Dynamic and integrated safety and security barrier management

A new framework to manage major event risks in chemical plants

Yuan, Shuaiqi; Reniers, Genserik; Yang, Ming

DOI
10.1016/j.jlp.2025.105632

Publication date
2025

Document Version
Final published version

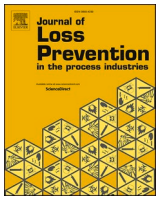
Published in
Journal of Loss Prevention in the Process Industries

Journal of Loss Prevention in the Process Industries 96 (2025) 105632

Contents lists available at ScienceDirect

Journal of Loss Prevention in the Process Industries

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



Dynamic and integrated safety and security barrier management: A new framework to manage major event risks in chemical plants

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:
Process safety
Process security
Risk assessment
Dynamic barrier management
Decision-making

ABSTRACT

Chemical process industries are threatened by accidental and intentional major events that may lead to catastrophic consequences due to hazardous materials' production, operation, and storage. Remarkably, the digitalization of industrial facilities brings emerging cyber-physical attack risks, which calls for a holistic and integrated safety and security risk assessment and management. Considering the dynamic aspects of risks, the continuous monitoring and assessment of risk-related variations plays a vital role in making timely adaptions to risk treatment strategies and, therefore, accommodating increasing risks. To this end, this study proposes a comprehensive framework for risk-based safety and security barrier management, handling challenges in assessing integrated safety and security risks and deriving timely and cost-efficient barrier improvement strategies in case undesired risks are increasing to unacceptable levels. The fundamental ideas and applicable procedures are elaborated before a case study is demonstrated to offer insights into its feasibility. The case study shows that implementing this framework holds advantages in managing safety and security risks in a unified way, considering the interplays between safety and security and making continuous risk-treatment adaptions to sustain the safety and security of digitalized chemical process systems. Furthermore, the principles and precautionary considerations pertinent to this new framework are discussed to foster its application in real-world settings.

1. Introduction

Chemical plants are exposed to multi-dimensional risks covering safety and security (S&S) aspects, which have attracted attention from academia and industry over the past decades (Reniers et al., 2008; Abdo et al., 2018; Yuan et al., 2024a). Notably, the digitalization of chemical process systems calls for assessing and managing cyber-physical (C2P) attack risks concerning the emerging cyber vulnerabilities brought into the systems. Although the happening of security incidents evidence that major events with possible high consequences may be triggered by either physical or cyber attacks (Iaiani et al., 2021a), the management of safety and security issues is still handled separately in most of the cases in Seveso sites (Ylönen et al., 2022). Meanwhile, the system complexity of the chemical process facilities with both the IT and OT infrastructures and the consideration of interplays between safety and security-associated events remain challenging, hindering the implementation of a holistic risk assessment and a unified safety and security risk management in real-world applications.

Researchers have endeavored to bridge the need to integrate safety and security risks in process industries. Remarkably, Song et al. (2019), Chen et al. (2019), and Casciano et al. (2019) have studied the integration of safety risks and physical-attack-associated risks in chemical process industries. Researchers have also investigated the consideration and incorporation of C2P attack risks into safety and security management. Abdo et al. (2018) integrate accidental and attack-related scenarios with a bow-tie analysis, which can serve as a basis for an integrated safety and security risk assessment and management. Guzman et al. (2021) proposed an integrated method for scenario identification and integrating industrial cyber-physical systems, considering both safety and security. Iaiani et al. (2021b) developed a dedicated methodology, named PHAROS, to identify major event scenarios triggered by malicious manipulations of IT and OT systems of an industrial facility. Then, the same research team combined past incident analysis (PIA) with the scenario identification methodologies and Bayesian

* Corresponding author.
E-mail address: cumtbyuanshuaiqi@163.com (S. Yuan).

https://doi.org/10.1016/j.jlp.2025.105632
Received 19 November 2024; Received in revised form 16 February 2025; Accepted 11 March 2025
Available online 14 March 2025

network analysis for cyber-risks identification and analysis concerning malicious interferences to the control and safety instrumented systems (Iaiani et al., 2023, 2024).

By contrast, our research team has investigated the integrated safety and security risk assessment and management under a barrier management umbrella. Firstly, the fundamental concepts and the state-of-the-art research status on safety barrier management have been reviewed (Yuan et al., 2022a). Followed by an approach for quantitative risk assessment of industrial cyber-physical systems concerning both safety and security-associated scenarios is proposed (Yuan et al., 2024a), and the treatment of uncertainties in the integrated risk assessment is investigated (Yuan et al., 2024b). By employing the quantitative risk assessment as the basis, safety and security barrier management strategy has been investigated by considering the maintenance optimization issues with cost-effectiveness analysis and genetic algorithms (Yuan et al., 2023a) and the implementation of dynamic management strategies to safety barriers considering barrier degradation (Yuan et al., 2023b). Based on our previous studies, we aim to propose a comprehensive framework for dynamic and integrated S&S barrier management with the wrap-up of our previous research endeavors and also with the newly added elements/aspects below:

- Multi-source data that can reveal risk-related variations covering both safety and security aspects is identified and characterized. The incorporation of time-varied security risks considering the variations in security threats and vulnerabilities is studied. A variety of methods are combined to update the integrated safety and security risks using the multi-source data.
- Integrating security risks (C2P and physical attack risks) into the dynamic barrier management (DBM) framework is researched. A systematic integration of the methodologies/approaches developed by the research team is conducted to wrap up all endeavors on this topic. A case study is provided to show the efficacy and advantages of the proposed framework, and the foundational principles and practical notes are discussed to foster the adoption of dynamic and integrated S&S barrier management in real-world practices.

To facilitate the readability and understanding of this study, we explained the key concepts and terminologies used in this paper in Table 1.

The remaining sections of this paper are organized as follows. Firstly, an overview of the proposed framework is provided in section 2. Following this, the systematic integration of the methodologies underpinning this framework is elaborated upon, with a focus on the pertinent data sources and models enabling dynamic barrier management, in section 3. Section 4 features a hypothetical case study, showcasing the application of the framework for dynamic and integrated S&S barrier management. Subsequently, discussions are presented in section 5, and conclusions are given in section 6.

2. An overview of the proposed framework

An overview of the proposed framework for dynamic and integrated S&S barrier management is shown in Fig. 1. This circular framework has three main elements to enable risk-based barrier management. Risk assessment quantifies and evaluates the integrated safety and security risks and provides baseline risk profiles for decision-making. Decision-making determines the optimal strategy for S&S barrier improvements in case of unacceptable risks. Variation monitoring takes responsibility for monitoring system performance after implementing barrier improvement strategies and revealing risk-related variations based on multiple-source data. Risk-related variations are incorporated into the next-round risk assessment, achieving dynamic risk assessment and dynamic barrier management.

Each step/component of the three elements in this framework is elaborated below:

Table 1
A summary of the key concepts used in this paper.

Concepts	Definitions/Descriptions
Safety barriers	Safety barriers present all kinds of measures/tools used to prevent the happening of accidental events or mitigate their corresponding consequences.
Security barriers	Security barriers are defined as all kinds of measures/tools used to protect vulnerable assets from intentional attacks/malicious acts (including deliberate physical and cyber acts) and/or mitigate the corresponding consequences.
Safety risks	Risks affiliated with safety hazards and unintentional causes, including accidental technical component failures, human errors, external interventions, etc. that may accidently lead to losses.
Security risks	Risks affiliated with intentional attacks/malicious acts aiming to deliberately exploit the vulnerability of specific targets to cause losses.
Major accidents	A major accident is an unplanned (unintentional) event that results in significant harm to people, property, or the environment. In process safety contexts, this often involves the release of hazardous substances, fires, or explosions.
Security incidents	Security incidents refer to a breach or compromise of a target's security, including unauthorized access, data breaches, physical sabotage, etc.
Adverse events	An adverse event is any unfavorable or harmful occurrence that negatively impacts a process, system, or individual. Adverse events can include both safety-related and security-related events. Similar terminology, "undesired events," is also used.
Major adverse events	Major adverse events include both safety-related and security-related events that result in significant harm to people, property, or the environment, for instance, a release of hazardous substances, fires, explosions, etc.

- R1. Scenario building: This step performs a safety and security analysis of the system of interest. This step identifies threatening safety hazards and security threats, and builds up adverse scenarios that potential safety causes and intentional attacks could induce while considering the intervention of safety and security barriers.
- R2. Risk analysis: This step quantifies the integrated safety & security risks based on the developed undesired scenarios in the last step. Both the likelihood and consequence severities of the undesired scenarios are assessed. Sensitivity analysis may also be performed to get insights into the criticality of each risk factor.
- R3. Risk evaluation: This step evaluates the acceptance of the adverse risks according to their corresponding thresholds and identifies unacceptable risks that must be managed.
- D1. Objectives & constraints: Risk treatment measures (S&S barrier improvements in this study) should be implemented if unacceptable risks exist. The objectives and constraints for barrier improvement are determined in this step to characterize the decision-making optimization problem.
- D2. Candidate strategies: This step proposes candidate strategies for S&S barrier improvements considering the effectiveness, economics, operability, alignment with laws/regulations, and other possible concerns (for instance, sustainability, societal concerns, etc.) related to the strategy implementation.
- D3. Optimization: This step determines the optimal strategy from a set of candidate strategies by solving the predefined decision-making optimization problem using tailored optimization algorithms.
- V1. Information updating: After implementing the optimal barrier improvement strategy, pertinent information must be updated to re-conduct scenario building and risk analysis.
- V2. Data collection & processing: This step collects multi-source data (incident data, condition-monitoring data, inspection data, etc.) that are capable of revealing risk-related variations. Data processing may be necessary to prepare further data analysis and risk variation quantification.
- V3. Data analysis: This step reveals and quantifies risk-related variations based on the collected multi-source data using a set of models.

- V4. Reporting risk-related variations: This step incorporates the quantified risk-related variations into the next-round risk assessment and decision-making process.

3. Methodologies

3.1. Overview of the integrated methodology

The proposed framework aims to achieve dynamic and integrated S&S barrier management based on multidisciplinary knowledge and techniques. As a result, methodologies from different scientific fields (process safety, process security, and cybersecurity) are integrated to implement this framework. Specific techniques/tools used in various phases of this framework are presented in Fig. 2, in which the workflows are also given to integrate those techniques/tools systematically. The methodology presented in Fig. 2 is a systematic reframing of the endeavors of our previous research and adds new elements to incorporating security-related data for dynamic S&S barrier management. Some techniques or approaches have already been explained and applied in our previous studies. Therefore, a brief introduction of each part of the proposed methodology is given in the following sub-sections, with an emphasis on the newly added elements.

3.2. Integrated safety and security risk assessment

3.2.1. Integrated risk assessment models

This section mainly adapts the approaches presented in (Yuan et al., 2024a, 2024b) for integrated risk assessment of industrial control systems (ICSs) considering safety causes, cyber-physical (C2P) attacks, and physical attacks. As shown in Fig. 2, scenario building starts with establishing a CPS master diagram (Guzman et al., 2021), which demonstrates the ICS in a multi-layered manner and serves as a basis for

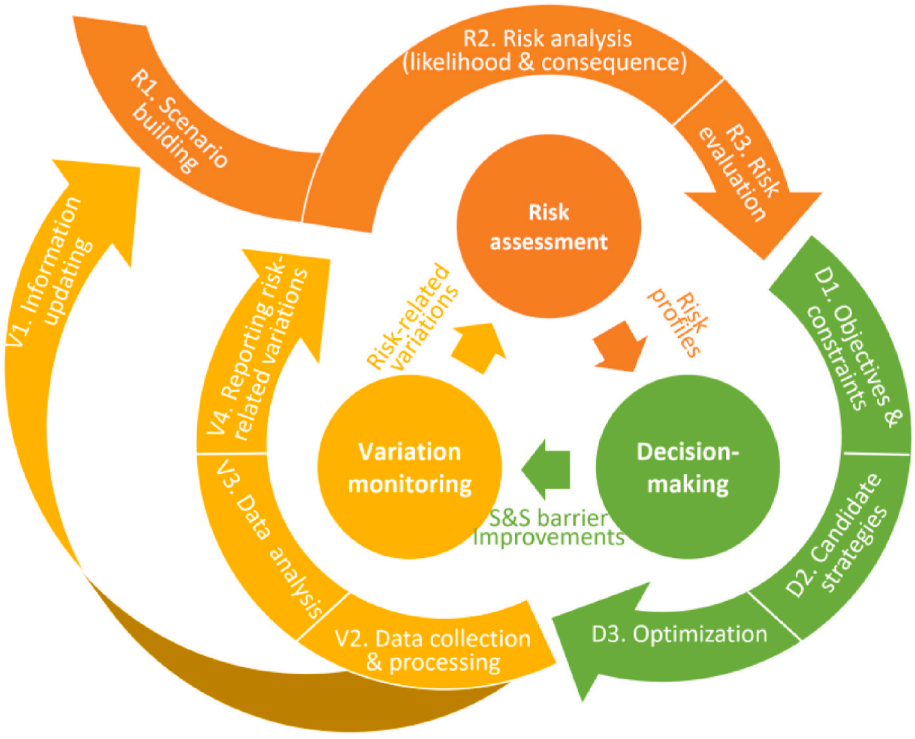


Fig. 1. A framework for dynamic and integrated S&S barrier management.

