Integrated management of safety and security barriers in chemical plants to cope with emerging cyber-physical attack risks under uncertainties

Yuan, Shuaiqi; Reniers, Genserik; Yang, Ming

DOI
10.1016/j.ress.2024.110320

Publication date
2024

Document Version
Final published version

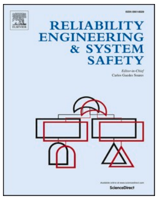
Published in
Reliability Engineering and System Safety

Reliability Engineering and System Safety 250 (2024) 110320

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress



Integrated management of safety and security barriers in chemical plants to cope with emerging cyber-physical attack risks under uncertainties

Shuaiqi Yuan [a,*], Genserik Reniers [a,b,c], Ming Yang [a]

[a] Safety and Security Science Section, Faculty of Technology, Policy and Management, TU Delft, Delft, The Netherlands
[b] Faculty of Applied Economics, Antwerp Research Group on Safety and Security (ARGoSS), Universiteit Antwerpen, 2000, Antwerp, Belgium
[c] CEDON, KULeuven, 1000 Brussels, Belgium

ARTICLE INFO

Keywords:
Safety and security risks
Cyber-physical attacks
Quantitative risk assessment
Safety barrier management
Bayesian networks
Cost-effectiveness analysis

ABSTRACT

Chemical facilities face threats from accidental and intentional events, including the rising concern of cyber-physical (C2P) attacks in the digitized industrial control system era. Addressing major accident risks from safety hazards and C2P attacks requires an immediate unified framework for safety and security barrier management. This study presents a systematic risk-based approach to integrate conventional safety risks with emerging C2P attack risks. Adverse scenarios are identified, integrated into an attack-tree-bow-tie diagram, and modelled using a Bayesian network (BN). A vulnerability assessment model is developed to quantify industrial control system vulnerability to C2P attacks, considering uncertainties in attackers' knowledge levels. Monte Carlo simulations are used to handle uncertainty propagation in risk assessment, allowing the use of probability distributions for BN root nodes. Sensitivity analysis identifies critical factors/events, guiding the proposal of candidate strategies for barrier improvements. Combining cost-effectiveness analysis with a risk matrix yields the optimal strategy for safety and security barrier enhancements based on risk estimations. A hypothetical case study demonstrates the proposed approach's effectiveness in integrated safety and security barrier management, considering security vulnerability patching and safety barrier maintenance scheduling from a cost-effective perspective.

1. Introduction

With the automation and digitization of chemical process facilities, industrial cyber-physical systems (ICPSs), also called industrial control systems (ICSs), are widely applied to chemical plants to integrate the operation of the physical process with computing and communication infrastructures [1–3]. Due to the vulnerabilities of ICSs to cyberattacks, dangerous failures of industrial facilities may be induced by either safety causes or cyberattacks [4]. The latter is known as cyber-physical (C2P) attacks. For instance, the malware "Stuxnet" is regarded as the world's first publically known digital weapon, which can target programmable logic controllers (PLCs) and induce physical damage [5]. In 2014, a series of alerts described the BlackEnergy malware that targeted human-machine interfaces (HMIs) in ICSs [6]. In 2017, FireEye claimed that industrial safety systems in the Middle East were targeted by the TRITON malware, which is capable of compromising safety instrument systems [7]. As a result, the security analysis of C2P attacks against ICSs has gained extensive attention from academia [8,9]. The traditional safety science domain is suggested to extend and incorporate possible C2P-attack-induced dangerous scenarios [10,11].

Some attempts have been made by researchers to quantify the vulnerabilities or security risks of ICSs considering C2P attacks. Remarkably, a time-to-compromise (TTC) based approach was proposed to quantify the vulnerability of ICSs to cyberattacks considering the number of known vulnerabilities on system components and attackers' skill levels [12,13]. Then, the TTC approach has been adapted for quantitative risk/reliability assessment of industrial cyber-physical systems regarding C2P attacks [14,15]. By contrast, some researchers used the exploitability subscores from the CVSS (Common Vulnerability Scoring System) approach to quantify the vulnerability of ICSs regarding cyberattacks [16,17]. More recently, the TTC approach was modified and combined with CVSS scores for the vulnerability assessment of ICSs [18], in which the exploitabilities of known vulnerabilities are also considered. Additionally, Markov decision process (MDP) [19], a game-theoretic methods [20] and the combination of a stochastic game with Markov processes [21] were employed to model cyber-physical attacks and support security assessment of ICSs. However, previous

[*] Corresponding author.
E-mail address: S.Yuan-2@tudelft.nl (S. Yuan).

https://doi.org/10.1016/j.ress.2024.110320
Received 1 February 2024; Received in revised form 8 May 2024; Accepted 1 July 2024
Available online 3 July 2024

Acronyms

BN	Bayesian network
C2P	cyber-physical
CDF	cumulative density function
CPS	cyber-physical system
CPTs	conditional probability tables
CSTR	continuous stirred tank reactor
CVSS	common vulnerability scoring system
DoS	denial-of-service
ESD	emergency shutdown system
FDI	false data injection
HMIs	human-machine interfaces
IT	information technology
ICPSs	industrial cyber-physical systems
ICSs	industrial control systems
MTTF	mean-time-to-failure
MTTSD	mean-time-to-shutdown
MOE	multiple occurring events
MTBV	mean-time-between-vulnerabilities
OT	operational technology
PLCs	programmable logic controllers
PFD	probability of failure on demand
SOC	security operations center
SV	safety relief valve
TTC	time-to-compromise

Notations

$P(X)$	a joint probability distribution in a BN
$Pa(X_i)$	the parent node set of X_i
E	evidence in a BN
$P(X\|E)$	posterior probabilities
$X(t)$	the degradation level at time t
L	a predefined failure threshold
$F_{X(t)}(x)$	the cumulative density function of $X(t)$
$A(t)$	the availability/reliability of a barrier component
$PFD_d(t)$	PFD considering barrier degradation
$PFD_{nd}(t)$	PFD without the consideration of barrier degradation
λ	failure rate
$Pr(i)$	the probability of attack path i being selected
$GTTC_i$	the global time-to-compromise of attack path i
L^i	the conditional probability of attack path i is executed successfully given an attack attempt
$MTTD_i$	mean-time-to-detect regarding attack path i
β_ia	coefficient describing the likelihood that a successful intrusion of attack path i induces a dangerous phenomenon
β_i^d	the likelihood that the attack-induced deviations escape the fault/anomaly detection algorithm successfully
β_i^r	the likelihood that the attack-induced deviations cause a dangerous phenomenon successfully
$X(k+1)$	the system state vector at time $k+1$
$\widetilde{U}(k)$	the control actions of actuators
w	process noise
$Y = \{y_1, \ldots, y_m\}$	the observation vector with m variables
v	observation noise
$\widetilde{S}(k)$	setpoint values
R	safety thresholds for the system state variables
$K = \{k_s,\cdots,k_e\}$	the attack duration from the start time k_s to the stop time k_e
I_n	the criticality of causal factor n
p_s	the probability of occurrence of the unwanted accident scenario
p_n	the probability of happening of causal factor n
C_i	the cost of strategy i
$Risk_i$	the risk value after implementing strategy i
$Risk_{threshold}$	risk threshold

methodologies rarely considered the uncertainties associated with attackers' knowledge levels in the vulnerability/risk assessment of ICSs. For instance, the CVSS approach assumes that attackers have sufficient knowledge of the weaknesses of the target system [17]. Due to the unpredictable nature of security threats and the inherent complexity of Industrial Control Systems (ICSs), the uncertainties associated with cyber-physical (C2P) attack risks can be substantial. Remarkably, the uncertainties concerning attackers' knowledge levels and attack path selections have not been well treated to support C2P risk assessment and management.

The consequence/impact analysis of C2P attacks against ICSs has also been investigated by previous studies. For instance, mean-time-to-failure (MTTF) [21] and mean-time-to-shutdown (MTTSD) [16,20] were used to describe the physical impact caused by C2P attacks on industrial systems. The potential economic consequences associated with production loss, operating cost, and loss of incidents caused by C2P attacks were also investigated by previous studies [22,23]. Patriarca et al. [24] developed a simulation-based methodology for modeling cyber resilience in a water treatment and distribution system. However, the effects of safety barriers (such as safety instrumented systems and manual shutdown) on C2P attack protection and the corresponding damage mitigation were rarely considered in the above-mentioned studies. Although the integrated safety and security risk analysis of ICSs regarding C2P attacks was suggested by researchers [10,25], the interactions and interdependence between accidental failures and C2P attacks were hardly investigated quantitatively in previous studies, which makes the risk analysis fail to capture more realistic scenarios.

The barrier concept originated from the safety science domain [26], and was also used in studies related to physical attacks [27,28] and cyberattacks [29]. Our previous study defines safety barriers as all kinds of measures/tools used to prevent the happening of undesired events or mitigate their corresponding consequences [30]. Similarly, security barriers represent all kinds of measures/tools used to protect vulnerable assets from intentional attacks/malicious acts or mitigate the corresponding consequences. Safety and security barriers are crucial for mitigating adverse risks, and the concept of barrier management is readily comprehensible to practitioners. Therefore, developing an integrated safety and security barrier framework is immensely beneficial in addressing the emerging risks associated with C2P attacks. The safety and security barrier integration also fits with the safety & security integration initiative suggested by process safety experts [31]. To achieve this, our research team conducted a series of studies on integrating safety and security barriers. Our research endeavors evolve from a safety barrier management approach [32] to cost-effective maintenance of safety and security barriers considering both safety failures and physical attacks [28]. Then, an exploratory study is conducted to support integrated risk assessment of ICSs considering safety causes, C2P attacks, and physical attacks [33]. In the same study, some conservative hypothesis is used (for instance, all cyber intrusions are assumed to be executed by high-skill-level attackers), and the uncertainties in the integrated risks have not been assessed and leveraged for decision-making. Some researchers stressed the necessity of uncertainty treatment in quantitative risk analysis [34,35]. Researchers have also highly emphasized the importance of using probability distributions rather than fixed-point probability values or expected values of probability distributions in risk modeling and decision-making processes [36].

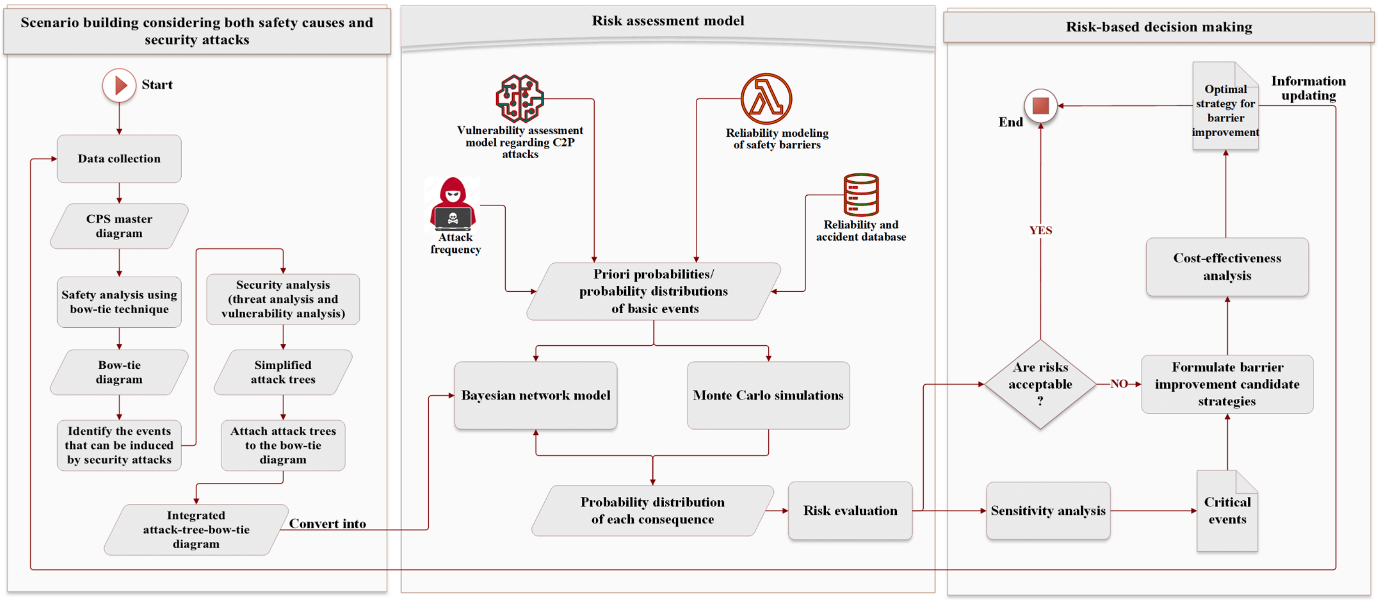

Scenario building considering both safety causes and security attacks
Start
Data collection
CPS master diagram
Safety analysis using bow-tie technique
Security analysis (threat analysis and vulnerability analysis)
Bow-tie diagram
Simplified attack trees
Identify the events that can be induced by security attacks
Attach attack trees to the bow-tie diagram
Integrated attack-tree-bow-tie diagram
Convert into
Risk assessment model
Vulnerability assessment model regarding C2P attacks
Reliability modeling of safety barriers
Attack frequency
Reliability and accident database
Priori probabilities/ probability distributions of basic events
Bayesian network model
Monte Carlo simulations
Probability distribution of each consequence
Risk evaluation
Risk-based decision making
End
Optimal strategy for barrier improvement
Information updating
YES
Cost-effectiveness analysis
Are risks acceptable ?
NO
Formulate barrier improvement candidate strategies
Sensitivity analysis
Critical events

Fig. 1. Overview of the proposed methodology.

Considering the noticeable uncertainties involved in the integrated safety and security risks, appropriate treatment of uncertainties in risk assessment helps decision-makers understand the strength of guidance for decision-making in practice and may benefit the establishment of an efficient safety/security barrier management methodology. Therefore, targeting the challenges in uncertainty treatment in the risk assessment of C2P attack scenarios and lifting our research endeavors on integrating safety and security barriers, this study aims to make methodological contributions from the following two aspects: i) improving uncertainty treatment in the integrated safety and security risks, particularly, assessing the uncertainties in C2P attack scenarios; ii) leveraging the risk with uncertainties for cost-effective barrier management decision-making.

This study introduces a systematic approach to integrated safety and security barrier management for C2P attack scenarios. It enhances risk-based decision-making by optimizing barriers through the integration of C2P attack risks with traditional safety risks. The remainder of this paper is organized as follows. Section 2 elaborates on the overall structure and operating procedures of each step of the proposed approach. An illustrative case study is employed to demonstrate the application of the proposed approach to cost-effective barrier optimization in Section 3. Discussions are presented in Section 4 before conclusions are given in Section 5.

2. Methodology

2.1. Overview of the proposed approach

This study proposes a systematic approach for integrated safety and security barrier management/optimization based on quantitative risks considering uncertainties. An overview of the proposed methodology is given in Fig. 1. The proposed methodology consists of three parts, which aim to address accident scenario integration considering both safety causes and C2P attacks, uncertainty handling and risk assessment, and risk-based decision-making, respectively. A detailed illustration of each part of the methodology is presented in the following sub-sections.

