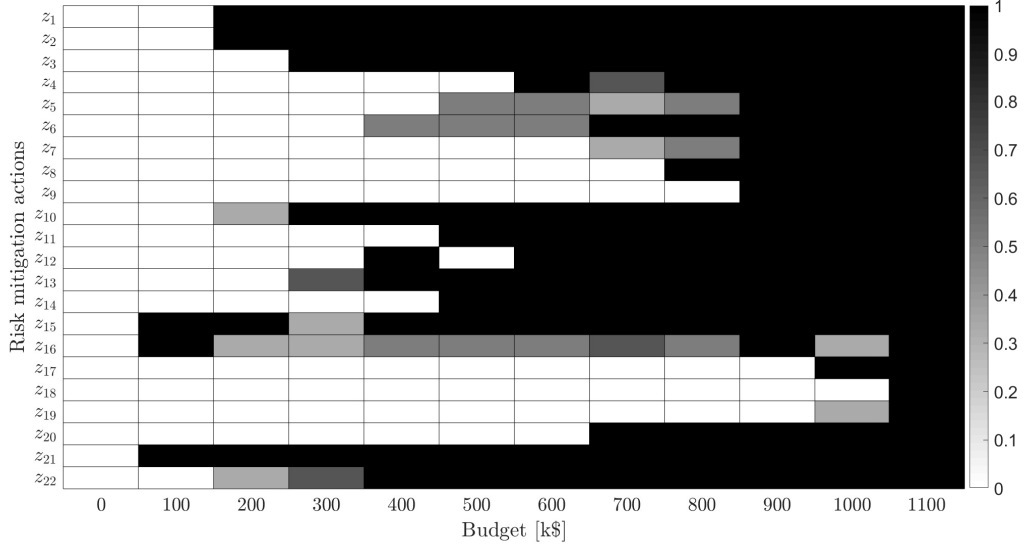


Figure 8: Core index map of mitigation actions for different budget levels.

Then, the core index of each mitigation action is computed to support the choice of actions that should be selected or rejected. Figure 8 maps the core index of each mitigation action through a gray scale. Specifically, a black square indicates that the action is included in every non-dominated portfolio, whereas a white square indicates that the action is not included in any non-dominated portfolio. Gray squares indicate a core index in the range  $(0, 1)$ , meaning that the mitigation action is included in some non-dominated portfolios, but not all.

As a result, the black-squared actions should be selected whereas the white-squared actions should be rejected. On the other hand, gray-squared actions need additional analyses to support the selection or rejection of the deployment on the AMI system. In this case study, the additional analyses would be necessary only for a limited number of mitigation actions for some budget levels. For instance, for budget  $B = 500\text{k\$}$  the mitigation actions  $z_5$ ,  $z_6$  and  $z_{16}$  belong to 50% of the non-dominated portfolios. The other mitigation actions belongs to either all or none of the non-dominated portfolios, so they do not require any additional analysis.

## 5 Discussion

The case study shows the potential of a comprehensive analysis of multiple cyber threats. Integrating the cyber threat scenarios into a Bayesian network facilitates the detection of system vulnerabilities and the definition of appropriate mitigation actions for protecting the cyber physical system. In this respect, actions affecting multiple cyber threats and synergies of actions affecting the same event(s) can be easily represented in a single model. This model results in a clear graphical representation of the possible cyber

threats to the system by erasing the redundancies deriving from equivalent events in multiple scenarios. The model relies on the definition of the occurrence probability of cyber threats, which could be a troublesome task. However, the decomposition of the cyber threat scenario into cascading events facilitates the definition of the occurrence probabilities of the single events. In addition, the collection of information on successful and unsuccessful cyber attacks could provide valuable data to estimate the occurrence probability of specific events [31]. These statistical analyses are not sufficient because the threat agents would exploit system vulnerabilities that were not necessarily available in past attacks, which by definition are not included in the existing data [32]. Specifically, a cyber threat may not be recognized until it manifests, thus it may be missed in threat scenarios that are examined as part of the risk assessment [33]. For this reason, it is necessary to integrate statistical analyses with information provided by experts based on investigations on possible system vulnerabilities.