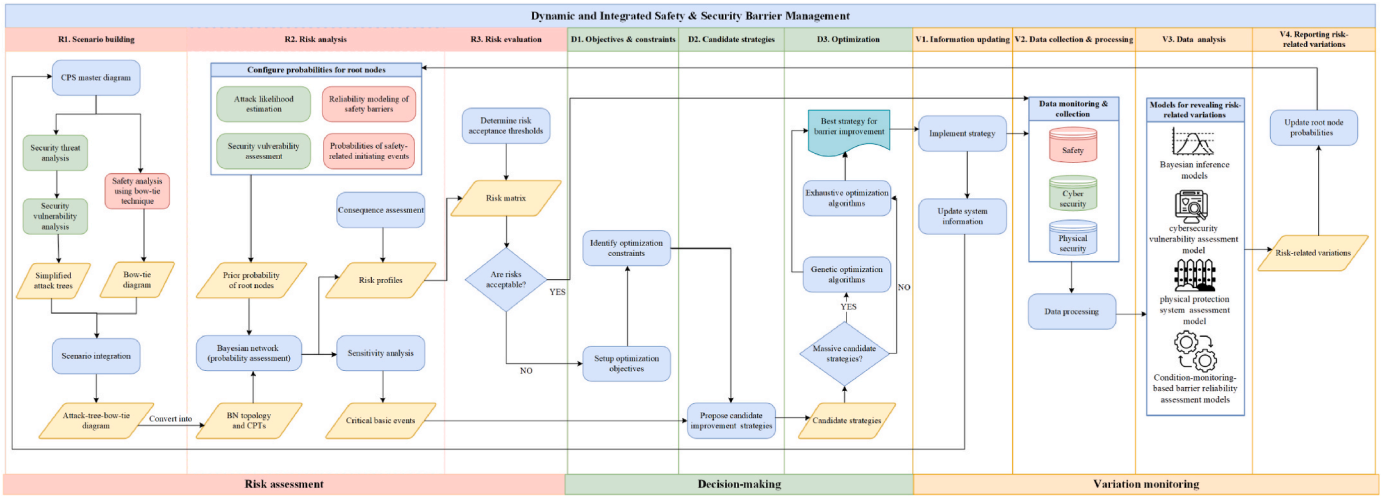


Fig. 2. An integrated methodology for dynamic and integrated S&S barrier management.

safety and security analysis. Then, the bow-tie technique is used to identify safety-related accidental scenarios. Security threat analysis and security vulnerability analysis are performed to identify credible attack modes and represent the attack modes in the form of simplified attack trees. Security threat analysis combines threat agent categories (Störfall Kommission, 2002), API threat levels (API, 2013), and cyber security incident data to characterize potential threat agents and approximate their attack likelihoods. Details on the threat analysis can be found in Section 3.4.1 in Yuan et al. (2024a). Adversary Sequence Diagram and Path Analysis helps conduct vulnerability analysis of physical attacks considering the protection of PPSs (physical protection systems) and identifies the credible attack paths of attackers (Norman, 2010). Attack/compromise graphs are used to conduct vulnerability analysis of C2P attacks and visualize the potential attack paths of attackers, considering the known vulnerabilities at each attack step (Semertzis et al., 2022). The obtained simplified attack trees are integrated with the bow-tie diagram to form an integrated attack-tree-bow-tie diagram. The integrated attack-tree-bow-tie diagram is then converted into a Bayesian Network (BN) model for quantitative risk assessment. The topology and conditional probability tables (CPTs) of the BN model are derived following the mapping algorithm proposed by Khakzad et al. (2013). In addition to the topology and CPTs, prior probabilities for the BN root nodes are required for the risk analysis. For the prior probabilities and probability distributions, different methods are applied based on the node types, as detailed in the following section. A more comprehensive description of the process for integrating attack trees and bow-tie diagrams, as well as developing a BN model for integrated risk assessment, can be found in Section 3.5 of Yuan et al. (2024a).

3.2.2. *Prior probabilities/probability distributions*

Regarding the prior probabilities/probability distributions for the BN model, different ways are used to determine prior probabilities or probability distributions for four types of root nodes: safety-related initiating events, safety barriers, attack likelihood, and conditional probabilities of successful attacks. In case probability distributions are used for root nodes, Monte Carlo simulations are combined with the BN to handle uncertainty propagations, as presented in Section 2.3.1 in Yuan et al. (2024b). It is important to note that, in principle, a larger number of Monte Carlo samples leads to more reliable probability bounds. To ensure the reasonableness of the generated probability bounds for the outcome events, sensitivity analysis can be conducted by progressively increasing the sample size until no significant differences are observed with further increases. Details on determining the prior probabilities or probability distributions are given below.

1) For safety-related initiating events, reliability databases (Hauge and Onshus, 2010), human reliability data (Kirwan, 2017), accident databases (Debray et al., 2004), or data available in the literature are used to derive the probabilities/probability distributions. Considering safety barriers, the probability of failure on demand (PFD) is used to quantify the reliability of safety barriers (IEC:61508, 2010). Reliability databases, accident databases, and human reliability analysis are helpful for determining the PFDs of safety barriers. The approaches presented in Section 2.3.2 in Yuan et al. (2024a) are used to calculate PFDs for technical safety barriers considering barrier maintenance.
2) Attack likelihood estimation relies more on expert judgment due to the lack of data found in industry and literature. Regarding physical attacks, a method proposed by Landucci et al. (2017) helps experts/stakeholders estimate the attack likelihood based on the API threat levels (API, 2013) and the facility's expected life, as presented in Table A1 in Appendix I. Considering C2P attacks, some incident statistics of comparable companies or in the same or similar sectors may help the attack likelihood estimation. For instance, the cyber security incident analysis conducted by Kuypers and Maillart (2018) may be used as a basis for attack likelihood estimation of C2P attacks, as presented in Table A2 in Appendix I.
3) For security vulnerability assessment, the time-to-compromise (TTC) approach presented in Section 2.3.3 in (Yuan et al., 2024b) is used to assess the vulnerability of industrial control systems to C2P attacks, considering the uncertainties associated with attackers' skill levels and attack path selection. This approach is explained briefly in Appendix II, and more details can be found in the original study (Yuan et al., 2024b).
4) Regarding physical attacks, Adversary Sequence Diagrams (ASD) and Path Analysis (Garcia, 2007), event tree analysis, and the benchmark data presented by Moreno et al. (2022) are combined to quantify the vulnerability of PPSs (physical protection systems). A data repository built by Moreno et al. (2022) is used for typical PPSs in chemical plants, as shown in Table 2.

For assessing the probability of the emergency response team successfully interrupting a physical attack, the EASI (Estimate of Adversary Sequence Interruption) model is employed (Garcia, 2007). The EASI

Table 2
A summary of the performance data of typical PPSs in the chemical process industry, adapted from Moreno et al. (2022).

PPS (physical protection system)	Type of Function	PFD (probability of failure on demand)	Effectiveness (η)	Calculation formulas
Entry gate	Delay	0.02	0.9975	$P_{fail} = PFD + (1-\eta) \times (1 - PFD)$
Entry control	Detection	0.40	0.80	
Fence	Delay	0.00	0.9968	
Closed Circuit TeleVision (CCTV)	Detection	0.205	0.97	$P_{success} = (1 - PFD) \times \eta$
Intrusion detection by site personnel	Detection	dayshift: 0.233 nightshift: 0.4	0.248 0.248	

Table 3
A summary of multi-source data and the models for revealing risk-related variations.

Data categories	Safety-related OR Security-related	Models for revealing risk-related variations	Revealed risk-related variations
Accident precursor data	safety	Bayesian updating (Gamma-Poisson model OR Beta-binomial model)	● Variations in the occurrence probabilities of initiating events. ● Variations in the reliability of safety barriers.
Condition-monitoring data	safety	Condition-based reliability analysis models	● Variations in the reliability of basic process control systems or safety barrier systems.
Security-related precursor data	security	Bayesian updating (Gamma-Poisson model OR Beta-binomial model); PPS assessment model	● Variations in attack likelihoods. ● Variations in the vulnerability of PPSs (physical protection systems).
CVE (Common Vulnerabilities and Exposures) data	security	C2P attack vulnerability assessment model	● Variations in the cyber vulnerability of ICSs (industrial control systems).
Cyber incident data for estimating MTTD (Mean-time-to-detect)	security	C2P attack vulnerability assessment model	● Variations in the cyber vulnerability of ICSs (industrial control systems).

model calculates the probability of adversary interruption (P_S) based on an analysis of the interactions of detection, delay, response, and communication, as follows (Garcia, 2007; Argenti et al., 2017):

$$P_S = P_D \times P_C \times P_T \tag{1}$$

$$P_T = \frac{1}{\sqrt{2\pi\sigma_t^2}} \int_0^{\infty} exp\left[-\frac{(t-\mu_t)^2}{2\sigma_t^2} \right] dt \tag{2}$$

$$t = ATT - RFT \tag{3}$$

where P_D is the probability of successful detection of the intrusion. P_C is the probability of successful communication to the response force to carry out the response. Based on the evaluation of many systems designed and implemented by Sandia National Laboratories, the value of P_C for most systems is at least 0.95. P_T is the probability of response force intervening in time to interrupt the adversary successfully. P_T is calculated by using a normal distribution considering two time parameters, adversary task time (*ATT*) remaining after detection, and response force time (*RFT*). μ_t is the mean value of t. σ_t^2 presents the variance of t. Standard deviations of the *RFT* and *ATT* may be obtained from field tests to decide σ_t^2. In case the specific data are unavailable, a conservative value of 30 % of the mean value is used based on the tests at Sandia National Laboratories (Garcia, 2007).

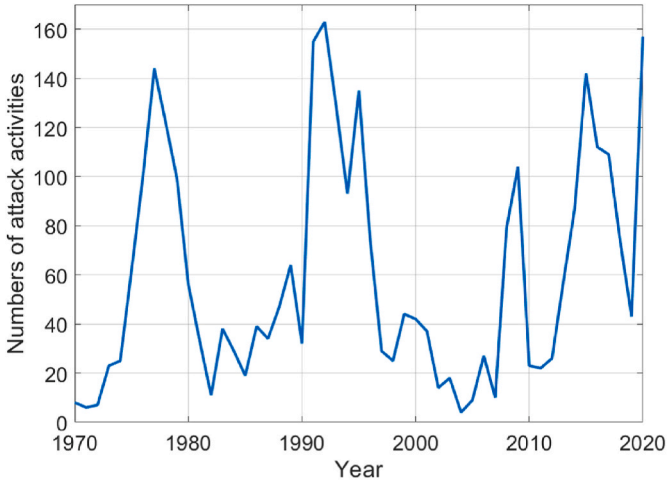


Fig. 3. The numbers of facility/infrastructure attacks in Western Europe, adapted from (START, 2022).

3.2.3. Consequence assessment and sensitivity analysis

A qualitative or quantitative consequence assessment may be performed. In the methodology presented in Fig. 2, a severity class regarding typical dangerous phenomena in chemical plants suggested by the ARAMIS project (Andersen et al., 2004) is used for qualitative consequence assessment and combined with a risk matrix to conduct risk evaluation. Regarding major events in chemical plants, such as toxic leakage, fires, and explosions, etc., it is also possible to combine physical effects modeling (using computational fluid dynamics simulations or empirical models) with damage analysis models for heat radiation, explosion effects, acute intoxication, etc. (Gubinelli et al., 2004; Cozzani et al., 2005) to quantify the consequence severity. For instance, Yuan et al. (2022b) combine computational fluid dynamics with a probit model to assess toxic gas leakage scenarios and represent the severity of the consequences in the form of fatality probability.

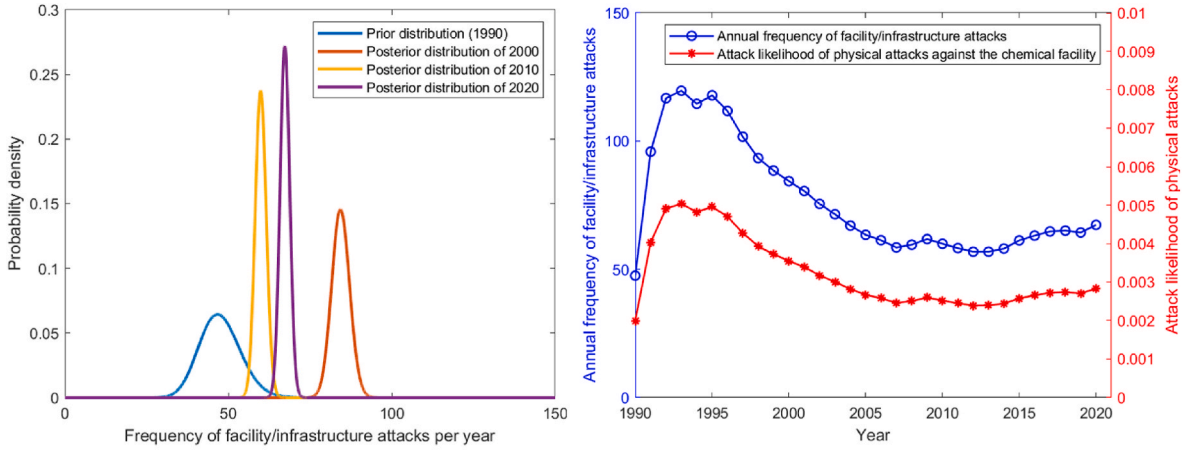
Additionally, sensitivity analysis plays a vital role in identifying critical basic events to unacceptable risks, facilitating the proposal of candidate risk-treatment strategies. Remarkably, the Birnbaum importance measure (Van der Borst and Schoonakker, 2001), risk reduction measure (Yazdi and Kabir, 2017), and ratio of variance (RoV) measure (Zarei et al., 2017) have been used for sensitivity analysis of fault-tree-like approaches. Because the Birnbaum importance measure can be applied to integrated models with good flexibility, this study uses the Birnbaum importance measure, which can be calculated below.

$$I_n = p_s(p_n = 1) - p_s(p_n = 0) \tag{4}$$

where I_n is the criticality of basic event n. p_s is the probability of occurrence of the unwanted accident scenario. p_n is the probability of happening of basic event n.

3.3. Cost-effective decision-making on barrier improvement

In case unacceptable risks are observed from the risk matrix, necessary actions should be taken to enhance the performance of S&S barriers. The decision-making step combines cost-effectiveness analysis (CEA) and optimization algorithms to decide the optimal strategy for S&S barrier improvements. It starts with the determination of optimization objectives and the configuration of optimization constraints. We elaborate on two typical practices when conducting CEA with constraints. The first applies to situations where a company has to reduce the risks below certain levels while using the minimum investment, as



(a) Probability distributions of the facility/infrastructure attack annual frequency.