2.2. Scenario integration considering safety causes and security threats

Bow-tie diagrams [37] and attack trees [38] are widely used for accident scenario building in terms of safety and security, respectively. ICSs are usually complex engineered systems with the integration of IT (information technology) and OT (operational technology) infrastructures. Although the combination of bow-tie and attack trees was suggested to demonstrate accident scenarios for ICSs [25], a systematic approach is needed to support the accident scenario integration due to the complexities of the ICSs. A tool named CPS (cyber-physical system) master diagram [10], which is capable of representing ICSs in a multi-layered manner with the demonstration of energy flows and information flows, can serve as a basis for accident scenario building. Our previous study developed a systematic approach for accident scenario building of ICSs considering both safety causes and security threats [33], in which the CPS master diagrams, bow-tie analysis, and attack trees are combined to identify and integrated safety-associated and security-associated adverse scenarios. Basically, the integration can be done by checking each event in the bow-tie diagram by asking if this event can also be induced by security attacks. If the answer is yes, the possible security-associated scenarios should be analyzed using attack trees, and the developed attack trees need to be attached to the corresponding places of the bow-tie diagram to form an integrated attack-tree-bow-tie diagram eventually. The workflow of this approach is presented in the left block of Fig. 1. Simplified attack trees composed of two basic events (representing attack likelihood/frequency and the conditional probability of successful attacks given an attack attempt, respectively) and one top event (representing a successful attack) are used in the approach. Attack likelihoods (attack attempt frequencies) and the conditional probabilities of successful attacks given an attack attempt serve as necessary quantitative data for security risk assessment [39]. Using the simplified attack trees helps reduce the complexity of the developed attack-tree-bow-tie diagram while retaining necessary quantitative data for risk assessment. More details on this approach can be found in the original study [33]. An example of the CPS master diagram can be found in Fig. 6(b) and the corresponding attack-tree-bow-tie diagram can be found in Fig. 7.

2.3. An integrated risk assessment model

A probabilistic risk assessment is performed using a Bayesian network (BN) model with the help of a vulnerability assessment model for C2P attacks, reliability modeling of safety barriers, and Monte Carlo simulations. The main procedures of the risk assessment and uncertainty propagation handling are illustrated in Section 2.3.1. Then, details on

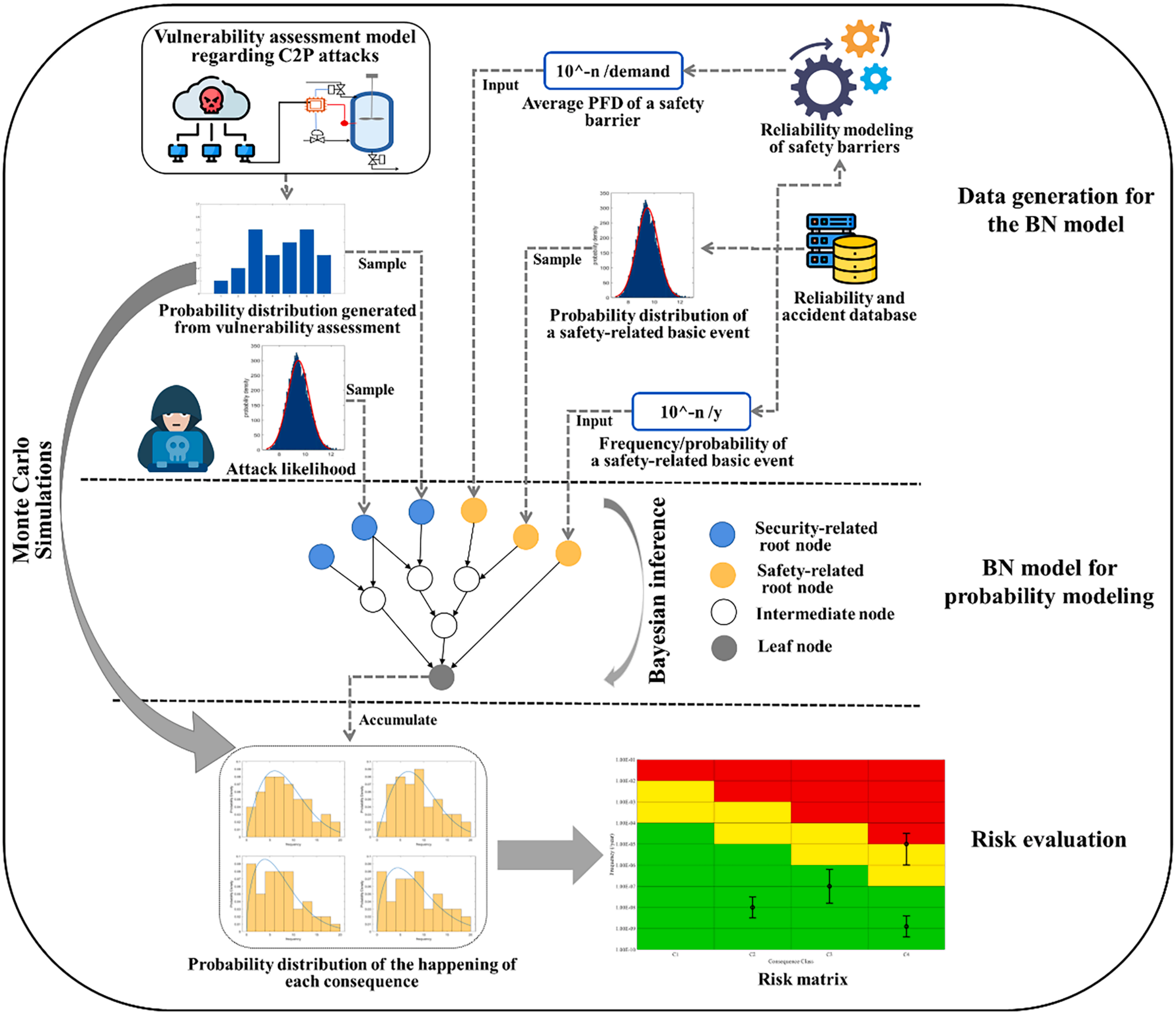

Vulnerability assessment model regarding C2P attacks
Input
10^-n /demand
Average PFD of a safety barrier
Reliability modeling of safety barriers
Data generation for the BN model
Sample
Probability distribution generated from vulnerability assessment
Sample
Probability distribution of a safety-related basic event
Reliability and accident database
Sample
Attack likelihood
Input
10^-n /y
Frequency/probability of a safety-related basic event
Monte Carlo Simulations
Bayesian inference
Security-related root node
Safety-related root node
Intermediate node
Leaf node
BN model for probability modeling
Accumulate
Probability distribution of the happening of each consequence
Risk matrix
Risk evaluation
C1
C2
C3
C4
Consequence Class

Fig. 2. Flowchart of the integrated risk assessment process.

the vulnerability assessment model and reliability modeling of safety barriers are presented in Sections 2.3.2 and 2.3.3, respectively.

2.3.1. Bayesian network model and uncertainty propagation handling

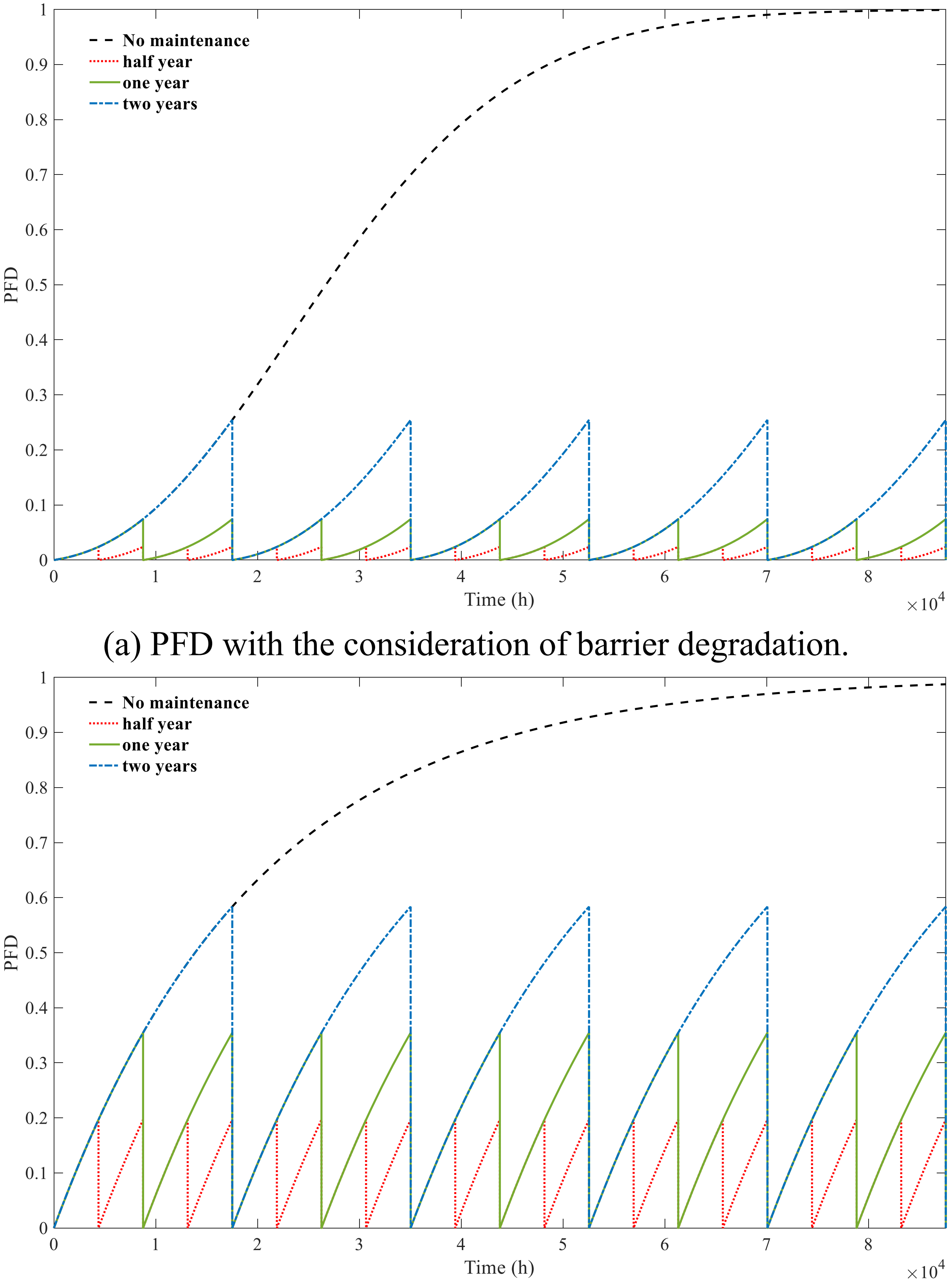
2.3.1.1. Convert the attack-tree-bow-tie diagram into a BN model. The obtained attack-tree-bow-tie diagram incorporates safety-associated and security-associated scenarios. However, a quantitative risk analysis is still challenging. Due to the complexity of the identified scenarios, multiple occurring events (MOE) are usually unavoidable. While those events are not allowed in fault/attack tree analysis when performing probability calculations [40]. To solve this problem, minimal cut sets must be determined and the tree should be translated into an equivalent set of Boolean equations for probability calculations. Alternatively, the fault/attack trees can be converted into BN models for probability modeling, which is capable of handling dependent basic events [41]. BNs are probabilistic graphical models and are widely used for risk assessments due to their advantage of representing random variables with their interdependencies [42,43]. In a BN, a joint probability distribution $P(X)$ of variables $X = \{X_1, \dots, X_n\}$ is presented as follows [44]:

$$P(X) = \prod_{i=1}^{n} P(X_i|Pa(X_i)) \tag{1}$$

$$P(X|E) = \frac{P(E|X)\cdot P(X)}{P(E)} = \frac{P(E,\ X)}{\sum_X P(E,\ X)} \tag{2}$$

where $Pa(X_i)$ is the parent node set of X_i. When evidence E becomes available, the posterior probabilities $P(X|E)$ can be derived based on Bayes theorem using Eq (2).

In this study, we convert the attack-tree-bow-tie diagram into a BN model for probability modeling because of its advantages of representing the dependencies of events, incorporating multi-state variables, and updating probabilities. A more detailed comparison of the bow-tie-based approach and BN can be found in [32]. The topology and conditional probability tables (CPTs) of the BN can be derived from the integrated attack-tree-bow-tie diagrams by following dedicated mapping algorithms. Previous studies have already given guidance on transforming attack trees [38] and bow-tie diagrams [45] into BNs. We avoid repeating illustrations here. After converting, the root nodes of the BN



(a) PFD with the consideration of barrier degradation.

(b) PFD without the consideration of barrier degradation.

Fig. 3. PFD calculation of technical safety barrier components with different preventive maintenance strategies.

model can be divided into safety-related and security-related nodes. Safety-related nodes are derived from basic events in the bow-tie diagram, which are the root accidental causes leading to the central event. Security-related nodes are derived from basic events in the attack trees, which include attack likelihood and the conditional probability of successful execution of each attack mode. Each state of the leaf node presents each outcome event (consequence) in the attack-tree-bow-tie diagram.

2.3.1.2. Determination of priori probabilities/probability distributions for root nodes. After the topology and CPTs of the BN model are determined, prior probabilities of the root nodes are required to perform the risk assessment. This study uses probability distributions or fixed-point probability values for the root nodes. Probability distributions are used to capture the uncertainties associated with the happening of some basic events. In case no data is available to form probability distributions, fixed-point probability values are used. Four ways are used to determine prior probabilities or probability distributions for different types of root nodes, as illustrated below.

- Regarding safety-related initiating events, for instance, critical failure of a technical component, human failure, external fires, etc., reliability databases [46], human reliability data [47], accident databases [48], or data available in the literature may be used to derive the occurrence probabilities. When a probability distribution of the occurrence of an event is available, it may be used instead of a fixed-point probability value.
- Regarding safety barriers, the probability of failure on demand (PFD) is used to quantify the reliability of safety barriers because they usually follow a low-demand mode [49,50]. For human components, human reliability data may be used to obtain the PFDs of human actions. The approach for reliability modeling of technical components of safety barriers considering different barrier maintenance strategies is illustrated in Section 2.3.2 in detail.
- Regarding security-associated basic events, attack likelihood/frequency and the conditional probabilities of successful attacks given attack attempts are needed. Attack likelihood of C2P attacks may be estimated according to incident statistics of the investigated chemical plants or comparable companies with the help of expert judgment. In case of lacking incident data, the estimation may mainly be performed based on expert judgment, which is one significant source of uncertainty.
- A vulnerability assessment should be performed to estimate the conditional probabilities of successful attacks given attack attempts. This is achieved by assessing the vulnerabilities of both IT systems and OT systems and considering the uncertainties associated with attackers' knowledge levels. Details about the vulnerability assessment are elaborated on in Section 2.3.3. The final results of the vulnerability assessment model are a series of probability distributions regarding different attack modes.

2.3.1.3. Uncertainty propagation handling. In this study, Monte Carlo simulations are used to handle uncertainty propagation in the risk assessment when probability distributions are used for root nodes. This is achieved by sampling point values from the probability distributions as inputs while accumulating the inferred probability of each leaf node state. Finally, probability distributions for each state (representing each consequence) of the leaf node can be obtained. Regarding the consequence assessment, a severity class for typical dangerous phenomena in chemical plants suggested by the ARAMIS project is used [51]. Then, a risk matrix is used to visualize and evaluate risk profiles by mapping the expected values and ranges of the probability distributions of the potential consequences into the risk matrix. A flowchart is presented in Fig. 2 to demonstrate the main procedures in the integrated risk assessment and uncertainty propagation handling.