The probabilistic representation of cyber threat scenarios provides a solid framework for the risk assessment of cyber physical system. It also enhances detailed analyses for risk management, in contrast to the binary representation through the attack graphs. Moreover, Bayesian networks make it possible to update the probability of the cascading events of cyber threat scenarios. As a result, the model represents the effect of the deployment of mitigation actions on the system, even considering intrusion detectors to tackle cyber threats that have not been examined for the risk assessment [34]. The evaluation of the risk for each impact criteria provides additional insights into risk management, which would not be possible with the additive model of scores proposed by the EPRI.

In the case study, the impacts of the cyber threats have been set according to the scoring system in Table 4. However, it is advisable to set different numeric scales based on the specificity of the impact criterion, for instance the criterion “Restoration costs” should be evaluated through a monetary scale. Note that the choice of the scale could lead to different solutions of the optimization model [35].

Finally, the Bayesian framework has broader applications than electric power systems to consider cyber threats on any cyber physical system. For instance, the National Vulnerability Database provides information about vulnerabilities of IT systems through the Common Vulnerability Scoring System [36].

## 6 Conclusions

In this paper, we have developed a Bayesian framework to analyze the vulnerabilities of cyber physical systems and optimize the resource allocation to protect the system from cyber threats. In particular, the selection of mitigation actions is based on the analysis of multiple outcomes of cyber attacks, including financial, safety and service impacts. Cyber threat scenarios are modeled through Bayesian networks to overview the alternative opportunities to pursue a cyber attack leading to such impacts. Thus, the mini-

mization of the expected impacts supports the choice of mitigation strategies based on a multi-objective optimization model.

The optimization model integrates budget and technical constraints that limit the set of feasible portfolios in order to select the optimal mitigation strategies. Specifically, the optimal mitigation strategies correspond to the portfolios that reduce the risk of cyber threats for any impact criterion without increasing the risk for other impact criteria. As a result, we have showed that a comprehensive analysis of the cyber threat scenarios leads to an optimal mitigation strategy for the system. The viability of the Bayesian framework has been illustrated through a case study concerning the Advanced Metering Infrastructure of an electric power system, which have raised several security concerns.

In conclusion, this framework can be introduced as a novel practice for assessing the risks of cyber threats and for supporting risk-based decisions on resource allocation to cyber physical systems. Possible extensions need to be investigated, such as modeling the objectives of the threat agent(s) through Adversarial Risk Analysis [37]. Future research will focus on the analysis of the cyber resilience [38], meaning the ability of the cyber physical system to continuously deliver the intended outcome despite adverse cyber events.

## Acknowledgments

The research was partly developed in the Young Scientists Summer Program at the International Institute for Applied Systems Analysis, Laxenburg (Austria) with financial support from the Academy of Finland. The research was partly supported by the European Union's Horizon 2020 project 740920 CYBECO.

## References

- [1] Lee, J., Bagheri, B. and Kao, H.A., 2015. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, pp.18-23.
- [2] Smith, M.D. and Paté-Cornell, M.E., 2018. Cyber Risk Analysis for a Smart Grid: How Smart is Smart Enough? A Multiarmed Bandit Approach to Cyber Security Investment. *IEEE Transactions on Engineering Management*.
- [3] Whitehead, D.E., Owens, K., Gammel, D. and Smith, J., 2017, April. Ukraine cyber-induced power outage: Analysis and practical mitigation strategies. In *Protective Relay Engineers (CPRE), 2017 70th Annual Conference for* (pp. 1-8). IEEE.

