Dynamic and Integrated Safety and Security Barrier Management in the Chemical Process Industries

Shuaiqi YUAN

Delft University of Technology

Dynamic and Integrated Safety and Security Barrier Management in the Chemical Process Industries

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Broad for Doctorates
to be defended publicly on
Tuesday, 26 November 2024, at 15:00 o' clock
by

Shuaiqi YUAN

Master of Engineering in Safety Engineering,
China University of Mining and Technology, Beijing, China
born in Jincheng, Shanxi, China

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairman
Prof.dr.ir. G.L.L.M.E. Reniers	Delft University of Technology, promotor
Dr. M. Yang	Delft University of Technology, copromotor

Independent members:

Prof.dr. V. Cozzani	University of Bologna, Italy
Prof.dr.ir. P.H.A.J.M. van Gelder	Delft University of Technology
Dr. G. Landucci	University of Pisa, Italy
Prof.dr.ir. J.R. van Ommen	Delft University of Technology
Prof.dr. J. Wu	China University of Mining and Technology, Beijing, China

This research is funded by Delft University of Technology (TU Delft) and China Scholarship Council (CSC) under Grant 202006430007.

Cover illustration by: Shuaiqi Yuan
Published and distributed by: Shuaiqi Yuan
Printed by: Ipskamp Printing
Copyright © 2024 by: Shuaiqi Yuan
ISBN: 978-94-6366-956-6

Printed in the Netherlands

To my family

Preface

Reflecting on the past four years, I see my PhD journey as an invaluable adventure. Although it presented challenges and moments of struggle, this experience has brought tremendous growth. I am deeply grateful to everyone who has accompanied, inspired, encouraged, and supported me along the way.

First and foremost, I would like to express my heartfelt gratitude to my promotor, Prof. Genserik Reniers. You played a pivotal role in my decision to pursue a PhD at TU Delft. Thank you for inspiring me with an exciting research topic and for guiding me throughout my PhD. You have been a consistent source of learning, and I aspire to become a researcher like you — someone who is interesting, open-minded, knowledgeable, productive, time-efficient, and, of course, handsome. I am also very grateful to my co-promotor, Dr. Ming Yang. Your patient guidance and insightful comments have significantly improved the quality of my research. I appreciate the immense support from both of you over the past four years.

I am also grateful to my independent doctoral committee members: Prof. V. Cozzani, Prof. P.H.A.J.M. van Gelder, Prof. J.R. van Ommen, Dr. G. Landucci, and Prof. J. Wu. It is a true honor to have esteemed researchers like you as part of my PhD committee. Thank you for agreeing to be involved, and for offering valuable evaluations and feedback on my dissertation.

I want to thank all my colleagues and PhD peers in our 3SG section: Prof. Pieter van Gelder, Prof. Jop Groeneweg, Dr. Perla Marang-van de Mheen, Dr. Eleonora Papadimitriou, Dr. Karolien van Nunen, Dr. Britte Bouchaut, Dr. Frank Guldenmund, Dr. Amir Pooyan Afghari, Dr. Oscar Oviedo-Trespalacios, Dr. Irene Grossmann, Astrid Pinzger, Nelleke van der Lee, Monika Polus, Dr. Rioshar Yarveisy, Dr. Yue Shang, Dr. Amna Chaudhry, Dr. Amir Hossein Kalantari, Jaco Tresfon, Rongxin Song, Dimmy van Dongen, Nienke Luijcks, Muhammad Bin Ab Rahim, Rustam Abubakirov, Jelle Knibbe, Gizem Ayerdem, Ajay Iyer, Louis Cleef, Mohammad Pashaee, and Mattia Zene. The engaging discussions and memorable activities we shared (the 3SG 40th-anniversary celebration, Christmas dinners, bowling evening, etc.) have enriched my PhD life. Special thanks to Prof. Pieter van Gelder for always warmly inviting us for lunch with relaxing talks and to Ms. Astrid Pinzger for always being warm-hearted and willing to help.

I am deeply grateful to my friends who brought joy to my PhD life: Prof. Chao Chen, Dr. Xin Ren, Dr. Sjonnie Boonstra, Rongxin Song, Dr. Yiping Bai, Dr. Tao Zeng, Weipeng Fang, Xinge Han, Jitao Cai, Zhuqiang Hu, Ruixue Li, Chunyang Dong, Dr. Mengxia Li, Prof. Liang Huang, Dr. Yunfei Huang, Xin Xiong, Hanting Ye, Siyuan Wang, Shengnan Zhang, Zichao Li, Shilun Zhang, and Dr. Mingxin Li. Whether through interesting conversations, sports, video games, table games, BBQs, or hotpot dinners, each of you has made my life more enjoyable and memorable.

I would like to acknowledge the financial support from the China Scholarship Council

(CSC) for funding my PhD research. I am grateful to my former supervisor, Prof. Jiansong Wu, for your guidance during my master's studies. You introduced me to the world of academia and empowered me to embark on this challenging and fascinating journey. I am thankful to the senior researchers and professors who have motivated and inspired me and who serve as role models in my professional career. Thanks to Prof. Rui Zhou and Prof. Xiaole Zhang for your guidance on my research and your support for my career development. A big thanks to Prof. Fuqiang Yang for inviting me for a guest lecture and a joyful trip to Fuzhou. Thanks to Prof. Yunxiao Fan, Prof. Jie Li, Dr. Laobin Zhang, Dr. Jie Xue, Prof. Yamin Huang, and Prof. Pengfei Chen. Your dedication and engagement in the 3SG unofficial Chinese group have been a source of inspiration and motivation for me to become a good researcher.

To my girlfriend, Yuanyuan, thank you for your unwavering support through all the highs and lows.

感谢我的家人，你们的支持是我无尽的力量源泉。

Research is an endless journey, and I am prepared for the challenges that lie ahead.

Shuaiqi Yuan

Delft, October 2024

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BN	Bayesian network
BPCS	Basic process control system
C2P	Cyber-to-physical
CEA	Cost-effectiveness analysis
CVEs	Common Vulnerabilities and Exposures
CSTR	Continuous stirred tank reactor
CBM	Condition-based maintenance
DBM	Dynamic barrier management
DBN	Dynamic Bayesian network
DRA	Dynamic risk assessment
DiD	Defence-in-depth
ESD	Emergency shutdown system
ETA	Event Tree Analysis
EASI	Estimate of Adversary Sequence Interruption
FTA	Fault Tree Analysis
GA	Genetic algorithm
GTD	Global Terrorism Database
HEP	Human error probability
HRA	Human Reliability Analysis
HMI	Human-machine interface
HAZID	HAZard Identification
ICS	Industrial control system
ICPS	Industrial cyber-physical system
IPL	Independent protection layer
IMSS	Integrated management of safety and security
LOPA	Layer of protection analysis
MTTD	Mean-time-to-detect
MDS	Management delivery system
MOE	Multiple occurring events
PPS	Physical protection system
PM	Preventive maintenance
PLC	Programmable logic controller
PFD	Probability of failure on demand
QRA	Quantitative risk assessment
RCM	Reliability centered maintenance
RBI	Risk based inspection
RBD	Reliability block diagram
S&S	Safety and security
SCADA	Supervisory control and data acquisition system
SIS	Safety instrumented system
SOC	Security operations center
SILs	Safety Integrity Levels
SSBM	Simulink-based safety barrier modeling
TTC	Time-to-compromise
VCE	Vapor Cloud Explosion

Chapter 1 Introduction

Due to the storage and usage of large quantities of hazardous substances, chemical process industries are subject to important safety and security issues. Particularly, with the advent of the digital age and Industry 4.0, new security threats and risks have emerged in the chemical process industries. Because both safety hazards and security threats may induce catastrophic accident scenarios, managing safety and security (S&S) barriers in an integrated manner to prevent the happening of major adverse events and mitigate their consequences is significantly important. This dissertation is devoted to performing a quantitative risk assessment of industrial control systems (ICSs) in chemical plants considering a variety of risk sources (safety causes and security threats) and further developing an approach for integrated and dynamic management of S&S barriers from a cost-effective point of view. This chapter aims to elaborate on the research background, motivations, research questions, and main contributions of this study. An outline of the dissertation and a summary of the publications generated from this PhD study are also presented to give insights into the overview of this dissertation.

1.1 Background

Accidental and intentional undesired events threaten the chemical process industries due to the potential damage to humans, assets, and the environment that may be caused if such events happen. Notably, major adverse events, such as toxic leakages, fires, and explosions triggered by either safety hazards or security threats, may induce catastrophic consequences. As critical infrastructures, the investigation of physical security issues in the chemical plants was promoted, especially after the 9/11 event happened in the United States of America (Reniers et al., 2008). To meet the development needs of Industry 4.0, industrial cyber-physical systems (ICPSs), also called industrial control systems (ICSs), are increasingly implemented to chemical facilities. The integration of digital technologies into chemical process operations increases the systems' complexity and introduces new security vulnerabilities in many cases (Flaus, 2019). With the automation and digitization of chemical process facilities, cyber-to-physical (C2P) attack-associated risks have emerged in the chemical process industries and also got attention because they may induce undesired events the same as those induced by critical safety failures (Kriaa et al., 2015). For instance, the malware "Stuxnet" is regarded as the world's first publically known digital weapon, which can target programmable logic controllers (PLCs) and induce physical damage (Kushner, 2013). As a result, chemical facilities are exposed to both safety and security risks, and the analogy between safety risk and security risk is shown in Figure 1.1.

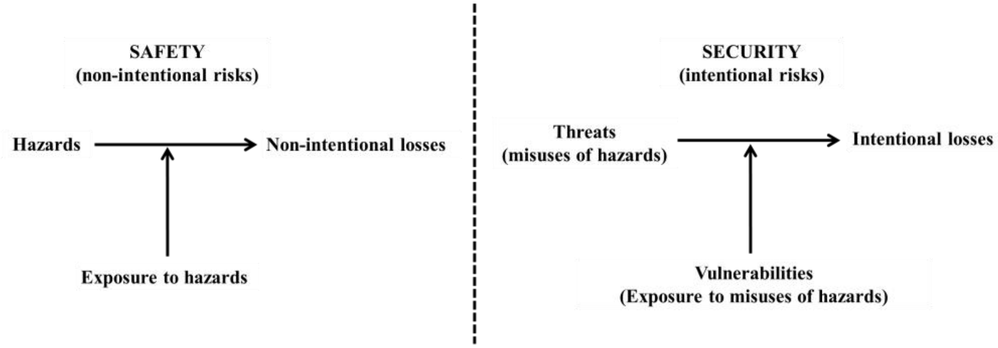


Figure 1.1. An analogy between safety risk and security risk, adapted from Landucci et al. (2020).

The multi-dimensional risks threatening chemical plants and associated with undesired dangerous scenarios (fires, explosions, toxic leakage, etc.) are summarized as follows.

- i) Safety risks affiliated with safety hazards/causes, including accidental technical component failures, human errors, human violations, external interventions, etc. that may accidentally lead to losses.

- ii) Physical security risks affiliated with intentional attacks/malicious acts aiming to exploit the vulnerability of physical elements (usually not including information systems) to cause losses deliberately.
- iii) Cyber-to-physical (C2P) risks affiliated with intentional attacks/malicious acts aiming to impact physical elements by exploiting the vulnerability of cyber elements (usually through attacks on information systems) to deliberately cause losses.

To get insights into the above-mentioned multi-dimensional risks, typical ICSs in chemical plants are selected as target systems under investigation in this thesis. For example, a chemical reactor with its supervisory control and data acquisition (SCADA) system is demonstrated in Figure 1.2. In this system, information and communication infrastructures are integrated with the physical process system to achieve the so-called digital control and operation. Both the basic process control system (BPCS) and the safety instrumented system (SIS) are controlled by programmable logic controllers (PLCs). Both PLCs are connected to the SCADA system and linked to the corporate network and the outside Internet/WAN. Site managers can access the information collected by the SCADA system and control the reaction process through the human-machine interface (HMI).

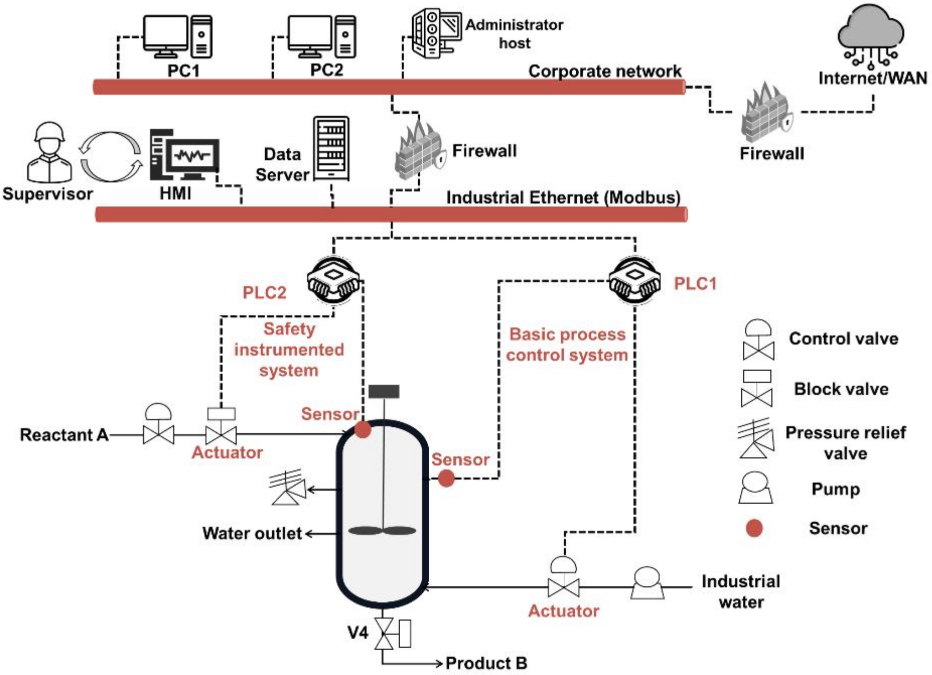


Figure 1.2. A demonstrative industrial control systems (ICS) in chemical plants.

Safety and security risks posed to ICSs inevitably interact with each other because security-related events may influence safety-related events and vice versa. For example, intentional attacks on the safety instrumented system (SIS) may increase safety risks due to the loss of function of the SIS. Also, specific safety protection equipment (for instance, pressure relief valves) may be regarded as mitigative barriers to security threats, which can mitigate consequences caused by intentional attacks. Due to the interdependency between safety and security risks, integrated safety and security management has been highly suggested by researchers (see, e.g. Reniers & Khakzad, 2017). Unfortunately, previous studies show integrated risks are poorly understood in practice because security risk analyses and safety risk analyses are often undertaken independently in Seveso sites (Ylönen et al., 2022). Some attempts at integrating safety and physical security risk assessment in chemical plants have been made by researchers (Song et al., 2019a; Chen et al., 2019; Casciano et al., 2019). Researchers also investigated the incorporation of C2P attack scenarios in the safety and security risk analysis of ICSs (Abdo et al., 2018; Guzman et al., 2020; Ji et al., 2021). However, a thorough quantitative risk assessment (QRA) of chemical facilities considering safety causes, physical attacks, and C2P attack scenarios has hardly been investigated in previous studies.

The barrier concept originated from the safety science domain (Haddon, 1973), and was also used in studies concerning physical attacks (Moreno et al., 2022) and cybersecurity (Øien et al., 2022). Safety barriers present all kinds of measures/tools used to prevent the happening of accidental events or mitigate their corresponding consequences (Yuan et al., 2022a). Some terms, like protection layers, defense, safety measures, safety functions, safeguards, etc., were also used to present a similar meaning of safety barriers in different industries worldwide. Similarly, security barriers can present all kinds of measures/tools used to prevent vulnerable assets from intentional attacks/malicious acts or mitigate the corresponding consequences. Examples of typical safety barriers and security barriers in chemical plants are listed in Table 1.1.

Table 1.1 Examples of typical safety and security (S&S) barriers in the chemical process industries.

Safety barriers	**Security barriers (physical security)**	**Security barriers (cyber security)**
Emergency shutdown system (ESD)	Entrance control	Network encryption
Manual shutdown	CCTV system	Security vulnerability patching
Fire protection system	Fence	Cyber intrusion detection
Pressure relief valve	Guard response	Anomaly detection and response

Regarding S&S barrier management, previous studies mainly focused on managing or optimizing safety barriers based on their reliability or risk-reduction performance. For instance, the European ARAMIS (Accidental Risk Assessment Methodology for

Industries) project suggested integrating add-on safety barriers into a QRA framework to facilitate safety barrier management (Andersen et al., 2004). A hybrid dynamic Bayesian network (DBN) embedding multiphase Markov (MSMM) process was developed for reliability modeling of safety barriers and supporting the decision-making on barrier maintenance (Wu et al., 2022). However, the interdependency between safety and security risks was rarely considered to support S&S barrier management in previous studies. Particularly, ICSs are exposed to multi-dimensional and integrated risks in chemical plants, and S&S barriers have synergistic effects on risk control. Separate safety or security risk assessment fails to capture the real interacting risks, and sole safety or security barrier management can hardly achieve effective and economical risk control. To enhance the readability of this thesis, we summarize and compare the key concepts/terminologies used in this thesis in Table 1.2.

Table 1.2 A summary of the key concepts used in this thesis.

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 accidentally lead to losses. (see Figure 1.1)
Security risks	Risks affiliated with intentional attacks/malicious acts aiming to deliberately exploit the vulnerability of specific targets to cause losses. (see Figure 1.1)
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 in this thesis.
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.

1.2 Research topic motivation and innovativeness

Although the barrier concept originated decades ago and has been applied in practice for many years, few studies have investigated the systematic framework for barrier management, particularly, considering the integration of S&S barriers. Additionally, the emerging C2P risks bring new challenges to barrier management due to the introduction of new complexities and the interactions between safety-associated and security-associated events. As important and practical measures to control adverse risks, S&S barriers are supposed to cope with the complex integrated risks in chemical plants. However, relevant theoretical and practical research is lacking to manage S&S barriers effectively and cost-efficiently. How to prevent and mitigate major adverse events effectively through integrated management of S&S barriers is worth investigating.