(b) Physical attack likelihood and annual frequency of facility/infrastructure attacks.

Fig. 4. Bayesian updating of attack likelihood estimations using precursor data.

demonstrated by Eq. (5). The second applies to situations where a company only has a limited budget and aims to mitigate undesired risks as much as possible, as demonstrated by Eq. (6). Considering the practical needs, some additional constraints may also be added to the optimization problems.

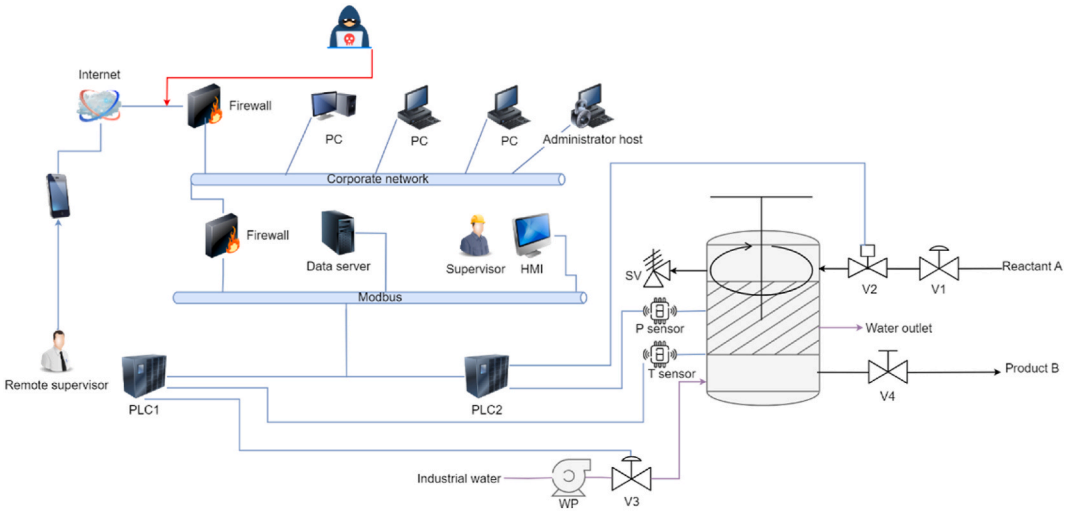
$$\begin{cases} Min(C_i) \\ Eff_i \geq Eff_{min} \\ i \in \{1, 2, 3, \cdots, N\} \end{cases} \tag{5}$$

or:

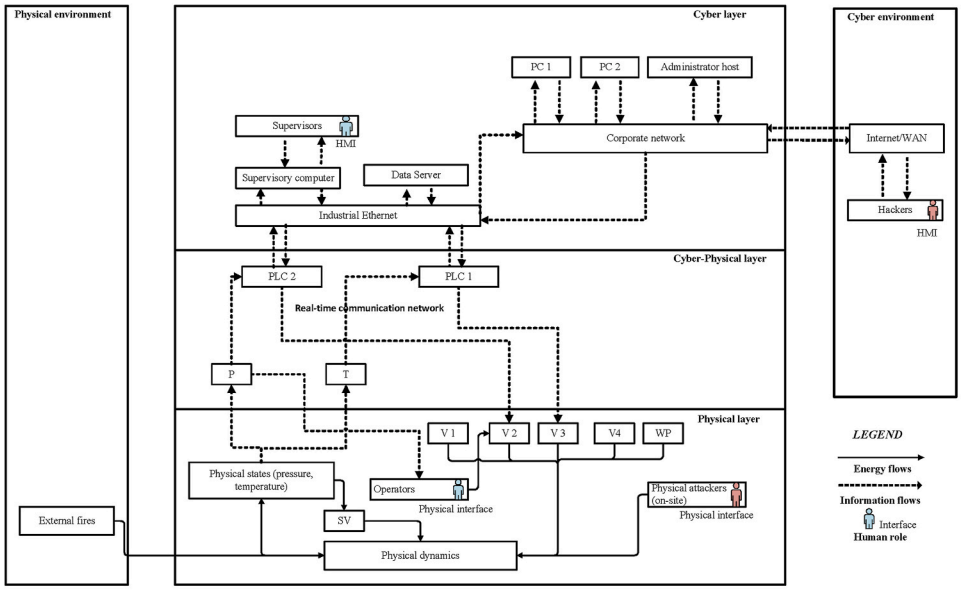
$$\begin{cases} Max(Eff_i) \\ C_i \leq Bu_{max} \\ i \in \{1, 2, 3, \cdots, N\} \end{cases} \tag{6}$$

where i means a strategy i from N possible strategies. C_i is the cost of the implementation of strategy i. Eff_i is the effectiveness after implementing strategy i. Eff_{min} is the minimum acceptable level of effectiveness. Bu_{max} is the maximum available budget that can be used for risk-treatment. The effectiveness is an indicator reflecting safety and/or security risks. The effectiveness of implementing a barrier maintenance strategy is measured by the corresponding risk reduction regarding specific accident scenarios.

To conduct barrier optimization, candidate strategies for barrier improvement should be proposed to form a strategy pool. The combinations of various activities/countermeasures may be considered candidate strategies. For instance, deploying new safety barriers, security vulnerability patching, shortening barrier maintenance intervals, etc. The results from the sensitivity analysis help the candidate strategy proposal since the enhancement of a more critical barrier is more likely to have a higher effectiveness regarding risk reduction. Additionally, the economics, operability, alignment with laws/regulations, and other possible concerns (for instance, sustainability, societal concerns, etc.) in relation to the strategy implementation may also be considered when proposing candidate strategies. After that, optimization algorithms are employed to search for the optimal strategy that can achieve the optimization objective best and meet the optimization constraints. Usually,



(a) The investigated industrial control system.



(b) The CPS master diagram.

Fig. 5. The investigated industrial control system and its CPS master diagram.

Table 4
A summary 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	
AT11	Physical attack on the shell	Induce shell rupture

the best strategy can be obtained through exhaustive search optimization, as demonstrated in the case study in (Yuan et al., 2023b). If a large amount of candidate strategies are proposed, it would be too vast to search for the optimal strategy exhaustively in a reasonable amount of time. In that case, evolutionary algorithms (for instance, genetic algorithms) are implemented to solve the optimization problem and approximate the optimal strategy, as demonstrated in the case study in (Yuan et al., 2023a).

3.4. Risk variation monitoring for dynamic barrier management

3.4.1. An overview of risk-related variations

After implementing the optimal barrier enhancement strategy, necessary information should be updated, and pertinent accidental scenarios may be modified for risk assessment. Then, risk-associated data is monitored and analyzed continuously to reveal possible risk variations and enable dynamic barrier management. Various data in relation to system safety and security may have the potential to reveal risk-related variations in a timely manner and drive a dynamic S&S barrier management. This study presents a preliminary attempt to monitor and quantify risk-related variations for dynamic risk assessment and dynamic barrier management based on data from multiple sources. Table 3 characterizes the multi-source data and the models used to reveal and quantify risk-related variations.

Yuan et al. (2023b) have already incorporated safety-related data for dynamic barrier management, incorporating both accident precursor data and condition-monitoring data. Based on accident precursor data, Bayes's theorem (for instance, Beta-binomial models) updates the failure probabilities of safety barriers. A similar Bayesian updating model (Gamma-Poisson model) can be employed to update the occurrence probabilities of initiating events in safety risk analysis (Siu and Kelly, 1998). Additionally, condition-monitoring data, including periodic inspection and continuous condition-monitoring data, has been used to update the failure probabilities of safety barrier systems or basic process control systems. More details and a demonstrative application can be found in our original study (Yuan et al., 2023b); the following sub-sections introduce incorporating security-related data for revealing risk-related variations and facilitating dynamic barrier management.

3.4.2. Attack likelihood updating using precursor data

To tackle the difficulties in attack likelihood estimation, Khakzad et al. (2018) suggested using precursor data (indirectly relevant data) for reasoning rare events when the amount of directly relevant data is insufficient. The precursor-data-based probability updating may be applied to security risk analysis as it has been applied in the safety science domain with good feasibility. The trends of similar terrorism activities in comparable sectors or in the same region may implicate the attack likelihood trends in chemical process industries. Therefore, the available terrorist attack data from a broader domain or similar sectors can be used as a valuable source of information for reasoning and updating attack likelihoods in chemical plants. To achieve this, we apply a Bayesian updating model (Gamma-Poisson model) to estimate the probability/frequency of comparable terrorism activities and assume the prior probability (λ) follows a gamma distribution as below.

$$g(\lambda) = \frac{\beta^{\alpha}\lambda^{\alpha-1}}{\tau(\alpha)} e^{-\beta\lambda} \tag{7}$$

where $g(\lambda)$ is a gamma distribution of λ. α and β are distribution parameters. $\tau(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}\,dt$ is a gamma function. A Poisson distribution is used to present the conditional probability of r terrorism events occurring in a period of time t, given the probability λ (Khakzad et al., 2012).

$$P(r \text{ events in } [0,t] \mid \lambda) = \frac{(\lambda t)^r}{r!} e^{-\lambda t} \tag{8}$$

When new precursor data becomes available, the prior probability distribution is updated using Bayes's theorem, as follows:

$$g(\lambda \mid Data) = \frac{P(Data \mid \lambda)g(\lambda)}{\int P(Data \mid \lambda)g(\lambda)d\lambda} \propto P(Data \mid \lambda)g(\lambda) \tag{9}$$

where P(Data | λ) is the likelihood function, and g(x | Data) presents the posterior distribution. The posterior distribution of λ can calculated as follows:

$$g(\lambda \mid r \text{ events in } [0,t]) = \frac{\beta'^{\alpha'}\lambda^{\alpha'-1}}{\tau(\alpha')} e^{-\beta'\lambda} \tag{10}$$

where $\alpha' = \alpha + r$ and $\beta' = \beta + t$. The mean values of the prior and posterior distributions of λ are calculated below.

$$E(\lambda) = \frac{\alpha}{\beta} \tag{11}$$

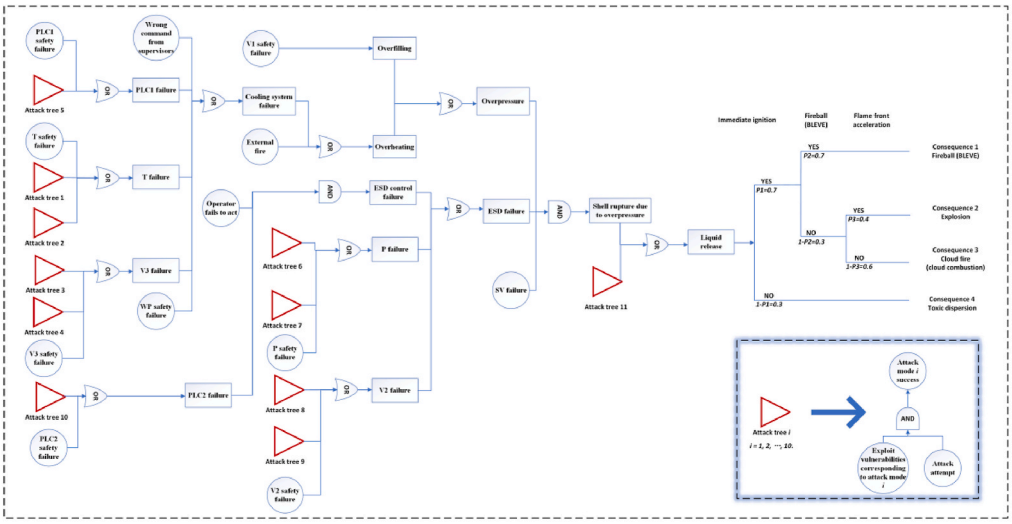
$$E(\lambda') = \frac{\alpha'}{\beta'} = \frac{\alpha + r}{\beta + t} \tag{12}$$

We assume that the attack likelihood of physical attacks against a specific chemical facility is linearly correlated with the occurrence frequency of the comparable terrorism activities. Then, the attack likelihood of physical attacks against this chemical facility can be estimated below.

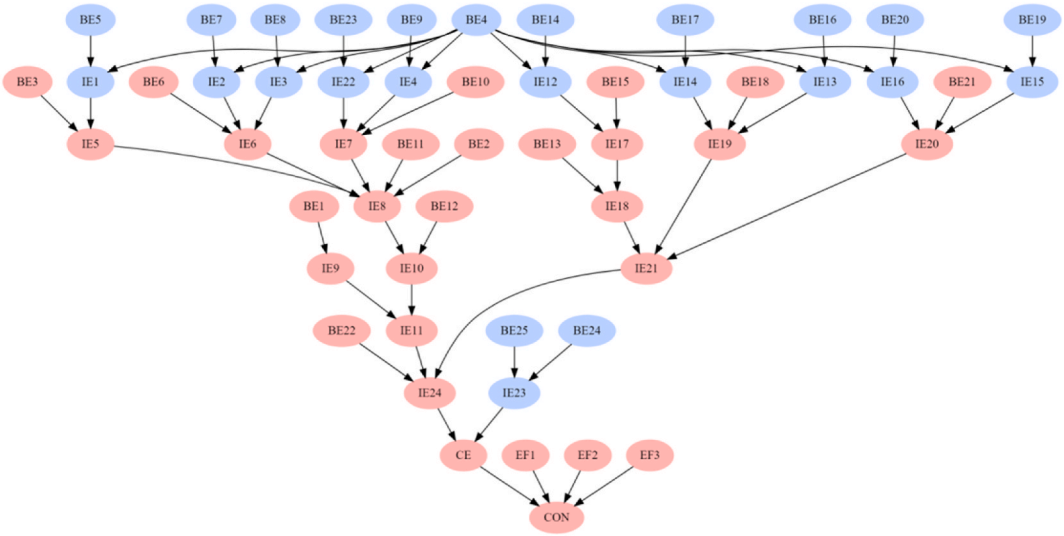
$$\frac{E(\lambda') - E(\lambda)}{E(\lambda)} = k\left(\frac{Pr' - Pr}{Pr}\right) \tag{13}$$

$$Pr' = Pr\left(\frac{E(\lambda') - E(\lambda) + E(\lambda)^{*}k}{E(\lambda)^{*}k}\right) \tag{14}$$

Where Pr is the prior probability of physical attacks against a chemical facility, which is estimated using the method presented in Table A1 considering the API threat levels (API, 2013) and the facility's expected life. Pr' is the posterior probability of physical attacks. k is a scale coefficient depicting the scaling correlation between the probability of physical attacks and the probability/frequency of the comparable terrorism activities. $k = 1$ means they have the same scaling trends (the probability of physical attacks doubles when the frequency of comparable terrorism activities doubles). Risk analysts may configure the value of k based on the judgment on the attractiveness of the investigated chemical facility. If the chemical facility has a relatively higher attractiveness than the average level of the facilities/infrastructures in the terrorist attack database, a value of more than one should be used for k, and vice versa.