- [4] Yaqoob, I., Ahmed, E., ur Rehman, M.H., Ahmed, A.I.A., Al-garadi, M.A., Imran, M. and Guizani, M., 2017. The rise of ransomware and emerging security challenges in the Internet of Things. *Computer Networks*, 129, pp.444-458.
- [5] Nourian, A. and Madnick, S., 2018. A systems theoretic approach to the security threats in cyber physical systems applied to stuxnet. *IEEE Transactions on Dependable and Secure Computing*, 15(1), pp.2-13.
- [6] Lee, R.M., Assante, M.J. and Conway, T., 2014. German steel mill cyber attack. *Industrial Control Systems*, 30, p.62.
- [7] Kshetri, N., 2010. *The global cybercrime industry: economic, institutional and strategic perspectives*. Springer Science and Business Media.
- [8] Poolsappasit, N., Dewri, R. and Ray, I., 2012. Dynamic security risk management using bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing*, 9(1), pp.61-74.
- [9] Shameli-Sendi, A., Louafi, H., He, W. and Cheriet, M., 2018. Dynamic optimal countermeasure selection for intrusion response system. *IEEE Transactions on Dependable and Secure Computing*, 15(5), pp.755-770.
- [10] Mancuso, A., Compare, M., Salo, A. and Zio, E., 2017. Portfolio optimization of safety measures for reducing risks in nuclear systems. *Reliability Engineering and System Safety*, 167, pp.20-29.
- [11] Barrett, M.P., 2018. *Framework for Improving Critical Infrastructure Cybersecurity Version 1.1* (No. NIST Cybersecurity Framework).
- [12] Zio, E., 2009. *Computational methods for reliability and risk analysis* (Vol. 14). World Scientific Publishing Company.
- [13] Lee, A., 2015. Analysis of selected electric sector high risk failure scenarios. National Electric Sector Cybersecurity Organization Resource (NESCOR) Technical Working Group 1.
- [14] Ciapessoni, E., Cirio, D., Kjølle, G., Massucco, S., Pitto, A. and Sforza, M., 2016. Probabilistic risk-based security assessment of power systems considering incumbent threats and uncertainties. *IEEE Transactions on Smart Grid*, 7(6), pp.2890-2903.
- [15] Shelar, D. and Amin, S., 2017. Security assessment of electricity distribution networks under DER node compromises. *IEEE Transactions on Control of Network Systems*, 4(1), pp.23-36.
- [16] Lee, A., 2015. Electric sector failure scenarios and impact analyses. National Electric Sector Cybersecurity Organization Resource (NESCOR) Technical Working Group 1.

- [17] Johnson, P., Vernotte, A., Gorton, D., Ekstedt, M. and Lagerström, R., 2016, October. Quantitative information security risk estimation using probabilistic attack graphs. In International Workshop on Risk Assessment and Risk-driven Testing (pp. 37-52). Springer, Cham.
- [18] Ni, H., Chen, A. and Chen, N., 2010. Some extensions on risk matrix approach. *Safety Science*, 48(10), pp.1269-1278.
- [19] Duijm, N.J., 2015. Recommendations on the use and design of risk matrices. *Safety science*, 76, pp.21-31.
- [20] Allodi, L. and Massacci, F., 2017. Security events and vulnerability data for cybersecurity risk estimation. *Risk Analysis*, 37(8), pp.1606-1627.
- [21] Liu, Y. and Man, H., 2005, March. Network vulnerability assessment using Bayesian networks. In Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security 2005 (Vol. 5812, pp. 61-72). International Society for Optics and Photonics.
- [22] Frigault, M. and Wang, L., 2008, July. Measuring network security using bayesian network-based attack graphs. In Annual IEEE International Computer Software and Applications Conference (pp. 698-703). IEEE.
- [23] Kordy, B., Pitre-Cambacds, L. and Schweitzer, P., 2014. DAG-based attack and defense modeling: Dont miss the forest for the attack trees. *Computer science review*, 13, pp.1-38.
- [24] Nielsen, T.D. and Jensen, F.V., 2009. Bayesian networks and decision graphs. Springer Science and Business Media.
- [25] Xie, P., Li, J.H., Ou, X., Liu, P. and Levy, R., 2010, June. Using Bayesian networks for cyber security analysis. In Dependable Systems and Networks (DSN), 2010 IEEE/IFIP international conference on (pp. 211-220). IEEE.
- [26] Couce-Vieira, A., Houmb, S.H. and Ros-Insua, D., 2017, August. CSIRA: A Method for Analysing the Risk of Cybersecurity Incidents. In International Workshop on Graphical Models for Security (pp. 57-74). Springer, Cham.
- [27] Liesiö, J., Mild, P. and Salo, A., 2008. Robust portfolio modeling with incomplete cost information and project interdependencies. *European Journal of Operational Research*, 190(3), pp.679-695.
- [28] Liesiö, J., 2014. Measurable multiattribute value functions for portfolio decision analysis. *Decision Analysis*, 11(1), pp.1-20.