2.3.2. PFD calculation of safety barriers under preventive maintenance

Safety barriers play important roles in protecting industrial systems from disastrous damage in case of dangerous failures/deviations. For a complex safety barrier system, fault tree analysis may be implemented to calculate the PFD of the whole barrier system. In practice, corrective maintenance and preventive maintenance are widely used for technical facilities in chemical plants. For the sake of safety, preventive maintenance is usually performed at specific intervals (e.g., once per year) for safety barriers, which is also known as periodic maintenance [28]. Because some safety barriers or barrier components are allocated under harsh environments, degradation inevitably happens and impacts the reliability of those safety barriers [52]. For the components subject to degradation, for instance, emergency shutdown valves, a Gamma degradation process is adapted to simulate the continuous aging degradation, as follows [32,52]:

$$X(t) \sim \Gamma(\alpha t, \beta) = f_{X(t)}(x) = \frac{\beta^{\alpha t}}{\Gamma(\alpha t)} x^{\alpha t - 1} e^{-\beta x}, \ \alpha, \beta > 0 \tag{3}$$

where $X(t)$ is the degradation level at time t. The mean and variance of $X(t)$ are $\alpha t/\beta$ and $\alpha t/\beta^2$, respectively. It is assumed that the component will fail when the degradation level reaches or overpasses a predefined failure threshold L. Then, the availability of the barrier component over time can be calculated below.

$$F_{X(t)}(x) = \int_0^x f_{X(t)}(x) dx \tag{4}$$

$$A(t) = Pr(X(t) < L) = F_{X(t)}(L) \tag{5}$$

where $F_{X(t)}(x)$ is the cumulative density function (CDF) of $X(t)$. $A(t)$ is the availability/reliability of the barrier component. Under the assumption that perfect maintenance is implemented at a periodic time interval, T, and with the ignorance of the maintenance time, PFD considering barrier degradation, $PFD_d(t)$, can be calculated as follows.

$$PFD_d(t) = 1 - A(t\%T) = 1 - F_{X(t\%T)}(L), \ nT \le t < (n+1)T \tag{6}$$

where $t\%T$ means the remainder when dividing t by T. n is an integer from 0 to positive infinity ($n = 0, 1, 2, \ldots, +\infty$). The calculated PFDs considering different maintenance intervals using Eq. (6) are compared in Fig. 3(a).

By contrast, for the components (e.g., programmable logic solvers) that are not obviously subjected to degradation, their PFDs are assumed to follow exponential distributions with constant failure rates [53,54]. With the assumption that perfect barrier maintenance with a time interval, T, is implemented and ignoring the time spent on maintenance, PFD can be calculated as below.

$$PFD_{nd}(t) = 1 - e^{-\lambda * (t\%T)}, \ nT \le t < (n+1)T \tag{7}$$

where $PFD_{nd}(t)$ is the PFD without the consideration of barrier degradation. λ is failure rate. By using Eq. (7), the calculated PFDs considering different maintenance intervals are shown in Fig. 3(b). The average PFD describes the reliability of safety barriers following a low-demand mode [49]. The average values of the PFDs over time are used as prior probabilities for the root nodes presenting safety barrier failures in the BN model.

2.3.3. Vulnerability assessment of ICS to C2P attacks considering uncertainties

A C2P attack process can be divided into two phases, intrusion into the IT (information technology) systems and manipulation of the OT (operational technology) systems. A vulnerability assessment model is developed in this study to assess both two phases. A time-to-compromise (TTC) based approach was proposed and combined with compromise graphs to quantify the vulnerability of IT infrastructures to cyberattacks

Table 1
Attack path selection mechanisms for attackers with different knowledge levels.

Attacker categories[1]	Likelihood of executing random attacks (a)	Likelihood of executing strategic attacks (b)
expert	0	1
intermediate	0.3	0.7
beginner	0.7	0.3
novice	1	0

[1] Attacker categories are adapted from the TTC-based approach [12].

[12,13]. Ling & Ekstedt [18] modified this TTC estimation approach and combined it with an ICS-specific vulnerability dataset [55] considering attackers' skill levels, the number of known vulnerabilities on a component, and the exploitability of the known vulnerabilities. This study combines the approach developed by Ling & Ekstedt [18] with compromise graphs to estimate the global TTC of each attack path. A detailed explanation of the TTC estimation approach can be found in Appendix I.

In previous studies, two approaches were used to address the attack path selection issue of attackers. McQueen et al. [13] and Semertzis et al. [14] assumed that all possible attack paths are executed by attackers in parallel, which is a conservative assumption. By contrast, Zhang et al. [15] related attack path selection to the exploitability of each attack path and it is assumed that a more exploitable attack path is more likely to be selected by attackers. However, the exploitability of unknown vulnerabilities and the uncertainty in attackers' knowledge levels are not well considered in previous studies. To address the uncertainties associated with attackers' knowledge levels in attack path selection, this study considers two different attack path selection mechanisms, which are random attacks and strategic attacks [56]. Random attack presents that an attacker selects one attack path from all possible attack paths randomly, which is applicable to attackers with low knowledge levels. A strategic attack presents that an attacker selects the attack path based on the exploitability ranking of all possible attack paths (a more exploitable attack path is more likely to be selected), which is applicable to attackers with advanced knowledge levels of the targeted system. We assign different probabilities of executing random attacks and strategic attacks for the attackers with different knowledge levels, as shown in Table 1.

Considering one attack target with n possible attack paths, the probability of attack path i being selected can be estimated as follows:

$$Pr(i) = \frac{a}{n} + \frac{b}{GTTC_i} \Bigg/ \sum_{j=1}^{n} \frac{1}{GTTC_j} \tag{8}$$

Where $Pr(i)$ is the probability of attack path i being selected. a and b are the likelihoods of executing a random attack and executing a strategic attack respectively, which are determined in Table 1. $GTTC_i$ is the global time-to-compromise of attack path i, which is calculated using the method presented in Appendix I. When executing a random attack, the likelihood of selecting each attack path from n possible attack paths is the same, which is $1/n$. By contrast, the attack path with a lower global time-to-compromise is more likely to be selected when executing a strategic attack. Based on the above two principles, $Pr(i)$ is calculated by summing up the probabilities of attack path i being selected under random attacks and strategic attacks. Then, the conditional probability of attack path i is executed successfully given an attack attempt (L^i) is estimated as follows [33]:

$$L^i = Pr(i) \times \frac{MTTD_i}{GTTC_i + MTTD_i} \times \beta_i \tag{9}$$

$$MTTD_i = \frac{\sum_{k=1}^{N} TTD_k}{N} \tag{10}$$

$$L^i_{at} = L^a + L^b, \ldots, L^n, \ (attack\ paths\ a\ to\ n\ lead\ to\ the\ same\ attack\ mode) \tag{11}$$

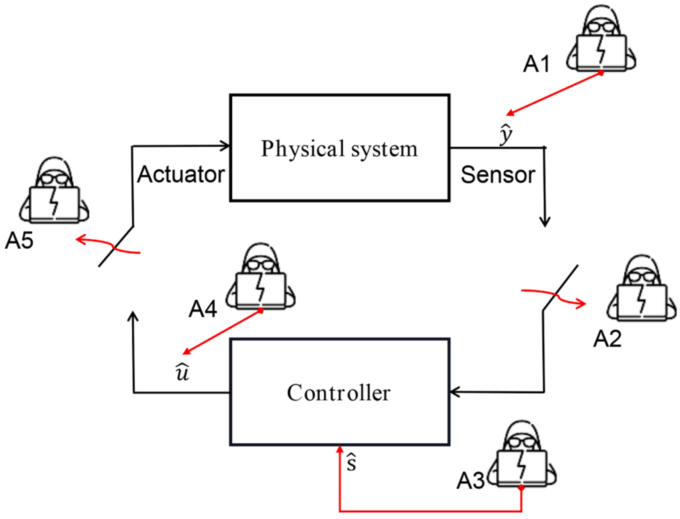


Fig. 4. Typical C2P attacks against industrial control systems, adapted from [22].

Table 2
Explanations of the C2P attacks in Fig. 4.

Marks	Attack types	Descriptions
A1	FDI attack against sensors	Maliciously manipulate the measurement data from sensors to the controller. Let $\hat{y} \neq y$, $\hat{y}$ is the manipulated data, and y is the true measurement.
A2	DoS attack against sensors	Maliciously prevent the controller from receiving sensor measurement data.
A3	Setpoint manipulation	Maliciously manipulate the setpoints configured in the controller. Let $\hat{s} \neq s$, $\hat{s}$ is the manipulated setpoint, and s is the predefined setpoint.
A4	FDI attack against actuators	Maliciously manipulate the control data from the controller to actuators. Let $\hat{u} \neq u$, $\hat{u}$ is the manipulated data, and u is the true control data.
A5	DoS attack against actuators	Maliciously prevent actuators from receiving control commands/data.

where $MTTD_i$ presents the mean-time-to-detect regarding attack path i. MTTD (mean-time-to-detect) is a widely used performance indicator describing the average time needed by the security operations center (SOC) to detect a cyber intrusion successfully (Mughal, 2022). Based on the principle that a cyber intrusion is detected with a likelihood of 50 % when its global time-to-compromise is equal to the MTTD, the probability of the SOC failing in detecting attack path i is calculated as $\frac{MTTD_i}{GTTC_i+MTTD_i}$. The MTTD regarding a specific intrusion type is calculated by averaging all incident detection times of this intrusion type, as presented in Eq. (10). For simplification, a reference value (14 days) from Semertzis et al. [14] is used as the MTTD for remote cyber intrusions. In practice, it may be determined based on actual incident data collected by SOCs. Because a successful cyber intrusion cannot always induce a physically dangerous scenario [22], coefficient β_i is defined to describe the likelihood that a successful intrusion of attack path i induces a dangerous phenomenon. The successful execution of attack path i depends on three conditions: i) attack path i is selected by attackers, ii) the SOC fails in detecting attack path i, and iii) the intrusion of attack path i induces a dangerous phenomenon successfully. Therefore, L^i is calculated as the product of the probabilities of the above-mentioned three conditions, as presented in Eq. (9). In case multiple attack paths lead to the same attack mode, the conditional probability of successful execution of the attack mode (L^i_{at}) is calculated by summing up the L^i values of those attack paths, as presented in Eq. (11). β_i depends on the fault detection capability and deviation suppression/tolerance capability of the OT system regarding the specific attack mode, and it is calculated as

Table 3
Configurations of β_i^d for attackers with different knowledge levels.

Attacker's knowledge levels	β_i^d for FDI attacks	β_i^d for DoS attacks	β_i^d for Setpoint manipulations
expert	1	1	1
intermediate	0.8	1	1
beginner	0.5	1	1
novice	0.2	1	1

$\beta_i = \beta_i^d \times \beta_i^r$. β_i^d presents the probability that the attack-induced deviations escape the fault/anomaly detection algorithm successfully. β_i^r presents the likelihood that the attack-induced deviations cause a dangerous phenomenon successfully. β_i^d and β_i^r are determined as below.

Regarding the manipulation of OT systems, five types of C2P attacks with representativeness are investigated, as illustrated in Fig. 4. Some basic descriptions of how FDI (false data injection) attacks, DoS (denial-of-service) attacks, and setpoint manipulations compromise the industrial control system are given in Table 2. More detailed illustrations can be found in [22] and [33].

The value of β_i^d should be determined considering both the attack mode and the fault detection algorithm of the OT system. In this study, we assumed that predefined ranges were applied for sensors' and actuators' signals as the fault detection method [22]. In that case, a FDI attack will be detected when the injected data is out of the scope of the predefined ranges, while DoS attacks and setpoint manipulations cannot be detected timely. Referenced β_i^d values, 1, 0.8, 0.5, and 0.2, are used for FDI attacks executed by attackers with expert, intermediate, beginner, and novice knowledge levels, respectively. The values of β_i^d considering different attack modes and different attackers' knowledge levels are configured in Table 3. In practice, the β_i^d values may be modified considering the specific fault detection algorithms the OT system uses.

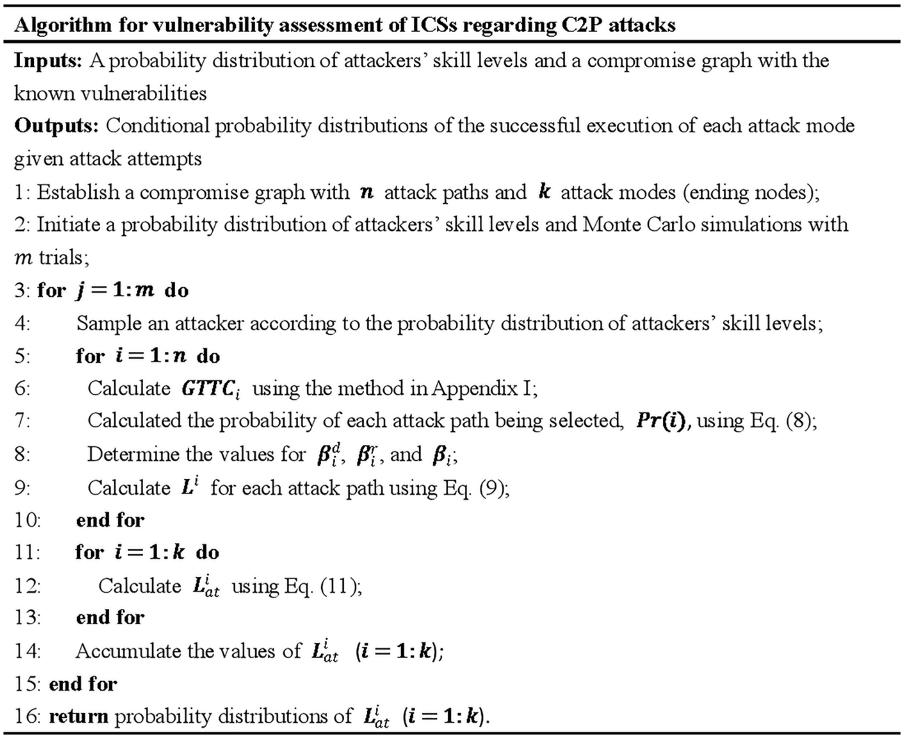
By integrating attack modeling with a numerical model of the investigated system, the deviations caused by C2P attacks can be assessed and the value of β_i^r can be determined. We use a generalized system to demonstrate this process. For a system represented by a system state vector with n variables ($X = \{x_1, \ldots, x_n\}$), the system states under the influence of C2P attacks can be estimated as below.

$$\begin{cases} X(k+1) = f(X(k),\ \widetilde{U}(k),\ w) \\ Y(k) = g(X(k),\ v) \\ U(k) = h(\widetilde{S}(k),\ \widetilde{Y}(k)) \end{cases} \tag{12}$$

where $X(k+1)$ is the system state vector at time $k+1$. $\widetilde{U}(k)=\{\widetilde{u}_1(k), \ldots, \widetilde{u}_l(k)\}$ presents the control actions of l actuators, w presents process noise. $Y = \{y_1, \ldots, y_m\}$ is the observation vector with m variables. $Y(k)$ depends on the system state vector, $X(k)$, and the observation noise, v. $U(k) = \{u_1(k), \ldots, u_l(k)\}$ is the control command for actuators, which depends on j setpoint values, $\widetilde{S}(k) = \{\widetilde{s}_1(k), \ldots, \widetilde{s}_j(k)\}$, and the observation data, $\widetilde{Y}(k)=\{\widetilde{y}_1(k), \ldots, \widetilde{y}_m(k)\}$. $\widetilde{Y}(k)$, $\widetilde{U}(k)$, and $\widetilde{S}(k)$ are derived from modeling of specific attack modes, as presented in Table 2. Usually, safety thresholds are defined for the system state variables and can be presented as $R = \begin{bmatrix} x_1^{min},\ x_1^{max} \\ \ldots \\ x_n^{min},\ x_n^{max} \end{bmatrix}$. A dangerous phenomenon occurs when $X(k) \notin R$. The coefficient β_i^r regarding a specific attack mode can be determined below.

$$\beta_i^r = \begin{cases} 1,\ if\ X_i(k) \notin R \\ 0,\ if\ X_i(k) \in R \end{cases},\ k \in K \tag{13}$$

where $K = \{k_s, \cdots, k_e\}$ represents the attack duration from the start time k_s to the stop time k_e. $X_i(k)$ is estimated by solving Eq. (12) regarding the attack mode of attack path i. Monte Carlo simulations are used to address the uncertainties in β_i^r associated with process noises and observation noises. To make the vulnerability assessment model structured, a pseudocode is presented in Fig. 5 to demonstrate the procedures of the vulnerability assessment model.