(a) An integrated attack-tree-bow-tie diagram.



(b) A Bayesian network model (nodes with pink and blue colors are derived from the bow-tie diagram and attack trees, respectively).

Fig. 6. The developed attack-tree-bow-tie diagram and BN model.

To demonstrate the proposed method, the Global Terrorism Database (GTD), which is an open-source database including information on terrorist events around the world from 1970 through 2020 (with annual updates planned), is used as the data source (START, 2022). Facility/infrastructure attacks in Western Europe are considered comparable terrorism activities, and the attack frequencies from 1991 to 2020 are regarded as hypothetical precursor data, as shown in Fig. 3. The prior annual frequency of facility/infrastructure attacks is initialized as a gamma distribution ($\Gamma(\alpha, \beta)$, $\alpha = 58.18$; $\beta = 1.22$) with a mean value of 47.52 per year and a standard deviation of 38.82 per year, based on the data from 1970 to 1990. The prior probability of physical attacks against a chemical process facility is configured according to the method in Table A1. Considering the threat level (configured as threat level 1) and the facility's expected life (configured as 50 years), the annual attack frequency is calculated as 2.0E-03. Then, the frequency of facility/infrastructure attacks and the annual frequency of physical attacks are updated using the precursor data, as shown in Fig. 4.

Fig. 4 (a) presents the prior probability distribution and selected posterior probability distributions of the annual frequency of facility/infrastructure attacks. When new precursor data becomes available yearly, the probability distribution is updated based on Bayes's theorem, as demonstrated in Eqs. (7)–(10). Fig. 4 (b) shows the mean value of the probability distribution of the facility/infrastructure attack annual frequency, which is updated yearly using the precursor data. Meanwhile, the likelihood of physical attacks against the investigated chemical facility is updated yearly using Eq. (13) ($k = 1$ is configured in Eq. (13)).

The present approach can also be applied to the likelihood estimation of C2P attack attempts based on the available cyber incident database. If the security operations center (SOC) has plant-specific data, a prior probability distribution of C2P attacks may be generated based on incident statistics, and the probability distribution is updated when new precursor data come, according to Eqs. (7)–(12). If no plant-specific data is available, the incident data from a broader source (the same sector or comparable sectors) may be used for the Bayesian updating. In that case, Eq. (13) is used to update the plant-specific C2P attack likelihood based on indirectly related cyber incident data and expert

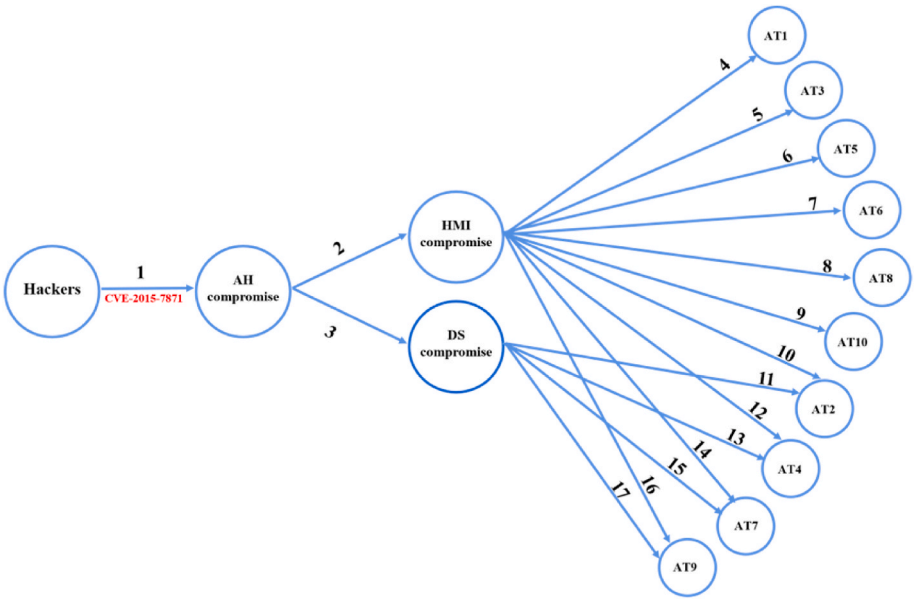


Fig. 7. A compromise graph with the known CVEs along each attack step.

judgment.

3.4.3. *Variations in security vulnerabilities*

Physical protection systems (PPSs) play an important role in protecting industrial facilities from intentional attacks and malicious acts (Garcia, 2007). The approach and the benchmark data provided in Table 2 can provide reference values for the vulnerability assessment of PPSs. More tailored data may be derived based on experts' judgments on the plant-specific PPSs. Additionally, Van Staalduinen & Khan (2015) suggest the application of Bayes's theorem, in which the Gamma-Poisson model updates the failure probabilities of PPSs based on hypothetical cumulative numbers of security incidents. Developing a physical-attack-related incident database is necessary to apply Bayes's theorem in the vulnerability assessment of PPSs. The prior failure probabilities of PPSs may be determined based on expert judgment or reference data (such as the data in Table 2). When new physical-attack-related precursor data is available, the failure probabilities of PPSs can be updated using Bayesian updating models. However, collecting the security precursor data related to PPSs is still challenging because physical attacks may rarely happen to a chemical plant with a relatively low threat level. As a result, the timely discovery of PPS's abnormal status, for instance, a breach in the fence, may be more useable evidence for experts to re-assess and update the failure probabilities of PPSs considering their effectiveness or availability.

Regarding the C2P attack vulnerability, risk-related variations mainly lie in the cyber vulnerability of the ICS (industrial control system) and the capability of intrusion detection systems. The exposure of new vulnerabilities may significantly change the difficulty in implementing a C2P attack and even create new attack paths for attackers. When the ICS has new vulnerabilities acknowledged, for instance, disclosed by the CVE (Common Vulnerabilities and Exposures) data (NVD, 2024), the new vulnerabilities need to be accommodated to re-perform the vulnerability assessment based on the C2P attack vulnerability assessment model, as provided in Appendix II. In that case, the attack/compromise graph regarding the ICS should be updated, the global TTCs of each attack path should be re-calculated, and the conditional probability of successful execution of each attack path should be re-assessed considering the newly acknowledged vulnerabilities.

Additionally, MTTD (mean-time-to-detect), which describes the average time needed by the security operations center (SOC) to detect a cyber intrusion successfully (Mughal, 2022), is also an important parameter used in the vulnerability assessment model needs to be updated based on incident data. The MTTD for a specific intrusion type is calculated by averaging all incident detection times of this intrusion type. The MTTD values may be calculated and updated in a timely manner based on actual incident data collected by SOCs in case noticeable variations appear in the performance of the intrusion detection systems.

All the above-mentioned variations would be quantified and incorporated into the risk assessment model based on monitoring risk-related variations. This is achieved by following the variation analysis models as summarized in Table 3 and utilizing those variations for probabilities updating of the BN root nodes. Then, the BN model performs the next-round risk assessment, followed by the next-round decision-making process, and achieves the circular loop.

4. Case study

A hypothetical case study is demonstrated in this section to show the advantages and feasibility of the proposed methodology. Implementing the proposed framework and methodology helps chemical plants shift into a new paradigm of dynamic and integrated S&S barrier management.

4.1. *System description and BN model development*

A typical industrial control system (ICS) is investigated in this case study. Considering potential safety failures, C2P attacks, and physical attacks, the proposed methodology is implemented to achieve dynamic and integrated S&S barrier management. Fig. 5 shows the basic information of the investigated ICS and its CPS master diagram. A typical continuous stirred tank reactor (CSTR) with its SCADA system is considered the system under investigation. It performs a hypothetical exothermic reaction $A \rightarrow B$ (Pilario and Cao, 2018). B is assumed to be a flammable liquid with toxicity. Reactant A is fed at a fixed flow rate with a control valve (V1). A jacketed cooling system composed of a water pump (WP), a control valve (V3), a temperature sensor (T), and a programmable logic controller (PLC1) is implemented to control the reactor temperature. An automatic emergency shutdown system (ESD) composed of a programmable logic controller (PLC2), a block/shutdown valve (V2), and a pressure sensor (P) is implemented to shut down the system in case of overpressure. Meanwhile, a safety relief valve (SV) is

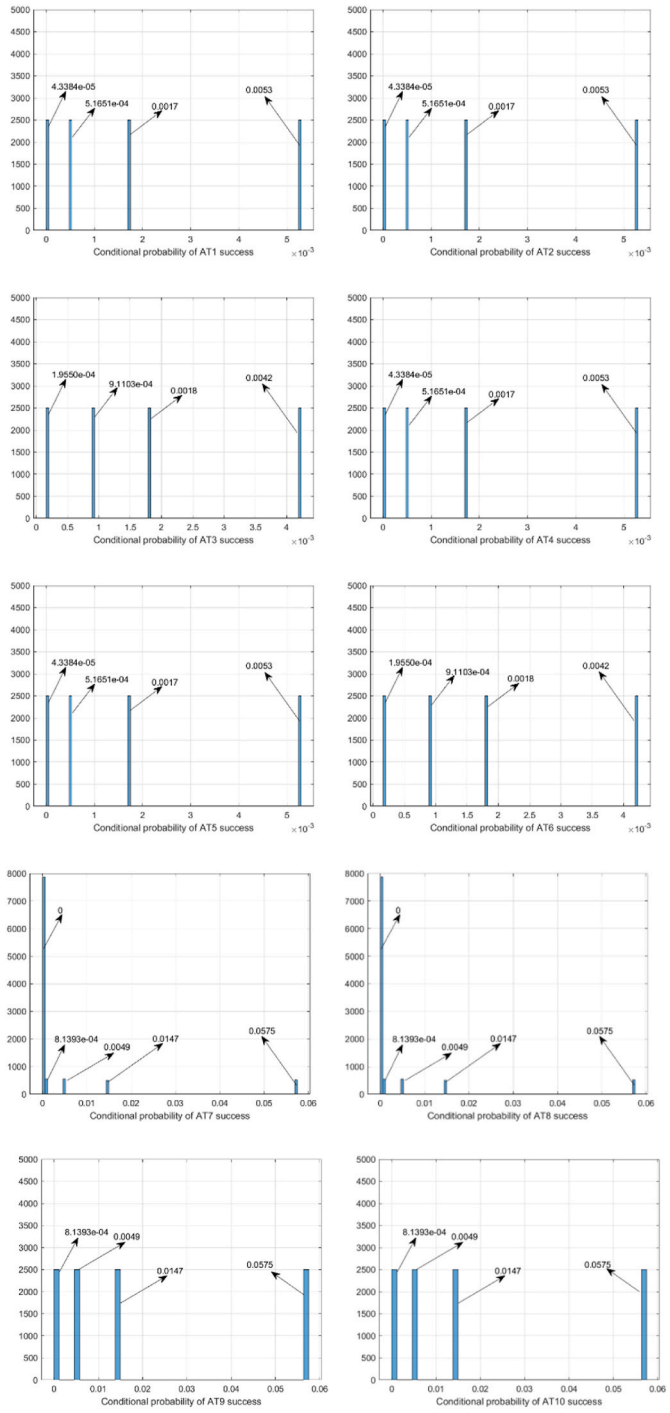


Fig. 8. Probability distributions of successful implementation of each attack mode.

installed to prevent potential scenario escalations triggered by overpressures. Both PLCs are connected to the SCADA system, and a CPS master diagram (Fig. 5(b)) is used to demonstrate the energy and information flow of the ICS in a multi-layered manner. Remote hackers with different knowledge levels and physical attackers with API threat level 1 (as illustrated in Table A1) are identified as potential threat agents, and their corresponding attack modes are explained in Table 4.

Following the integrated safety and security risk assessment method, as presented in Section 3.2, an integrated attack-tree-bow-tie diagram is developed and then converted into a BN model (as shown in Fig. 6). Explanations of the BN nodes are presented in Table 5. All BN nodes, except the consequence node, have two states (happening and not happening), while the consequence node has five states (no consequence, fireball, explosion, cloud fire, and toxic dispersion).

Table 5
[1]Explanations of the BN nodes.

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	BE24	External physical attacks
BE25	Exploit vulnerabilities corresponding to AT11	CE	Central event (Liquid leakage)
CON	Consequences	EF1	Immediate ignition
EF2	Fireball (BLEVE)	EF3	Flame front acceleration
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
IE23	AT11 success	IE24	Shell rupture induced by overpressure

4.2. Configurations of prior probabilities (probability distributions)

Configurations of the root nodes in the BN model are summarized in Table 6. Most nodes' prior probabilities are derived from reliability or historical accident databases. The sources of the prior probabilities are also given in the table. Some root nodes' prior probabilities or probability distributions are configured using different approaches, as explained below.