Moreover, the dynamic barrier management concept was suggested to determine near-real-time barrier status using multiple data and to ensure undesired risks at acceptable levels by continuously maintaining barrier performance (Pitblado et al., 2016). Dynamic barrier management has the advantages of capturing dynamic variations in barrier performance, maintaining risks at the target levels throughout the installation's life, and making timely and cost-efficient decisions. Considering the safety and security of complex ICSs, more data sources have the potential to reveal barrier status and update risk profiles. However, few studies have investigated the application of the dynamic barrier management concept in the process security domain. Making good use of multiple data to reveal barrier performance and improving barrier performance continuously are important to ensure undesired risks are at acceptable levels despite dynamic risk variations.

To meet new requirements regarding the safety and security of ICSs in chemical plants, this study aims to develop a systematic approach for dynamic and integrated management of S&S barriers considering major adverse events. To achieve this objective, a quantitative risk assessment of ICSs considering multiple risk sources (safety causes, physical attacks, and C2P attacks) is investigated to support risk-based barrier management. New approaches are developed to update risk profiles based on multiple data and to serve dynamic barrier management. Additionally, the development of dedicated decision-making methods for S&S barrier optimization is important for preventing and mitigating adverse events effectively and cost-efficiently.

1.3 Research questions

To achieve the research objectives of this study, main issues related to this research are presented as research questions. The main research question of this study is proposed as follows.

Main question: How to integrate and manage safety and security barriers effectively and cost-efficiently with respect to major adverse events in the chemical process industries?

To answer this question, a list of sub-questions has been identified. The proposed research questions are structured and demonstrated in Figure 1.3. The main research question is answered by addressing all sub-questions.

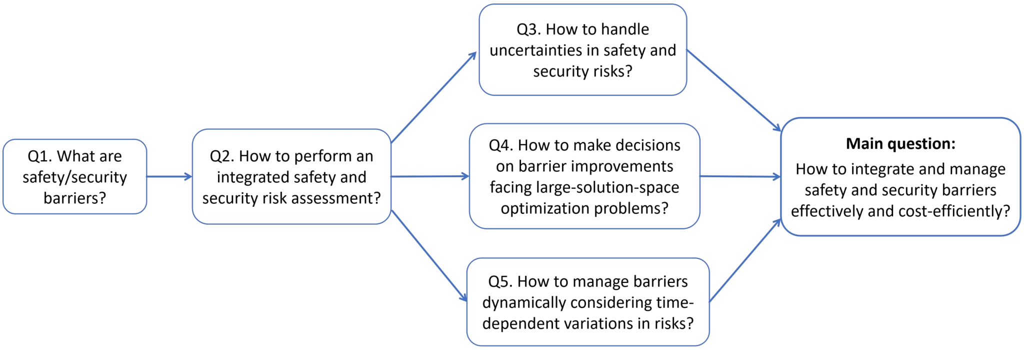

Q1. What are safety/security barriers?
Q2. How to perform an integrated safety and security risk assessment?
Q3. How to handle uncertainties in safety and security risks?
Q4. How to make decisions on barrier improvements facing large-solution-space optimization problems?
Q5. How to manage barriers dynamically considering time-dependent variations in risks?
Main question:
How to integrate and manage safety and security barriers effectively and cost-efficiently?

Figure 1.3. Structure of the proposed research questions.

Q1. What are safety/security barriers?

Although the barrier concept appeared already at the beginning of the 1970s (Haddon, 1973) and has been continuously further developed, a universally accepted definition of safety barriers or security barriers has never been achieved. A thorough literature review on safety and security (S&S) barriers is essential to get a deep understanding of the fundamental concepts of barriers and tackle the challenges in improving their management. Therefore, a systematic literature review of the definition, classification, performance assessment, and management of S&S barriers is performed to answer this research question and identify current research gaps in assessing and managing S&S barriers. This research question is answered by publications 1 and 2 in Table 1.2.

Q2. How to perform an integrated and quantitative risk assessment of chemical facilities considering both safety hazards/causes and security threats?

The safe and sustainable operation of chemical facilities is threatened by both accidental and intentional adverse events. Particularly, cyber-physical (C2P) attacks became a significant concern with the digitization of industrial control systems. As a result, chemical facilities are exposed to multi-dimensional risks that may be induced by safety causes, physical attacks, or C2P attacks. However, to the best of the authors' knowledge, the research on quantitative risk assessment of chemical facilities considering the interdependency between safety risks, physical security risks, and C2P risks has rarely been investigated in previous studies. A systematic approach to risk assessment of chemical facilities considering the interdependency and interactions between safety-hazard-induced adverse events and security-threat-induced adverse events is urgently needed. An integrated risk assessment approach based on attack-tree-bow-tie diagrams and Bayesian networks is developed to answer this research question (see publication 3 in Table 1.2).

Q3. How to handle uncertainties in safety and security risks properly and make decisions on barrier optimization?

Security attacks are difficult to predict, and the complexity of the industrial control systems (ICSs) makes security analysis harder. Uncertainties involved in the integrated safety and security risk assessment are significant. Appropriate treatment of the uncertainties in security and safety risks becomes necessary for implementing risk-based or risk-informed decision-making in practice. A systematic approach is lacking to support integrated safety and security management based on risk assessment considering uncertain parameters. This research question is answered by combining Bayesian network and Monte Carlo simulations for risk assessment and managing S&S barriers (see publication 4 in Table 1.2).

Q4. How to make decisions on safety and security barrier improvements effectively and cost-efficiently when facing large-solution-space optimization problems?

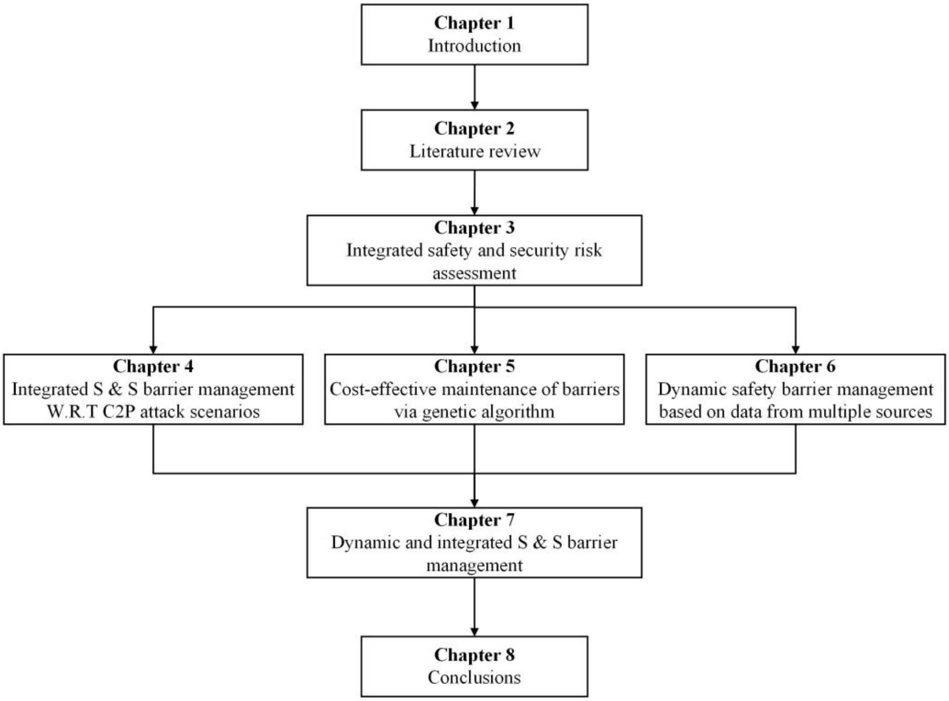
Safety and security barriers are employed widely to protect chemical plants from accidental and intentional undesired events and mitigate consequences. Managing S&S barriers effectively and cost-efficiently is a research topic with practical significance. Regarding barrier maintenance and optimizations, barrier aging, degradation, and the influence of human and organizational factors should be considered (Fiorentini and Marmo, 2018). Meanwhile, the economic issues of barrier maintenance play an indispensable role in the decision-making process for safety and security management since companies usually face budget limitations. The trade-off between accident risk levels and barrier maintenance costs is vital concerning cost-efficient barrier maintenance (Zhen et al., 2021). However, for a complex system with many safety and security barriers, it is difficult to determine a specific optimal strategy because the solution space is too large. New approaches should be developed to obtain the optimal barrier improvement strategy from a cost-effective perspective when facing large-solution-space optimization problems. An approach combining cost-effectiveness analysis (CEA) and genetic algorithms is developed to answer this research question (see publication 5 in Table 1.2).

Q5. How to assess dynamic variations in safety and security risks and achieve integrated safety and security barrier management dynamically?

The dynamic barrier management concept was introduced by Pitblado et al. (2016), who suggested using multiple data sources to reveal near-real-time barrier status. Regarding integrated safety and security risks, the integration of the dynamic risk assessment (DRA) with barrier management procedures needs to be enhanced because more data sources may be useful for revealing the status of barriers. For instance, real-time monitoring data, barrier inspection data, accident precursor data, security incident data, etc. Making good use of those data sources to update risk profiles when new data becomes available and making continuous adaptations to barrier management strategies are important for ensuring the undesired risks are up-to-date and acceptable. New approaches are developed to achieve dynamic S&S barrier management by incorporating data from multiple sources into the risk assessment and barrier management frameworks, and therefore, answer this research question (see publications 6 and 7 in Table 1.2).

1.4 Research plan and outline of the dissertation

This study investigates the methodology for integrated risk assessment of chemical facilities considering multiple risk sources (safety causes and security threats). Systematic approaches are developed for dynamic and integrated management of S&S barriers from a cost-effective perspective. This dissertation consists of 8 chapters, and its structure is shown in Figure 1.4.



Chapter 1
Introduction
Chapter 2
Literature review
Chapter 3
Integrated safety and security risk assessment
Chapter 4
Integrated S & S barrier management W.R.T C2P attack scenarios
Chapter 5
Cost-effective maintenance of barriers via genetic algorithm
Chapter 6
Dynamic safety barrier management based on data from multiple sources
Chapter 7
Dynamic and integrated S & S barrier management
Chapter 8
Conclusions

Figure 1.4. Outline of the dissertation.

Chapter 1 illustrates the research background, motivations, research questions, main contributions, and structure of this dissertation.

Chapter 2 provides a systematic literature review on S&S barriers in the chemical process industries and answers sub-question Q1. This chapter reviews the current status of classification, assessment, and management of S&S barriers. The current research status is deeply discussed to identify research gaps.

Chapter 3 presents an approach for integrated process safety and process security risk assessment of industrial control systems in chemical plants considering safety causes, C2P attacks, and physical attacks. Meanwhile, the interdependency between safety-associated events and security-related events is addressed in the risk assessment. This chapter aims to answer sub-question Q2.

Chapter 4 presents an approach for integrated management of safety and security barriers with respect to cyber-physical attack risks under uncertainties. Uncertainties from multiple sources are considered in the risk assessment, and a decision-making approach for barrier improvements is developed. This chapter aims to answer sub-question Q3.

Chapter 5 provides an approach for cost-effective maintenance of safety and security barriers. A combination of cost-effectiveness analysis (CEA) and genetic algorithm (GA) is employed to support the decision-making on barrier maintenance optimization. An illustrative case is employed to validate the feasibility of the proposed approach. This chapter aims to answer sub-question Q4.

Chapter 6 provides an approach to facilitate dynamic-risk-informed safety barrier management. Data from multiple sources are used to update risk profiles and support decision-making on safety barrier optimization. This chapter aims to answer sub-question Q5.

Chapter 7 incorporates all methodologies proposed in this study systematically. The integration of dynamic risk assessment (DRA) with the barrier management framework is explored in this chapter. A new paradigm is suggested to manage S&S barriers effectively, cost-efficiently, and dynamically to cope with complex major adverse risks in chemical plants. This chapter aims to answer sub-question Q5.

Chapter 8 concludes the dissertation and discusses future research directions on the management of S&S barriers in the chemical process industries. This chapter answers the main research question.

1.5 Publications and main contributions

This PhD thesis is organized as a paper-based thesis, and several publications should answer the proposed research questions point-by-point. A summary of all publications generated from this PhD study is shown in Table 1.3, in which the connections of each publication to the research questions are presented. The linkages between the publications and the remaining chapters are also given. Additionally, the main contributions of each publication are summarized below.

1) A systematic review of research on safety and security barriers in the chemical process industries is conducted. The research status of classification, assessment, and management of S&S barriers is deeply discussed to identify research gaps. Recommendations for future research are given at the end of this thesis.

2) An approach is proposed to integrate accident scenarios concerning both safety hazards and security threats by combining attack trees and bow-tie diagrams. A quantitative risk assessment of industrial control systems is performed using a Bayesian network (BN) model considering the interdependency between safety-associated events and security-associated events.

3) A framework for integrated safety and security barriers management is suggested to cope with safety risks and emerging cyber-physical attack risks considering uncertainties. A vulnerability assessment model is developed to quantify the vulnerability of ICSs to C2P attacks. The combination of Monte Carlo simulations and BN helps to handle uncertainty propagation in risk assessment. A cost-effectiveness analysis helps to make optimal decisions on safety and security barrier improvements.

4) A novel approach for optimizing safety and security barrier maintenance strategy considering economic constraints is proposed. A combination of cost-effectiveness analysis (CEA) and genetic algorithm (GA) is employed to support the decision-making on barrier maintenance optimization in case of large solution spaces.

5) A simulation approach is proposed in this study to conduct a dynamic risk assessment of chemical facilities based on multiple data (periodic proof test data, continuous condition-monitoring data, and accident precursor data). The degradation of safety barriers is modeled, and the performance of safety barriers is assessed dynamically to support the decision-making on safety barrier establishments and improvements.

6) A systematic framework is proposed to promote a paradigm shift towards managing safety and security (S&S) barriers in a more effective, cost-efficient, and dynamic manner, addressing the complex risks faced by the chemical process industries. Furthermore, all methodologies developed in this study are incorporated and structured as an example to demonstrate the implementation of this framework for dynamic and integrated management of safety and security barriers in chemical plants.

Table 1.3 A summary of the publications generated from this PhD study.

No.	Publications	Methodologies	Answer which research questions	Linked to which contributions	Support which chapters
1	Yuan, S., Yang, M., Reniers, G., et al. (2022). Safety barriers in the chemical process industries: A state-of-the-art review on their classification, assessment, and management. Safety Science, 148, 105647.	Literature review, Theoretical analysis and discussions	Q1	Contribution 1	Chapter 2
2	Yuan, S., Reniers, G., & Yang, M. (2022). The Necessity of Integrating Safety and Security Barriers in the Chemical Process Industries and its Potential Framework. Chemical Engineering Transactions, 91, 13-18.	Literature review, Theoretical analysis and discussions	Q1	Contribution 1	Chapter 2
3	Yuan, S., Yang, M., & Reniers, G. (2024). Integrated process safety and process security risk assessment of industrial cyber-physical systems in chemical plants. Computers in Industry, 155, 104056.	Attack-tree-bow-tie diagram, CPS master diagram, Bayesian network, Sensitivity analysis	Q2	Contribution 2	Chapter 3
4	Yuan, S., Reniers, G., Yang, M. (2024). Integrated management of safety and security barriers in chemical plants to cope with emerging cyber-physical attack risks under uncertainties. Reliability Engineering & System Safety, 250, 110320.	Bayesian network, C2P attack modeling, Monte Carlo simulations, Cost-effectiveness analysis	Q3	Contribution 3	Chapter 4
5	Yuan, S., Reniers, G., Yang, M., et al. (2023). Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. Process Safety and Environmental Protection, 170, 356-371.	Bow-tie diagram, Simulink-based barrier modeling, Cost-effectiveness analysis, Genetic algorithm	Q4	Contribution 4	Chapter 5

6	Yuan, S., Reniers, G., & Yang, M. (2023). Dynamic-risk-informed safety barrier management: An application to cost-effective barrier optimization based on data from multiple sources. Journal of Loss Prevention in the Process Industries, 83, 105034.	Simulink-based barrier modeling, Degradation modeling, Bayesian updating, Cost-effectiveness analysis	Q5	Contribution 5	Chapter 6
7	Yuan, S., Reniers, G., Yang, M. (2024). Dynamic and integrated safety and security barrier management in chemical plants: a new paradigm to manage complex major adverse risks. Process Safety and Environmental Protection, (To be submitted).	Attack-tree-bow-tie diagram, Bayesian network, Bayesian updating, Cost-effectiveness analysis, et al.	Q5	Contribution 6	Chapter 7

Chapter 2 Safety and security barriers in the chemical process industries: A literature review

This chapter provides a systematic literature review on safety and security (S&S) barriers in chemical process industries. Specifically, the definitions of safety barriers and security barriers, classifications of safety barriers and security barriers, methodologies for performance assessment of barriers, and research on safety and/or security barrier management are reviewed thoroughly. The current research status is deeply discussed to identify vital research gaps. Accordingly, integrated and cost-efficient management of safety and security barriers in chemical plants is suggested to fill the gaps.

This chapter is drafted with modifications based on the following publications:

- Yuan, S., Yang, M., Reniers, G., Chen, C., & Wu, J. (2022). Safety barriers in the chemical process industries: A state-of-the-art review on their classification, assessment, and management. *Safety Science, 148,* 105647.
- Yuan, S., Reniers, G., & Yang, M. (2022). The Necessity of Integrating Safety and Security Barriers in the Chemical Process Industries and its Potential Framework. *Chemical Engineering Transactions, 91,* 13-18.

2.1 Introduction

As a commonly used term to present preventive measures, safeguards, mitigation measures, and protective layers to prevent or mitigate accidents, “safety barrier” is generally used. The barrier concept originated from the energy model (Gibson, 1961). The term “safety barrier” first appeared in 1973 (Haddon, 1973). Then Johnson (1975) integrated the safety barrier concept into the MORT (Management Oversight and Risk Tree) technique, in which the barrier analysis was investigated. Different terms like “safety barrier” were used in various industries and organizations. As an example, the term “protection layer” was used in the chemical process industry as a similar representation to “safety barrier” at an early age (CCPS, 1993). Svenson developed an accident evolution and barrier function (AEB) model that is capable of performing accident evolution analysis (Svenson, 1991). Then, the concept and functions of the so-called “safety barrier” were elaborated by Hollnagel (1999) before some researchers tried to interpret and define safety barriers clearly to reduce misconceptions in work related to risk management and accident prevention (Duijm et al., 2004). Additionally, the European ARAMIS (Accidental Risk Assessment Methodology for Industries) project developed an integrated approach for modeling and managing risks concerning major accident scenarios, in which the concept of safety barrier was applied and highly recommended. (Andersen et al., 2004; De Dianous & Fiévez, 2006). Furthermore, a systematic literature review presented how safety barriers and similar concepts were interpreted and used in various industries (Sklet, 2006).