- [29] Mancuso, A., Compare, M., Salo, A. and Zio, E., 2019. Portfolio optimization of safety measures for the prevention of time-dependent accident scenarios. *Reliability Engineering and System Safety* 190 (106500), pp. 1-9.
- [30] Coello, C.A.C., Lamont, G.B. and Van Veldhuizen, D.A., 2007. *Evolutionary algorithms for solving multi-objective problems* (Vol. 5). New York: Springer.
- [31] Holm, H., 2014. A large-scale study of the time required to compromise a computer system. *IEEE Transactions on Dependable and Secure Computing*, 11(1), pp.2-15.
- [32] Paté-Cornell, M.E., Kuypers, M., Smith, M. and Keller, P., 2018. Cyber risk management for critical infrastructure: a risk analysis model and three case studies. *Risk Analysis*, 38(2), pp.226-241.
- [33] Linkov, I., Eisenberg, D.A., Plourde, K., Seager, T.P., Allen, J. and Kott, A., 2013. Resilience metrics for cyber systems. *Environment Systems and Decisions*, 33(4), pp.471-476.
- [34] Modelo-Howard, G., Bagchi, S. and Lebanon, G., 2008, September. Determining placement of intrusion detectors for a distributed application through bayesian network modeling. In *International Workshop on Recent Advances in Intrusion Detection* (pp. 271-290). Springer, Berlin, Heidelberg.
- [35] Hämäläinen, R.P. and Lahtinen, T.J., 2016. Path dependence in Operational ResearchHow the modeling process can influence the results. *Operations Research Perspectives*, 3, pp.14-20.
- [36] Zhang, Y., Wang, L., Xiang, Y. and Ten, C.W., 2015. Power system reliability evaluation with SCADA cybersecurity considerations. *IEEE Transactions on Smart Grid*, 6(4), pp.1707-1721.
- [37] Banks, D.L., Aliaga, J.M.R. and Insua, D.R., 2015. *Adversarial risk analysis*. Chapman and Hall/CRC.
- [38] Gisladottir, V., Ganin, A.A., Keisler, J.M., Kepner, J. and Linkov, I., 2017. Resilience of cyber systems with over and underregulation. *Risk Analysis*, 37(9), pp.1644-1651.

# Supplementary tables

Table 3: Likelihood Criteria With Scoring System [16].

Likelihood criterion	Scoring system
Skill required	0: Deep domain/insider knowledge and ability to build custom attack tools; 1: Domain knowledge and cyber attack techniques; 3: Special insider knowledge needed; 9: Basic domain understanding and computer skills.
Accessibility (physical)	0: Inaccessible; 1: Guarded, monitored; 3: Fence, standard locks; 9: Publicly accessible.
Accessibility (logical, assume have physical access)	0: High expertise to gain access; 1: Not readily accessible; 3: Publicly accessible but not common knowledge; 9: Common knowledge or none needed.
Attack vector (assume have physical and logical access)	0: Theoretical; 1: Similar attack has been described; 3: Similar attack has occurred; 9: Straightforward, for example script or tools available.
Common vulnerability among others	0: Isolated occurrence; 1: More than one utility; 3: Half or more of power infrastructure; 9: Nearly all utilities.

Table 4: Impact Criteria With Scoring System [16].

<b>Impact criterion</b>	<b>Scoring system</b>
Public safety concern	0: none; 1: 10-20 injuries possible; 3: 100 injured possible; 9: one death possible.
Workforce safety concern	0: none; 3: any possible injury; 9: any possible death.
Ecological concern	0: none; 1: logical ecological damage such as localized fire or spill, repairable; 3: permanent local ecological damage; 9: widespread temporary or permanent damage to one or more ecosystems.
Financial impact of compromise on utility	0: petty cash or less; 1: up to 2% of utility revenue; 3: up to 5 %; 9: greater than 5 %.
Restoration costs	0: petty cash or less; 1: up to 1% of utility organization O&M budget; 3: up to 10%; 9: greater than 10%.
Negative impact on generation capacity	0: no effect; 1: small generation facility off-line or degraded operation of large facility; 3: more than 10% loss of generation capacity for 8 hours or less; 9: more than 10% loss of generation capacity for more than 8 hours.
Negative impact on the energy market	0: no effect; 1: localized price manipulation, lost transactions, loss of market participation; 3: price manipulation. lost transactions, loss of market participation impacting a large metro area; 9: market or key aspects of market non operational.
Negative impact on the bulk transmission system	0: no; 1: loss of transmission capability to meet peak demand or isolate problem areas; 3: major transmission system interruption; 9: complete operational failure or shut down of the transmission system.
Negative impact on customer service	0: no; 1: up to 4 hour delay in customer ability to contact utility and gain resolution, lasting one day; 3: up to 4 hour delay in customer ability to contact utility and gain resolution, lasting a week; 9: complete operational failure or shut-down of the transmission system.
Negative impact on billing functions	0: none; 1: isolated recoverable errors in customer bills; 3: widespread but correctible errors in bills; 9: widespread loss of accurate power usage data.
Damage to goodwill toward utility	0: no effect; 1: negative publicity but this does not cause financial loss to utility; 3: negative publicity causing up to 20% less interest in programs; 9: negative publicity causing more than 20% less interest in programs.
Immediate macro economic damage	0: none; 1: local businesses down for a week; 3: regional infrastructure damage; 9: widespread runs on banks.
Long term economic damage	0: none; 3: several years of local recession; 9: several years of national recession.
Loss of privacy	0: none; 1: 1000 or less individuals; 3: thousands of individuals; 9: millions of individuals.