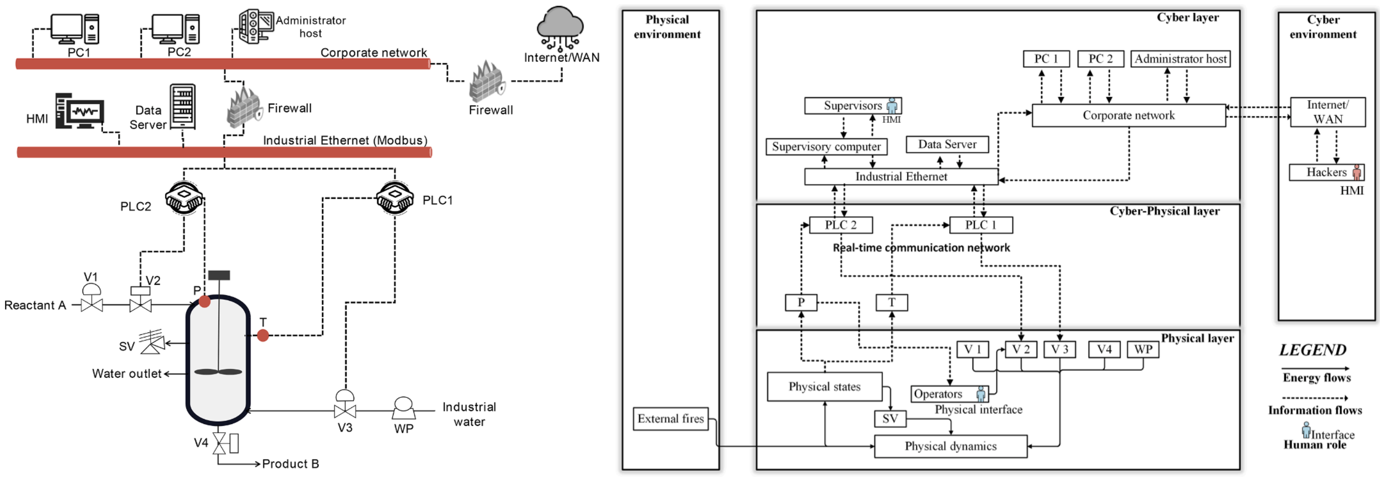
Algorithm for vulnerability assessment of ICSs regarding C2P attacks

Inputs: A probability distribution of attackers' skill levels and a compromise graph with the known vulnerabilities

Outputs: Conditional probability distributions of the successful execution of each attack mode given attack attempts

```
Establish a compromise graph with $n$ attack paths and $k$ attack modes (ending nodes);
Initiate a probability distribution of attackers' skill levels and Monte Carlo simulations with $m$ trials;
for $j = 1:m$ do
    Sample an attacker according to the probability distribution of attackers' skill levels;
    for $i = 1:n$ do
        Calculate $GTTC_i$ using the method in Appendix I;
        Calculated the probability of each attack path being selected, $Pr(i)$, using Eq. (8);
        Determine the values for $\beta_i^d$, $\beta_i^r$, and $\beta_i$;
        Calculate $L^i$ for each attack path using Eq. (9);
    end for
    for $i = 1:k$ do
        Calculate $L_{at}^i$ using Eq. (11);
    end for
    Accumulate the values of $L_{at}^i$ $(i = 1:k)$;
end for
return probability distributions of $L_{at}^i$ $(i = 1:k)$.
```

Fig. 5. Pseudocode of the vulnerability assessment model.



(a) (b)

Fig. 6. The investigated industrial control system (a) and its CSP master diagram (b), adapted from [61] and [33].

2.4. Sensitivity analysis and risk-based decision-making

Sensitivity analysis of BN root nodes helps identify critical basic events (causal factors) leading to undesired scenarios and, therefore, supports effective risk treatment. Typically, the Birnbaum importance measure [57], risk reduction measure [58], and ratio of variance (RoV) measure [59] are the widely used measurements to rank the importance of basic events to the happening of the undesired event in fault tree analysis or BN networks. This study uses the Birnbaum importance measure because it can be easily implemented into the integrated risk assessment model. Using the Birnbaum importance measure, the importance/criticality of a basic event (causal factor) to the occurrence of a major accident scenario is presented as follows:

$$I_n = p_s(p_n = 1) - p_s(p_n = 0) \tag{14}$$

where I_n is the importance/criticality of basic event (causal factor) n to the happening of undesired scenario S. p_s is the probability of occurrence of undesired scenario S. p_n is the probability of occurrence of basic event (causal factor) n. I_n is calculated as the difference in the occurrence probabilities of the undesired scenario under basic events occurring and not occurring, as shown in Eq (14). With the identification of critical basic events, corresponding risk treatment strategies may be proposed to reduce undesired risks. In practice, the objectives of risk management usually include ensuring the risks are at acceptable levels, saving the costs used for risk reduction, reducing production losses resulting from downtime, meeting legislation requirements, etc. It is crucial to make decisions on risk reduction while considering the trade-offs between multiple objectives, for instance, the trade-off between safety and costs. Among risk-based decision analysis methods, risk matrix is one of the widely-used tools because it is straightforward and user-friendly. Particularly, the combination of cost-effectiveness analysis and a risk matrix helps to investigate the trade-off between safety and costs [60]. An optimization problem under constraints is formulated to characterize the decision analysis for barrier improvements, as follows:

$$\begin{cases} Min(C_i) \\ Risk_i \le Risk_{threshold} \\ i \in \{1, 2, 3, \cdots, N\} \end{cases} \tag{15}$$

where C_i means the cost of strategy i regarding barrier improvements. $Risk_i$ is the risk estimation after implementing strategy i. $Risk_{threshold}$ is the risk threshold that could be the risk acceptable level in the risk matrix. In case probability distributions are used to represent risks considering uncertainties, thresholds may be used to constrain the expected values of probability distributions, and other constraints can also be applied to the probability distributions (for instance, setting up thresholds for the boundary values of the probability distributions). Eq (15) can be solved by using exhaustive search algorithms. In case of massive candidate strategies are proposed, evolutionary algorithms (for instance, genetic algorithms) may be used to obtain approximately optimal strategy while saving computation efforts.

Table 4
Descriptions of the identified attack modes.

Attack mode marks	Attack modes	Attack objectives
AT1	FDI attack against sensor T	Compromise PLC1 (cooling system) and trigger dangerous deviations.
AT2	DoS attack against sensor T	
AT3	FDI attack against actuator V3	
AT4	DoS attack against actuator V3	
AT5	Setpoint manipulation of temperature threshold of PLC1	
AT6	FDI attack against sensor P	Compromise PLC2 (ESD system) and trigger dangerous leakage scenarios.
AT7	DoS attack against sensor P	
AT8	FDI attack against actuator V2	
AT9	DoS attack against actuator V2	
AT10	Setpoint manipulation of overpressure threshold of PLC2	

3. Case study

3.1. System description and scenario building

In this case study, a continuous stirred tank reactor (CSTR) with its SCADA system is investigated, as demonstrated in Fig. 6(a). The CSTR model performs a hypothetical exothermic reaction A→B [61]. Product B is assumed to be a flammable liquid with toxicity. The reactor temperature is controlled using a jacketed cooling system with a water pump (WP), a control valve (V3), a temperature sensor (T), and a programmable logic controller (PLC1). Reactant A is fed at a fixed flow rate using a control valve (V1). An emergency shutdown system (ESD) with a programmable logic controller (PLC2), a block/shutdown valve (V2), and a pressure sensor (P) is deployed to block feeding in case of overpressure. Additionally, a safety relief valve (SV) is installed. Both PLCs are connected to the SCADA system and linked to the corporate network and the outside Internet/WAN. A CPS master diagram considering the

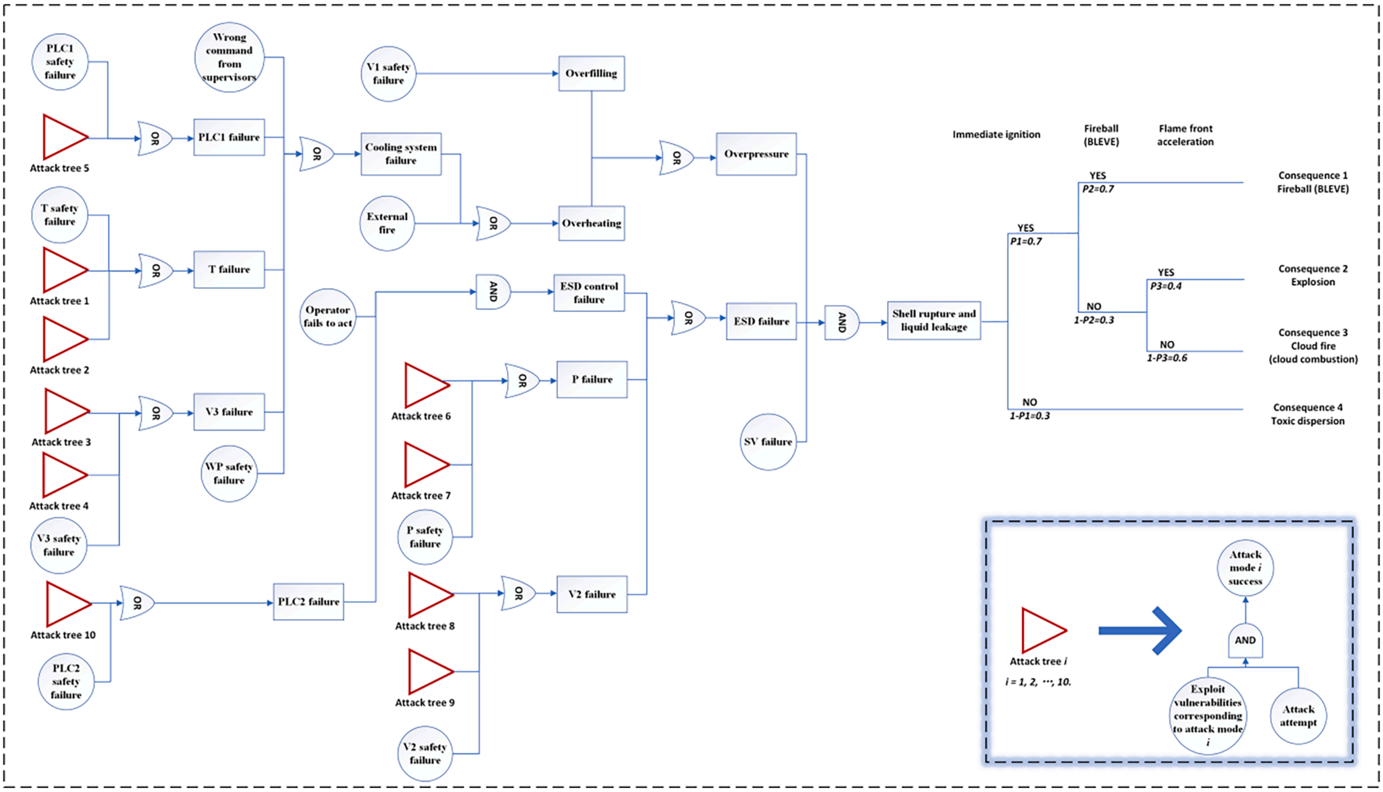


Fig. 7. An integrated attack-tree-bow-tie diagram considering safety causes and C2P attacks, adapted from [33] and [62].

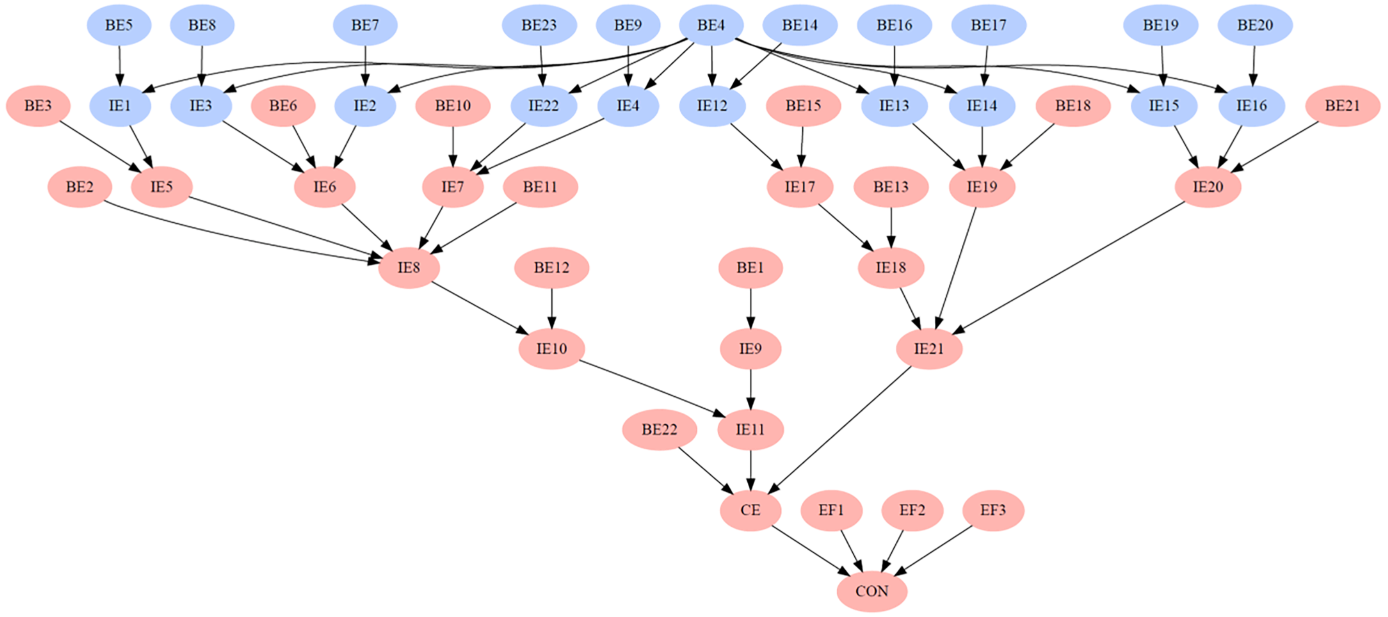


Fig. 8. Bayesian network model for integrated risk assessment (nodes with pink and blue colors are derived from the bow-tie diagram and attack trees, respectively).

multi-layered structure of the ICS is constructed, as shown in Fig. 6(b). Remote hackers (probably have different knowledge levels) are identified as threat agents, and ten specific attack modes are identified, as illustrated in Table 4. Based on the scenario integration approach presented in Section 2.2, an attack-tree-bow-tie diagram was developed to integrate accidental scenarios considering both safety causes and C2P attacks, as shown in Fig. 7.