The Petroleum Safety Authority Norway (PSA) presented the principles for barrier management in the petroleum industry (PSA, 2013). Meanwhile, the concept of safety barrier was also mentioned and stressed in other European regulations, national-level regulations, and international standards since the 1990s (EC., 1996; EC., 1998; IEC:61508., 1998; IEC:61511., 2002). Furthermore, ISO standards (ISO:13702., 1999; ISO:17776., 2000; ISO:13702., 2015; ISO:16530., 2017) implied the requirement for safety barriers regarding the prevention and mitigation of major accidents in the petroleum and natural gas industries. The literature above indicates that the concept of safety barrier was widely used in process industries and played an essential role in hazard control and risk management-related policies and strategies.

With the emergence and popularization of the safety barrier concept in chemical process industries, the assessment of safety barrier performance became a crucial issue. Typically, the safety barrier performance constituted three dimensions: functionality/efficiency, availability/reliability, and robustness (Sklet, 2006). A set of properties, including effective, independent, and auditable, is also proposed as the requirements for safety barriers from a barrier management perspective (CCPS/EI, 2018). Additionally, many studies involved safety barriers in the risk assessment framework and assessed safety barrier performance by measuring the importance/effects of safety barriers in risk reductions. For instance, Landucci et al. (2015) developed a method to assess safety barrier performance in preventing domino scenarios triggered by fires. Bayesian Networks were also used in combination with safety barriers for dynamic assessment of the escalation scenario in offshore Oil&Gas (Bubbico et al., 2020) and fire-caused domino effects with good effectiveness (Khakzad et al., 2017a; Zeng et al., 2020). In terms of safety barrier management, the graph theory (Khakzad et al., 2017b) and Bayesian networks (Khakzad et al., 2018a) were combined with cost-effectiveness analysis to support decision-making on the safety barrier allocation concerning fire-induced domino effects. An agent-based model was proposed to assess complex domino events and support safety barriers allocation in chemical plants (Ovidi et al., 2021).

A timeline regarding the development of the literature on safety barriers in the chemical process industries is presented in Figure 2.1. The first period is named "the origin period" when the specific term "safety barrier" did not appear yet. Still, similar concepts already existed. In the second stage, some primary contents about safety barriers were investigated, and the concept of safety barriers was used in some studies. The research work in this period focused on the concept, definition, classification, function, and performance criteria of safety barriers. In the next stage, the performance assessment of safety barriers became the focal point. Recent years (after 2020) have seen more studies on the performance optimization and management of safety barriers considering cost-benefit analysis.

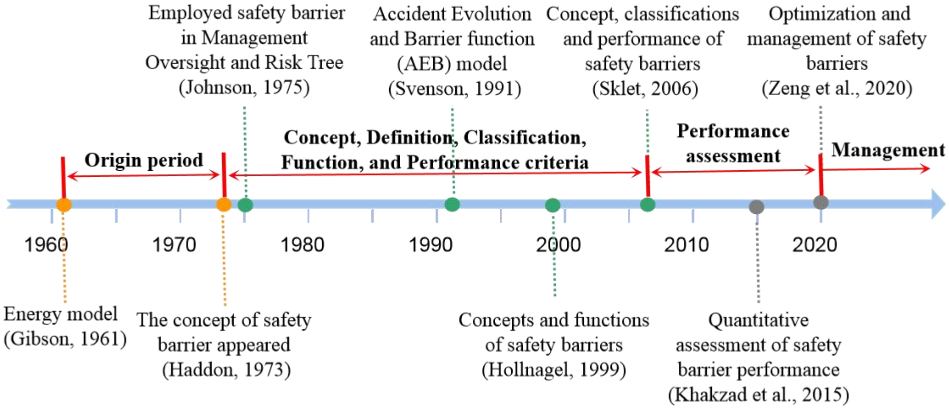

Employed safety barrier in Management Oversight and Risk Tree (Johnson, 1975)
Accident Evolution and Barrier function (AEB) model (Svenson, 1991)
Concept, classifications and performance of safety barriers (Sklet, 2006)
Optimization and management of safety barriers (Zeng et al., 2020)
Origin period
Concept, Definition, Classification, Function, and Performance criteria
Performance assessment
Management
1960
1970
1980
1990
2000
2010
2020
Energy model (Gibson, 1961)
The concept of safety barrier appeared (Haddon, 1973)
Concepts and functions of safety barriers (Hollnagel, 1999)
Quantitative assessment of safety barrier performance (Khakzad et al., 2015)

Figure 2.1. History and highlight publications of safety barriers in the chemical process industry.

By contrast, the security barrier concept has not been widely recognized in the chemical process industry because the integration of safety risks and security risks is poorly understood in practice (Ylönen et al., 2022). The necessity of involving security risks in the traditional safety risk assessment and performing unified safety and security management was suggested by researchers (Aven, 2007; Reniers & Khakzad, 2017). Regarding physical security, some approaches were proposed to facilitate security risk assessment or vulnerability assessment of chemical facilities, for instance, the VAM-CF methodology (Jaeger, 2002), the CCPS methodology (CCPS, 2003), and the API 780 approach (API, 2013). Some attempts have been made to perform assessment and optimization of safety and security barriers regarding physical-attack-induced domino effects (Reniers et al., 2008; Chen et al., 2020; Moreno et al., 2022). Moreover, the concept and terminology of "cybersecurity barrier" were also used in the chemical process industries (Øien et al., 2022) after the cyberattack issues were explicitly noticed (Kornecki & Zalewski 2010; Thomas et al., 2015). Some approaches were developed to perform integrated safety and security analysis of industrial control systems (ICSs) emphasizing the interventions of safety and security barriers (Ji et al., 2021; Guzman et al., 2021). However, the understanding of the roles of safety and security (S&S) barriers in safety and security risk integration and unified safety and security management needs to be enhanced to facilitate S&S barrier management.

Therefore, a systematic literature review and analysis of the concepts, classifications, performance assessment, and management of S&S barriers in the process industries is performed to identify the knowledge and technical gaps of current studies. The terms that have similar meanings to "safety barrier" such as "safety function", "safety measure", "safeguards", and "protective layer", etc. were all considered as review terms in this study. Relevant studies on "security barriers" are also reviewed and discussed to reveal the necessity of integrating safety and security barriers. Only the references relevant to the aims of this study are selected and discussed. Thus, to ensure a suitable scope for this review and to achieve our research objectives.

2.2 Definitions of safety barriers and security barriers

Although the safety barrier concept appeared already at the beginning of the 1970s and has been continuously further developed, a universally accepted definition of safety barrier has never been achieved. In the classical definitions, a safety barrier is regarded as a physical obstacle, obstruction, or hindrance to protecting "a target" from "hazards" (Sklet, 2006). For instance, Holland (1997) viewed a safety barrier as a physical protection barrier. On the one hand, a barrier should "reduce the probability of an accident" or "reduce the consequences of an accident" (ISO:17776, 2000). On the other hand, a barrier should "prevent the flow" (Holland, 1997) and should be capable of preventing a scenario from escalating to the undesired consequences (CCPS, 2001). In conventional definitions, a safety barrier is related to a hazard, an energy source, or an event sequence. This indicates that safety barriers should be related to a specific hazard to specify their functions and locations. As a physical structure or obstacle, a safety barrier can be used to prevent or delay the occurrence of accidents and/or mitigate the severity of their consequences.

Apart from the conventional definition, the concept of safety barrier was extended to have a broader scope to include non-physical barriers. Defence-in-depth (DiD), known as a military strategy (Fleming & Silady, 2002), was applied to expand the definition of the safety barrier. The modern principle of DiD combines different types of barriers, from physical obstacles, and protection measures to strategies and safety policies. Similarly, some broader definitions of safety barriers were proposed by researchers, such as Schupp et al. (2004) defined safety barriers as the combination of technical, human, and organizational measures that prevent or protect against an adverse effect. Johnson (2003) defined safety barriers as the diverse physical and organizational measures taken to prevent a target from being affected by a potential hazard. The concept of "protection layer", whereby a device, system, or human action is provided to reduce the likelihood and/or severity of a specific loss event, has a similar meaning as the concept of "safety barrier" (CCPS, 2001). CCPS (USA) and Energy Institute (UK) defined a barrier based on the bow-tie diagram as a risk reduction measure that on its own can prevent a threat from developing into a top event or can mitigate the consequences of a top event (CCPS/EI, 2018). In the same study, a barrier is considered a complete system fulfilling the criteria of being effective, independent, and auditable, similar to the characteristics defined in LOPA for an IPL (independent protection layer). Under this definition, active barriers are considered must-have separate elements of 'detect-decide-act'. To compare the conventional and extended concepts of safety barriers, we conclude the features of different definitions in Table 2.1.

Table 2.1 Comparison of conventional and extended definitions of safety barriers.

Aspects of the definition	**Conventional definition**	**Extended definition**	
Dimensions	Physical protection	Physical protection	Non-physical protection
Means of implementation	Obstacle, obstruction, hindrance, fence, structure, etc.	Same as the classical definition	Strategies, human action, socio-technical system, organizational measures, etc.
Objectives	To prevent accidents from taking place, delay the occurrence of accidents, prevent or mitigate the impact of the accident consequences	To prevent, control, or mitigate undesired events or accidents, including reducing the risk of undesired events or accidents, limit the extent and/or duration of undesired events or accidents from escalation, and mitigate the impacts of undesired events or accidents	
Application domain	Should be related to a specific hazard, can be applied to the physical protection of fires, explosions, etc.	Can be applied to events or accidents caused by errors from technical facilities, human actions, inherent designs, organizations, external events, and a combination of those	

By reviewing the existing definitions of safety barriers, we define a safety barrier as a physical or non-physical tool planned to prevent, control, or mitigate undesired events or accidents. The means of safety barriers can vary from a technical facility or human action to a complex socio-technical system. The purposes of the safety barrier are to reduce the risk of an undesired event by reducing the occurrence probability of this event, limiting the extent and/or duration of the undesired event from escalation, or mitigating the impacts of the undesired event.

Regarding security barriers, limited literature has used this terminology currently. In the physical security domain, security barriers were used to represent similar meanings of physical protection systems (PPS), for instance in the study of Moreno et al. (2022). A physical protection system (PPS) is defined as a collection of system elements strategically combined to achieve protection according to a predefined plan (U.S. Congress, 1992). A PPS may integrate people, procedures, and equipment for the protection of assets or facilities against theft, sabotage, or other malevolent human attacks (John & David, 2018). A security barrier is also defined as a locked, impenetrable wall, fence, or structure that completely seals an area from unauthorized entry or trespass (LawInsider, 2024). Concerning the security of process control environments, cybersecurity barriers were used to present specific security measures to protect process control systems from cyberattacks, for example, access control or firewalls (Øien et al., 2022). However, an explicit definition of cybersecurity barriers has not been given in previous studies. In this study, we adapt the safety barrier definition to the security domain. 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.

2.3 Typical classifications of safety and security barriers

In this section, the previous classifications of safety barriers are summarized first. The potential classifications of security barriers are also discussed to get insights into the functionalities of security barriers in intentional attack protection. According to previous studies, most safety barrier classifications are based on the extended definition of safety barrier, which involves physical and non-physical objects. Typically, passive and active barriers were widely used to present barriers not required to be activated to achieve their function and the barriers necessary to move from one state to another in response to a change or a signal to fulfill their role, respectively. Some classifications of safety barriers were proposed based on the bow-tie model. Safety barriers used on the left-hand side of the bow-tie are called “preventive” or “proactive” barriers, which are used to reduce the likelihood of occurrence of the central event. By contrast, the safety barriers used on the right-hand side of the bow-tie are called “reactive” or “mitigating” barriers, which are used for mitigating the consequences of the central event. Typical classifications of safety barriers are compared in Table 2.2.

Table 2.2 Typical classifications of safety barriers.

Classifications	Classification basis	Examples	References
Physical and non-physical barriers	Physical or non-physical	Physical barrier: fireproofing material Non-physical barrier: emergency team	(Johnson, 1980)
Physical, technical, and administrative barriers	Physical or non-physical	Physical barrier: fireproofing material Technical barrier: water spray system Administrative barrier: safety training on employees	(Wahlstrom & Gunsell, 1998)
Hardware and behavioral	Physical or non-physical	Hardware barrier: fireproofing coating Behavioral barrier: emergency evacuation	(Hale, 2003)
Permanent and temporary barriers	Natural attributes	Permanent barrier[1]: corrosion prevention system Temporary barrier[2]: foam-water sprinkler system	(Hollnagel, 2004)
Static and dynamic barriers	Natural attributes	Static barrier[3]: well packer Dynamic barrier[4]: stabbing valve	(Holland, 1997)
Prevention, protection, and mitigation barriers	Operational features	Prevention barrier: good engineering practice Protection barrier: rupture discs Mitigation barrier: fire brigade	(Markowski & Kotynia, 2011)
Passive and active barriers	Operational features	Passive barrier: pressure safety valve Active barrier: foam-water sprinkler system	(Kjellén, 2000)
Passive barriers, active barriers, and procedural and emergency measures	Operational features	Passive barrier: pressure safety valve Active barrier: water spray system	(CCPS, 2001), (Landucci et al., 2015), (Khakzad et al., 2017a), et

[1] Permanent barriers are usually part of the design base, for instance, as a response to an accident (Hollnagel, 2004).

[2] Temporary barriers are restrictions that apply for a limited period of time only, typically referring to a change in external conditions (Hollnagel, 2004).

[3] Static barriers are available over a "long" period of time. This situation applies during production/injection or when the well is temporary closed in (Holland, 1997).

[4] Dynamic barriers vary over time. This applies to drilling, workover, and completion operations (Holland, 1997).

Classifications	Classification basis	Examples	References
		Procedural and emergency measures: emergency measures	al.
Inherent design and add-on barriers	Operational features	Inherent design: land use planning Add-on barrier: pressure safety valve	(Schupp, 2004)
Passive barriers, activated barriers, human actions, and symbolic barriers	Operational features	Passive barrier: retention bund Activated barrier: emergency blowdown system Human action: emergency team Symbolic barrier: passive warnings	(De Dianous & Fievez, 2006), (Guldenmund, et al., 2006)
Passive hardware, active hardware, active hardware + human, active human, and continuous hardware	Operational features and bow-tie model	Passive hardware: blast wall Active hardware: process control systems and safety instrumented systems Active hardware + human: operator-activated ESD valve Active human: operator detection and response Continuous hardware: ventilation system	(CCPS/EI, 2018)
Technical safety barriers, non-technical safety barriers, and management delivery systems	Operational features and bow-tie model	Technical safety barrier: emergency stop on a pallet mover Non-technical safety barrier: manual removal of leaking containers Management delivery system: training of pallet mover operators on removing leaking containers	(Van Nunen et al., 2019)

Concerning intentional attacks, Norman (2015) classified security barriers (countermeasures) into natural countermeasures and man-made security countermeasures. Man-made security countermeasures are further categorized into physical countermeasures, electronic countermeasures, and operational countermeasures. In the same study, the functionalities/goals of security barriers are identified as access control, deterrence, detection, assessment, response, evidence gathering, compliance with the business culture, minimizing impediments to normal business operations, and safe and secure environment. Garcia (2007) classified physical protection systems (PPSs) into detection, delay, and response according to their functionalities. A complete function of PPSs is described as being effectively aware of attacks (detection), slowing down adversary progress to the targets (delay), and allowing the response force enough time to interrupt or stop the adversary (response). Concerning cybersecurity issues of ICSs, countermeasures are categorized into 18 themes in the NIST SP 800-53 guide (NIST, 2014), which are access control, awareness and training, audit and accountability, security assessment and authorization, configuration management, contingency planning, identification and authentication, incident response, maintenance, media protection, physical and environmental protection, planning, personnel security, risk assessment, system and services acquisition, system and communications protection, system and information integrity, organization-wide security management, and privacy controls. Similarly, a number of security requirements, called Security Program Elements (SPE), are classified into eight groups by the IEC 62443 standard (IEC:62443-2-1, 2010), which are organizational security. configuration management, network security, component security, protection of data, user security, event and incident management, and system integrity and availability. Security barriers (countermeasures) are also categorized as organizational measures, technical security measures, and security-related processes by previous studies (Flaus, 2019).

According to existing classifications of S&S barriers, operational or functional features of barriers are widely used as the criteria for barrier classification. For instance, the classification of passive barriers, active barriers, and procedural barriers is widely used for safety barriers. By contrast, the classification of detection, delay, and response is suitable for security barriers. In practice, different classification methods may be applied to different operational stages and serve different purposes. For instance, functionality-based classifications may be applied to the technical assessment of barriers while task-based classifications are more practical in barrier management in general.

2.4 Performance assessment of safety and security barriers

2.4.1 Safety barrier performance assessment

Assessment of the barriers is critical to identify the risks of an accident scenario under the protection of barriers. The evaluation criteria for barrier assessment should have the ability to reflect how well the barriers perform their functionalities regarding specific accidental scenarios. Typically, effectiveness is widely regarded as an evaluation criterion to measure safety barrier performance. For instance, Kang et al. (2016) proposed a quantitative assessment approach for safety barriers considering three perspectives: confidence, effectiveness, and economic impact, in which the effectiveness was defined as how well a safety barrier can prevent accidents and reduce the risk to an expected level. The terms efficiency and sufficiency were also used to describe safety barrier performance with similar implications as effectiveness (Liu, 2020). Availability is also widely used to measure safety barrier performance. Availability is defined as the capacity of a barrier to perform its function effectively at a particular time, which is a time-dependent evaluation criterion (Liu, 2020). Availability was also combined with effectiveness to evaluate safety barrier performance, especially for the active safety barriers (Landucci et al., 2015). In the same study, the availability was quantified by using the probability of failure on demand (PFD). The effectiveness was quantified as the probability that a safety barrier will prevent accident escalation after being activated successfully.