1) For BE15, BE18, and BE22, which present the failures of technical components of safety barriers, PDFs (probabilities of failure on demand) are calculated below.

$$PFD(t) = 1 - e^{-\lambda^*(t\%T)}, nT \le t < (n+1)T \quad (15)$$

where $PFD(t)$ is the PFD over time. λ is the failure rate. Perfect barrier maintenance with a time interval, T, is assumed with the ignorance of the time spent on maintenance. λ values of those components are derived from reliability databases. The average values of the PFDs over time are used as prior probabilities. The configurations of λ and T for those components are also given in Table 6.

2) The PFD calculation model from Zhang et al. (2020) is used for BE21, considering the degradation process (Gamma process) of the shutdown valve. The configuration of the parameters used in this model is given in Table 6. The time interval for proof tests is configured as

Table 6
Configurations of the root nodes.

Symbols	Prior probabilities (probability distributions)	Symbols	Prior probabilities (probability distributions)
BE1	4.00E-02 (Taylor, 2010)	BE2	1.00E-02 (Andersen et al., 2004)
BE3	4.38E-02 (Hauge and Onshus, 2010)	BE4	Gamma distribution ($\Gamma(\alpha, \beta)$, $\alpha = 6.08$; $\beta = 5$)
BE5	Probability distribution for AT5, as shown in Fig. 8.	BE6	2.13E-02 (Hauge and Onshus, 2010)
BE7	Probability distribution for AT1, as shown in Fig. 8.	BE8	Probability distribution for AT2, as shown in Fig. 8.
BE9	Probability distribution for AT3, as shown in Fig. 8.	BE10	4.00E-02 (Taylor, 2010)
BE11	3.125E-02 (OREDA, 2002)	BE12	5.52E-02 (Debray et al., 2004)
BE13	1.00E-02 (Andersen et al., 2004)	BE14	Probability distribution for AT10, as shown in Fig. 8.
BE15	average PFD = 4.37E-03, λ = 1.0E-06 (Hauge and Onshus, 2010); $T = 1$ year.	BE16	Probability distribution for AT6, as shown in Fig. 8.
BE17	Probability distribution for AT7, as shown in Fig. 8.	BE18	average PFD = 6.57E-04, λ = 1.5E-07 (Hauge and Onshus, 2010); $T = 1$ year.
BE19	Probability distribution for AT8, as shown in Fig. 8.	BE20	Probability distribution for AT9, as shown in Fig. 8.
BE21	Initial average PFD = 3.5E-09; ($\alpha = 1.02$E-04, $\beta =$ 1.2E04, $L = 1.25$E-03, $\tau =$ 4380 h).	BE22	average PFD = 5.47E-04, λ = 5E-07 (HSE, 2012); $T = 3$ months.
BE23	Probability distribution for AT4, as shown in Fig. 8.	BE24	Attack frequency estimations of physical attacks from 2016 to 2020 (as shown in Fig. 4 (b)).
BE25	5.60E-02, calculated from vulnerability assessment of PPSs.	EF1	7.00E-01 (Vílchez et al., 2011)
EF2	7.00E-01 (Vílchez et al., 2011)	EF3	4.00E-01 (Vílchez et al., 2011)

six months. With proof testing on the shutdown valve's health status, the PFD of the shutdown valve is updated using the periodic proof test data. Details on this model can be found in Section 2.3.3 in (Yuan et al., 2023b).

3) Regarding C2P attacks, the recurrence interval of attack attempts is estimated at approximately 150~465 days based on the data from Kuypers and Maillart (2018). A Gamma distribution ($\Gamma(\alpha, \beta)$, $\alpha = 6.08$; $\beta = 5$) with a mean value of 1.22/year and a standard deviation of 0.49/year (corresponding to a mean recurrence interval of 300 days) is used to depict the frequency of C2P attack attempts (BE4). When new cyber incident data becomes available, the distribution can be updated based on Bayes' theorem, as explained in Section 3.4.2.
4) A compromise graph is developed for the investigated ICS, and the known CVEs (Common Vulnerabilities and Exposures) at each attack step are also demonstrated, as shown in Fig. 7. The vulnerability assessment model presented in Appendix II is used to assess the system's vulnerability to C2P attacks, considering uncertainties in attackers' knowledge levels. A uniform distribution (a ratio 1:1:1:1) is configured for attackers with different skill levels (expert, intermediate, beginner, and novice). Monte Carlo simulations with 10,000 trials are conducted to obtain the probability distributions of successful execution of each attack mode. The obtained vulnerability assessment results, which are a set of probability distributions (as presented in Fig. 8), are used as prior probability distributions for BN nodes: BE5, BE7, BE8, BE9, BE14, BE16, BE17, BE19, BE20, and BE23.
5) The annual frequency of physical attacks is initialized as 3.30E-03/year based on the method presented in Table A1, considering the API threat levels (API, 2013) and the facility's expected life. Then, the attack frequency is updated using precursor data, as presented in Section 3.4.2. The period from 2016 to 2020 is assumed to be the case study's investigated time region. Therefore, the attack frequencies of physical attacks from 2016 to 2020 (as shown in Fig. 4 (b)) are used for the BN node BE6.
6) Regarding the vulnerability of physical protection systems (PPSs), the combination of Adversary Sequence Diagrams and Path Analysis and an event tree analysis is used, as presented in Section 3.2.2. A layout of the investigated chemical plant and its adversary sequence diagram are demonstrated in Fig. 9(a) and (b), respectively.

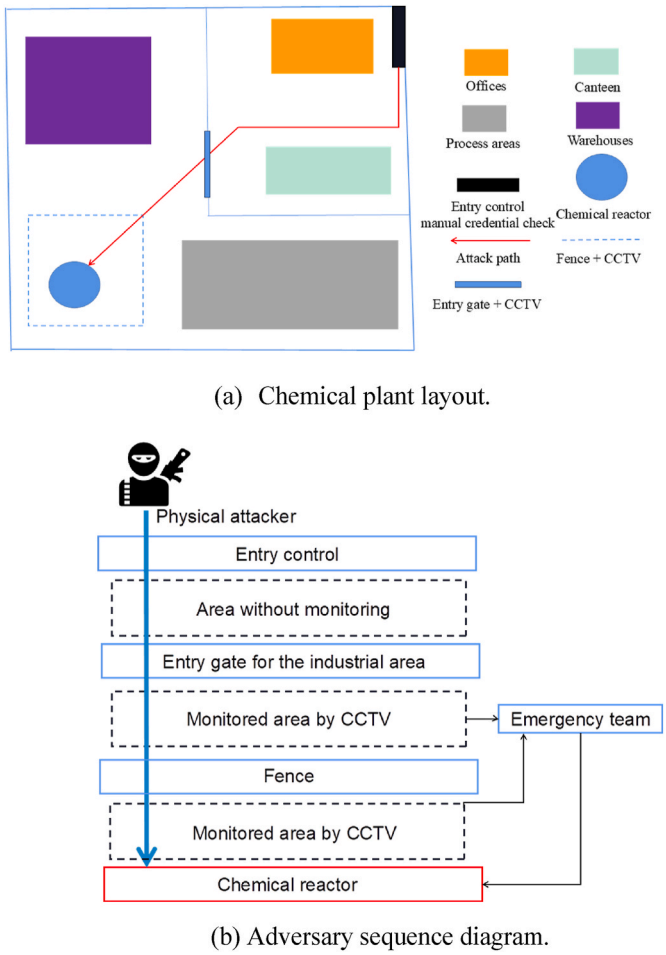


Fig. 9. Chemical plant layout and the adversary sequence diagram considering external physical attacks.

Five types of PPS are considered in this case study: entry control (manual credential check), entry gate, fence, CCTV, and emergency response team. The benchmark data provided by Moreno et al. (2022) are adapted to quantify the failure probabilities of entry control, entry gate, fence, and CCTV, as shown in Table 2. The EASI model presented in Eqs (1)–(3) assesses the probability of the emergency team successfully interrupting the physical attack. In practice, the adversary task time (ATT) and response force time (RFT) should be estimated based on filed trails. The adversary task time (ATT) remaining after detection is calculated considering the delay time caused by the delay elements along the path. In this case study, the response force time (RFT) is assumed to be 200s ± 30 %. The time for an intruder to overcome each physical barrier (fence and gates) is estimated as 90s, and the time to complete a deliberate operation (damage an instrument, operate on an arson device, etc.) is estimated as 20s (Garcia, 2007; Moreno et al., 2022). An event tree is used to assess the vulnerability of the whole PPS system given an attack attempt, as presented in Fig. 10. According to Fig. 10, the conditional probability of a successful physical attack given

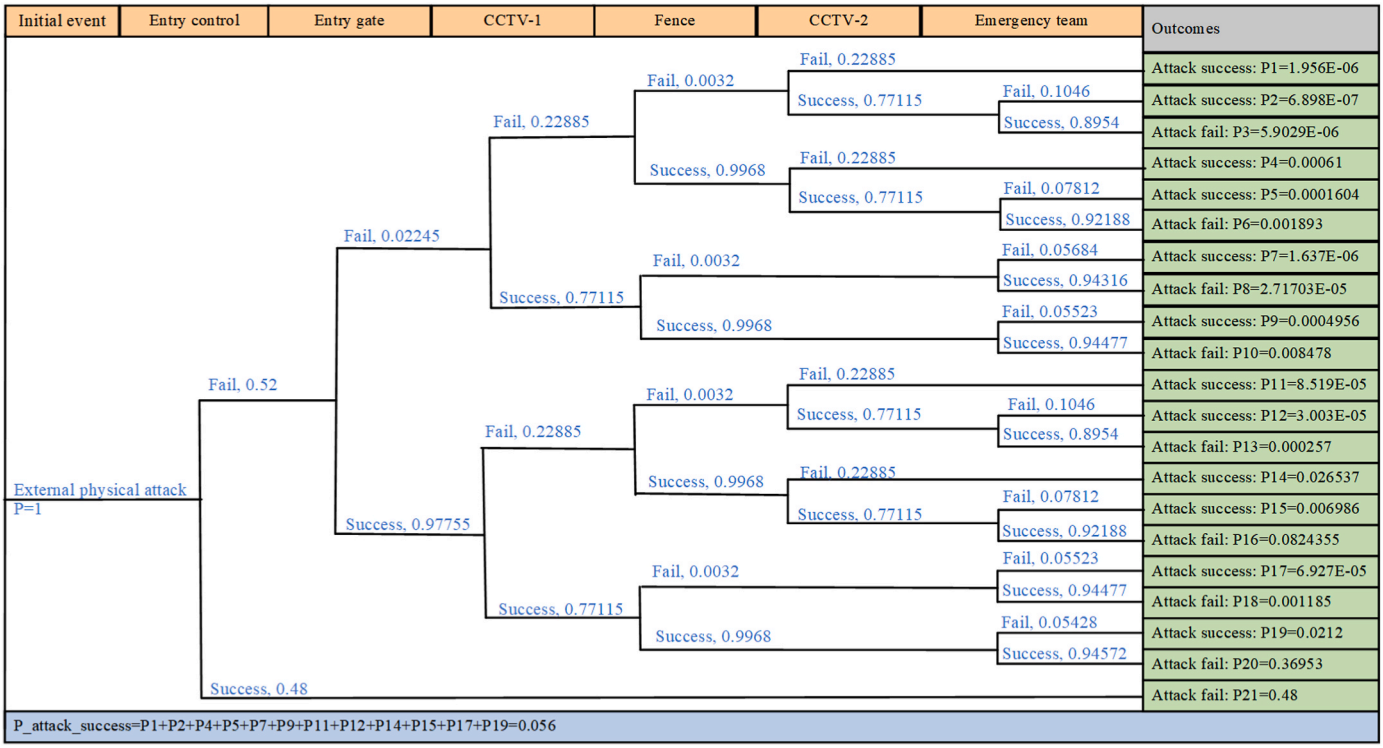


Fig. 10. Vulnerability assessment of PPSs regarding external physical attacks using an event tree.

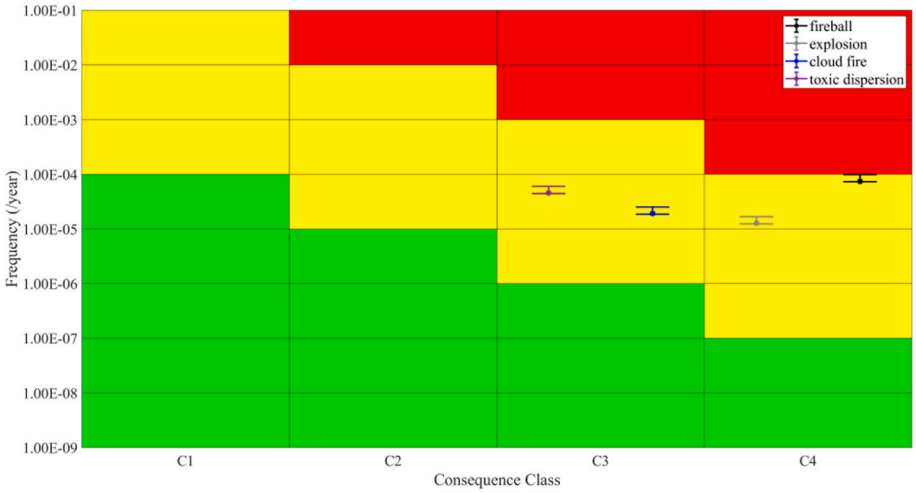


Fig. 11. Initial risk profiles demonstrated in a risk matrix.

an attack attempt is calculated as 0.056. This result is used as prior probability for root node BE25.