3.2. Security vulnerability assessment and risk assessment

Based on the developed attack-tree-bow-tie diagram, a BN model was developed following the mapping algorithm suggested by Khakzad et al. [45] using the Bayes net MATLAB toolbox [63]. The BN nodes with pink color are derived from the bow-tie diagram, while the nodes with blue color are derived from attack trees, as shown in Fig. 8. The explanation of the abbreviations of the BN nodes is given in Table 5. All BN nodes, except the consequence node, have two states (happening and not

Table 5
Explanations of the BN nodes in Fig. 8.

Symbols	Node names	Symbols	Node names	Symbols	Node names	Symbols	Node names
BE1	V1 safety failure	BE2	Human error in giving commands	BE3	PLC1 safety failure	BE4	C2P attack attempts
BE5	Exploit vulnerabilities corresponding to AT5	BE6	T safety failure	BE7	Exploit vulnerabilities corresponding to AT1[1]	BE8	Exploit vulnerabilities corresponding to AT2
BE9	Exploit vulnerabilities corresponding to AT3	BE10	V3 safety failure	BE11	WP safety failure	BE12	External fire
BE13	Operator fails to shutdown	BE14	Exploit vulnerabilities corresponding to AT10	BE15	PLC2 safety failure	BE16	Exploit vulnerabilities corresponding to AT6
BE17	Exploit vulnerabilities corresponding to AT7	BE18	P safety failure	BE19	Exploit vulnerabilities corresponding to AT8	BE20	Exploit vulnerabilities corresponding to AT9
BE21	V2 safety failure	BE22	SV safety failure	BE23	Exploit vulnerabilities corresponding to AT4	EF1	Immediate ignition
EF2	Fireball (BLEVE)	EF3	Flame front acceleration	CON	Consequences	CE	Central event (Liquid leakage)
IE1	AT5 success	IE2	AT1 success	IE3	AT2 success	IE4	AT3 success
IE5	PLC1 failure	IE6	T failure	IE7	V3 failure	IE8	Cooling system failure
IE9	Overfilling	IE10	Overheating	IE11	Overpressure	IE12	AT10 success
IE13	AT6 success	IE14	AT7 success	IE15	AT8 success	IE16	AT9 success
IE17	PLC2 failure	IE18	ESD control failure	IE19	P failure	IE20	V2 failure
IE21	ESD failure	IE22	AT4 success	/	/	/	/

[1] AT1 means attack mode 1, and the explanation of each attack mode can be found in Table 4.

Table 6
Probabilities/probability distributions of the root BN nodes.

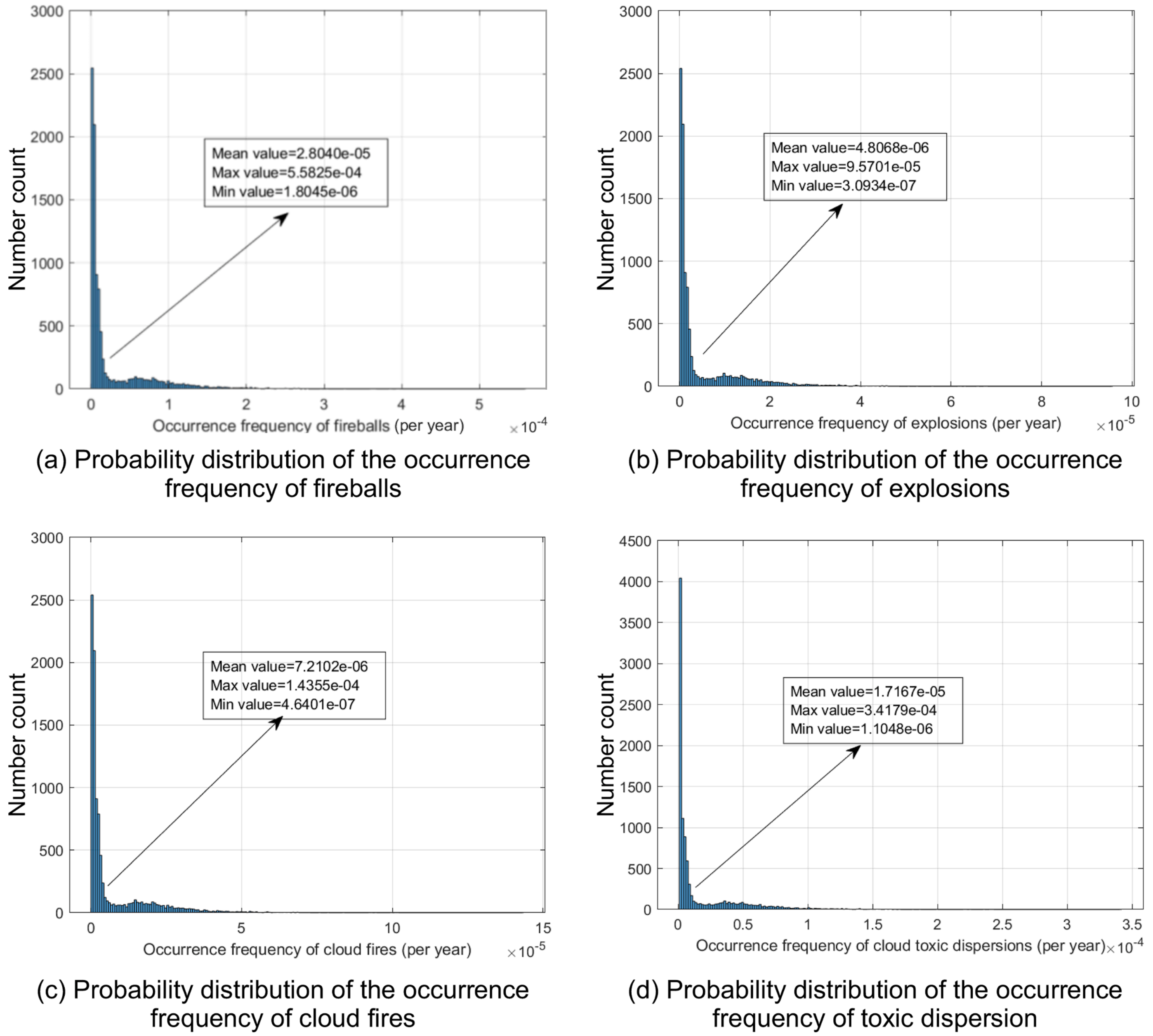
Nodes	Prior probabilities (probability distributions)	Sources	Nodes	Prior probabilities (probability distributions)	Source	Nodes	Prior probabilities (probability distributions)	Source
BE1	4.00E-02	[66]	BE2	1.00E-02	[51]	BE3	4.38E-02	[46]
BE4	Gamma distribution ($\Gamma(\alpha,\beta)$, $\alpha = 6.08$; $\beta = 5$)	Assumed based on the data from [64]	BE5	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE6	2.13E-02	[46]
BE7	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE8	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE9	As shown in Fig. A4	Calculated from the vulnerability assessment model
BE10	4.00E-02	[66]	BE11	3.125E-02	[67]	BE12	5.52E-02	[48].
BE13	Beta distribution ($Beta(a,b)$, $a = 32.3$; $b = 137.7$)	[68]	BE14	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE15	average PFD=4.37E-03; λ=1.0E-06	λ is from Hauge & Onshus [46]; Eq (7) is used to calculate the PFD.
BE16	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE17	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE18	average PFD=6.57E-04; λ=1.5E-07	λ is from Hauge & Onshus [46]; Eq (7) is used to calculate the PFD.
BE19	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE20	As shown in Fig. A4	Calculated from the vulnerability assessment model	BE21	average PFD=7.63E-03; α=1.02E-04, β=1.2E04, L=3E-04	α and β are from Zhang et al. [52]; Eq (6) is used to calculate the PFD.
BE22	average PFD=2.19E-03; λ=5E-07	λ is from [69]; Eq (7) is used to calculate the PFD.	BE23	As shown in Fig. A4	Calculated from the vulnerability assessment model	EF1	7.00E-01	[62]
EF2	7.00E-01	[62]	EF3	4.00E-01	[62]	/	/	/

happening), while the consequence node has five states (no consequence, fireball, explosion, cloud fire, and toxic dispersion).

According to a data analysis of cyber security incidents in a large American organization [64], the recurrence intervals of severe cyber incidents remain overall stable, and the recurrence interval of C2P attacks may be estimated as approximately 150~465 days. Due to the lack of actual incident data, a Gamma distribution ($\Gamma(\alpha,\beta)$, $\alpha = 6.08$; $\beta = 5$) with a mean value of 1.22 (corresponds to a recurrence interval of 300 days) and a variance value of 0.24 is used to depict the C2P attack annual frequency in this case study. When new incident data becomes available, the Gamma distribution may be updated based on Bayes' theorem [65].

Regarding vulnerability assessment, the approach presented in Section 2.3.3 was used to quantify the conditional probabilities of successfully executing each attack mode. A compromise graph was constructed to demonstrate all possible attack paths according to the identified attack modes and the IT structure of the ICS, as presented in Fig. A1 in Appendix I. A MATLAB/Simulink model was developed based on the CSTR model from Pilario & Cao [61] to assess the physical effects of different attack modes and to determine the value of β_i^r based on Eq (13). The developed MATLAB/Simulink model and selected simulation results are presented in Appendix II. According to the simulation results, the ICS's process and observation noises may decide if a DoS attack on PLC1 can succeed (if $\beta_7^r = 1$ and $\beta_8^r = 1$). Considering the uncertainties associated with attackers' skill levels and process and observation noises, Monte Carlo simulations with 10,000 trials were performed to obtain the probability distribution of successful execution of each attack mode following the algorithm presented in Fig. 5. A uniform distribution (a ratio 1:1:1:1) was used for potential attackers with different skill levels (expert, intermediate, beginner, and novice). In practice, it may be configured based on expert judgment considering possible threat agents. The calculation results of the vulnerability assessment model are presented in Fig. A4.

Safety barrier maintenance time intervals were initialed as one year in PFD calculations by using the reliability models in Section 2.3.2.



(a) Probability distribution of the occurrence frequency of fireballs

(b) Probability distribution of the occurrence frequency of explosions

(c) Probability distribution of the occurrence frequency of cloud fires

(d) Probability distribution of the occurrence frequency of toxic dispersion

Fig. 9. Probability distributions of the occurrence frequency of each consequence.

Then, the prior probabilities/probability distributions for all root nodes are summarized in Table 6, based on which the Bayesian inference was performed to obtain the probability distributions of the occurrence of each consequence, as shown in Fig. 9. Finally, the mean values and ranges of the obtained probability distributions are visualized in a risk matrix to present risk profiles considering parameter uncertainties, as shown in Fig. 10.

3.3. Decision making on safety and security barrier improvements

Because the calculated "risk ranges" overlap with the red region in the risk matrix (in Fig. 10), necessary improvements must be made to safety and security barriers. Based on the developed BN model, a sensitivity analysis of the basic events (root nodes) regarding the happening of disastrous consequences was performed, and it may help to propose candidate strategies for barrier improvements. The results obtained from the sensitivity analysis are shown in Fig. 11, in which BE22 (SV safety failure) has the dominant importance (with an importance measure value of 0.0207). BE16 to BE 21 also have relatively high sensitivities, with importance measure values around 4E-4. Other basic events have importance measure values below 2E-4. BE16, BE17, BE19, and BE20 are related to C2P attacks on the ESD system, and BE18 and BE21 are related to safety failures of the ESD system components. Potential strategies may be proposed for the safety or security protection of the safety relief valve (SV) and the ESD system.

Vulnerability patching is a crucial way to remove vulnerabilities from an IT system by delivering security patches [70], and it is regarded/called a security barrier in this study. It is assumed that the security management team has the capability to patch vulnerabilities CVE-2016–2200, which is closely related to DoS attacks on PLCs. Meanwhile, considering the maintenance scheduling of the ESD system and the safety relief valve, a cost analysis of the potential barrier improvement actions is given in Table 7. Considering also the technical constraints and practicability of the possible strategies, 18 schemes are proposed as candidate strategies for barrier improvements, as concluded in Table 8. According to Eq (15), the optimization problem is

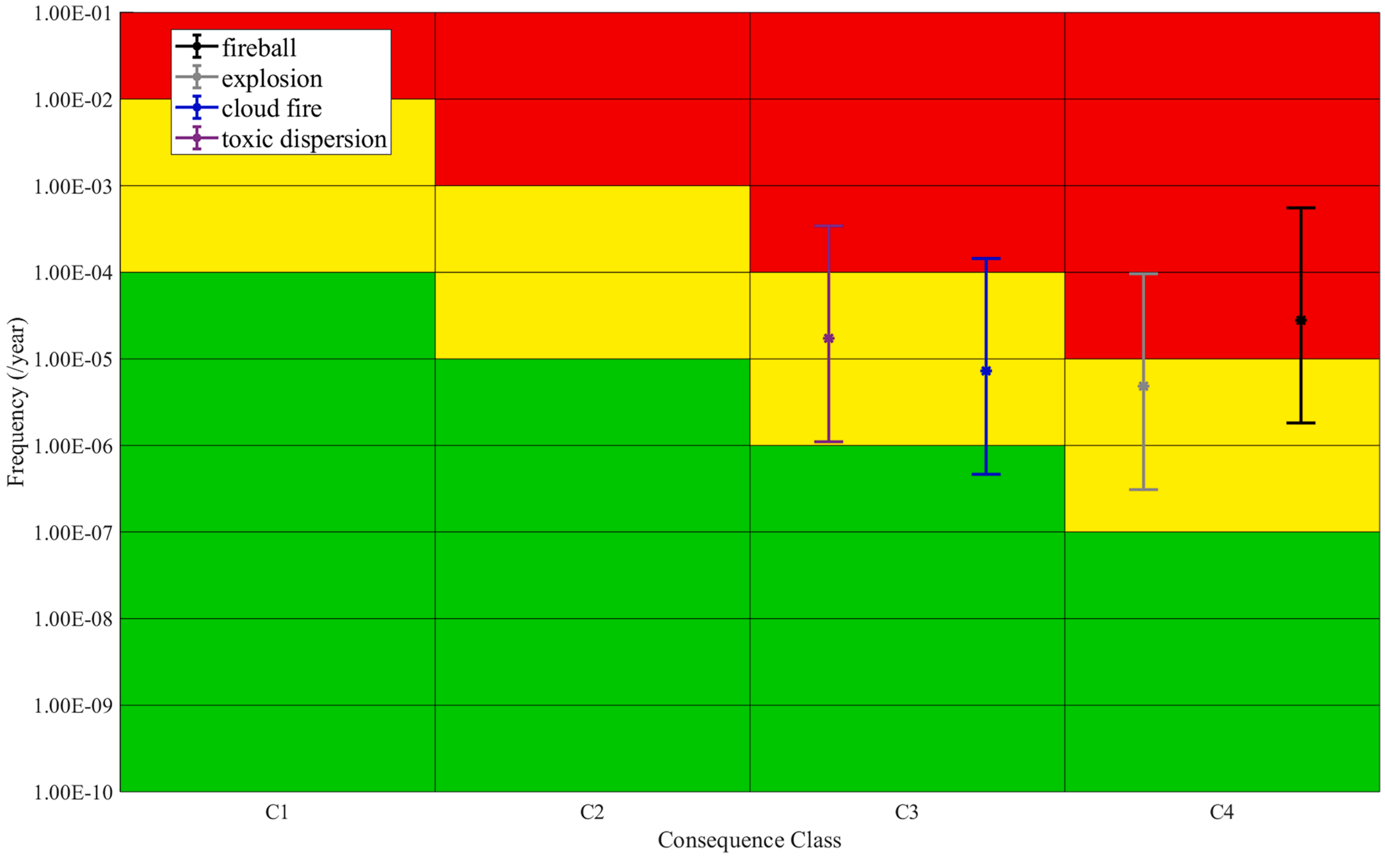


Fig. 10. Estimated risks with ranges are demonstrated in a risk matrix (the risk matrix is adapted from Andersen et al. [51]).