Regarding the specific methods/techniques, event tree analysis (Landucci et al., 2016), LOPA (Basheer et al., 2019), bow-tie diagrams (CCPS/EI, 2018), Bayesian networks (Ding et al., 2020), Petri net (PN) models (De Souza et al., 2017), FRAM (Functional Resonance Analysis Method) method (Sultana & Haugen, 2023), and the combination of those were widely used for safety barrier assessment. For instance, a modified event tree analysis was developed to perform quantitative assessment of safety barrier regarding the prevention of domino scenarios (Landucci et al., 2015; Landucci et al., 2016). The same method was also modified to be applied in Natech scenarios (Misuri et al., 2020) and assessing safety barrier performance in harsh environments (Bucelli et al., 2018). A LOPA-based approach called the cloud model layer of protection analysis (CM-LOPA) was proposed to assess gas leakage risk in a biomass gasification station (Yan et al., 2017). The performance of safety-instrumented systems was assessed by considering different proof-testing strategies using Petri nets (Liu & Rausand, 2016). Ding et al. (2020) combined bow tie and Bayesian network models to investigate the relationships among accident causes, safety barriers, and the possible consequence of a cotton storage fire. Sultana & Haugen (2023) assessed the adequacy of safety barriers with respect to the safety of a socio-technical system by using an extended FRAM model. A list of publications related to performance assessment of safety barriers is summarized in Table 2.3.

Table 2.3 A list of current publications on safety barrier performance assessment in the chemical process industry (before 2024, February).

Publications	Methodologies	Target objects or scenarios	Investigated barriers	Other keywords
Yun et al. (2009)	LOPA and Bayesian estimation	LNG importation terminals	one independent protection layer (IPL) involving the temperature safety valve (TSV)	risk assessment, failure data quality
Rathnayaka et al. (2011a, 2011b)	process accident model and Bayesian updating mechanism	LNG processing facilities	release prevention barrier, dispersion prevention barrier, ignition prevention barrier, escalation prevention barrier, damage control and emergency management barrier, human factor barrier, management and organizational barrier	accident precursors, predictive model
Cai et al. (2012)	Bayesian networks	petroleum drilling rig explosion and oil spill	subsea blowout preventer control system	reliability, common cause failure, imperfect coverage
Pitblado & Nelson (2013)	Bow-tie diagram	major accidents in Oil and Gas industries	relief valve, ESD valve, gas detection device, training course, fatigue management, et al.	/
Xue et al. (2013)	barrier-based accident model and event tree analysis	offshore drilling blowouts	primary well, well monitoring, secondary well, ignition prevention, escalation prevention, emergency response, blowout control, oil spill control	accident model, active failures
Myers (2013)	LOPA (layer of protection analysis)	accident scenarios in process industries	active protection layers involving human IPLs (independent protection layers)	process safety, human reliability analysis
Badreddine et al. (2014)	Bow-tie diagram	unconfined vapor cloud explosion	employee training, storage tank inspection, gas leak detection, gas detection and protection system, fire detection and protection system, carbon monoxide alarms	propagation algorithms, multi-objectives influence diagrams
Innal et al. (2014)	fault tree and Markov model	steam boiler breakup	safety instrumented systems	failure frequency
Landucci et al. (2015)	LOPA and event tree analysis	fire-triggered domino scenarios	water deluge system, pressure safety valve, fireproofing, emergency teams	quantitative risk assessment

Publications	Methodologies	Target objects or scenarios	Investigated barriers	Other keywords
Ramzali et al. (2015)	event tree analysis, reliability block diagram, and fault tree analysis	leakage in drilling well of offshore drilling system	barriers in operational phase	barrier analysis
Villa & Cozzani (2016)	Bayesian networks	major accidents in the process industry	fixed foam system, rim seal fire extinguisher	/
Landucci et al. (2016)	LOPA and event tree analysis	fire-triggered domino scenarios	foam-water sprinkler system, water deluge system, emergency shutdown system, pressure safety valve, fireproofing coating, external emergency intervention	performance analysis, escalation frequency
Kang et al. (2016)	fuzzy mathematic theory and incident process analysis	oil storage facility explosion	personnel barriers, organizational barriers, technological barriers	accident evolution, failure mechanism
De Souza et al. (2017)	Petri nets	modern production systems	safety instrumented system	risk analysis, scenarios of faults
Bucelli et al. (2017)	Bow-tie and risk barometer methodology	major accidents in the oil and gas industry	limit hydrocarbon leak	risk assessment
Khakzad & Reniers (2017)	Bayesian networks and limited memory influence diagram (LIMID)	fire-induced domino effects	fireproofing of storage tanks	multi-attribute decision analysis
Khakzad et al. (2017a)	event tree analysis and dynamic Bayesian network (DBN)	fire-induced domino effects	fireFigurehting systems, emergency isolation and depressurization systems, passive fire protections, emergency response	quantitative risk assessment
Khakzad et al., (2017b)	graph theory	fire-induced domino effects	fireproofing protection (passive fire protection) and active protection systems	multicriteria decision making
Yan et al. (2017).	LOPA and cloud model	gas leakage in a biomass gasification station	ventilation and alarms	Randomness, fuzziness, normal cloud
Bucelli et al. (2018)	LOPA and event tree analysis	fire-induced cascading events of offshore facilities in harsh environments	water deluge system, emergency shut down, pressure safety valve, passive fire protection, emergency response and rescue	major accident hazard

Publications	Methodologies	Target objects or scenarios	Investigated barriers	Other keywords
Tsunemi et al. (2019)	event tree analysis	hydrogen leaks in the hydrogen refueling station	excess flow stop valve, leak detector and shutoff valve, manual operation	quantitative risk assessment
Sobral & Soares (2019)	Bow-tie diagram and LOPA	a fire pumping system	sensor system, logic system, actuator system	PFD, safety integrity level
Simon et al. (2019)	dynamic Bayesian network (DBN)	chemical reactor protection system	safety instrumented systems	proof tests, test strategy, test duration
Zhang et al. (2019)	mathematical modeling	offshore facilities	safety instrumented systems	degradation, redundant structure
Bubbico et al. (2020)	Bayesian networks	process leak-fire/explosion-escalation in Oil&Gas platforms	leak detection, blowdown, deluge system, hydrocarbon inflow shut-off, ignition prevention, escalation prevention, passive fire protection, depress pressure safety valve	dynamic risk analysis, extreme environment
Ding et al. (2020)	bow-tie and Bayesian network	cotton storage fire	detection and extinguishment, fire brigade	criticality analysis, risk control strategies
Zhu et al. (2020)	LOPA and dynamic simulations	batch reactor systems	human actions and response	human error probability, human reliability analysis
Misuri et al. (2021)	LOPA and event tree analysis	domino scenarios caused by Natech events	pressure safety valve, foam-water system, water deluge system, passive fire protection (fireproofing), external emergency intervention	escalation, quantitative risk assessment
Park et al. (2021)	LOPA	hydrogen refueling stations	passive IPLs: dike, underground draining system, et al.; active IPLs: gas detector and emergency shutoff valve, basic process control system, et al.	individual risk, societal risk, F–N curve, IPLs
Ovidi et al. (2021)	agent-based modelling	domino effects	fireproofing, foam/water system (FWS), water deluge system (WDS), external emergency intervention (EEI)	computational experiments, chemical tank farm

Publications	Methodologies	Target objects or scenarios	Investigated barriers	Other keywords
Di Maio et al. (2021)	Bayesian networks	major accident scenarios (fire, explosion, toxic dispersion)	process safety management system, task management, design integrity, operating integrity, process control system, pressure protection system, isolation & depressurization, fire management, emergency response system, spill containment system	key performance indicator, probabilistic safety margins
Hosseinnia et al. (2021)	bow-tie analysis	floating, production, storage, and offloading unit (FPSO) process	a series of active barriers, passive barriers, human actions, and symbolic barriers	offshore platforms, dynamic risk analysis, risk-based inspection, accident prevention
Yuan et al. (2022b)	computational fluid dynamics, evacuation modeling, and event tree analysis	ammonia leakage in a chemical plant	emergency shutdown system (ESD), gas detection & alarm system, emergency evacuation	toxic gas leakage, individual fatality risks
Wu et al. (2022)	multiphase Markov process and hybrid DBN	subsea downhole leakage incidents	production packer, completion string, deep set tubing plug, casing cement	process demand, dynamic risk characteristics
Di Maio et al. (2023a)	value of information, conditional value at risk, Beta measure	NaTech scenarios in chemical plants	water deluge system, passive fire protection material, pressure safety valve, emergency team intervention, protective vessel	sensitivity analysis
Sultana & Haugen (2023)	extended FRAM method	LNG ship-to-ship transfer process	Preventive Safety functions: condition monitoring of pipe, do regular inspection and maintenance, et al.; Mitigative Safety functions: to repair, mitigate leak, et al.	/
Yuan et al. (2023a)	bow-tie diagram and MATLAB/Simulink simulations	a chemical reactor with its SCADA (supervisory control and data acquisition) system	fire protection system, emergency shutdown system (ESD), manual shutdown, burst disk, et al.	barrier maintenance, cost-effectiveness analysis, Genetic algorithm

Publications	Methodologies	Target objects or scenarios	Investigated barriers	Other keywords
Yuan et al. (2023b)	Simulink-based Safety Barrier Modeling (SSBM) and Monte Carlo simulations	a continuous stirred tank reactor (CSTR)	fire protection system, ESD (emergency shutdown system), manual shutdown and pressure relief valve	dynamic risk assessment, cost-effectiveness analysis, Bayesian updating, condition monitoring
Tamascelli et al. (2024)	hazard identification techniques and data-driven regression models	environmentally critical facilities	dolomitic lime injection and automatic backup feeder	digital model, dynamic performance assessment

2.4.2 Security barrier performance assessment

Compared to safety barrier assessment, the studies on security barriers are relatively less. Concerning adverse scenarios that intentional attacks may induce, some guidelines were suggested to facilitate security risk/vulnerability assessment of chemical facilities considering the intervention of security barriers. For instance, the VAM-CF methodology incorporates the evaluation of PPSs (physical protection systems) and safety/mitigation measures into the security risk assessment framework (Jaeger, 2002). The assessment of the effectiveness and reliability of security countermeasures against security threats and asset vulnerabilities is regarded as an important part of the CCPS Security Vulnerability Analysis (SVA) methodology (CCPS, 2003). The IEC 62443 standard, which is dedicated to the security management of industrial control systems (ICSs), presents an approach for lifting and ensuring the security level of ICSs by managing security measures according to fundamental security requirements (IEC:62443-2-1, 2010). However, the above guidelines can hardly address the quantitative performance assessment of security barriers.

Some studies investigated security barrier performance in the prevention of domino effects. For instance, Reniers et al. (2008) developed a computer-automated tool for the prevention and mitigation of intentional-attack-induced domino effects considering the effects of security measures. Chen et al. (2019) proposed a dynamic graph approach for reducing man-made domino effect risk by integrating security measures and safety barriers. Moreno et al. (2022) investigated the prevention of cascading events triggered by intentional attacks, considering the combined contribution of PPSs and safety barriers. Additionally, Argenti et al. (2017) carried out a quantitative performance assessment of PPSs by combining effectiveness metrics and expert judgment, in which the probability of successful action of security barriers was assessed. Conventional safety risk analysis approaches were also adapted for security risk analysis and performance assessment of security barriers. For instance, previous studies show that bow-tie diagrams have the potential to demonstrate and assess intentional-attack-induced accident scenarios considering the intervention of S&S barriers. Ji et al. (2021) integrated safety-associated and security-associated scenarios using bow-tie diagrams with the identification of S&S barriers. Guzman et al. (2021) developed a practical toolkit for integrated safety and security risk analysis of adverse scenarios of ICSs and provided a database of checklists for building bow-tie diagrams. Other approaches, such as attack trees, cyber PHA (Preliminary Hazard Analysis), and cyber HAZOP (Hazard and Operability Study), were also suggested for security analysis or integrated safety and security analysis of ICSs (Flaus, 2019). Moreover, some dedicated models were developed with respect to cyberattacks on ICSs and can assess the performance of security countermeasures. For instance, the Network Security Risk Model (NSRM) was developed to evaluate the efficacy of candidate risk management policies regarding cyberattacks against process control networks (Henry & Haimes, 2009). Roy et al. (2010) proposed attack countermeasure trees (ACT), which can apply defense mechanisms at any node of the tree and perform qualitative and probabilistic analysis of cyberattacks on SCADA systems. A BDMP (Boolean logic Driven Markov Processes) approach was developed to model risk scenarios considering both safety and security aspects, which is also able to model repair actions, attack detections, and other reactions (Piètre-Cambacédès & Bouissou, 2010). However, a quantitative performance assessment of barriers regarding the integrated safety and security risks has rarely been achieved in previous studies.

2.5 Safety and security barrier management

In this section, the research status of safety and/or security barrier management in the chemical process industry is presented, followed by an elaboration on the necessity of integrated safety and security barrier management.

2.5.1 Safety barrier management

In previous studies, the management of safety barriers was widely investigated from both the conceptual perspective and operational perspective. Bow-tie analysis was widely used to demonstrate the safety barrier concept and facilitate barrier management. For instance, Duijm (2009) extended bow-tie diagrams as safety-barrier diagrams and suggested them as a tool to support the audit, maintenance, and management of safety barrier systems. Hudson & Hudson (2015) suggested integrating cultural and regulatory factors in bow-tie analysis and facilitating barrier management with the integration of incident analysis and reporting systems. The challenges and clarifications of the central concepts and steps in barrier management were discussed (Johansen & Rausand, 2015). The principles for barrier status monitoring in the operational phase were outlined in a handbook (Hauge et al., 2015). Meanwhile, barrier validity, barrier audit, barrier degradation controls, and the involvement of human and organizational factors were also discussed to achieve sound safety barrier management based on bow-tie analysis (Fiorentini & Marmo, 2018; CCPS/EI, 2018). The influence of human factors on barrier management was investigated and recommendations for good practice in developing and managing human factors of barrier systems were given by previous studies (McLeod, 2016, CIEHF, 2016). A practical guidance on barrier management with a focus on maintaining barriers throughout the lifetime of an offshore or onshore petroleum facility was proposed (Hauge & Øien, 2016). Van Nunen et al. (2019) suggested combining safety barriers with safety indicators to support occupational safety management. The guidance for managing safety-critical elements (SCEs) was proposed to prevent and limit the effects of major accident hazards (EI, 2020). In the same study, the use of safety integrity level (SIL) was suggested and the guidance on

SCE continual improvement, aging management, obsolescence, and life extension was given.

Moreover, the integration of quantitative risk assessment (QRA) and safety barrier management has also been investigated in previous studies. For instance, the ARAMIS project suggested integrating safety barriers into the QRA framework to facilitate barrier management concerning major accident scenarios (Andersen et al., 2004). A risk barometer methodology was suggested to support safety barrier management by reflecting barrier status concerning technical features and operational and organizational activities related to barriers (Bucelli et al., 2017). The dynamic barrier management (DBM) concept was also proposed to support continuous barrier improvements by revealing near-real-time barrier status based on multiple data sources and achieve better safety at lower cost (Pitblado et al., 2016, Hosseinniaa et al., 2019). A multi-objective optimization approach was developed to improve performance parameters of safety barriers considering NaTech scenarios (Di Maio et al., 2023b). Additionally, the optimization of safety barriers concerning domino effect scenarios has been widely investigated in previous studies. For instance, a decision model was proposed to allocate protective safety barriers with limited budgets and mitigate domino effects (Janssens et al., 2015). A methodology combining graph theory and multi-criteria decision analysis was proposed to support decision-making on allocating fire-protection barriers considering the availability and degradation of safety barriers (Khakzad et al., 2017b). An approach was developed based on BN and limited memory influence diagrams to ac0hieve a cost-effective allocation of add-on safety barriers with respect to fire-induced domino effects (Khakzad et al., 2018a). Ovidi et al. (2021) developed an agent-based model that can be employed to assess complex domino events and support decision-making on allocating safety barriers. Moreover, the reliability analysis and degradation modeling of safety barriers were also investigated to facilitate barrier maintenance and management. For instance, Zhen et al. (2021) proposed a multi-objective optimization approach for decision-making on preventive maintenance of safety barriers considering safety risks and maintenance costs. The maintenance and proof test strategies for safety instrumented systems (SISs) have been studied considering element degradation (Zhang et al., 2022) and imperfect detection of degraded states (Zhang et al., 2021).

2.5.2 Integrated S&S barrier management

With security threats gradually attracting attention, the integrated management of safety and security (IMSS) has been promoted in chemical process industries. However, a study shows that security threat analysis and process safety analysis are usually undertaken independently in Seveso sites (Ylönen et al., 2022). The studies investigating integrated risk assessment and management (considering both safety hazards and security threats) are lacking. More attention should be paid to cyber-attack-induced loss of production (process shutdown) and major accidents. As the main ways to prevent and mitigate undesired events, safety and security (S&S) barriers play important roles in controlling integrated safety and security risks and achieving integrated risk management. We conclude the reasons for integrated S&S barrier management as below.

- Safety and security risks posed to chemical facilities inevitably interact with each other because security-related events may influence safety-related events and vice versa. Due to the interdependency between safety and security risks, it is necessary to conduct a risk analysis to consider both safety and security risks, and the obtained integrated risks can be regarded as the “real risks”. Because safety and security barriers have a synergistic effect on controlling the “real risks”, it is necessary to integrate safety and security barriers during the risk assessment process.
- The common goal of safety and security barriers is to reduce integrated safety and security risks. Considering the economic constraints in relation to barrier management actions (such as barrier allocations, operations, inspections, and maintenance), sole safety or security barrier management can hardly achieve the most cost-efficient risk control strategy. Therefore, integrated safety and security barrier management and investment can be more reasonable, effective, and economical with the help of safety and security barrier co-analysis and integrated investment.
- Compared to safety-barrier-related studies, the studies on security barriers are relatively less. Particularly, emerging C2P risks bring new challenges to barrier management, and related research is lacking. Because many similarities exist in safety and security-related problems, security-related research can learn a lot from the theories and models adopted in safety science. Developing systematic approaches to deal with integrated safety and security risks and achieving unified barrier management is possible.

Current research on S&S barrier integration and management is still in its early stages. Some attempts have been made to incorporate physical attack risks with conventional safety risks and support integrated risk management. For instance, a unified framework for risk analysis and management covering both safety and security was developed (Aven, 2007). Ayyub et al. (2007) developed a quantitative framework for critical asset and portfolio risk analysis (CAPRA) considering natural and human-caused hazards. Song et al. (2019b) employed an influence diagram to support decision-making on management actions considering accidental and intentional risks. Chen et al. (2020) proposed a methodology for decision-making on managing safety and security resources with respect to the prevention and mitigation of intentional domino effects. Khakzad et al. (2018b) suggested using precursor-based methodologies and data mining techniques to infer security risks of chemical plants and update security risk levels based on new data or observations. The suggested precursor-based dynamic risk assessment technique is aligned with the dynamic barrier concept (Pitblado et al., 2016), and is helpful for achieving dynamic S&S barrier management.