Table 5: Mitigation Actions for Scenario “Authorized Employee Brings Malware Into System or Network”.

Index	Mitigation actions	Cost [k\$]	Affected event(s)
1	Train personnel on possible paths for infection	30	Compromised mobile device
			Compromised computer peripherals
			Unintentional installation of malware
2	Maintain patches and anti-virus	70	Compromised mobile device
			Compromised computer peripherals
			Unintentional installation of malware
3	Test for malware before connection	50	Intentional installation of malware
			Compromised mobile device
			Compromised computer peripherals

Table 6: Mitigation Actions for Scenario “Threat Agent Exploits Firewall Gap”.

Index	Mitigation actions	Cost [k\$]	Affected event(s)
4	Implement configuration management	40	Intentional set of firewall rule that permits access between two networks
			Accidental set of firewall rule that permits access between two networks
5	Verify all firewall changes	60	Intentional set of firewall rule that permits access between two networks
			Accidental set of firewall rule that permits access between two networks
6	Require intrusion detection	30	Intentional set of firewall rule that permits access between two networks
			Accidental set of firewall rule that permits access between two networks
7	Require authentication to access firewall	50	Intentional set of firewall rule that permits access between two networks
			Accidental set of firewall rule that permits access between two networks

Table 7: Mitigation Actions for Scenario “Threat Agent Uses Social Engineering”.

Index	Mitigation actions	Cost [k\$]	Affected event(s)
8	Conduct penetration testing periodically	70	Info from Internet
			Info from dumpster diving
			Info from other means
			Threat agent posing as trustworthy party
9	Train personnel on social engineering attacks	40	Threat agent posing as trustworthy party

Table 8: Mitigation Actions for Scenario “Threat Agent Obtains Credentials for System or Function”.

Index	Mitigation actions	Cost [k\$]	Affected event(s)
10	Strong passwords	30	Crack of passwords
11	Encrypt communication paths	80	Capture of passwords on network or through keystroke logger
12	Protect against replay	60	Capture of passwords on network or through keystroke logger
13	Strong security questions	30	Reset passwords
14	Require multi-factor authentication	50	No assistance from authorized user
15	Use a token with PIN	20	Theft of an authentication token

Table 9: Mitigation Actions for Scenario “Threat Agent Gains Access to Network”.

Index	Mitigation actions	Cost [k\$]	Affected event(s)
16	Limit individuals with privilege	30	Having privilege to access network hosting disconnect function
			Privilege to access a network connected to network hosting disconnect function
17	Isolate network	90	Privilege to access a network connected to network hosting disconnect function
18	Enforce restrictive firewall rules to access connected network	70	Path to gain privilege to access network hosting disconnect function
19	Require authentication to access connected network	40	Path to gain privilege to access network hosting disconnect function

Table 10: Mitigation Actions for Scenario “Reverse Engineering of AMI Equipment Allows Unauthorized Mass Control”.

Index	Mitigation actions	Cost [k\$]	Affected event(s)
20	Remove unsecure development features	80	Reverse engineering of AMI meters
21	Include credentials in equipment design	50	Control of many devices simultaneously
22	Configure for least functionality	30	Control of many devices simultaneously