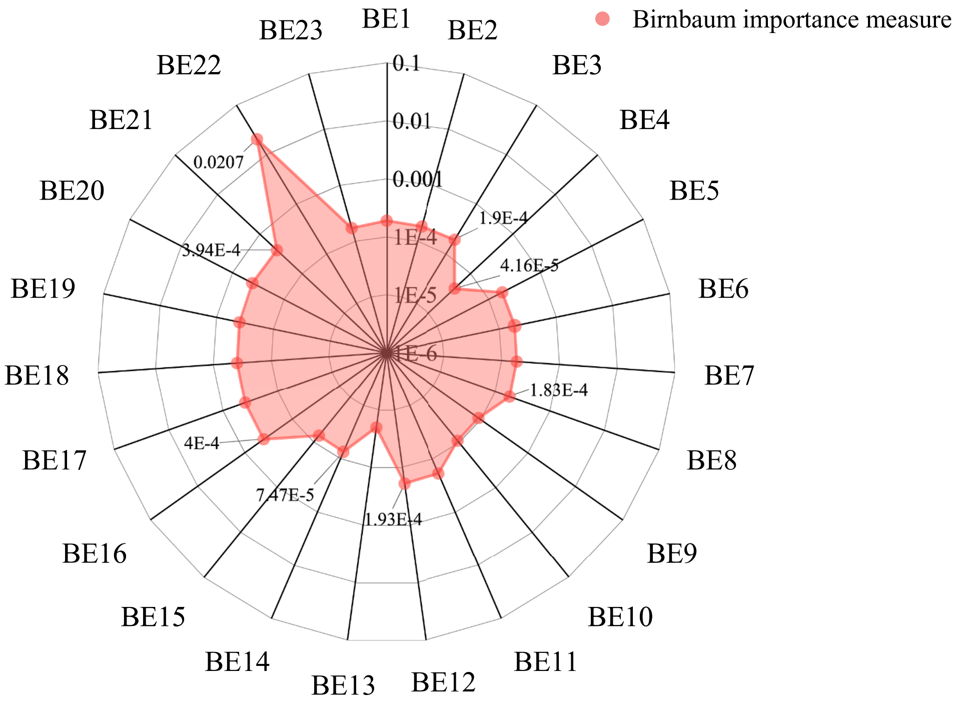


Fig. 11. Sensitivity analysis of root nodes.

Table 7
Cost analysis of safety and security barrier improvement actions.

Improvement actions	Cost analysis[1]	Total costs
Maintenance of the ESD system	10,000€ (one-time maintenance cost); 100,000€×2 days (downtime cost)	210,000€
Maintenance of the safety relief valve	2000€(one-time maintenance cost); 100,000€×1 day (downtime cost)	102,000€
Patch vulnerability CVE-2016–2200	20,000€(patching cost); 100,000€×14 days (downtime cost)[2]	1420,000€

[1]Approximate costs are used for the perfect maintenance (replacement) because the cost varies for specific equipment.
[2]The time needed for vulnerability patching also varies. The downtime is estimated referring to an optimistic Mean Time To Patch (MTTP) for cost saving. https://purplesec.us/learn/average-time-patch-vulneraiblity/.

Table 8
Proposed candidate strategies for improving the performance of safety and security barriers.

No.	Strategy details	No.	Strategy details	No.	Strategy details
1 (no improvement)	• No vulnerability patching • Maintenance interval for ESD: one year • Maintenance interval for SV: one year	2	• No vulnerability patching • Maintenance interval for ESD: six months • Maintenance interval for SV: one year	3	• No vulnerability patching • Maintenance interval for ESD: three months • Maintenance interval for SV: one year
4	• No vulnerability patching • Maintenance interval for ESD: one year • Maintenance interval for SV: six months	5	• No vulnerability patching • Maintenance interval for ESD: six months • Maintenance interval for SV: six months	6	• No vulnerability patching • Maintenance interval for ESD: three months • Maintenance interval for SV: six months
7	• No vulnerability patching • Maintenance interval for ESD: one year • Maintenance interval for SV: three months	8	• No vulnerability patching • Maintenance interval for ESD: six months • Maintenance interval for SV: three months	9	• No vulnerability patching • Maintenance interval for ESD: three months • Maintenance interval for SV: three months
10	• Patch CVE-2016–2200 • Maintenance interval for ESD: one year • Maintenance interval for SV: one year	11	• Patch CVE-2016–2200 • Maintenance interval for ESD: six months • Maintenance interval for SV: one year	12	• Patch CVE-2016–2200 • Maintenance interval for ESD: three months • Maintenance interval for SV: one year
13	• Patch CVE-2016–2200 • Maintenance interval for ESD: one year • Maintenance interval for SV: six months	14	• Patch CVE-2016–2200 • Maintenance interval for ESD: six months • Maintenance interval for SV: six months	15	• Patch CVE-2016–2200 • Maintenance interval for ESD: three months • Maintenance interval for SV: six months
16	• Patch CVE-2016–2200 • Maintenance interval for ESD: one year • Maintenance interval for SV: three months	17	• Patch CVE-2016–2200 • Maintenance interval for ESD: six months • Maintenance interval for SV: three months	18	• Patch CVE-2016–2200 • Maintenance interval for ESD: three months • Maintenance interval for SV: three months

characterized as minimizing the cost of barrier improvements while ensuring the accident risks are within the acceptable thresholds. Because the threshold for fireball risks is the most difficult one to meet, the estimated mean values and maximum values of the occurrence frequency of fireballs after implementing each candidate strategy are compared in Fig. 12. Because the proposed approach can present accident risks in the form of risk ranges, two thresholds (threshold A and threshold B) are used for the expected/mean value and maximum value of the estimated fireball risks respectively. Combined with the comparison of the cost of each strategy in Fig. 12, the optimal strategy may be determined. As shown in Fig. 12, the expected values of fireball risks after implementing strategies 7 to 9 and 16 to 18 are below threshold A, and strategy 7 has the lowest cost. Therefore, strategy 7 is most cost-effective when only the expected/mean value of fireball risks is

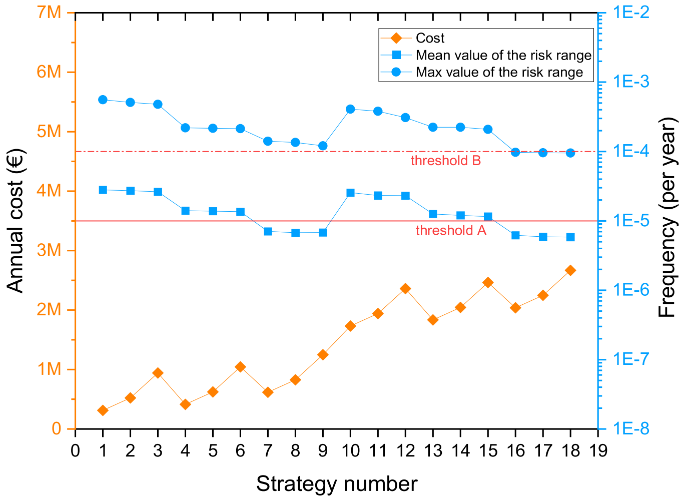


Fig. 12. Risk profiles after implementing candidate strategies (fireball risk as example).

considered decision-making criteria. By contrast, max values of fireball risks meet threshold B only after implementing strategies 16 to 18, and strategy 16 has the lowest cost. As a result, strategy 16 is the optimal strategy when both the mean value and maximum value of uncertain risks are configured as decision-making criteria.

4. Discussions

4.1. How uncertainty treatment helps decision-making on barrier management

This study proposed a systematic approach for cost-effective safety and security barrier management based on integrated safety and security risks. Considering the noticeable uncertainties in the integrated safety and security risks, multiple uncertain parameters are considered. For instance, uncertainties associated with attackers' skill levels in cyber intrusion risks are assessed using Monte Carlo simulations. The combination of a Bayesian network and Monte Carlo simulations enables the use of probability distributions and handles uncertainty propagation in the risk assessment. As a result, the risk assessment results, which are in the form of probability distributions regarding the happening of each possible consequence, can provide more insights for the decision-making on barrier management. As illustrated in Fig. 10, the obtained risk profiles considering parameter uncertainties can be mapped in a risk matrix demonstrating not only the expected risk values but also the risk ranges. Subsequently, risks can be evaluated according to the expected values and the risk ranges, and different thresholds may be set up for the expected/mean values and maximum values of the risk ranges to decide if the risks are acceptable. During the decision-making process, both the expected/mean values and the maximum values of the risk ranges can be used as criteria to determine appropriate strategies for safety and security barrier improvement. The expected/mean values of risk ranges correspond to conventional risk values usually used in risk management decision-making. By contrast, the maximum values of risk ranges reflect worst-case risks considering parameter uncertainties, and they may be used as criteria when the worst cases are concerned by decision-makers. Therefore, with appropriate treatment and handling of parameter uncertainties, more criteria (optimization objectives) can be used to guide decision-making on safety and security barrier management according to the needs and interests of decision-makers.

4.2. Limitations and recommendations for future studies

The developed approach incorporates a relatively thorough list of parameters/factors in the risk assessment. Some approximate assumptions or referenced values are currently used for some variables due to the lack of data. For instance, a reference MTTD value was used for all C2P attack modes, and values of $\beta_{\mathrm{i}}^{\mathrm{d}}$ in Table 3 were configured based on the assumption of a simple fault/anomaly detection scheme. The configuration of those parameters may be improved according to the actual intrusion detection capability of the security operations center. Some technical parameters (true positive rate, false alarm rate, etc.) of the deployed fault/anomaly detection algorithm may help determine those parameters when related data is available. Additionally, the configuration of the attack likelihood and the probability distribution of attackers with different skill levels depends on subjective judgment in the present study. The incorporation of more incident data and evidence in the configuration of those parameters helps to obtain more credible risk assessment results.

Regarding the vulnerability assessment of C2P attacks, a modified TTC (time-to-compromise) estimation approach proposed by Ling & Ekstedt [18] is employed with the help of the ICS-specific vulnerability dataset [55]. Compared to the original TTC estimation approach [12], the modified TTC method has the advantage of specializing in the vulnerability assessment of industrial control systems (ICSs), and it leverages the exploitability scores of CVEs (Common Vulnerabilities and Exposures) into the vulnerability assessment. However, the ICS-specific vulnerability dataset is currently not maintained and updated in a timely manner, which hinders the leverage of newly published CVEs into the vulnerability assessment model. The timely updating and maintenance of the dataset helps incorporate newly published CVEs associated with C2P attacks on ICSs and enhances the practicality of the proposed approach.

Moreover, several typical C2P attack modes (as shown in Table 2) are investigated in this study. Considering the emergence of new attack modes, the characterization of those attack modes and the accommodation of the new attack modes in the vulnerability assessment are necessary. It should also be noted that the TTC-based method used in this study is only capable of assessing the ICS vulnerability to cyber intrusions. Some attack modes cannot be handled by the TTC-based method, for instance, the physical attacks on the IT system. The vulnerability and risk assessment of more C2P attack modes is worth investigating in future studies.

5. Conclusions

An integrated approach is proposed to bolster integrated safety and security barrier management for C2P attack risks. A case study featuring a prototypical industrial control system is executed to demonstrate the efficacy of the proposed approach. Major accident risks emanating from the ICS, attributed to safety-related factors and C2P attacks, are evaluated, considering multiple uncertain parameters. The outcomes are visually represented in a risk matrix. Conducting a sensitivity analysis on fundamental events reveals safety relief valves and emergency shutdown (ESD) systems as pivotal components. Safety relief valves are only subject to safety failures, whereas critical failure of ESD systems may result from either safety failures or C2P attacks. Consequently, eighteen potential strategies for enhancing safety and security barriers are formulated, incorporating considerations of maintenance scheduling for safety barriers and security vulnerability patching. Conducting a cost-effectiveness analysis with the help of the derived risk matrix, the optimal strategy is discerned, taking into consideration the expected values and maximum values of risk estimations, along with the associated costs. The optimization results reveal that the expected values of risk estimations not only form a foundational element for risk-based decision-making but also that the risk ranges provide supplementary insights, facilitating decision-making that accounts for inherent uncertainties in risk assessments.

CRediT authorship contribution statement

Shuaiqi Yuan: Writing – original draft, Visualization, Methodology, Investigation, Data curation, Conceptualization. **Genserik Reniers:** Writing – review & editing, Validation, Supervision, Formal analysis. **Ming Yang:** Writing – review & editing, Supervision, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work is supported by the China Scholarship Council (Grant No: 202006430007).

Appendix I (Method for Time-to-compromise (TTC) estimation)

A random attack process composed of three subprocesses is assumed in the TTC estimation approach: i) at least one vulnerability is known, and the attacker has the capability to exploit the known vulnerability successfully, ii) at least one vulnerability is known, but the attacker needs develop an exploit for it, and iii) the attacker has to find and exploit new vulnerabilities because either no known vulnerabilities exist or the attacker is unable to exploit the known vulnerabilities. TTC of an attack step is estimated as follows [12,18]:

$$TTC = t_1 P_1 + t_2(1 - P_1)(1 - u) + t_3 u(1 - P_1) \tag{A1}$$

where t_i is the expected time spent in subprocess i (i= *1, 2, 3*) in days. P_1 is the probability of being in subprocess 1. u is the probability that subprocess 2 fails. The probabilities of an attacker being in subprocess 1 and subprocess 2 are calculated as follows [18].

$$P_1 = 1 - e^{-vm/k} \tag{A2}$$

$$P_2 = e^{-vm/k} = 1 - P_1 \tag{A3}$$

where v is the number of vulnerabilities at an attack step and m is the number of exploits readily available to the attacker. k is the total number of vulnerabilities in the database. The value of k is 2740 according to the ICS vulnerability dataset [55] available on October 5th, 2023. Subprocess 3 is assumed running in parallel to subprocess 1 and 2. The time spent to complete each subprocess is estimated as below [18].

$$t_1 = 1 * ((10 / C_2 + 3.9 / C_3)2) \tag{A4}$$

$$t_2 = 37 \text{ (novice), } 27 \text{ (beginner), } 16 \text{ (intermediate), or } 6 \text{ (expert)} \tag{A5}$$

$$t_3 = (f' - 0.5) * b + t_2 \tag{A6}$$

where C_2 is the average base score of the vulnerabilities derived from CVSS v2.0[1] and C_3 is the average exploitability score of the vulnerabilities derived from CVSS v3.0.[2] Considering the determination of t_2, 37 days, 27 days, 16 days, and 6 days are used for novice, beginner, intermediate, and expert attackers respectively. b is the Mean-Time-Between-Vulnerabilities (MTBV) in days as calculated from the ICS advisory creation date [55]. f is the fraction of vulnerabilities that are exploitable to the attacker, which is determined in Table A1. The probability that subprocess 2 is unsuccessful (u) is calculated as $u = (1 - f)^v$. An Excel tool[3] developed by Thomas & Chothia [55] was used to conduct the TTC estimations.

Table A1
The number and fraction of exploitable vulnerabilities to attackers with different skill levels, adapted from [18].