Regarding the incorporation of cyberattack risks, some tools or methodologies have been developed to facilitate integrated safety and security risk assessment and management. Kosmowski et al. (2015) proposed an integrated safety and security analysis approach for hazardous industrial plants based on the rings of protection concept. The proposed approach is similar to the layer of protection analysis (LOPA) method, incorporating functional safety concepts with cyber security aspects. Abdo et al. (2018) combined bow-tie analysis with attack tree analysis to represent risk scenarios in terms of safety and security. In the same study, a qualitative risk analysis was performed to prioritize system weaknesses and help decision-makers propose appropriate countermeasures. Similarly, Guzman et al. (2021) developed an integrated approach for safety and security analysis, in which the bow-tie model is used to develop an ontology of accident scenarios, and a database of risk sources and S&S barriers is given. Amin et al. (2022) developed a Bayesian network (BN) for performing a holistic safety and security risk analysis of chemical facilities with the consideration of intentional and unintentional risk sources. However, the research on quantitative risk assessment (QRA) of chemical facilities considering the interdependency between safety risks, physical security risks, and C2P risks has rarely been investigated in previous studies. Current studies cannot give enough theoretical and methodological support to achieve integrated S&S barrier management.

2.6 Discussions

Based on the literature review, we found that the barrier concept has been applied in chemical process industries for a long time, but challenges still exist in the management of S&S barriers. The primary findings are discussed as follows.

2.6.1 Lacks in integrating safety and security barriers

Previous studies mainly investigated safety issues and security issues in chemical plants separately (Ylönen et al., 2022). The interdependency between safety risks and security risks has not been well addressed in previous risk assessment studies. The emergency of cyber-physical (C2P) risks brings new challenges to the integrated risk assessment. Although some studies have investigated the optimization, maintenance, and management of safety barriers in chemical plants, the research on security barriers and the integration of safety and security barriers is lacking. Safety and security barriers may have synergistic effects on risk control of undesired scenarios in case chemical facilities are exposed to multi-dimensional risks that consider both safety and security aspects. The consideration of purely safety barriers or security barriers and the sole safety or security barrier management can hardly achieve real cost-effective decision-making. As important measures to prevent, control, or mitigate undesired events, safety and security (S&S) barriers are necessary to be integrated and managed in a unified framework to cope with integrated safety and security risks.

2.6.2 Potentials of data-driven dynamic barrier management

Quantitative risk assessment (QRA) is regarded as a valuable tool to improve risk management by identifying major risk contributors and proposing various risk reduction countermeasures in process industries (Freeman, 1990). QRA makes it possible to quantify the effects of the S&S barriers on risk reduction. It may integrate complex optimization algorithms to empower the decision-making process. However, difficulties exist in the quantitative assessment of barrier performance and in handling barrier dependency in the case of complex barrier systems. Additionally, the dynamic barrier management (DBM) concept was proposed by (Pitblado et al., 2016) to emphasize the importance of determining near-real-time barrier status and making decisions on barrier maintenance/improvements based on multiple data sources. The DBM methodology has the advantages of capturing dynamic variations in barrier performance, maintaining risks at the target levels throughout the installation's life, and possibly deriving timely and cost-efficient decisions. However, the application of the DBM concept can rarely be found in existing literature.

With the development of data-driven and artificial intelligence techniques, the landscape of safety and security in the chemical and process industry has been dramatically changed. On the one hand, the implementation and operation of automatic or intelligent control systems bring new challenges to safety and security management, such as cybersecurity issues becoming new threats to process safety. On the other hand, data-driven approaches have the potential to support S&S barrier management by revealing or estimating the status of barriers based on a large amount of data in relation to the performance of barriers. For instance, the data-driven approach helps plant operators and engineers deal with complex tasks like process monitoring, fault detection and diagnosis, and maintenance optimization (Stluka & Mařík, 2007; Jain et al., 2019). By incorporating data-driven techniques into barrier management frameworks, data from multiple sources may be employed to determine near-real-time barrier status, and therefore achieve dynamic barrier management.

2.7 Conclusions

This chapter presents a systematic review of the definition, classification, performance assessment, and management of safety and security (S&S) barriers in chemical process industries. Although the literature has contributed tremendously to implementing and managing S&S barriers, many endeavors are needed to develop a dynamic and integrated S&S barrier management system. Previous studies mainly addressed safety risks and security risks in chemical plants separately. An integrated assessment of safety and security risks is needed considering their dependency.

Particularly, the involvement of cyber-physical risks in the assessment is necessary. The management and optimal allocation of safety and/or security barriers have been focused on in previous studies, but a comprehensive safety and security barrier management framework has not yet been developed. Challenges in safety and security barrier integration, monitoring and revealing near-real-time barrier status, and cost-effective decision-making should be addressed to facilitate dynamic and integrated safety and security barrier management.

Chapter 3 Integrated safety and security risk assessment of industrial control systems in chemical plants

Aligned with the development needs of Industry 4.0, industrial cyber-physical systems (ICPSs) are widely applied to chemical facilities to achieve the so-called intelligent production processes. Meanwhile, emerging cyber-to-physical (C2P) risks are introduced due to the vulnerability of ICPSs to cyberattacks. An integrated safety and security risk assessment of chemical facilities equipped with industrial cyber-physical systems becomes challenging, particularly in performing a probabilistic/quantitative risk assessment. Targeting this gap, this study develops a systematic approach to construct accident scenarios concerning both safety hazards and security threats and performs a probabilistic risk assessment of industrial control systems considering the interdependency between safety risks and security risks. An illustrative case study is used to give guidance on performing integrated safety and security risk assessment of ICPSs and validate the feasibility of the proposed approach.

This chapter is drafted with modifications based on the following publication:

◆ Yuan, S., Yang, M., & Reniers, G. (2024). Integrated process safety and process security risk assessment of industrial cyber-physical systems in chemical plants. *Computers in Industry, 155,* 104056.

3.1 Introduction

With the advent of the digital age and Industry 4.0, new threats and risks have emerged in the chemical process industries (Kriaa et al., 2015; Ji et al., 2021). Previous studies show that a cyberattack on an ICPS (industrial cyber-physical system) may adversely impact physical components and further cause damage to humans, assets, and the environment (Kriaa et al., 2015; Huang et al., 2018). The corresponding risks are known as cyber-to-physical (C2P) risks (Yampolskiy et al., 2013). As a result, chemical facilities are exposed to multi-dimensional risks, and managing those risks will inevitably involve both safety and security barriers.

For effective S&S barriers management, an integrated safety and security risk assessment of chemical facilities should serve as the basis. Different researchers have already made some attempts to analyze the integrated safety and security risks. For instance, Abdo et al. (2018) combined bow-tie diagrams and attack trees to demonstrate adverse scenarios of industrial control systems (ICSs), considering safety hazards and security threats. Guzman et al. (2020) suggested using a multi-layered representation for safety and security analysis of CPSs considering information flows and energy flows. Additionally, as an extension of the system theoretic process analysis (STPA) approach, STPA-SafeSec was developed and implemented for the safety and security analysis of cyber-physical systems (Friedberg et al., 2017). Alanen et al. (2022) proposed an ontology-based approach for cybersecurity risk analysis of ICSs. Huang et al. (2018) combined a Bayesian network (BN) and a stochastic hybrid system (SHS) to quantify the physical impact of cyberattacks on ICPSs.

However, to the best of the authors' knowledge, research on quantitative risk assessment of chemical facilities that consider the interdependency between safety, physical security, and C2P risks is still lacking. Because both safety hazards and security threats could lead to major adverse consequences, the consideration of only safety hazards or security threats could lead to a risk underestimation. Additionally, separate assessments of safety-associated scenarios and security-associated scenarios cannot reveal the real risks due to the ignorance of the interdependency between safety and security, which hinders the implementation of integrated S&S barrier management.

To this end, this chapter provides a systematic approach to risk assessment of industrial control systems in chemical plants, considering the interactions between safety-hazard-induced adverse events and security-threat-induced adverse events. The remainder of this chapter is organized as follows. Firstly, the research scope of interest is well identified and the safety and security risk calculations are presented in Section 3.2. Then, the proposed approach is illustrated in Section 3.3 while an illustrative case study is used to theoretically validate the feasibility of the proposed approach in Section 3.4. A sensitivity analysis of the basic events is presented in Section 3.5 before conclusions are given in Section 3.6.

3.2 Theoretical background

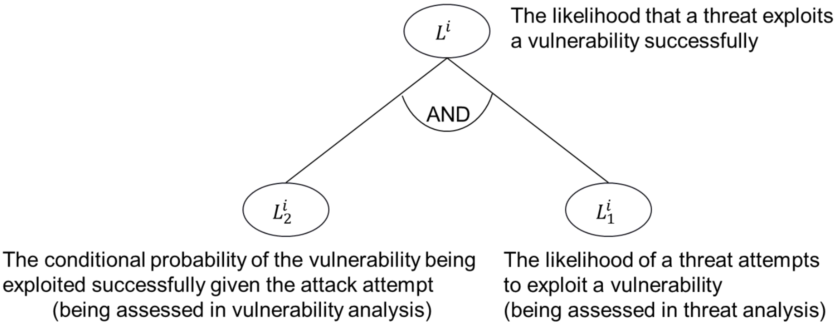
As already indicated in the introduction chapter, studies associated with unintentional or random losses (due to hazards) are considered to belong to the safety domain. In contrast, studies related to intentional losses and with deliberate nature (deliberate misuse of hazards) belong to the security domain (Landucci et al., 2020). With cyber-physical (C2P) attacks getting more and more attention, the investigation of the physical damages induced by cyberattacks becomes even more important (Flaus, 2019). Aligned with the promotion of safety and security integration, all kinds of causes (safety hazards, physical attacks/acts, and cyber-physical attacks) leading to major adverse scenarios (fires, explosions, toxic leakage, etc.) should be covered in the risk assessment exercise to generate a variety of dangerous scenarios and to lead to more thorough risk analysis. In the safety science domain, risk is widely presented as a function of the likelihood of an unwanted scenario i (presented by L^i) and its expected consequence severity, S^i, as follows (Freeman, 1990; Meyer & Reniers, 2022):

$$R^i_{safety} = L^i \times S^i \tag{3.1}$$

Regarding security risks, the API standard 780 (API, 2013) defined security risk as a function of the consequences (C^i) of a successful attack scenario i and the likelihood (L^i) of the happening of this successful attack scenario. The likelihood is further defined as a function of the attractiveness (A^i) to the adversary of the asset, the degree of threat (T^i) posed by the adversary, and the degree of vulnerability (V^i) of the asset. According to the IEC 62443-3-2 standard (IEC, 2020), which particularly serves the information security of ICSs, the cybersecurity risk is expressed as the likelihood (L^i) that a particular threat will exploit a particular vulnerability with a particular consequence (C^i). With the consideration of the above two definitions, security risks can be calculated as follows (Landucci et al., 2020):

$$R^i_{security} = L^i \times C^i = (A^i \times T^i) \times V^i \times C^i = L^i_1 \times L^i_2 \times C^i \quad (3.2)$$

where L^i_1 is the likelihood of an attempt to exploit a vulnerability. L^i_2 presents the conditional probability of the vulnerability being exploited successfully given the attack attempt. Usually, L^i_1 should be evaluated in the threat analysis. L^i_2 reflects the vulnerability of the targeted system posed to attacks, and it is usually addressed in the vulnerability assessment. We demonstrate how to calculate security risks based on an attack tree analysis, as shown in Figure 3.1.



L^i
The likelihood that a threat exploits a vulnerability successfully
AND
L^i_2
L^i_1
The conditional probability of the vulnerability being exploited successfully given the attack attempt (being assessed in vulnerability analysis)
The likelihood of a threat attempts to exploit a vulnerability (being assessed in threat analysis)

Figure 3.1. Probabilistic calculation based on an attack tree analysis.

3.3 Methodology

3.3.1 Overview of the proposed approach

The workflow of the proposed approach is organized in a systematic manner, as presented in Figure 3.2. The approach starts with system representation using a CPS (cyber-physical system) master diagram, and then, safety analysis and security analysis are conducted to generate an integrated attack-tree-bow-tie diagram. Furthermore, a BN model is developed based on the attack-tree-bow-tie diagram to perform quantitative risk assessment. The following subsections illustrate the details of each step.

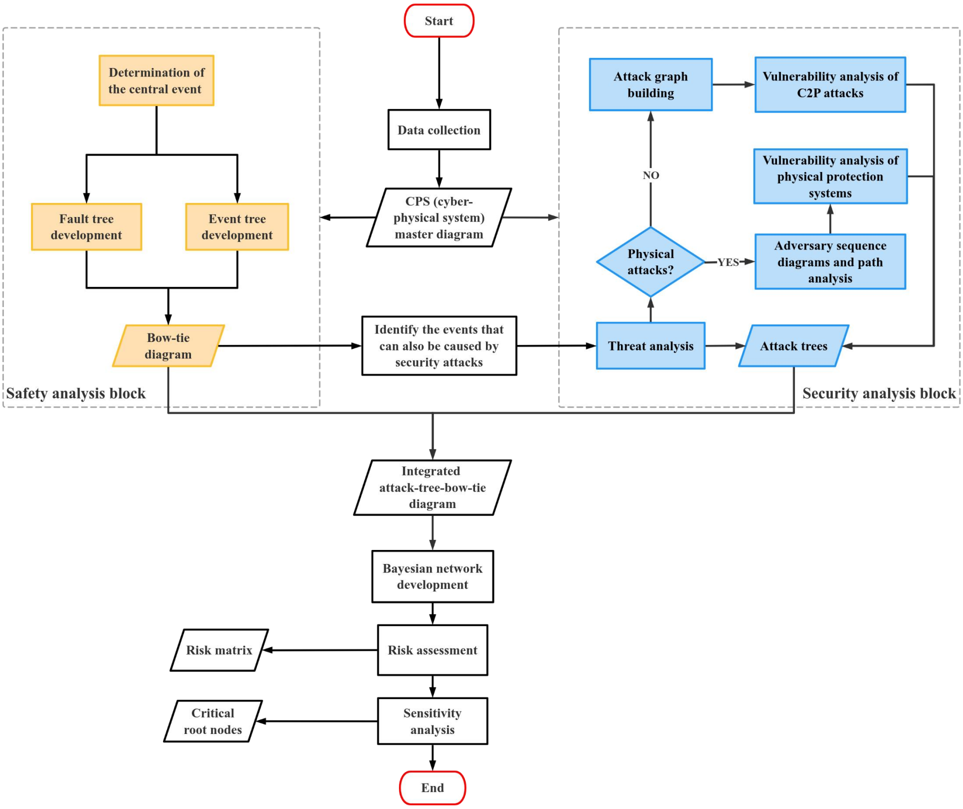

Start
Data collection
CPS (cyber-physical system) master diagram
Determination of the central event
Fault tree development
Event tree development
Bow-tie diagram
Safety analysis block
Identify the events that can also be caused by security attacks
Attack graph building
Vulnerability analysis of C2P attacks
Vulnerability analysis of physical protection systems
NO
Physical attacks?
YES
Adversary sequence diagrams and path analysis
Threat analysis
Attack trees
Security analysis block
Integrated attack-tree-bow-tie diagram
Bayesian network development
Risk matrix
Risk assessment
Critical root nodes
Sensitivity analysis
End

Figure 3.2. Flowchart of the proposed methodology.

3.3.2 System representation and CPS master diagrams

The system complexities of ICPS bring enormous difficulties to safety analysis and security analysis. As a result, an appropriate representation of the ICPSs should serve as a basis for safety analysis, security analysis, and their corresponding scenario building. Guzman et al. (2020) suggested using a multi-layered representation of CPSs for an integrated safety and security analysis. A tool named CPS master diagram was proposed by the same study, in which the CPS is represented by three layers (physical layer, cyber-physical layer, and cyber layer) with the illustration of the information/data flow and energy flow between different components and between the CPS and external environments. An exemplary CPS master diagram is presented in Figure 3.3.

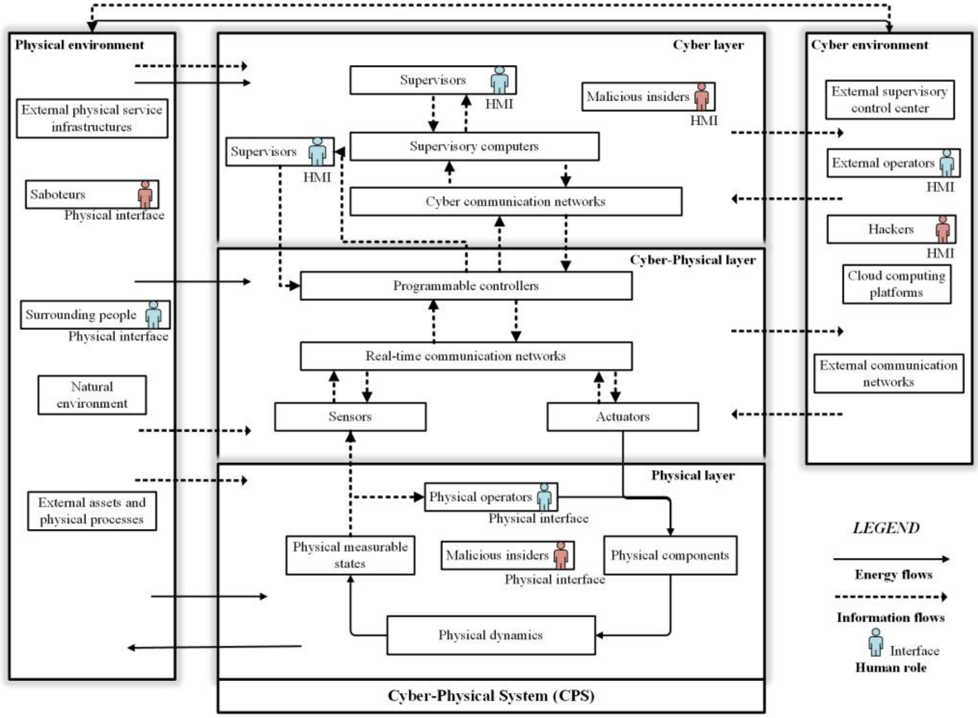


Figure 3.3. An exemplary CPS master diagram, adapted from (Guzman et al., 2020).