4.3. Case study results

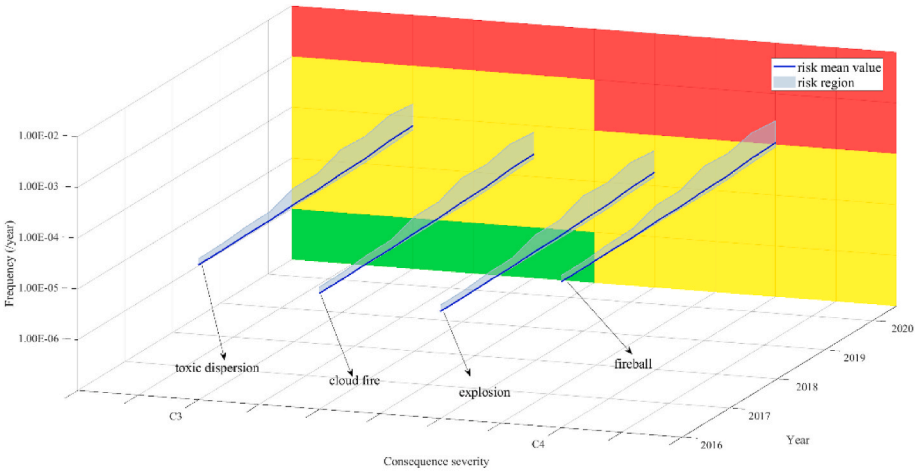
The BN model is developed and solved using the Bayes net MATLAB toolbox (Murphy, 2001). Because probability distributions are used for some root nodes, Monte Carlo simulations with 10,000 trials are conducted to handle uncertainty propagation in the risk assessment. The calculated mean values and ranges of the probability distributions for each possible consequence are visualized in a risk matrix, as shown in Fig. 11. In the risk matrix, the consequence severity classes are determined according to the European ARAMIS project (Andersen et al., 2004). Because the thresholds for integrated safety and security risks have been rarely investigated previously, we made modifications to the safety risk thresholds used in the ARAMIS project. We configured the thresholds a bit looser, considering the incorporation of security risks.

To demonstrate the capability of the proposed methodology in dynamic risk assessment and dynamic barrier management, we use some hypothetical data in relation to risk variations, as explained in Table 7. The risk profiles are updated based on the hypothetical data, as shown in Fig. 12.

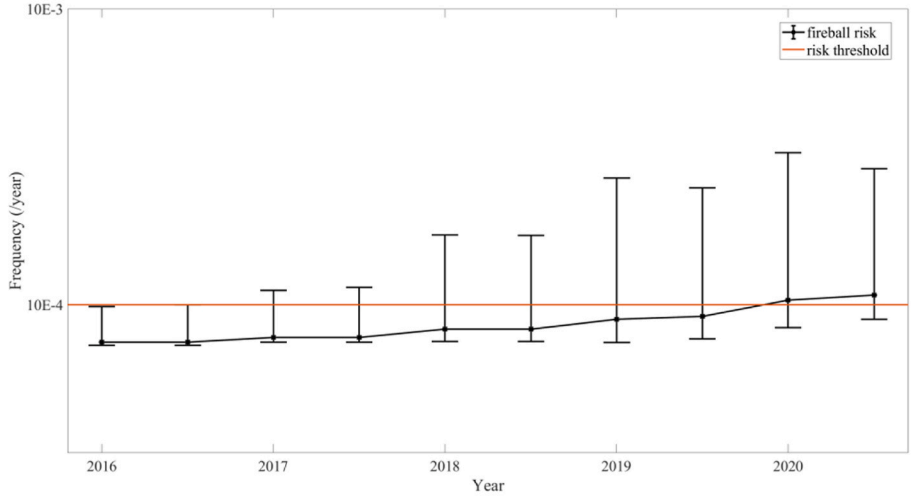
Fig. 12(a) shows the risk evolutions over time, obtained by updating risk profiles based on incorporating risk-related variations. Both the mean values and ranges of the risk profiles are demonstrated in the figure, considering risk uncertainties. Because only the fireball risk exceeds the risk threshold, the mean values and ranges of the fireball risk over time are highlighted in Fig. 12(b). In this case study, we assume that the decision-makers aim to ensure the maximum values of the risk ranges are below the risk thresholds from a conservative perspective.

Table 7
[12]Hypothetical data for updating risks.

Data types	Descriptions										
Periodic proof test data on the shutdown valve	Time	2016 Test 1	Test 2	2017 Test 1	Test 2	2018 Test 1	Test 2	2019 Test 1	Test 2	2020 Test 1	Test 2
	Degradation level[2]	1.5E-04	2.3E-04	3.5E-04	5.8E-04	8.5E-04	9.7E-04	1.15E-03	1.2E-03	1.23E-03	/
Cyber incident data (C2P attack attempts)	Time	2016		2017		2018		2019		2020	
	Cumulative attack attempts	2		7		15		25		36	
Acknowledgment of new CVEs	Time	2016		2017		2018		2019		2020	
	New known CVEs	CVE-2016-2200 (at attack steps: 10–17).		CVE-2017-2683 (at attack steps: 1); CVE-2017-13997 (at attack steps: 2).		CVE-2018-13799 (at attack steps: 3); CVE-2018-5459 (at attack steps: 6, 9).		/		/	
Physical attack precursor data[3]	Time	2016		2017		2018		2019		2020	
	Physical attack likelihood	0.002656		0.002724		0.002738		0.002707		0.002832	



(a) Dynamic risk profiles.



(b) Fireball risk over time.

Fig. 12. Dynamic risk profiles without timely S&S barrier improvement.

Therefore, necessary actions must be taken when the fireball risk exceeds its threshold in 2017.

A sensitivity analysis of basic events (root nodes of the BN model) is conducted based on the Birnbaum importance measure to identify critical events and help with the candidate strategy proposal. As shown in Fig. 13, BE24 and BE25 have the dominant sensitives, followed by BE16 to BE22. Among them, BE24 and BE25 are physical-attack-related nodes. BE18 (P safety failure), BE21 (V2 safety failure), and BE22 (SV

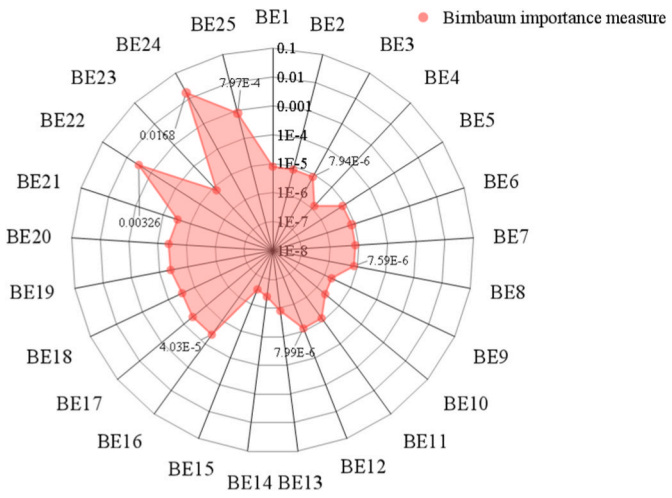


Fig. 13. Sensitivity analysis of BN root nodes.

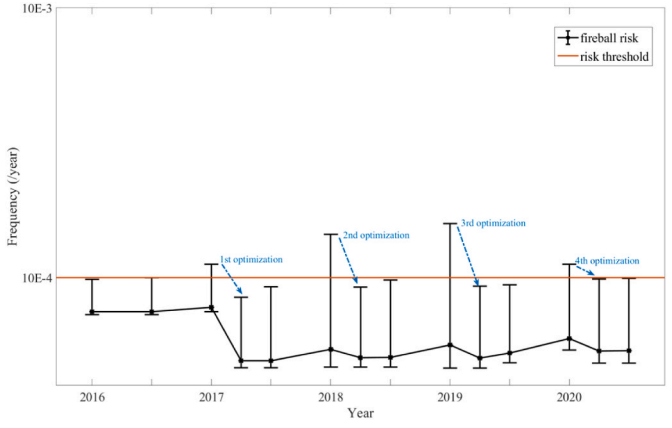


Fig. 14. Dynamic risk profiles with timely S&S barrier improvement.

safety failure) are related to safety barrier failures. BE17, BE19, and BE20 are associated with C2P attacks.

Accordingly, a group of candidate strategies are proposed considering their feasibility and effectiveness, as demonstrated in Table A3 in Appendix I. A cost-effectiveness analysis (CEA) of the candidate strategies is conducted to determine the strategy with the minimum cost while reducing the undesired risks below the risk thresholds. Based on the CEA, strategy No.4 is optimal for the first-round barrier improvement. Following the same procedures, three additional rounds of barrier optimization have been performed based on CEA when the undesired risks are unacceptable, as demonstrated in Table A3. Implementing the dynamic and integrated S&S barrier methodology allows cost-effective barrier improvement strategies to be derived timely whenever risk profiles are predicted to be unacceptable based on the new evidence. Therefore, it helps to make continuous barrier improvements and ensure the undesired risks are acceptable considering multiple dynamic risk variations. The dynamic risk profiles with timely S&S barrier improvement are demonstrated in Fig. 14.

As shown in Fig. 14, four rounds of barrier optimization are performed to reduce the fireball risk when it exceeds the risk threshold. The CEA-based barrier optimization process can be found in Table A3 in Appendix I. Compared to Fig. 12, the fireball risk is effectively mitigated by S&S barrier improvement when its maximum value is beyond the threshold in Fig. 14, which demonstrates the advantage of dynamic barrier management in continuous and timely risk treatment.

5. Discussions

5.1. Notes when applying the proposed framework

i) Sources of uncertainties

In the present methodology, models/methods from different domains (chemical process safety, physical security, and industrial cybersecurity) are leveraged to quantify the integrated safety and security risks based on multi-source data and experts' knowledge. On the one hand, the proposed methodology has the advantage of managing S&S barriers based on the integrated safety and security risks, which reveal the risks more realistically concerning the interactions between safety-associated events and security-associated events. On the other hand, more uncertainties are inevitably involved due to the extension of the risk assessment scopes and the integration of different methods/models with various natural features. Some assumptions are made due to the lack of background knowledge or pertinent data. For instance, rough reference values are used for human error probabilities, perfect barrier inspection and maintenance are assumed, rough reference values are used for the performance assessment of physical protection systems, etc. Those assumptions may hide or camouflage the pertinent uncertainties. Additionally, the use of expert knowledge, for instance, in the attack likelihood estimation, also brings subjective uncertainties. As a result, the derived optimal barrier improvement strategy may not definitely ensure perfect safety and security while saving costs due to the uncertainties involved.

Therefore, practitioners must be aware of the uncertainties when applying the proposed methodology. It is essential to state that the decision-making suggestions provided by the approach are subject to model uncertainties and the input data. It can only give valuable references for decision-making. Identifying uncertainty sources, alleviating uncertainties, and properly treating uncertainties in the decision-making process help to derive a barrier management solution with higher confidence.

ii) Alignment with relevant standards/regulations

Integrated S&S barrier management is a topic across multiple domains (chemical process safety, chemical process security, and industrial cybersecurity). Those domains are guided by different national, international, or industry standards/regulations. For instance, the IEC 62443 standard guides the industrial cybersecurity of ICSs (IEC:62443-2-1, 2010). The ANSI/API Standard 780 provides guidelines for the security risk assessment of petrochemical plants considering physical attacks (API, 2013). The IEC 61508 standard is dedicated to the functional safety of electronic safety-related systems (IEC:61508, 2010). A fundamental principle of the proposed framework lies in making decisions based on assessing integrated S&S risks. Without additional precautions, the results derived from the proposed methodology may conflict with other standards because existing standards/regulations solely emphasize safety risks or security risks. Therefore, the decision-making phase should carefully consider the alignment with pertinent standards/regulations to avoid possible conflicts.

As a part of the decision-making process, practitioners are supposed to propose candidate strategies for barrier improvements in case risk profiles are unacceptable. The candidate strategy proposal may consider the alignment with other pertinent standards/regulations to accommodate the requirements of relevant standards/regulations or legislation authorities. For example, the IEC 61508 standard determines Safety Integrity Levels (SILs) for safety-critical systems with different demand modes. If a degraded safety barrier fails to meet the requirement of the SIL, the maintenance/replacement of this barrier should be proposed as part of the barrier improvement strategy. The optimization functions can also consider the SIL requirements by adding corresponding technical constraints to accommodate the IEC 61508 standard. The

consideration and accommodation of pertinent standards/regulations in the candidate strategy proposal and strategy optimization help to derive more tailored solutions for barrier improvement.

iii) The determination of thresholds for integrated S&S risks

In the present study, risks of major event scenarios in chemical plants that safety causes and security causes may induce are assessed in a unified manner to support barrier management. The risks are called "integrated safety and security risks". The risk evaluation is conducted based on a risk matrix, in which the corresponding risk thresholds regarding typical disastrous phenomena in the chemical process industry are adapted from the European ARAMIS project (Andersen et al., 2004). The ARAMIS project provides an accidental risk assessment methodology for Seveso sites, and only safety issues were considered in the project. This study promotes a paradigm shift from managing safety risks and security risks separately to integrated safety and security risk approaches. Adjusting the current risk thresholds to accommodate the integrated safety and security risks is worthy of investigation in future studies. To address this issue, different stakeholders may be involved in determining the appropriate thresholds for integrated safety and security risks.

5.2. Establish a unified S&S risk management system

Current practice regarding safety and security risk management in chemical plants lacks in considering the interdependency between safety risks and security risks, and therefore can hardly achieve cost-effective risk management. The transition from handling safety risks and security risks separately to integrated safety and security risk management is urgently needed. On the one hand, the emerging cybersecurity risks threatening digitalized chemical facilities should be given more attention to help this paradigm shift. On the other hand, the establishment of a unified S&S risk management system helps barrier management in several aspects. i) Establishing a unified S&S risk management team helps information transfer and knowledge learning between the risk analysts and practitioners from either safety or security science domains. The integration of the knowledge from safety science and security science and the cooperation between experts and practitioners from both domains are necessary to achieve integrated S&S risk management. ii) Some subjective uncertainties involved in the risk assessment may be alleviated based on the analysis of pertinent data. A unified center for collecting and processing the multi-source data in relation to safety or security helps reduce subjective uncertainties in the risk assessment and manage S&S barriers in a dynamic manner based on new evidence. iii) A unified S&S management team helps with the systematic handling and coordination of various tasks and activities in relation to barrier management. Risk assessment, risk treatment (barrier improvements), and risk monitoring may be operated smoothly based on a unified risk management system.