Skill level	CVSS exploitability range	Exploitable vulnerabilities	Fraction of exploitable vulnerabilities
Expert	0.1–3.9	1916	1
Intermediate	0.1–3	966	0.50
Beginner	0.1–2.1	455	0.24
Novice	0.1–1.2	105	0.05

After calculating the TTC of each attack step, the global time-to-compromise of each attack path ($GTTC_i$) may be calculated based on a compromise graph. For instance, a compromise graph regarding the ICS in Fig. 6 is established and presented in Fig. A1. The vulnerabilities at each attack step are presented in Table A2. Each ending node (AT1-AT9) of the compromise graph presents an attack mode, as illustrated in Table 4. For attack path 1, which is highlighted in red color, the global time-to-compromise is calculated as: $GTTC_1 = TTC_1 + TTC_2 + TTC_4$. (Fig. A1)

[1] CVSS v2.0 user guide. (n.d.). Retrieved October 06, 2023, from https://www.first.org/cvss/v2/guide.

[2] CVSS v3.0 user guide. (n.d.). Retrieved October 06, 2023, from https://www.first.org/cvss/v3.0/user-guide.

[3] TTC-ICS. Retrieved October 06, 2023, from https://github.com/EngLi/ttc-ics.

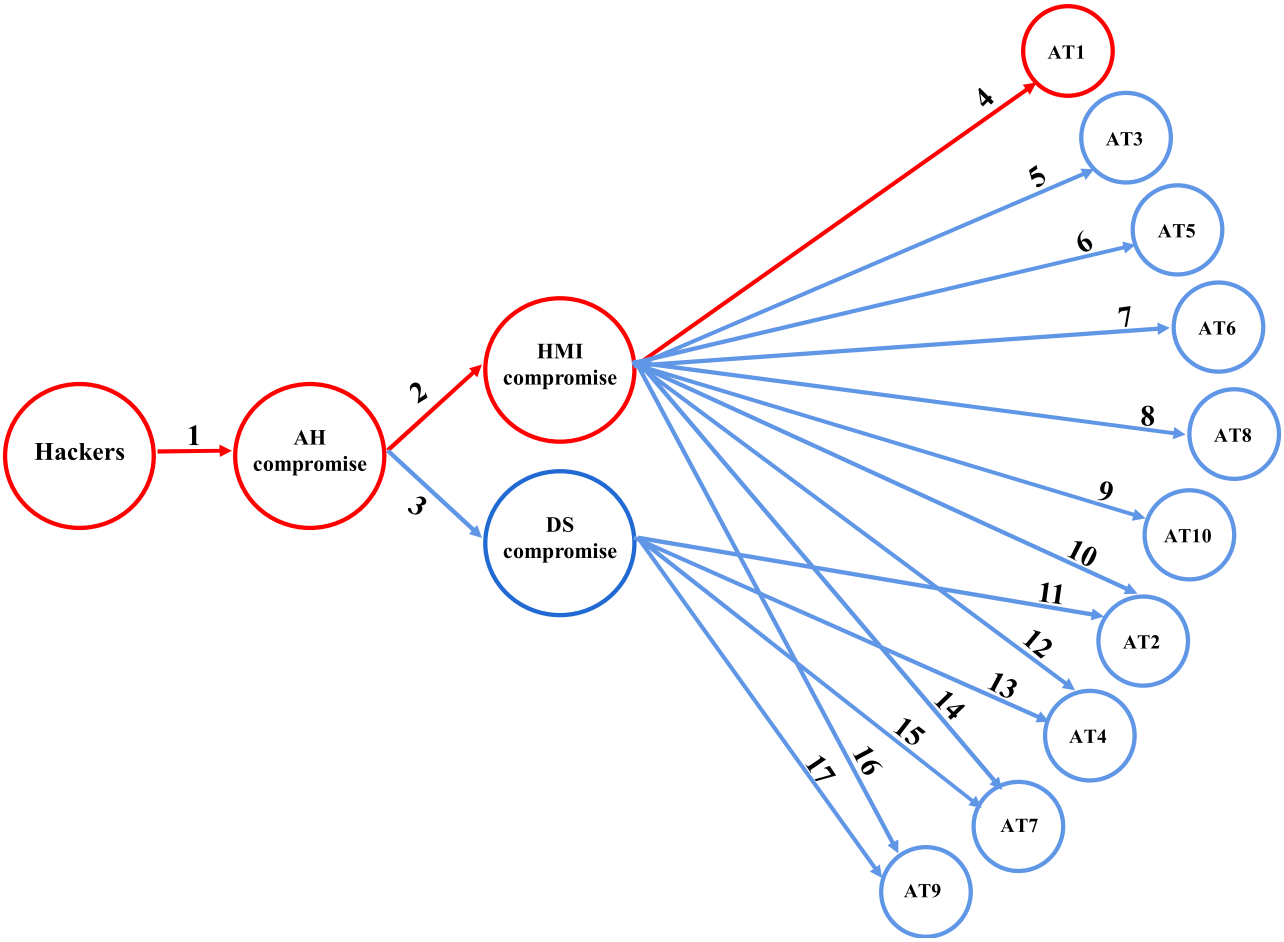


Fig. A1. A compromise graph regarding the investigated ICS.

Table A2
Known vulnerabilities at each attack step.

Attack step number	Known vulnerabilities (cve_id[1])	Attack step number	Known vulnerabilities (cve_id)
1	CVE-2015–7871; CVE-2017–2683	2	CVE-2017–13,997
3	CVE-2018–13,799	4	no
5	no	6	CVE-2018–5459
7	no	8	no
9	CVE-2018–5459	10	CVE-2016–2200
11	CVE-2016–2200	12	CVE-2016–2200
13	CVE-2016–2200	14	CVE-2016–2200
15	CVE-2016–2200	16	CVE-2016–2200
17	CVE-2016–2200	/	/

[1] A cve_id uniquely identifies one vulnerability from the Common Vulnerabilities and Exposures (CVE) database [71].

Appendix II (C2P attack modeling)

Fig. A2 shows the C2P attack assessment model, which is developed based on the MATLAB/Simulink platform. The deviations caused by different C2P attacks on reactor temperature are demonstrated in Fig. A3. It is considered a dangerous overheating scenario when the reactor temperature overpasses 450 K. Consequently, the β_i^r value for each C2P attack can be determined ($\beta_i^r = 1$ when $T(k) > 450\ K$, otherwise, $\beta_i^r = 0$). Regarding C2P attacks against ESD systems, $\beta_i^r = 1$ is always the case because a manipulation of the ESD system can stop the ESD system from performing its functionality on demand, no matter the process noise and observation noise [33]. Monte Carlo simulations are used to assess the probability distribution of successful execution of each attack mode based on this C2P attack model following the algorithm presented in Fig. 5. The calculation results are presented in Fig. A4.

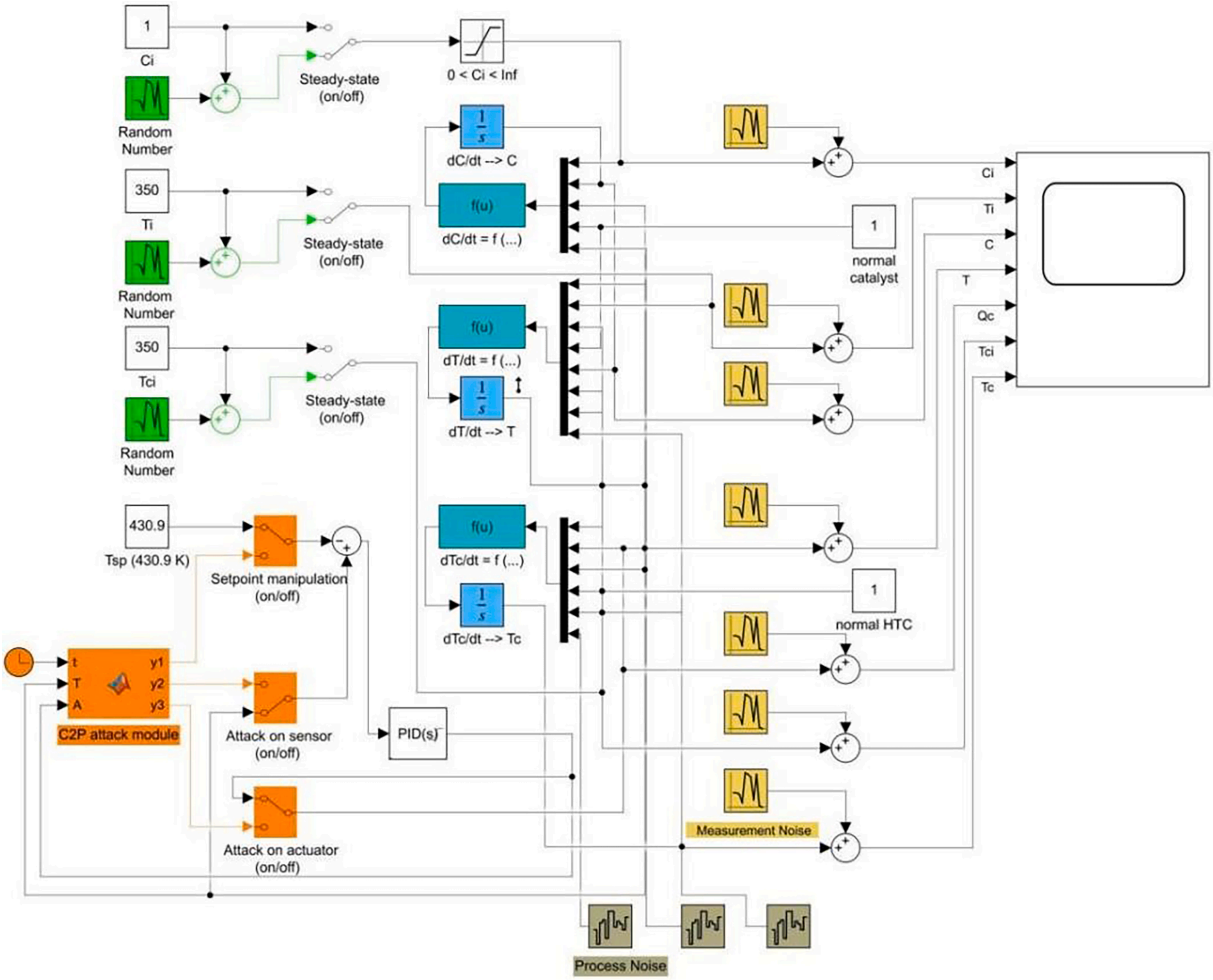
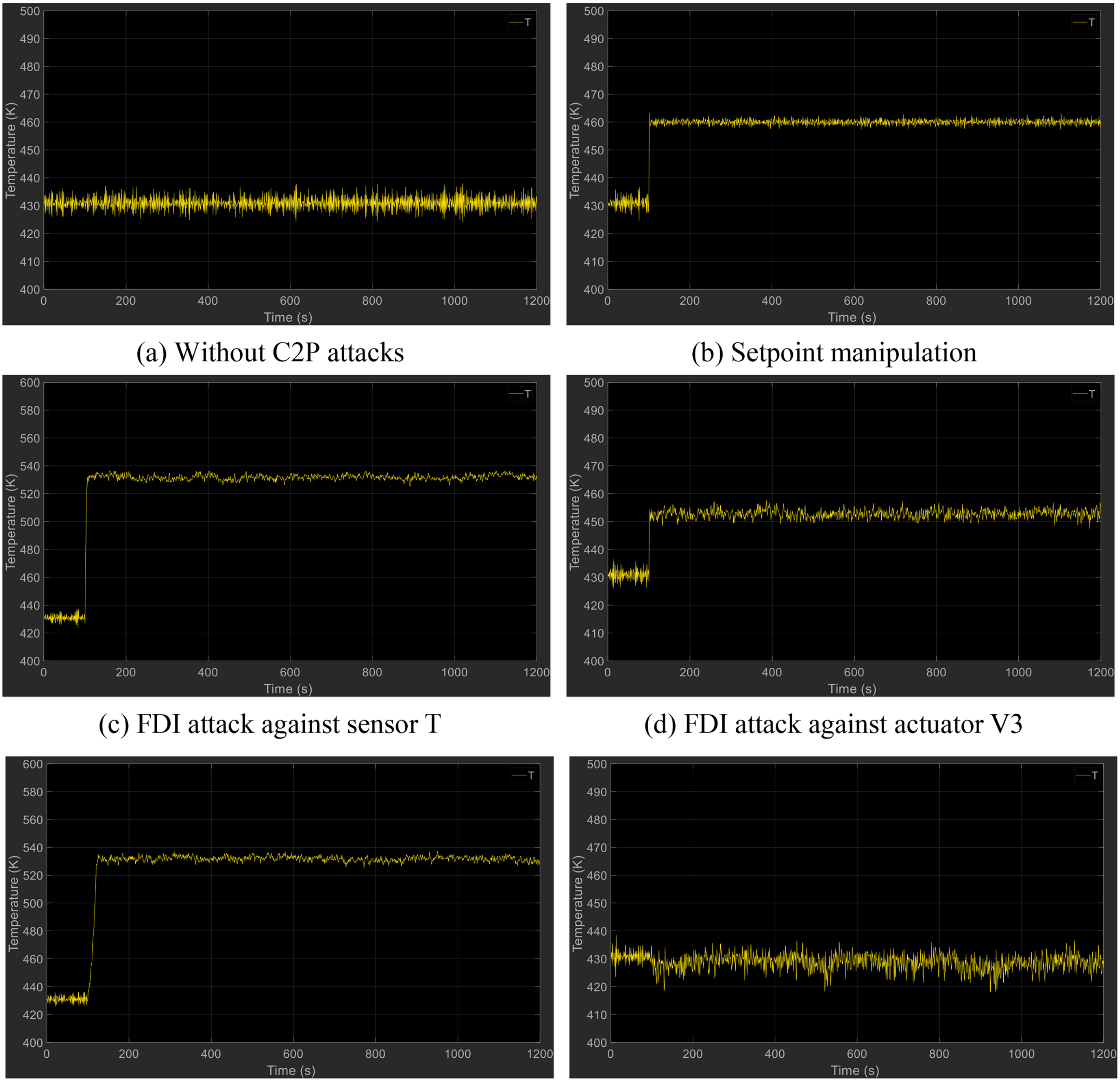


Fig. A2. A C2P attack assessment model developed based on MATLAB/Simulink.



(a) Without C2P attacks

(b) Setpoint manipulation

(c) FDI attack against sensor T

(d) FDI attack against actuator V3

(e) DoS attack against sensor T

(f) DoS attack against actuator V3

Fig. A3. C2P attacks' effects on the reactor temperature (attacks start from 100 s).