3.3.3 Safety risk analysis based on bow-tie diagrams

As a graphical tool, bow tie diagrams are widely used for accident scenario identification and visualization due to their advantage of being straightforward to communicate to a wide range of audiences (CCPS/EI, 2018). The development of a bow-tie diagram usually begins with the determination of the central event. Then, a fault tree considering the possible causes of the central events should be constructed based on the energy flows and information flows presented in the CPS master diagram. Typically, the basic events in the fault tree include technical component failures, human errors, external interventions, etc. The occurrence probabilities of those events can be derived from reliability databases (OREDA, 2002; Hauge & Onshus, 2010), human reliability data (Kirwan, 2017), accident databases (Debray et al., 2004), or data available in the literature. Meanwhile, an event tree considering the possible consequences after the occurrence of the central event should be developed with the help of the available guidelines. For instance, the ARAMIS project (Andersen et al.,

2004) provided methods for constructing event trees with respect to major accident hazards in chemical plants. Vílchez et al. (2011) provided a set of generic event trees considering the release of hazardous materials in chemical plants.

3.3.4 Security risk analysis

3.3.4.1 Threat analysis

Security risk analysis begins with a threat analysis aiming to identify threat agents who may execute physical attacks or C2P attacks. It is suggested to identify potential threat agents of the targeted chemical facilities by asking who can conduct attacks and why. The approach suggested by SFK (2002) is adopted to identify threat agents' categories (TAC), as shown in Table 3.1.

Table 3.1 Definitions of threat agent categories (TAC) according to SFK (2002).

Features	**TAC1: threat agent moved by contingent intent**	**TAC2: threat agent moved by direct intent**	**TAC3: terrorists and extremists**
Agents	Individuals or small groups	Small network of activists, members of organized crime, individuals, radical political groups	Extremist and terrorist individuals and groups
Aim	Limited damage; possible unawareness of attack escalation into major adverse events	Major damage; escalation into a major adverse event/scenario may be a possible objective	Massive terrorist attack, armed action, causing the maximum possible damage, without regard to people's life (own or others)
Motivation	Revenge, frustration, prove existence of deficits, achieve social effects	Revenge, political radicalism, gaining financial/competitive advantages	Religion related motives, anarchy, "punishing companies"
Potentiality	Limited potentiality, dependent on the motive	Above average criminal energy, average communication capability, medium level of organizational support, poor financial backing	Extremely great criminal energy, highly developed communication capability, high level of organizational support, high financial backing
Tools and means	Simple or major tools, possibly simple incendiary	Simple and specialized tools, incendiary devices, home-made explosives	Simple and heavy tools, weapons, explosives, incendiary devices

	devices

Threat analysis also addresses the estimation of attack likelihoods, which correspond to L_1 in Figure 3.1. It is suggested to estimate attack likelihoods according to the actual annual frequency of attacks in the investigated chemical plants or refer to similar companies in the same/similar sector. However, not many companies revealed the security attack information due to confidential issues. Alternatively, a simplified frequency estimation of physical attacks can be implemented based on the API threat levels and facility expected life according to (API, 2013) and (Landucci et al., 2017), as presented in Table 3.2.

Table 3.2 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}

Regarding C2P attacks, attack likelihood may be estimated by analyzing cyber incident data. According to the statistical analysis of 60,767 cyber security incidents that occurred in the US from November 2008 to January 2015 (Kuypers & Maillart, 2018), it was observed that the recurrence intervals of larger events remain overall stable. The recurrence intervals of cyber security incidents with different severities, which are measured in the form of efforts (man-hours) spent remediating the incidents, are given in Table 3 (Kuypers & Maillart, 2018). With reference to the results presented in Table 3.3, the recurrence interval of C2P attacks is estimated at approximately 150~465 days. In case no incident data is available, security experts may estimate the attack likelihood based on their own knowledge and experience.

Table 3.3 Recurrence intervals of cyber security incidents with different severities, adapted from (Kuypers & 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

3.3.4.2 Vulnerability analysis with respect to physical attacks

Vulnerability assessment aims to identify credible attack paths and estimate the conditional probability of successful attacks given the attack attempts (corresponding to L_2 in Figure 3.1). Physical attacks are usually subject to PPSs (physical protection systems). A systematic approach should be implemented to identify credible attack paths considering associated PPSs, for instance, fences, entry control, closed circuit television (CCTV), emergency teams, and so on (Reniers et al., 2017). An Adversary Sequence Diagram (ASD) is a graphical representation of PPS elements along attack paths that adversaries may follow to accomplish their objectives. This study adapts Adversary Sequence Diagrams and Path Analysis to identify the credible attack paths of physical attackers. Details of Adversary Sequence Diagrams and Path Analysis can be found in Norman (2010). After the credible attack paths are identified, an event tree analysis and the benchmark data presented by Moreno et al. (2022) are used for the vulnerability assessment of PPSs.

3.3.4.3 Vulnerability analysis with respect to C2P attacks

Based on the analysis of 82 cybersecurity-related incidents in the process industry, gathered from various sources, Iaiani et al. (2021) reported that petrochemical and energy production facilities are the most affected industrial sectors by cyberattacks. These sectors are particularly attractive targets due to the potential for severe consequences resulting from such attacks. While it was found that most cyberattacks primarily impacted IT systems, several cases demonstrated that cyberattacks can also manipulate OT systems, leading to major adverse events. For example, two major incidents (explosions) were caused by remote manipulations of OT systems, where infections in OT components (such as HMI workstations, OT servers, etc.) ultimately resulted in physical damage to the CPS and severe consequences.

Given this, it is essential to investigate potential attack modes targeting OT systems in industrial CPS, as they may be the ultimate targets for C2P attacks. Industrial PLCs (programmable logic controllers), which play a critical role in controlling the physical dynamics of chemical facilities, often become primary targets for C2P attacks. Figure 3.4 illustrates six typical types of attacks against industrial PLCs, with explanations provided in Table 3.4.

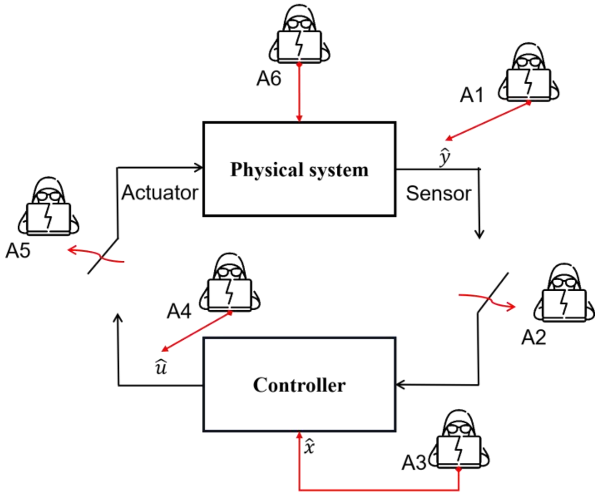


Figure 3.4. Typical attacks against industrial PLCs, adapted from (Huang et al., 2009; Wen et al., 2023).

Table 3.4 Explanation of typical C2P attacks against industrial cyber-physical systems, adapted from (Huang et al., 2009; Orojloo et al., 2017; Wen et al., 2023).

Marks	Attack types	Descriptions
A1	FDI (false data injection) attack against sensors	Maliciously manipulate the measurement data from sensors to the controller. Let $\hat{y} \neq y$, $\hat{y}$ is the manipulated data, and y is the true measurement.
A2	DoS (denial-of-service) attack against sensors	Maliciously prevent the controller from receiving sensor measurement data.
A3	Setpoint manipulation	Maliciously manipulate the setpoints configured in the controller. Let $\hat{x} \neq x$, $\hat{x}$ is the manipulated setpoint, and x is the predefined setpoint.
A4	FDI attack against actuators	Maliciously manipulate the control data from the controller to actuators. Let $\hat{u} \neq u$, $\hat{u}$ is the manipulated data, and u is the true control data.

A5	DoS attack against actuators	Maliciously prevent actuators from receiving control commands/data.
A6	Physical attack	Physical attacks against actuators or direct physical attacks on the vessels.

The execution of C2P attacks cannot always induce dangerous scenarios, instead, some of those attacks only cause some deviations that the control system can suppress (Huang et al., 2009; Cárdenas et al., 2011). As a result, an attack modelling approach is developed based on MATLAB/Simulink platform to assess the deviations caused by C2P attacks and to identify dangerous attack modes. Details of the developed attack modelling approach can be found in Appendix I.

Regarding the identified dangerous attack modes, the CPS master diagram and attack/compromise graph are combined to identify and visualize the credible attack paths. This process starts with the identification of possible PoAs (points of access), which are usually the interfaces between the attackers and the investigated cyber-physical system. Then, each attack step executed by the attackers starting from the PoAs to achieve their final attack objectives (which is usually the compromise of physical components) may be analyzed based on the information flows and control flows demonstrated in the CPS master diagram. Additionally, ICS vulnerability databases, for instance, an ICS-specific vulnerability dataset (Thomas & Chothia, 2020), may be used to identify the known vulnerabilities that may be exploited by attackers at each attack step. Finally, the implementation of an attack/compromise graph helps to visualize the attack paths considering the known vulnerabilities at each attack step (Semertzis et al., 2022). An example of the attack/compromise graph can be found in Figure 3.12.

Time-to-compromise (TTC) was defined as the time needed for an attacker to gain some levels of privilege on a system component by McQueen et al. (2006a). In the same study, a TTC estimation approach was developed based on the Common Vulnerabilities and Exposures (CVEs) database (NVD, 2023). Then, the TTC approach has been extended and applied to quantitative risk/reliability assessment of process control networks (Henry & Haimes, 2009) and power systems (Zhang et al., 2015; Semertzis et al., 2022). More recently, an augmented TTC approach was developed particularly for ICSs (Ling & Ekstedt, 2022; 2023) based on an ICS-specific vulnerability dataset (Thomas & Chothia, 2020). We adapt the approach developed by Ling & Ekstedt (2022) to estimate the TTC of each attack step considering attackers' skill levels, the number of known vulnerabilities, and the exploitabilities of the known vulnerabilities. Details of the TTC estimation approach can be found in Appendix II. More details can be found in the original study (Ling & Ekstedt, 2022).

After the estimation of the TTC of each attack step, the global TTC of an attack path can be calculated by summing the TTC of each attack step along the attack path. For the target that can be accessed through multiple attack paths, the attack path with the shortest global TTC is used from a conservative point of view. Then, the conditional probability of a C2P attack inducing physically dangerous scenarios successfully, L^i, can be estimated based on the global TTC ($TTC_G(i)$) and the mean-time-to-detect ($MTTD_i$) regarding this attack scenario (Semertzis et al., 2022), as presented below.

$$TTC_G(i) = \sum_{j=1}^{n} TTC_j \tag{3.3}$$

$$L^i = \frac{MTTD_i}{TTC_G(i)+MTTD_i} \times \beta_i \tag{3.4}$$

$$MTTD_i = \frac{\sum_{k=1}^{N} TTD_k}{N} \tag{3.5}$$

where $TTC_G(i)$ is the global TTC of an attack path i. TTC_j is the local TTC of attack step j. n is the number of attack steps along this attack path. Mean-time-to-detect (MTTD) measures the average time it takes for the security operations center (SOC) to detect a security incident, which is one of the key metrics used to measure SOC performance (Mughal, 2022). The MTTD regarding a specific intrusion type is the sum of all incident detection times of this intrusion type ($\sum_{k=1}^{N} TTD_k$) divided by the total incident number of this intrusion type (N), as shown in Eq (3.5). The MTTD values can be estimated based on the analysis of security incident data in practice. β_i is a coefficient depicting the likelihood of a physically dangerous scenario that may be induced by a successful intrusion of attack path i. β_i depends on the vulnerability of the OT (operational technology) system regarding specific C2P attack modes, and it is determined based on the approach presented in Appendix I. Regarding the attacks that are not subject to intrusion detection systems (IDS), for instance, stealthy attacks (Hu et al., 2019), $L^i \approx \beta_i$ ($MTTD \approx + \infty$ in case of stealthy attacks) may be used because stealthy attacks can evade the detection of the IDS and inject manipulated data into the control system.

3.3.5 Integrated safety and security risk analysis

3.3.5.1 Integrating attack trees into the bow-tie diagram

After threat analysis and vulnerability analysis, a simplified attack tree (like the attack tree in Figure 3.1) for each attack mode should be developed, to incorporate the results from the threat analysis and vulnerability analysis. The simplified attack trees employ attack likelihoods (derived from threat analysis) and the conditional probabilities of successful attacks given attack attempts (derived from vulnerability analysis) to calculate the probability of successful execution of each attack mode without the demonstration of detailed attack paths and attack steps. Then, the developed attack trees are integrated into the bow-tie diagram for developing a BN model and for integrated safety and security risk assessment. This simplification of the attack trees helps to reduce the number of BN nodes effectively and meanwhile retain the necessary quantitative data for risk assessment. Particularly, regarding the assessment of large-scale facilities, complex attack paths may make the integrated safety and security risk analysis unachievable/unmanageable using BN models. The simplification of the attack trees makes the BN model developing process easier and makes it possible to perform a risk assessment of large-scale facilities considering both safety-related and security-related scenarios. A systemic workflow is implemented to conduct the scenario integration, as presented in Figure 5. Basically, this can be done by checking if each event in the bow-tie diagram can also be induced by security attacks. If the answer is yes, the corresponding attack trees should be attached to the event.

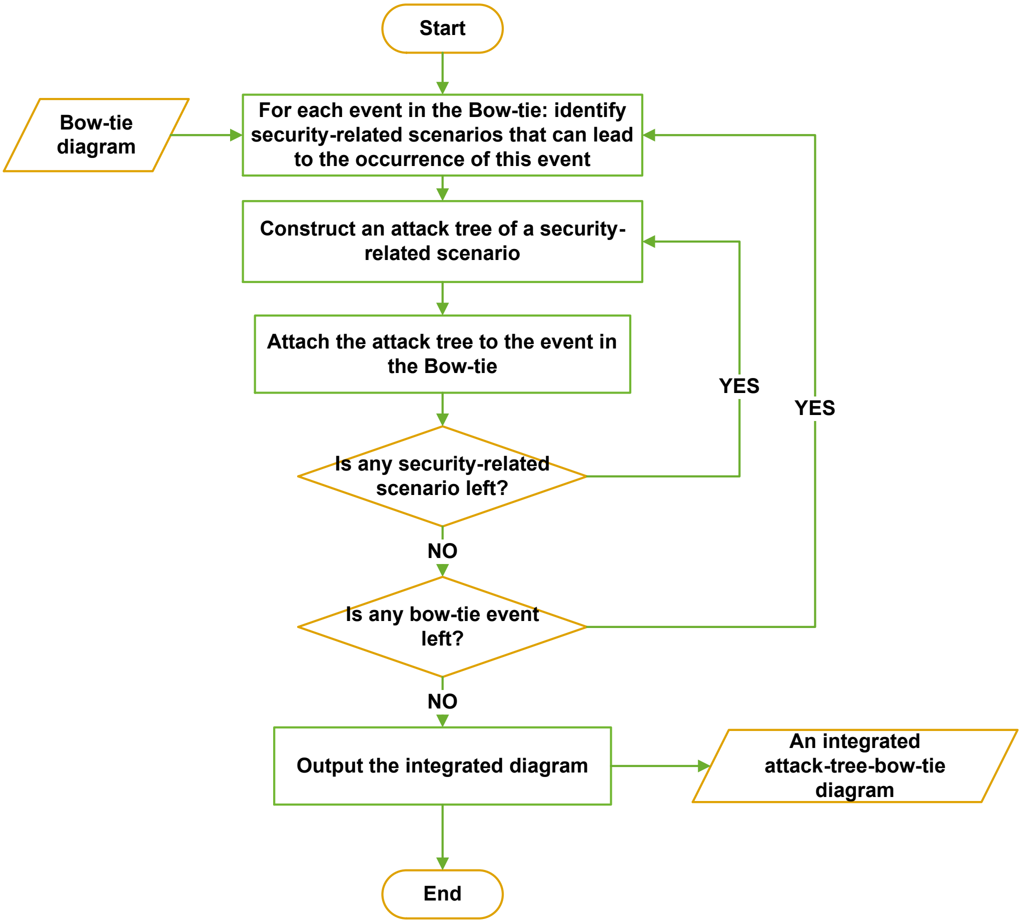

Start
Bow-tie diagram
For each event in the Bow-tie: identify security-related scenarios that can lead to the occurrence of this event
Construct an attack tree of a security-related scenario
Attach the attack tree to the event in the Bow-tie
YES
YES
Is any security-related scenario left?
NO
Is any bow-tie event left?
NO
Output the integrated diagram
An integrated attack-tree-bow-tie diagram
End

Figure 3.5. Flowchart of integrating attack trees into the bow-tie diagram, adapted from (Abdo et al., 2018).

3.3.5.2 Bayesian networks

Bayesian Networks (BNs) are widely-used to perform safety or security risk assessments (Tong et al., 2018; George & Renjith, 2021). Compared to conventional bow-ties and fault/attack trees, BN has the advantage of backward diagnostic analysis and handling dependent basic events and multiple occurrence events (Yuan et al., 2023b). Therefore, it is suggested to transform the obtained attack-tree-bow-tie diagram into a BN model for integrated safety and security risk analysis. A BN consists of a set of nodes, their correlations (represented by directed arcs), prior probabilities, and conditional probability tables (CPTs). A joint probability distribution $P(X)$ of variables $X = \{X_1, ..., X_n\}$ is presented in a BN as follows (Jensen & Nielsen, 2007):

$$P(X) = \prod_{i=1}^{n} P(X_i|Pa(X_i)) \tag{3.6}$$

where $Pa(X_i)$ is the parent node set of X_i. When evidence E becomes available, the posterior probabilities $P(X|E)$ can be calculated based on Bayes theorem as follows (Jensen & Nielsen, 2007):

$$P(X|E) = \frac{P(E|X) \cdot P(X)}{P(E)} = \frac{P(E, X)}{\sum_X P(E, X)} \tag{3.7}$$

Both the topology and CPTs of the BN model can be derived from the integrated attack-tree-bow-tie diagram. Previous studies have already illustrated the mapping process for transforming fault trees (Bobbio et al., 2001), attack trees (Gribaudo et al., 2015), and bow-tie diagrams (Khakzad et al., 2013) into BNs. Detailed procedures and guidelines can be found in related studies, this paper avoids repeating illustrations here.