Meanwhile, several challenges and knowledge gaps exist regarding unified safety and security (S&S) risk management, which may be explored in future studies: i) The performance of S&S barriers may vary across different domains, depending on safety-related and security-related scenarios. For example, the effectiveness of a manual shutdown might differ between accidental emergency scenarios and attack-induced emergencies due to variations in responders' cognition and awareness of accidental events versus intentional threats. Therefore, this variability in barrier performance under different conditions should be investigated within the context of integrated safety and security management. ii) The required levels of protection for safety and security events should also be a focus of future research on unified S&S risk management. Although this study primarily emphasizes potential common adverse scenarios or losses caused by both accidental and intentional events, the ultimate consequences may differ when considering factors such as ethics, reputation, and social impact. Consequently, different acceptable risk levels may need to be adopted for similar adverse scenarios depending on whether they arise from accidental events or intentional attacks. The distinction in protection levels needed for safety and security is a crucial area for investigation under the unified S&S risk management framework.

6. Conclusions

This study proposes a systematic framework for dynamic and integrated S&S barrier management in chemical process industries. The methodologies derived from various domains (process safety, physical security, and cybersecurity) are integrated to implement this framework in a hypothetical case study. The results demonstrate that multiple data about the system's safety and security can reveal risk-related variations and may be incorporated to enable dynamic risk assessment and further drive dynamic barrier management. Implementing dynamic and integrated S&S barrier management has the advantage of making timely adaptations according to the new evidence (risk-related variations) and, therefore, sustaining the safety and security of critical chemical facilities. Finally, practical notes for implementing this framework are discussed, and establishing a unified S&S risk management system is suggested to foster the implementation of dynamic and integrated S&S barrier management in practices.

CRediT authorship contribution statement

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

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.

Acknowledgments

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

Appendix I

Table A1
Attack annual probability estimation based on the API threat level and facility expected life (Λ, in year), adapted from (Landucci et al., 2017).

API threat level	Description	Attack annual probability
1	Little or no credible evidence of capability or intent, and no history of actual or planned threats against the facility.	$10^{-1} \times 1/\Lambda$
2	Low threat against the facility, few known adversaries would pose a threat to the asset.	$1/\Lambda$
3	Medium threat level, possible threat's desire to compromise similar assets, but no specific threat exists for the facility under analysis.	1×10^{-1}
4	A credible threat exists against the facility based on the knowledge of the threat's capability and intent to attack similar assets and some indication exists of the threat specific to the company, facility or asset.	2×10^{-1}
5	Some credible threat exists against the facility and the threat demonstrates the capability and intent to launch an attack; similar assets are attacked on a frequently recurring base and the frequency of attack is very high.	6×10^{-1}

Table A2
Recurrence intervals of cyber security incidents with different severities, adapted from (Kuypers and Maillart, 2018).

Effort spent to remediate incident (man-hours)	Recurrence intervals (days)	Effort spent to remediate incident (man-hours)	Recurrence intervals (days)
>6	2.99	>48	41.87
>12	8.02	>168	153.91
>24	24.17	>720	465.97

Table A3
Candidate barrier improvement strategies.

1st optimization	Strategy number	Candidate strategies	Cost analysis	Meet risk thresholds?	Optimal strategy?
	No.1	● Maintain V2 immediately	2,000€ (one-time maintenance cost)+ 100,000€ × 1 day (downtime cost) = 102,000€	No	/
	No.2	● Patch security vulnerability: CVE-2016-2200	20,000€ (patching cost)+100,000€ × 14 days (downtime cost) = 1,420,000€	No	/
	No.3	● Change maintenance interval for P to Six months	(5,000€ (one-time maintenance cost)+ 100,000€ × 2 day (downtime cost)) × 1 (annual increase in maintenance frequency) = 205,000€	No	/
	No.4	● Deploy one additional CCTV system to monitor the industrial area	2,000€ (installation cost)[4]+ 24,000€ (annual operation cost) = 26,000€	Yes	✓
	No.5	● Maintain V2 immediately ● Change maintenance interval for P to Six months	102,000€ (V2 maintenance cost)+ 205,000€ (P maintenance cost) = 307,000€	No	/
	No.6	● Patch security vulnerability: CVE-2016-2200 ● Change maintenance interval for P to Six months	1,420,000€ (patching CVE-2016-2200)+ 205,000€ (P maintenance cost) = 1,625,000€	No	/
	No.7	● Maintain V2 immediately ● Patch security vulnerability: CVE-2016-2200	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2016-2200) = 1,522,000€	No	/
	No.8	● Maintain V2 immediately ● Deploy one additional CCTV system to monitor the industrial area	102,000€ (V2 maintenance cost)+26,000€ (CCTV system) = 128,000€	Yes	/
	No.9	● Patch security vulnerability: CVE-2016-2200 ● Deploy one additional CCTV system to monitor the industrial area	1,420,000€ (patching CVE-2016-2200)+ 26,000€ (CCTV system) = 1,446,000€	Yes	/
	No.10	● Change maintenance interval for P to Six months ● Deploy one additional CCTV system to monitor the industrial area	205,000€ (P maintenance cost)+26,000€ (CCTV system) = 231,000€	Yes	/
	No.11	● Patch security vulnerability: CVE-2016-2200 ● Maintain V2 immediately ● Change maintenance interval for P to Six months	1,420,000€ (patching CVE-2016-2200)+ 102,000€ (V2 maintenance cost)+ 205,000€ (P maintenance cost) = 1,727,000€	No	/
	No.12	● Patch security vulnerability: CVE-2016-2200 ● Maintain V2 immediately ● Deploy one additional CCTV system to monitor the industrial area	1,420,000€ (patching CVE-2016-2200)+ 102,000€ (V2 maintenance cost)+26,000€ (CCTV system) = 1,548,000€	Yes	/

(continued on next page)

[4] Information from https://reolink.com/blog/security-camera-installation-cost/.

Table A3 (*continued*)

	No.13	● Patch security vulnerability: CVE-2016-2200 ● Change maintenance interval for P to Six months ● Deploy one additional CCTV system to monitor the industrial area	1,420,000€ (patching CVE-2016-2200)+ 205,000€ (P maintenance cost)+26,000€ (CCTV system) = 1,651,000€	Yes	/
	No.14	● Maintain V2 immediately ● Change maintenance interval for P to Six months ● Deploy one additional CCTV system to monitor the industrial area	102,000€ (V2 maintenance cost)+ 205,000€ (P maintenance cost)+ 26,000€ (CCTV system) = 333,000€	Yes	/
	No.15	● Patch security vulnerability: CVE-2016-2200 ● Maintain V2 immediately ● Change maintenance interval for P to Six months0 ● Deploy one additional CCTV system to monitor the industrial area	1,420,000€ (patching CVE-2016-2200)+ 102,000€ (V2 maintenance cost)+ 205,000€ (P maintenance cost)+ 26,000€ (CCTV system) = 1,753,000€	Yes	/
2nd optimization	**Strategy number**	**Candidate strategies**	**Cost analysis**	**Meet risk thresholds?**	**Optimal strategy?**
	No.1	● Change maintenance interval for P to Six months	(5,000€ (one-time maintenance cost)+ 100,000€ × 2 day (downtime cost)) × 1 (annual increase in maintenance frequency) = 205,000€	No	/
	No.2	● Maintain V2 immediately	2,000€ (one-time maintenance cost)+ 100,000€ × 1 day (downtime cost) = 102,000€	No	/
	No.3	● Patch CVE-2017-2683	20,000€ (patching cost)+100,000€ × 14 days (downtime cost) = 1,420,000€	No	/
	No.4	● Patch CVE-2017-13997	20,000€ (patching cost)+100,000€ × 14 days (downtime cost) = 1,420,000€	No	/
	No.5	● Change maintenance interval for P to Six months ● Maintain V2 immediately	205,000€ (P maintenance cost)+102,000€ (V2 maintenance cost) = 307,000€	No	/
	No.6	● Change maintenance interval for P to Six months ● Patch CVE-2017-2683	205,000€ (P maintenance cost)+ 1,420,000€ (patching CVE-2017-2683) = 1,625,000€	No	/
	No.7	● Change maintenance interval for P to Six months ● Patch CVE-2017-13997	205,000€ (P maintenance cost)+ 1,420,000€ (patching CVE-2017-13997) = 1,625,000€	No	/
	No.8	● Maintain V2 immediately ● Patch CVE-2017-2683	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2017-2683) = 1,522,000€	No	/
	No.9	● Maintain V2 immediately ● Patch CVE-2017-13997	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2017-13997) = 1,522,000€	No	/
	No.10	● Patch CVE-2017-2683 ● Patch CVE-2017-13997	1,420,000€ (patching CVE-2017-2683)+ 1,420,000€ (patching CVE-2017-13997) = 2,840,000€	Yes	✓
	No.11	● Change maintenance interval for P to Six months ● Maintain V2 immediately ● Patch CVE-2017-2683	205,000€ (P maintenance cost)+102,000€ (V2 maintenance cost)+1,420,000€ (patching CVE-2017-2683) = 1,727,000€	No	/
	No.12	● Change maintenance interval for P to Six months ● Maintain V2 immediately ● Patch CVE-2017-13997	205,000€ (P maintenance cost)+102,000€ (V2 maintenance cost)+1,420,000€ (patching CVE-2017-13997) = 1,727,000€	No	/
	No.13	● Change maintenance interval for P to Six months ● Patch CVE-2017-2683 ● Patch CVE-2017-13997	205,000€ (P maintenance cost)+ 1,420,000€ (patching CVE-2017-2683)+ 1,420,000€ (patching CVE-2017-13997) = 3,045,000€	Yes	/
	No.14	● Maintain V2 immediately ● Patch CVE-2017-2683 ● Patch CVE-2017-13997	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2017-2683)+ 1,420,000€ (patching CVE-2017-13997) = 2,942,000€	Yes	/
	No.15	● Change maintenance interval for P to Six months ● Maintain V2 immediately ● Patch CVE-2017-2683 ● Patch CVE-2017-13997	205,000€ (P maintenance cost)+102,000€ (V2 maintenance cost)+1,420,000€ (patching CVE-2017-2683)+1,420,000€ (patching CVE-2017-13997) = 3,147,000€	Yes	/
3rd optimization	**Strategy number**	**Candidate strategies**	**Cost analysis**	**Meet risk thresholds?**	**Optimal strategy?**
	No.1	● Maintain V2 immediately	2,000€ (one-time maintenance cost)+ 100,000€ × 1 day (downtime cost) = 102,000€	No	/
	No.2	● Patch security vulnerability: CVE-2016-2200	20,000€ (patching cost)+100,000€ × 14 days (downtime cost) = 1,420,000€	No	/
	No.3	● Patch security vulnerability: CVE-2018-13799	20,000€ (patching cost)+100,000€ × 14 days (downtime cost) = 1,420,000€	No	/
	No.4	● Patch security vulnerability: CVE-2018-5459	20,000€ (patching cost)+100,000€ × 14 days (downtime cost) = 1,420,000€	No	/
	No.5	● Maintain V2 immediately ● Patch security vulnerability: CVE-2016-2200	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2016-2200) = 1,522,000€	No	/

(*continued on next page*)

Table A3 (*continued*)

	No.6	● Maintain V2 immediately ● Patch security vulnerability: CVE-2018-13799	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2018-13799) = 1,522,000€	No	/
	No.7	● Maintain V2 immediately ● Patch security vulnerability: CVE-2018-5459	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2018-5459) = 1,522,000€	No	/
	No.8	● Patch security vulnerability: CVE-2016-2200 ● Patch security vulnerability: CVE-2018-13799	1,420,000€ (patching CVE-2016-2200)+ 1,420,000€ (patching CVE-2018-13799) = 2,840,000€	No	/
	No.9	● Patch security vulnerability: CVE-2016-2200 ● Patch security vulnerability: CVE-2018-5459	1,420,000€ (patching CVE-2016-2200)+ 1,420,000€ (patching CVE-2018-5459) = 2,840,000€	No	/
	No.10	● Patch security vulnerability: CVE-2018-13799 ● Patch security vulnerability: CVE-2018-5459	1,420,000€ (patching CVE-2018-13799)+ 1,420,000€ (patching CVE-2018-5459) = 2,840,000€	No	/
	No.11	● Maintain V2 immediately ● Patch security vulnerability: CVE-2016-2200 ● Patch security vulnerability: CVE-2018-13799	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2016-2200)+ 1,420,000€ (patching CVE-2018-13799) = 2,942,000€	No	/
	No.12	● Maintain V2 immediately ● Patch security vulnerability: CVE-2016-2200 ● Patch security vulnerability: CVE-2018-5459	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2016-2200)+ 1,420,000€ (patching CVE-2018-5459) = 2,942,000€	No	/
	No.13	● Maintain V2 immediately ● Patch security vulnerability: CVE-2018-13799 ● Patch security vulnerability: CVE-2018-5459	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2018-13799)+ 1,420,000€ (patching CVE-2018-5459) = 2,942,000€	No	/
	No.14	● Patch security vulnerability: CVE-2016-2200 ● Patch security vulnerability: CVE-2018-13799 ● Patch security vulnerability: CVE-2018-5459	1,420,000€ (patching CVE-2016-2200)+ 1,420,000€ (patching CVE-2018-13799)+ 1,420,000€ (patching CVE-2018-5459) = 4,260,000€	Yes	✓
	No.15	● Maintain V2 immediately ● Patch security vulnerability: CVE-2016-2200 ● Patch security vulnerability: CVE-2018-13799 ● Patch security vulnerability: CVE-2018-5459	102,000€ (V2 maintenance cost)+ 1,420,000€ (patching CVE-2016-2200)+ 1,420,000€ (patching CVE-2018-13799)+ 1,420,000€ (patching CVE-2018-5459) = 4,362,000€	Yes	/
4th optimization	**Strategy number**	**Candidate strategies**	**Cost analysis**	**Meet risk thresholds?**	**Optimal strategy?**
	No.1	● Change maintenance interval for P to Six months	(5,000€ (one-time maintenance cost)+ 100,000€ × 2 day (downtime cost)) × 1 (annual increase in maintenance frequency) = 205,000€	No	/
	No.2	● Change maintenance interval for P to Three months	(5,000€ (one-time maintenance cost)+ 100,000€ × 2 day (downtime cost)) × 3 (annual increase in maintenance frequency) = 615,000€	No	/
	No.3	● Maintain V2 immediately	2,000€ (one-time maintenance cost)+ 100,000€ × 1 day (downtime cost) = 102,000€	Yes	✓
	No.4	● Change maintenance interval for P to Six months ● Maintain V2 immediately	102,000€ (V2 maintenance cost)+ 205,000€ (P maintenance cost) = 307,000€	Yes	/
	No.5	● Change maintenance interval for P to Three months ● Maintain V2 immediately	102,000€ (V2 maintenance cost)+ 615,000€ (P maintenance cost) = 717,000€	Yes	/

Appendix II

Considering the uncertainties associated with attackers' knowledge levels, we involve two attack path selection mechanisms in the vulnerability assessment. They are random attacks and strategic attacks, respectively. Random attack presents an attacker who selects one attack path from all credible attack paths. A strategic attack means an attacker selects the attack path based on the exploitability of all credible attack paths, and a more exploitable attack path is more likely to be selected. Based on that, we have assigned different probabilities of executing random and strategic attacks for the attackers with different knowledge levels, as shown below.