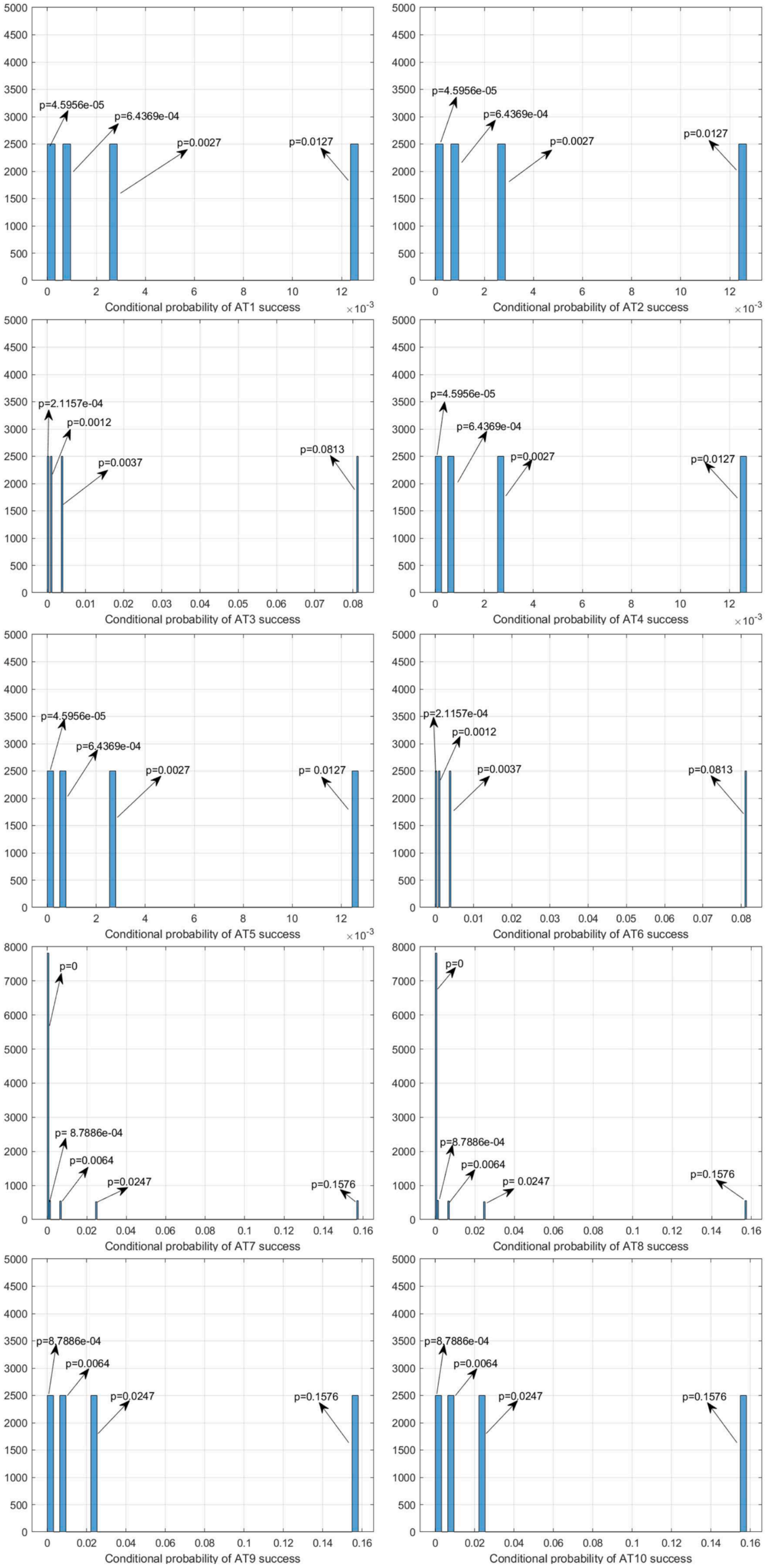


Fig. A4. Conditional probabilities of successful execution of each attack mode (p in the figures means conditional probability).

References

[1] Derler P, Lee EA, Vincentelli AS. Modeling cyber-physical systems. Proc. IEEE 2011;100(1):13–28.

[2] Ji X, He G, Xu J, Guo Y. Study on the mode of intelligent chemical industry based on cyber-physical system and its implementation. Adv. Eng. Softw. 2016;99:18–26.

[3] Xing L, Distefano S. Reliability and performance of cyber-physical systems. Reliab. Eng. Syst. Saf. 2022;225:108642.

[4] Flaus JM. Cybersecurity of Industrial Systems. John Wiley & Sons; 2019.

[5] Kushner D. The real story of stuxnet. IEEE Spectr. 2013;50(3):48–53.

[6] Hemsley, K.E., & Fisher, E. (2018). History of industrial control system cyber incidents (No. INL/CON-18-44411-Rev002). Idaho National Lab.(INL), Idaho Falls, ID (United States).

[7] Di Pinto A, Dragoni Y, Carcano A. TRITON: the first ICS cyber attack on safety instrument systems. Proc. Black Hat USA 2018;2018:1–26.

[8] Monzer MH, Beydoun K, Ghaith A, Flaus JM. Model-based IDS design for ICSs. Reliab. Eng. Syst. Saf. 2022;225:108571.

[9] Wu S, Jiang Y, Luo H, Zhang J, Yin S, Kaynak O. An integrated data-driven scheme for the defense of typical cyber–physical attacks. Reliab. Eng. Syst. Saf. 2022;220:108257.

[10] Guzman NHC, Kozine I, Lundteigen MA. An integrated safety and security analysis for cyber-physical harm scenarios. Saf. Sci. 2021;144:105458.

[11] Paul, S. (2015). On the meaning of security for safety (s4s). In: Safety and Security Engineering Vi, pp. 379–89. https://doi.org/10.2495/safe150321.

[12] McQueen MA, Boyer WF, Flynn MA, Beitel GA. Time-to-compromise model for cyber risk reduction estimation. Quality of Protection: Security Measurements and Metrics. US: Springer; 2006. p. 49–64.

[13] McQueen MA, Boyer WF, Flynn MA, Beitel GA. Quantitative cyber risk reduction estimation methodology for a small SCADA control system. In: Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06). 9. IEEE; 2006. 226-226.

[14] Semertzis I, Rajkumar VS, Ştefanov A, Fransen F, Palensky P. Quantitative risk assessment of cyber attacks on cyber-physical systems using attack graphs. 2022 10th workshop on modelling and simulation of cyber-physical energy systems (MSCPES). IEEE; 2022. p. 1–6.

[15] Zhang Y, Wang L, Xiang Y, Ten CW. Power system reliability evaluation with SCADA cybersecurity considerations. IEEe Trans. Smart. Grid. 2015;6(4):1707–21.

[16] Huang K, Zhou C, Tian YC, Yang S, Qin Y. Assessing the physical impact of cyberattacks on industrial cyber-physical systems. IEEE Trans. Industr. Electr. 2018;65(10):8153–62.

[17] Poolsappasit N, Dewri R, Ray I. Dynamic security risk management using Bayesian attack graphs. IEEE Trans. Dependable Secure Comput. 2011;9(1):61–74.

[18] Ling, E.R., & Ekstedt, M. (2022). Estimating the Time-To-Compromise of Exploiting Industrial Control System Vulnerabilities. In ICISSP (pp. 96–107).

[19] Chen Y, Hong J, Liu CC. Modeling of intrusion and defense for assessment of cyber security at power substations. IEEe Trans. Smart. Grid. 2016;9(4):2541–52.

[20] Orojloo H, Azgomi MA. A game-theoretic approach to model and quantify the security of cyber-physical systems. Comput. Ind. 2017;88:44–57.

[21] Lalropuia KC, Gupta V. Modeling cyber-physical attacks based on stochastic game and Markov processes. Reliab. Eng. Syst. Saf. 2019;181:28–37.

[22] Huang YL, Cárdenas AA, Amin S, Lin ZS, Tsai HY, Sastry S. Understanding the physical and economic consequences of attacks on control systems. Int. J. Crit. Infrastruct. Protect. 2009;2(3):73–83.

[23] Li X, Zhou C, Tian YC, Xiong N, Qin Y. Asset-based dynamic impact assessment of cyberattacks for risk analysis in industrial control systems. IEEe Trans. Industr. Inform. 2017;14(2):608–18.

[24] Patriarca R, Simone F, Di Gravio G. Modelling cyber resilience in a water treatment and distribution system. Reliab. Eng. Syst. Saf. 2022;226:108653.

[25] Abdo H, Kaouk M, Flaus JM, Masse F. A safety/security risk analysis approach of Industrial control systems: a cyber bowtie–combining new version of attack tree with bowtie analysis. Comput. Secur. 2018;72:175–95.

[26] Haddon Jr W. Energy damage and the ten countermeasure strategies. Hum. Factors 1973;15(4):355–66.

[27] Moreno VC, Marroni G, Landucci G. Probabilistic assessment aimed at the evaluation of escalating scenarios in process facilities combining safety and security barriers. Reliab. Eng. Syst. Saf. 2022;228:108762.

[28] Yuan S, Reniers G, Yang M, Bai Y. Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. Process Safety Environ. Protect. 2023;170:356–71.

[29] Øien K, Hauge S, Jaatun MG, Flå L, Bodsberg L. A Survey on Cybersecurity Barrier Management in Process Control Environments. In: 2022 IEEE International Conference on Cloud Computing Technology and Science (CloudCom). IEEE; 2022. p. 113–20.

[30] Yuan S, Yang M, Reniers G, Chen C, Wu J. Safety barriers in the chemical process industries: a state-of-the-art review on their classification, assessment, and management. Saf. Sci. 2022;148:105647.

[31] Reniers G, Khakzad N. Revolutionizing safety and security in the chemical and process industry: applying the CHESS concept. J. Integrated Secur.Safety Sci. 2017;1(1):2–15.

[32] Yuan S, Reniers G, Yang M. Dynamic-risk-informed safety barrier management: an application to cost-effective barrier optimization based on data from multiple sources. J. Loss. Prev. Process. Ind. 2023;83:105034.

[33] Yuan S, Yang M, Reniers G. Integrated process safety and process security risk assessment of industrial cyber-physical systems in chemical plants. Comput. Ind. 2024;155:104056.

[34] Meng X, Zhu J, Chen G, Shi J, Li T, Song G. Dynamic and quantitative risk assessment under uncertainty during deepwater managed pressure drilling. J. Clean. Prod. 2022;334:130249.

[35] Xu Y, Reniers G, Yang M, Yuan S, Chen C. Uncertainties and their treatment in the quantitative risk assessment of domino effects: classification and review. Process Safety Environ. Protect. 2023;172:971–85.

[36] Bier VM, Lin SW. On the treatment of uncertainty and variability in making decisions about risk. Risk Anal. 2013;33(10):1899–907.

[37] CCPS/EI. Bow Ties in Risk Management. Center for Chemical Process Safety and Energy Institute (UK), Wiley - AIChE; 2018.

[38] Gribaudo M, Iacono M, Marrone S. Exploiting Bayesian networks for the analysis of combined attack trees. Electron. Notes. Theor. Comput. Sci. 2015;310:91–111.

[39] Landucci G, Khakzad N, Reniers G. Physical Security in the Process Industry: Theory with Applications. Elsevier; 2020.

[40] Ericson, C.A. (2005), Hazard Analysis Techniques for System Safety. Published by John Wiley & Sons, Inc. https://doi, 10, 0471739421.

[41] Khakzad N, Khan F, Amyotte P. Safety analysis in process facilities: comparison of fault tree and Bayesian network approaches. Reliab. Eng. Syst. Saf. 2011;96(8):925–32.

[42] Chen XL, Lin WD, Liu CX, Yang FQ, Guo Y, Li X, Reniers G. An integrated EDIB model for probabilistic risk analysis of natural gas pipeline leakage accidents. J. Loss. Prev. Process. Ind. 2023;83:105027.

[43] Tong X, Fang W, Yuan S, Ma J, Bai Y. Application of Bayesian approach to the assessment of mine gas explosion. J. Loss. Prev. Process. Ind. 2018;54:238–45.

[44] Jensen FV, Nielsen TD. Bayesian Networks and Decision Graphs (Vol. 2). New York: Springer; 2007.

[45] Khakzad N, Khan F, Amyotte P. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. Process Safety Environ. Protect. 2013;91(1–2):46–53.

[46] Hauge S, Onshus T. Reliability data for safety instrumented systems PDS data handbook, 2010 edition. SINTEF Report A 2010:13502.

[47] Kirwan B. A Guide to Practical Human Reliability Assessment. CRC press; 2017.

[48] Debray, B., Piatyszek, E., Cauffet, F., & Londiche, H. (2004). Frequencies and Probabilities Data for the Fault Tree Accidental Risk Assessment Methodology for Industries in the Framework of seveso ii directive (ARAMIS), armines, École Nationale Supérieure de Mines de Saint Etienne, France, 100.

[49] IEC, 2010. IEC 61508 Functional safety of electrical/electronic/programmable electronic safety-related systems. IEC Standards Online.

[50] Yuan S, Cai J, Reniers G, Yang M, Chen C, Wu J. Safety barrier performance assessment by integrating computational fluid dynamics and evacuation modeling for toxic gas leakage scenarios. Reliab. Eng. Syst. Saf. 2022;226:108719.

[51] Andersen, H., Casal, J., Dandrieux, A., Debray, B., De Dianous, V., Duijm, N., Gowland, R. (2004). ARAMIS user guide. *EC Contract number EVG1-CT-2001-00036*.

[52] Zhang A, Zhang T, Barros A, Liu Y. Optimization of maintenances following proof tests for the final element of a safety-instrumented system. Reliab. Eng. Syst. Saf. 2020;196:106779.

[53] IEC, 2016. Functional Safety – Safety Instrumented Systems for the Process Industry Sector, Gen`eve, Switzerland (IEC).

[54] Schmitz P, Swuste P, Reniers G, van Nunen K. Predicting major accidents in the process industry based on the barrier status at scenario level: a practical approach. J. Loss. Prev. Process. Ind. 2021;71:104519.

[55] Thomas RJ, Chothia T. Learning from vulnerabilities - categorising, understanding and detecting weaknesses in industrial control systems. Computer security. Cham: Springer International Publishing; 2020.

[56] Bier V, Gutfraind A. Risk analysis beyond vulnerability and resilience–characterizing the defensibility of critical systems. Eur. J. Oper. Res. 2019;276(2):626–36.

[57] Van der Borst M, Schoonakker H. An overview of PSA importance measures. Reliab. Eng. Syst. Saf. 2001;72(3):241–5.

[58] Yazdi M, Kabir S. A fuzzy Bayesian network approach for risk analysis in process industries. Process safety Environ. Protect. 2017;111:507–19.

[59] Zarei E, Azadeh A, Khakzad N, Aliabadi MM, Mohammadfam I. Dynamic safety assessment of natural gas stations using Bayesian network. J. Hazard. Mater. 2017;321:830–40.

[60] Reniers GL, Van Erp HN. Operational safety economics: a practical approach focused on the chemical and process industries. John Wiley & Sons; 2016.

[61] Pilario KES, Cao Y. Canonical variate dissimilarity analysis for process incipient fault detection. IEEe Trans. Industr. Inform. 2018;14(12):5308–15.

[62] Vílchez JA, Espejo V, Casal J. Generic event trees and probabilities for the release of different types of hazardous materials. J. Loss. Prev. Process. Ind. 2011;24(3):281–7.

[63] Murphy K. The bayes net toolbox for matlab. Comput. Sci. Stat. 2001;33(2):1024–34.

[64] Kuypers M, Maillart T. Designing organizations for cyber security resilience. In: Proceedings of the 2018 The Workshop on the Economics of Information Security (WEIS); 2018. p. 18–9. Innsbruck, Austria.

[65] Eide, S.A., Wierman, T.E., Gentillon, C.D., Rasmuson, D.M., & Atwood, C.L. Industry-Average Performance for Components and Initiating Events at US Commercial Nuclear Power Plants; NUREG/CR-6928; Nuclear Regulatory Commission: Washington, DC, USA, 2007.

[66] Taylor, J.R. (2010). The QRAQ Project Volume 4: frequency of Releases and Accidents.https://www.academia.edu/35376294/The_QRAQ_Project_Volume_4_Frequency_of_Releases_and_Accidents. (accessible 2023, November).

[67] OREDA. Offshore Reliability Data Handbook. Trondheim, Norway: DNV; 2002.

[68] Roy A, Srivastava P, Sinha S. Dynamic failure assessment of an ammonia storage unit: a case study. Process Safety Environ. Protect. 2015;94:385–401.

[69] HSE, U. (2012). Failure Rate and Event Data for use within Risk Assessments (28/06/2012).

[70] Hong JB, Kim DS, Haqiq A. What vulnerability do we need to patch first?. In: 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. IEEE; 2014. p. 684–9.

[71] National Vulnerability Database (NVD). (n.d.). Retrieved November 24, 2023, from https://nvd.nist.gov/.