3.3.5.3 Risk evaluation and sensitivity analysis

Risk evaluation considers both the occurrence probabilities and severities of the undesired consequences. The BN model takes the responsibility to estimate the occurrence probabilities of the undesired consequences. In this study, severities of the undesired consequences are determined based on qualitative severity classifications, for instance, the severity classes of typical dangerous phenomena defined in the ARAMIS project (Andersen et al., 2004). Then, implementing a risk matrix helps visualize risk profiles considering both probabilities and consequence severities. The acceptance of risk may be decided by comparing the occurrence probability of each consequence to its threshold defined by experts or stakeholders. In this study, a risk matrix adapted from the ARAMIS project (Andersen et al., 2004) is used to perform the risk visualization and risk evaluation.

Regarding sensitivity analysis, the ratio of variance (RoV) measure was introduced by Zarei et al. (2017) to identify critical root nodes of BNs. Both prior and posterior probabilities of the BN nodes are required in the calculation of RoV, as follows:

$$RoV_N = \frac{P'(N) - P(N)}{P(N)} \tag{3.8}$$

where $P^{'}(N)$ is the posterior probability of node N and $P(N)$ is the prior probability of node N. By changing the state of the leaf/intermediate node (that denotes the undesired event) into “happening”, the sensitivity of each root node to the happening of the undesired event can be analyzed. The root node with a higher RoV value is more sensitive/critical.

3.4 Case study

3.4.1 System representation and scenario building

In this case study, an integrated safety and security risk analysis of a continuous stirred tank reactor (CSTR) with its SCADA system is investigated. This CSTR model runs a hypothetical exothermic first-order reaction A → B, and it is adapted from (Pilario & Cao, 2018). Product B is assumed to be a flammable liquid with toxicity. The CSTR with its SCADA system and safety instrumented system is shown in Figure 3.6. A jacketed tank is deployed to maintain the temperature inside the reactor with industrial water provided by a water pump (WP). A control valve (V1) is implemented to feed reactant A at a fixed flow rate, and two PLCs (programmable logic controllers) are implemented to achieve automatic process control and emergency shutdown. PLC1 controls the coolant flow rate by regulating a control valve (V3) based on the measurement of a temperature sensor (T). PLC2 serves the emergency shutdown system (ESD) by activating the block/shutdown valve (V2) in case overpressure is detected by the pressure sensor (P). Both of the PLCs are supervised by site managers through the HMI (human-machine interface) of the SCADA system, which is linked to the corporate network and, further, the outside Internet/WAN. A safety relief valve (SV) is installed to ensure the safety of the chemical reactor in case of overpressure. V4 is a block valve and will be activated in abnormal situations manually to stop outputting the product and isolate the CSTR from other downstream facilities. With the analysis of the information flows and energy flows, a CPS master diagram is developed for this chemical reactor with the consideration of the human roles and the interactions between the CPS and external environments. The developed CPS master diagram is demonstrated in Figure 3.7.

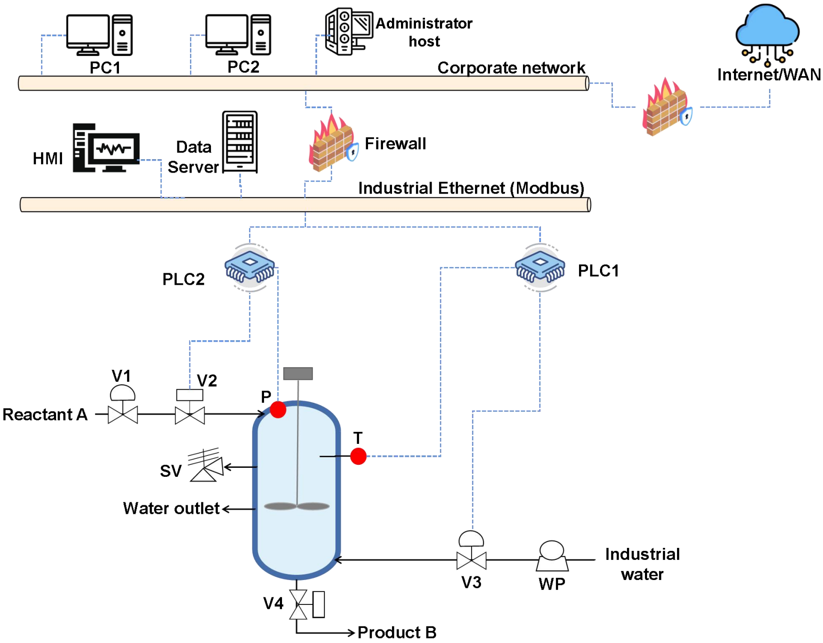


Figure 3.6. The investigated chemical reactor with its SCADA system.

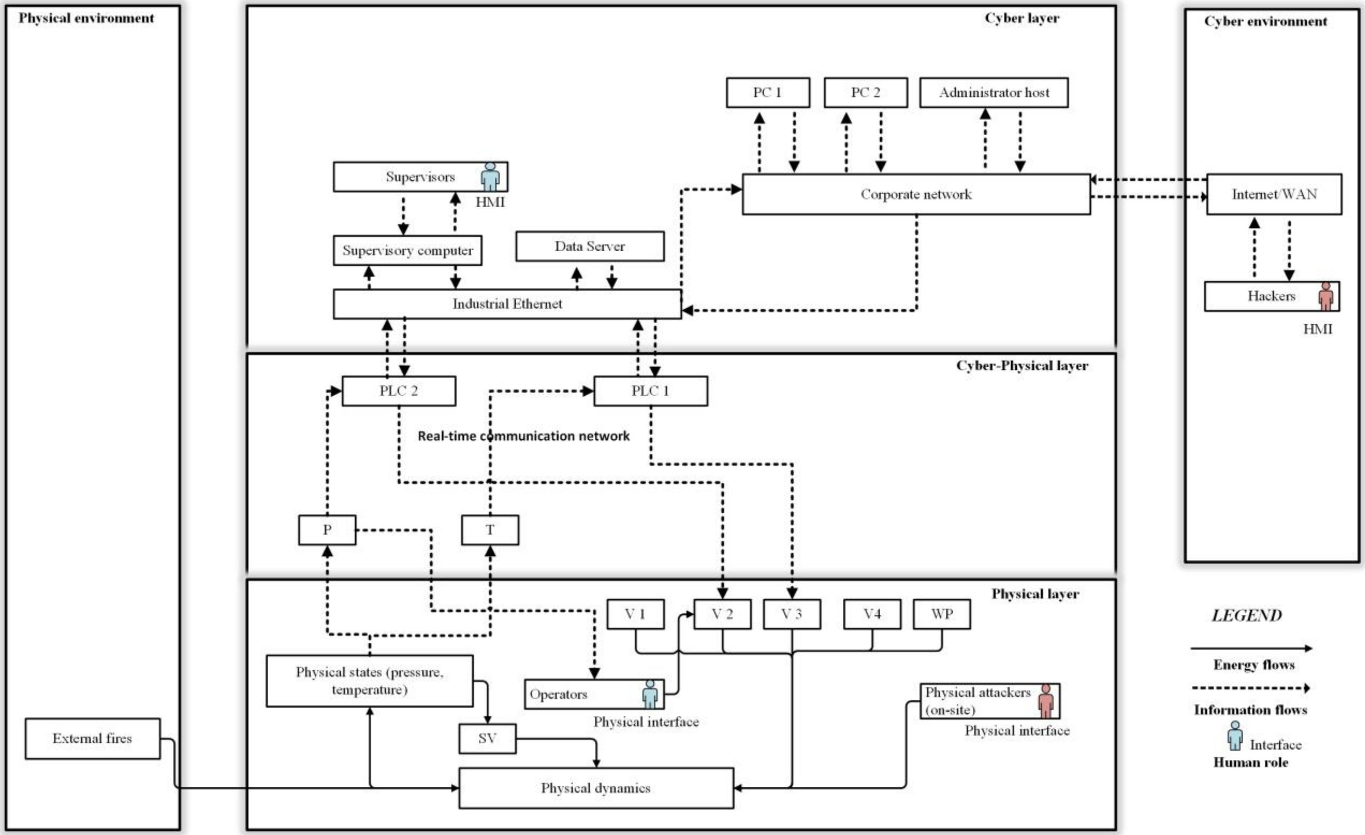


Figure 3.7. CPS master diagram of the investigated chemical reactor.

3.4.2 Risk assessment model development

3.4.2.1 The developed bow-tie diagram

Based on the CPS master diagram presented in Figure 3.7, a bow-tie diagram for representing accident scenarios was constructed, as shown in Figure 3.8. A release of toxic and flammable liquid was decided as the central event. Two safety barriers, which are an ESD system (emergency shutdown system) and a safety relief valve (SV), are deployed to prevent shell rupture in case of overpressure. At the left-hand side of the bow-tie diagram, a fault tree analysis was performed to identify the possible causes of a liquid release. At the right-hand side of the bow-tie diagram, a generic event tree for the release of toxic and flammable liquids adapted from Vílchez et al. (2011) is used.

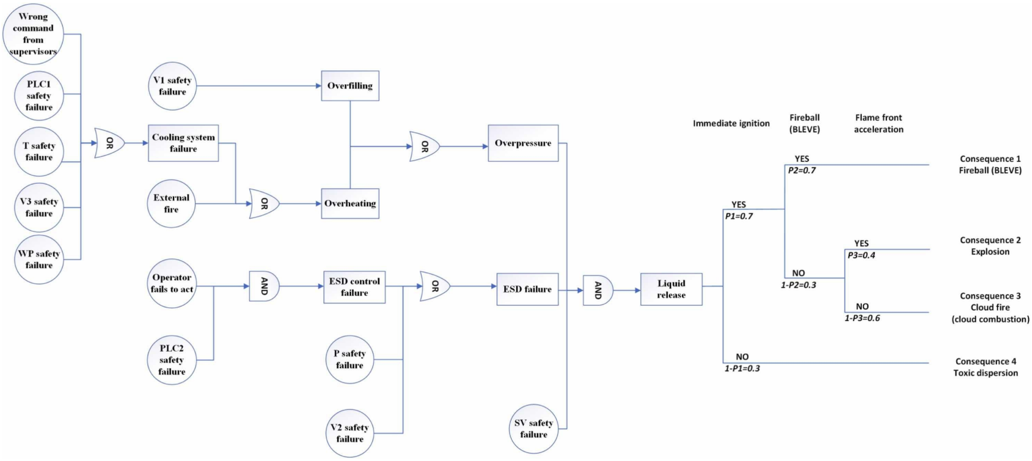


Figure 3.8. Bow-tie diagram with a toxic and flammable liquid release as the central event, adapted from (Vílchez et al., 2011).

3.4.2.2 Threat analysis and attack impact analysis

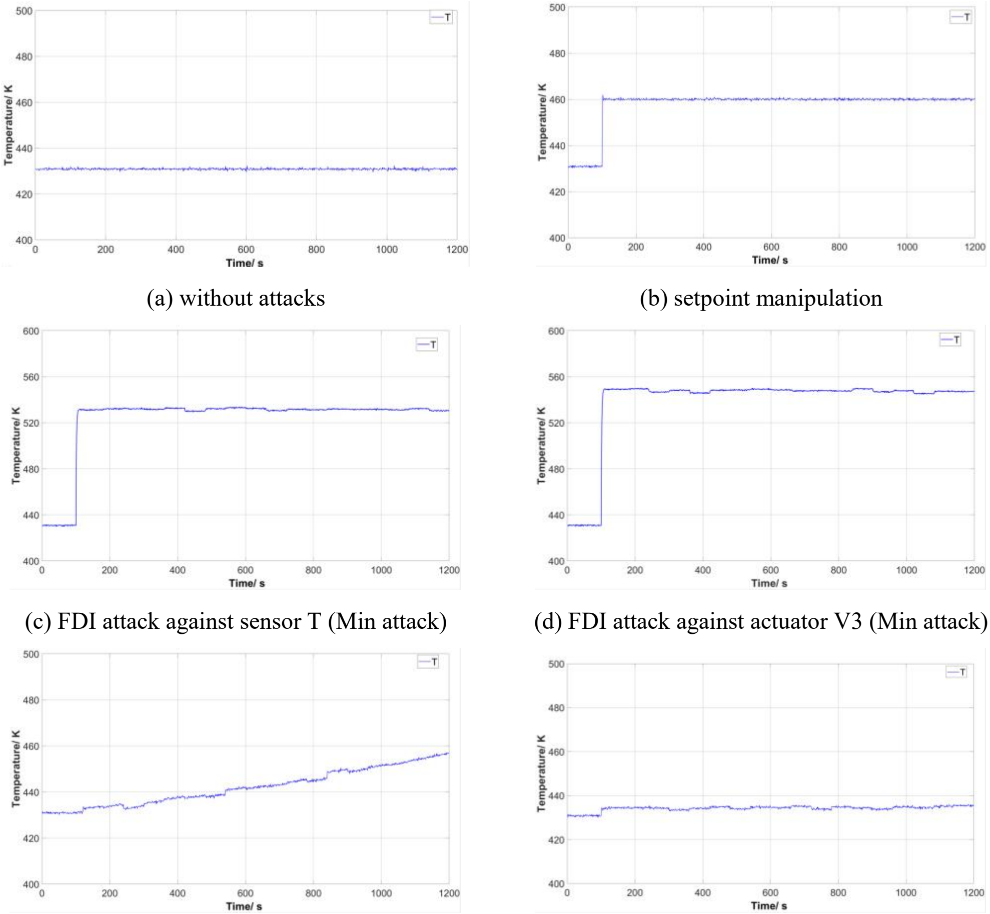
In this step, a threat analysis was performed first to identify threat agents and their corresponding PoAs (points of access), attack targets, and attack modes. From the conservative point of view, hackers with high-skill levels are identified as potential attackers implementing C2P attacks. Individuals or small groups driven by contingent intent with simple or major tools (TAC1) are identified as potential external physical attackers (SFK, 2002). According to the accident data analysis of security-related events in chemical plants (Landucci et al., 2020), terrorism mainly causes explosions as final scenarios, thefts and vandalisms are more likely to result in the release of hazardous chemicals, and C2P attacks mainly result in the loss of control of process systems. Therefore, the attack objectives of the physical attackers and hackers are identified as triggering the release of hazardous chemicals and triggering the loss of control of the cooling system and ESD system, respectively. The identified attack modes are given in Table 3.5. It is important to note that OT safety barriers play a crucial role in preventing the success of C2P attacks. In attacks targeting the jacket cooling system, which could lead to overpressure scenarios, OT safety barriers such as the emergency shutdown system, manual shutdown, and safety relief valve are expected to activate in time to prevent shell rupture and liquid release. However, C2P attacks may also compromise digitized barriers, causing them to fail to activate successfully. Consequently, attacks targeting the emergency shutdown system controlled by PLC2 are also identified as dangerous attack modes (AT5-AT9). Table 3.5 doesn't provide a thorough list of security threats and the case study is only used for demonstration purposes. For instance, stealthy attacks and APTs (advanced persistent threats) are not considered in the case study. In practice, more security threats may exist, and it is possible to consider more security attack scenarios with credibility and perform an integrated safety and security risk assessment based on the proposed framework.

Table 3.5 Identified threat agents and their corresponding attack modes.

Threat agents	**Points-of-access (PoAs)**	**Attack targets or attack objectives**	**Attack modes**	**Is it capable to induce dangerous scenarios?**	β **coefficient**[5]	**Marks**
Hackers (with high-skill levels)	Device that is connected to external Internet/WAN	Compromise PLC1 (cooling system) and PLC2 (ESD system), trigger dangerous overpressure scenarios	FDI attack against sensor T	YES	$\beta = 1$	AT1
			DoS attack against sensor T	YES	$\beta = 0.5$	AT2
			FDI attack against actuator V3	YES	$\beta = 1$	AT3
			DoS attack against actuator V3	NO	$\beta = 0$	/
			Setpoint manipulation of temperature threshold of PLC1	YES	$\beta = 1$	AT4
			FDI attack against sensor P	YES	$\beta = 1$	AT5
			DoS attack against sensor P	YES	$\beta = 1$	AT6
			FDI attack against actuator V2	YES	$\beta = 1$	AT7
			DoS attack against actuator V2	YES	$\beta = 1$	AT8
			Setpoint manipulation of overpressure threshold of PLC2	YES	$\beta = 1$	AT9
External physical attackers	Physical protection systems	Induce shell rupture and the release of hazardous chemicals	Physical attack with simple or major tools	YES	/	AT10

[5] β coefficient is determined using the approach presented in Appendix I.

For C2P attacks against PLC1, it is considered a dangerous overheating scenario when the temperature inside the reactor overpasses 450 K. The physical impacts induced by different C2P attack modes were analyzed and the β coefficient for each attack mode was determined using the approach presented in Appendix I. Some selected results of the attack impact modeling are presented in Figure 3.9.



(a) without attacks

(b) setpoint manipulation

(c) FDI attack against sensor T (Min attack)

(d) FDI attack against actuator V3 (Min attack)

(e) DoS attack against sensor T (last received measurement is below the setpoint value)

(f) DoS attack against actuator V3

Figure 3.9. Temperature inside the reactor under different C2P attack modes against PLC1 (attacks start from 100 s).