Table A4
Attack path selection mechanisms for attackers with different knowledge levels (Yuan et al., 2024b).

Attacker categories[5]	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

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

$$Pr(i) = \frac{a}{n} + \frac{b}{GTTC_i} \Big/ \sum_{j=1}^{n} \frac{1}{GTTC_j} \tag{A1}$$

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 can refer to Table A4. $GTTC_i$ is the global time-to-compromise of attack path i. It is calculated by summing the TTC (time-to-compromise) of each attack step along attack path i. The TTC of each attack step is calculated using the method Appendix I in Yuan et al. (2024b).

The conditional probability of attack path i is executed successfully given an attack attempt (L^i) is estimated as follows:

$$L^i = Pr(i) \times \frac{MTTD_i}{GTTC_i + MTTD_i} \times \beta_i \tag{A2}$$

$$MTTD_i = \frac{\sum_{k=1}^{N} TTD_k}{N} \tag{A3}$$

$$L_{at}^i = L^a + L^b, \ldots, L^n \tag{A4}$$

where $MTTD_i$ presents the mean-time-to-detect for attack path i. The MTTD regarding a specific intrusion type is calculated by averaging all incident detection times of this intrusion type, as shown in Eq. (A3). In this case study, a reference value (14 days) from Semertzis et al. (2022) is used for the MTTD. In practice, it should be determined based on actual incident data. In cases where multiple attack paths lead to the same attack mode, the conditional probability of successful execution of the attack mode (L_{at}^i) is calculated by summing up the L^i values of those attack paths, as shown in Eq. (A4). Coefficient β_i describes the likelihood that a successful intrusion of attack path i induces a dangerous phenomenon. β_i depends on the fault detection capability and deviation tolerance capability of the OT system. β_i is calculated as $\beta_i = \beta_i^d \times \beta_i^r$. β_i^d presents the probability that the attack-induced deviations escape the anomaly detection algorithm successfully. β_i^r presents the likelihood that the attack-induced deviations cause a dangerous phenomenon successfully. β_i^d should be determined considering both the attack mode and the fault detection algorithm of the system. For simplicity, we assumed predefined ranges for sensors' and actuators' signals as the fault detection scheme (Huang et al., 2009). 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. Reference β_i^d values are summarized in the table below. In practice, the β_i^d values should be modified considering the specific performance of the fault detection algorithms. β_i^r is determined using a stochastic attack modeling approach, as illustrated in Appendix II in Yuan et al. (2024b).

Table A5
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

Data availability

Data will be made available on request.

References

Abdo, H., Kaouk, M., Flaus, J.M., Masse, F., 2018. A safety/security risk analysis approach of Industrial Control Systems: a cyber bowtie–combining new version of attack tree with bowtie analysis. Comput. Secur. 72, 175–195.

American Petroleum Institute (API), 2013. ANSI/API Standard 780-Security Risk Assessment Methodology for the Petroleum and Petrochemical Industry. American Petroleum Institute, Washington, DC.

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.

Argenti, F., Landucci, G., Cozzani, V., Reniers, G., 2017. A study on the performance assessment of anti-terrorism physical protection systems in chemical plants. Saf. Sci. 94, 181–196.

Casciano, M., Khakzad, N., Reniers, G., Cozzani, V., 2019. Ranking chemical industrial clusters with respect to safety and security using analytic network process. Process Saf. Environ. Prot. 132, 200–213.

[5] Attacker categories are adapted from the TTC-based approach (McQueen et al., 2006).

Cozzani, V., Gubinelli, G., Antonioni, G., Spadoni, G., Zanelli, S., 2005. The assessment of risk caused by domino effect in quantitative area risk analysis. J. Hazard Mater. 127 (1–3), 14–30.
Chen, C., Reniers, G., Khakzad, N., 2019. Integrating safety and security resources to protect chemical industrial parks from man-made domino effects: a dynamic graph approach. Reliab. Eng. Syst. Saf. 191, 106470.
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 100. France.
Garcia, M.L., 2007. Design and Evaluation of Physical Protection Systems, second ed. Elsevier.
Guzman, N.H.C., Kozine, I., Lundteigen, M.A., 2021. An integrated safety and security analysis for cyber-physical harm scenarios. Saf. Sci. 144, 105458.
Gubinelli, G., Zanelli, S., Cozzani, V., 2004. A simplified model for the assessment of the impact probability of fragments. J. Hazard Mater. 116 (3), 175–187.
Hauge, S., Onshus, T., 2010. Reliability Data for Safety Instrumented Systems PDS Data Handbook, 2010 Edition. SINTEF Report A, 13502.
HSE, U., 2012. Failure Rate and Event Data for Use within Risk Assessments, 28/06/2012).
Huang, Y.L., Cárdenas, A.A., Amin, S., Lin, Z.S., Tsai, H.Y., Sastry, S., 2009. Understanding the physical and economic consequences of attacks on control systems. International Journal of Critical Infrastructure Protection 2 (3), 73–83.
IEC:61508, 2010. Functional Safety of Electrical/electronic/programmable Electronic Safety-Related Systems. IEC Standards Online.
IEC:62443-2-1, 2010. Industrial communication networks - network and system security - Part 2-1: establishing an industrial automation and control system security program. Retrieved from. https://webstore.iec.ch/publication/7030.
Iaiani, M., Tugnoli, A., Bonvicini, S., Cozzani, V., 2021a. Analysis of cybersecurity-related incidents in the process industry. Reliab. Eng. Syst. Saf. 209, 107485.
Iaiani, M., Tugnoli, A., Bonvicini, S., Cozzani, V., 2021b. Major accidents triggered by malicious manipulations of the control system in process facilities. Saf. Sci. 134, 105043.
Iaiani, M., Tugnoli, A., Cozzani, V., 2023. Identification of cyber-risks for the control and safety instrumented systems: a synergic framework for the process industry. Process Saf. Environ. Prot. 172, 69–82.
Iaiani, M., Fazari, G., Tugnoli, A., Cozzani, V., 2024. Identification of reference security scenarios from past event datasets by Bayesian Network analysis. Reliab. Eng. Syst. Saf. 254, 110615.
Khakzad, N., Khan, F., Amyotte, P., 2012. Dynamic risk analysis using bow-tie approach. Reliab. Eng. Syst. Saf. 104, 36–44.
Khakzad, N., Khan, F., Amyotte, P., 2013. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. Process Saf. Environ. Prot. 91 (1–2), 46–53.
Khakzad, N., Martinez, I.S., Kwon, H.M., Stewart, C., Perera, R., Reniers, G., 2018. Security risk assessment and management in chemical plants: challenges and new trends. Process Saf. Prog. 37 (2), 211–220.
Kirwan, B., 2017. A Guide to Practical Human Reliability Assessment. CRC press.
Kuypers, M., Maillart, T., 2018. Designing organizations for cyber security resilience. In: Proceedings of the 2018 the Workshop on the Economics of Information Security (WEIS), pp. 18–19. Innsbruck, Austria.
Landucci, G., Argenti, F., Cozzani, V., Reniers, G., 2017. Assessment of attack likelihood to support security risk assessment studies for chemical facilities. Process Saf. Environ. Prot. 110, 102–114.
McQueen, M.A., Boyer, W.F., Flynn, M.A., Beitel, G.A., 2006. Time-to-compromise model for cyber risk reduction estimation. In: Quality of Protection: Security Measurements and Metrics. Springer US, pp. 49–64.
Moreno, V.C., Marroni, G., Landucci, G., 2022. Probabilistic assessment aimed at the evaluation of escalating scenarios in process facilities combining safety and security barriers. Reliab. Eng. Syst. Saf. 228, 108762.
Murphy, K., 2001. The bayes net toolbox for matlab. Computing Science and Statistics 33 (2), 1024–1034.
Mughal, A.A., 2022. Building and securing the modern security operations center (SOC). International Journal of Business Intelligence and Big Data Analytics 5 (1), 1–15.
Norman, T.L., 2010. Risk Analysis and Security Countermeasure Selection. CRC press, Boca Raton/London/New York.
National Vulnerability Database (NVD). (n.d.). Retrieved November 17, 2024, from https://nvd.nist.gov/.
OREDA, 2002. Offshore Reliability Data Handbook. DNV, Trondheim, Norway.
Pilario, K.E.S., Cao, Y., 2018. Canonical variate dissimilarity analysis for process incipient fault detection. IEEE Trans. Ind. Inf. 14 (12), 5308–5315.
Reniers, G., Dullaert, W., Audenaert, A., Ale, B., Soudan, K., 2008. Managing domino effect-related security of industrial areas. J. Loss Prev. Process. Ind. 21 (3), 336–343.
Song, G., Khan, F., Yang, M., 2019. Probabilistic assessment of integrated safety and security related abnormal events: a case of chemical plants. Saf. Sci. 113, 115–125.
Semertzis, I., Rajkumar, V.S., Ştefanov, A., Fransen, F., Palensky, P., 2022. Quantitative risk assessment of cyber attacks on cyber-physical systems using attack graphs. In: 2022 10th Workshop on Modelling and Simulation of Cyber-Physical Energy Systems (MSCPES). IEEE, pp. 1–6.
START (National Consortium for the Study of Terrorism and Responses to Terrorism), 2022. Global terrorism database 1970 - 2020 [data file]. https://www.start.umd.edu/gtd.
Störfall Kommission (SFK), 2002. SFK-GS-38 Report.
Siu, N.O., Kelly, D.L., 1998. Bayesian parameter estimation in probabilistic risk assessment. Reliab. Eng. Syst. Saf. 62 (1–2), 89–116.
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 2024, November).
Van der Borst, M., Schoonakker, H., 2001. An overview of PSA importance measures. Reliab. Eng. Syst. Saf. 72 (3), 241–245.
Van Staalduinen, M., Khan, F., 2015. A barrier based methodology to assess site security risk. In: SPE Health, Safety, Security, Environment, & Social Responsibility Conference-North America, pp. SPE–173561. SPE.
Vílchez, J.A., Espejo, V., Casal, J., 2011. Generic event trees and probabilities for the release of different types of hazardous materials. J. Loss Prev. Process. Ind. 24 (3), 281–287.
Yazdi, M., Kabir, S., 2017. A fuzzy Bayesian network approach for risk analysis in process industries. Process Saf. Environ. Prot. 111, 507–519.
Ylönen, M., Tugnoli, A., Oliva, G., Heikkilä, J., Nissilä, M., Iaiani, M., et al., 2022. Integrated management of safety and security in Seveso sites-sociotechnical perspectives. Saf. Sci. 151, 105741.
Yuan, S., Yang, M., Reniers, G., Chen, C., Wu, J., 2022a. Safety barriers in the chemical process industries: a state-of-the-art review on their classification, assessment, and management. Saf. Sci. 148, 105647.
Yuan, S., Cai, J., Reniers, G., Yang, M., Chen, C., Wu, J., 2022b. Safety barrier performance assessment by integrating computational fluid dynamics and evacuation modeling for toxic gas leakage scenarios. Reliab. Eng. Syst. Saf. 226, 108719.
Yuan, S., Reniers, G., Yang, M., Bai, Y., 2023a. Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. Process Saf. Environ. Prot. 170, 356–371.
Yuan, S., Reniers, G., Yang, M., 2023b. Dynamic-risk-informed safety barrier management: an application to cost-effective barrier optimization based on data from multiple sources. J. Loss Prev. Process. Ind. 83, 105034.
Yuan, S., Yang, M., Reniers, G., 2024a. Integrated process safety and process security risk assessment of industrial cyber-physical systems in chemical plants. Comput. Ind. 155, 104056.
Yuan, S., Reniers, G., Yang, M., 2024b. Integrated management of safety and security barriers in chemical plants to cope with emerging cyber-physical attack risks under uncertainties. Reliab. Eng. Syst. Saf. 250, 110320.
Zarei, E., Azadeh, A., Khakzad, N., Aliabadi, M.M., Mohammadfam, I., 2017. Dynamic safety assessment of natural gas stations using Bayesian network. J. Hazard Mater. 321, 830–840.
Zhang, A., Zhang, T., Barros, A., Liu, Y., 2020. Optimization of maintenances following proof tests for the final element of a safety-instrumented system. Reliab. Eng. Syst. Saf. 196, 106779.