It was found that setpoint manipulations can induce overheating scenarios in a short time no matter the influence of process noise and observation noise, as shown in Figure 3.9 (b). Therefore, a successful setpoint manipulation has an extremely high likelihood of inducing a dangerous scenario (β=1). Similarly, FDI attacks (Min or

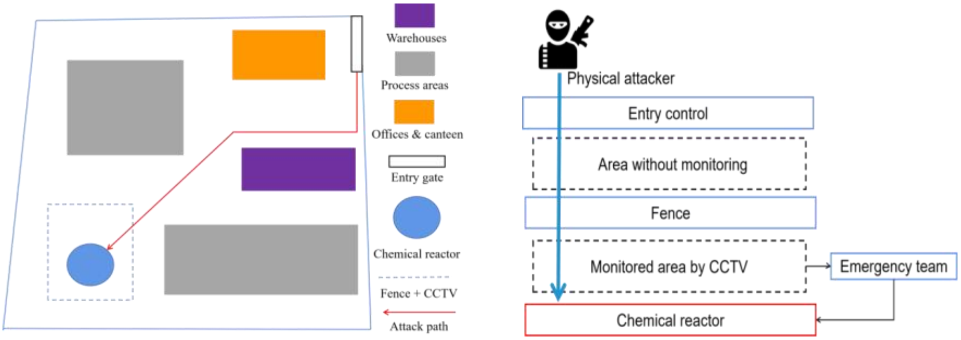
Max Attacks) can induce overheating scenarios no matter the process noise and observation noise using the least attack time, as shown in Figure 3.9 (c) and (d). This finding is consistent with the results from Huang et al. (2009). By contrast, the DoS attacks cannot always induce overheating scenarios. Because the last received signal will be used under DoS attacks, the attack impact depends on both the attack duration and the last received signal before the attack. In this case study, both the measurement signal of sensor P and the control signal of actuator V3 follow normal distributions considering process noises and observation noises. It was found that overheating can be induced by DoS attacks against sensor T only if the last received temperature signal is below the temperature setpoint (430.9 *K*), as shown in Figure 3.9 (e). The temperature measurement signal fluctuates around the setpoint and has a 50% probability of being below the setpoint. Under the assumption that the attacker is able to perform an attack with an enough long duration, a successful DoS attack against sensor T has a 50% probability of inducing a dangerous scenario (β=0.5). Regarding DoS attacks against actuator V3, it was found that overheating scenarios cannot be induced when the control signal of V3 is fluctuating within its operating range, as shown in Figure 3.9 (f). Therefore, β coefficient for DoS attacks against actuator V3 is 0 and this attack mode is not considered a dangerous attack mode.

For C2P attacks against PLC2, all the attack modes are considered dangerous (β=1 for all attack modes against PLC2) because they are capable of making the ESD system fail to perform its functionality on demand no matter the process noise and observation noise. For instance, FDI attacks can inject malicious measurement data or control data to prevent the ESD system from being activated successfully on demand. DoS attacks can block the data flow and force the ESD system into an unactivated state. Setpoint manipulation of the overpressure threshold is capable of hindering the activation of shutdown actions even if the pressure already overpassed the pre-defined threshold. The determined β coefficient values for each attack mode are also given in Table 3.5.

3.4.2.3 Vulnerability analysis results

In this study, the Adversary Sequence Diagrams and Path Analysis was employed to identify credible attack paths for physical attacks considering the deployment of PPSs. The obtained site-specific adversary sequence diagram is shown in Figure 3.10. For simplicity, an event tree analysis was used to perform the vulnerability assessment of PPSs, as shown in Figure 3.11. The PFDs (probability of failure on demand) of the

PPSs are determined by using the method and benchmark data introduced by Moreno et al. (2022).



(a) Layout of the chemical plant. (b) Adversary sequence diagram.

Figure 3.10. Site-specific adversary sequence diagram considering external physical attacks.

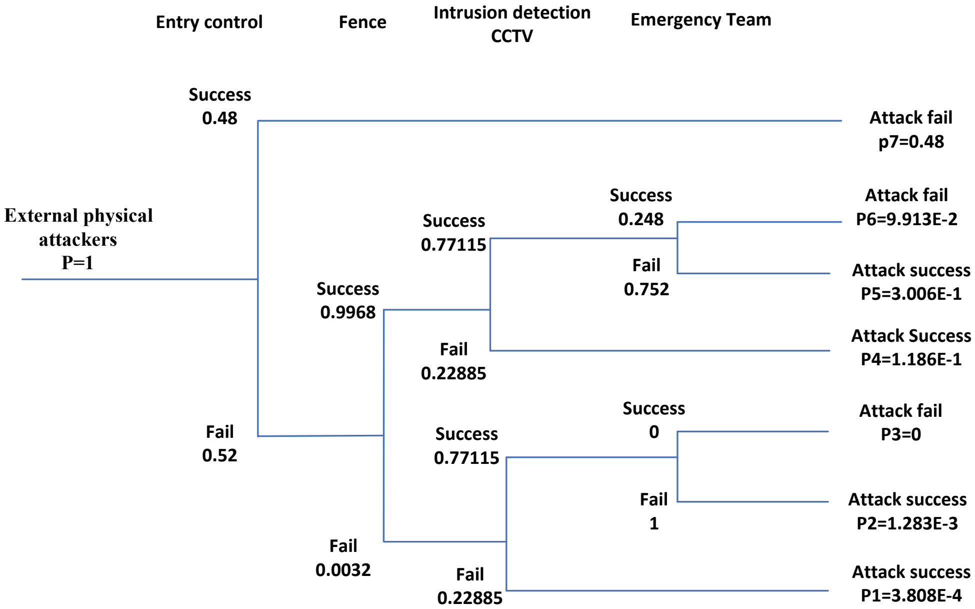


Figure 3.11. Event tree analysis of an external physical attack.

Four types of PPSs are considered in this case study, as presented in Table 3.6. AIT (adversary intrusion time) and ERT (emergency response time) are employed to evaluate the effectiveness of the emergency team. It is assumed that when the fence works for delaying the attackers effectively, AIT > ERT. Otherwise, AIT < ERT and the emergency team cannot prevent the attack effectively. According to Figure 3.11, the conditional probability of a successful physical attack given an attack attempt is calculated as: $P_s = P_1 + P_2 + P_4 + P_5 = 0.42$.

Table 3.6 Fail probabilities and success probabilities of PPSs, adapted from (Moreno et al., 2022).

PPS (physical protection system)	PFD (probability of failure on demand)	Effectiveness (η)	Calculation formulas
Entry control	0.40	0.80	$\mathrm{P_{fail}} = PFD + (1 - \eta) \times (1 - PFD)$ $P_{success} = (1 - PFD) \times \eta$
Fence	0.00	0.9968	
Closed Circuit TeleVision (CCTV)	0.205	0.97	
Emergency Team	0.752	1 if AIT > ERT; 0 if AIT < ERT	$P_{success} = (1 - PFD) \times \eta$ $P_{fail} = 1 - (1 - PFD) \times \eta$

Regarding C2P attacks, the information provided by the CPS master diagram helps to identify each attack step of the attacker. An ICS-specific vulnerability dataset (Thomas & Chothia, 2020) was used to identify the known vulnerabilities that may be exploited by attackers at each attack step. Then, an attack graph was constructed to demonstrate the attack paths, as shown in Figure 3.12. Local TTC (time-to-compromise) of each attack step is estimated using the approach presented in Appendix II, and the results are given in Table 3.7. The global TTC of each attack path was calculated by summing up the local TTCs of the attack steps along the attack path. For the attack modes with multiple attack paths, the shortest global TTC of each mode is used for the security vulnerability quantification. Because all the dangerous C2P attack modes in the case study are assumed to be executed by remote attackers through network intrusions, network detection and response (NDR) is the main technology used to detect C2P attacks through the monitoring of network traffic (Pérez et al., 2021). A reference value (14 days) from Semertzis et al. (2022) is used as the MTTD value for all C2P attack modes. In practice, the MTTD value may be determined based on incident data analysis regarding specific intrusion types. The calculated shortest global TTCs and the conditional probabilities of successful execution of each attack mode are given in Table 3.8.

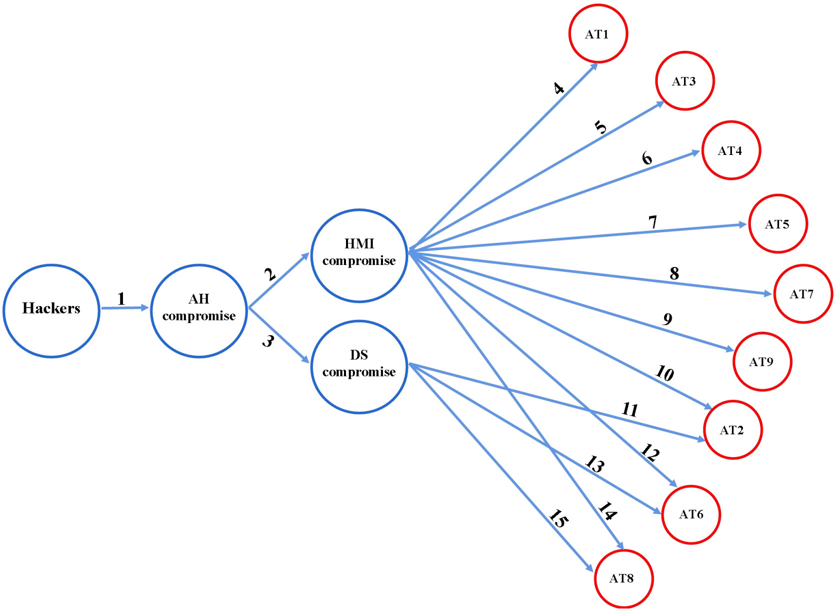


Figure 3.12. Attack/compromise graph considering C2P attacks against the investigated ICS with reference to Zhang et al., (2017) (explanations of AT1~AT9 can be found in Table 3.5).

Table 3.7 Time-to-compromise of each attack step.

Attack step number	**Vulnerabilities (cve_id[6])**	**Average base score of CVSS v2.0**	**Average exploitability score of CVSS v3.0**	**TTC (days)**	**Attack step number**	**Vulnerabilities (cve_id)**	**Average base score of CVSS v2.0**	**Average exploitability score of CVSS v3.0**	**TTC (days)**
1	CVE-2015-7871; CVE-2017-2683	9.00	3.35	5.28	2	CVE-2017-13997	9.80	3.90	5.94
3	CVE-2018-13799	9.10	3.90	5.94	4	no	/	/	40.01
5	no	/	/	40.01	6	CVE-2018-5459	9.80	3.90	5.64
7	no	/	/	40.01	8	no	/	/	40.01
9	CVE-2018-5459	9.80	3.90	5.64	10	CVE-2016-2200	7.50	3.90	5.99
11	CVE-2016-2200	7.50	3.90	5.99	12	CVE-2016-2200	7.50	3.90	5.99
13	CVE-2016-2200	7.50	3.90	5.99	14	CVE-2016-2200	7.50	3.90	5.99
15	CVE-2016-2200	7.50	3.90	5.99	/	/	/	/	/

[6] cve_id: A cve_id uniquely identifies one vulnerability from the Common Vulnerabilities and Exposures (CVE) database (NVD, 2023).

Table 3.8 Estimation of shortest global time-to-compromise for each attack mode.

Attack mode	The shortest global TTC (days)	The conditional probability of successful attacks[7]	Attack mode	The shortest global TTC (days)	The conditional probability of successful attacks
1	51.23	0.21	2	17.21	0.22
3	51.23	0.21	4	16.86	0.45
5	51.23	0.21	6	17.21	0.45
7	51.23	0.21	8	17.21	0.45
9	16.86	0.45	/	/	/

3.4.2.4 BN model development

For each attack mode in Table 3.5, a simplified attack tree was developed and attached to appropriate places in the bow-tie diagram. For instance, the attack tree for attack mode 1 (AT1 in Table 3.5) is composed of two basic events (the frequency of attack attempts to execute AT1 and the conditional probability of the corresponding vulnerability being exploited successfully) and one top event (AT1 is executed successfully). In this case study, all the attack trees were integrated into the left-hand-side of the bow-tie diagram (fault tree), as shown in Figure 3.13. Then, a BN topology was developed based on the attack-tree-bow-tie diagram, as shown in Figure 3.14. All the BN nodes have two states (happening and not happening), except the consequences node, which is composed of five states (no consequence, fireball, explosion, cloud fire, and toxic dispersion). Table 3.9 gives prior probabilities of the root nodes. The abbreviations of other BN nodes are explained in Table 3.10.

[7] The conditional probability of successful attacks for each attack mode is calculated using Eq. 3.4.

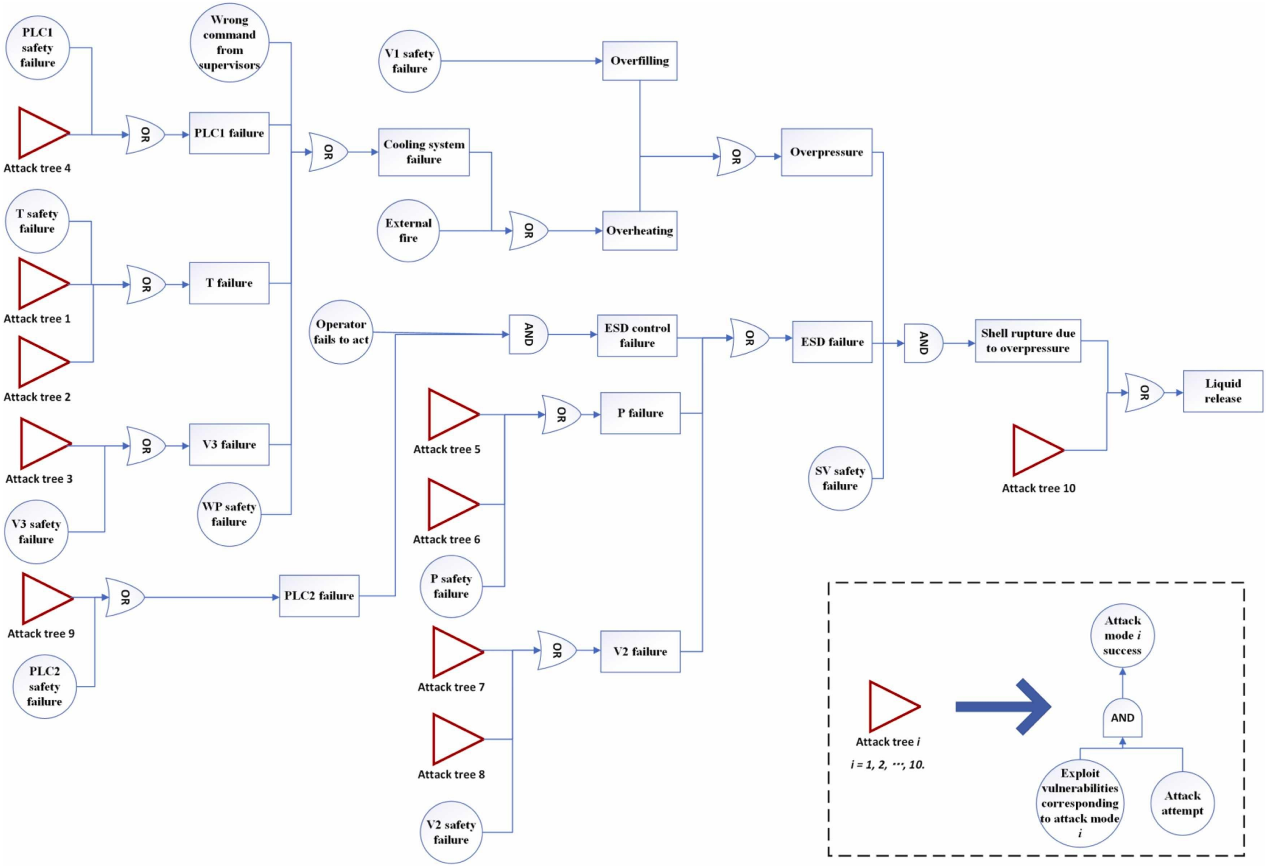


Figure 3.13. The integration of the fault tree and attack trees.

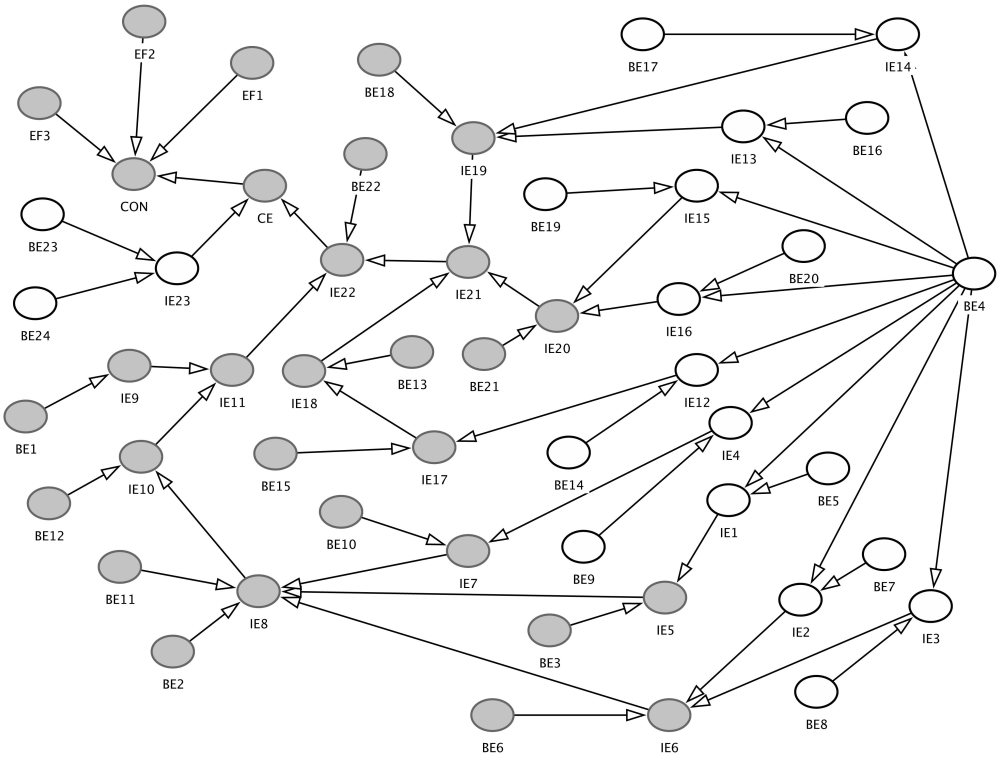


Figure 3.14. Topology of the BN model (nodes without fillings are derived from attack trees and nodes with grey fillings are derived from the bow-tie diagram).

Table 3.9 Prior probabilities of the root nodes in the BN model.

Symbols	Root nodes	Prior probabilities	Sources	Symbols	Root nodes	Prior probabilities	Sources
BE1	V1 safety failure	4.00E-02	(Taylor, 2010)	BE2	Wrong command from supervisors	1.00E-02	(Andersen et al., 2004)
BE3	PLC1 safety failure	4.38E-02	(Hauge & Onshus, 2010)	BE4	C2P attack attempts	7.85E-01	Estimated with Table 3.3.
BE5	Exploit vulnerabilities corresponding to AT4	4.50E-01	Estimated by vulnerability assessment (Table 3.8)	BE6	T safety failure	2.13E-02	(Hauge & Onshus, 2010)
BE7	Exploit vulnerabilities corresponding to AT1	2.10E-01	Estimated by vulnerability assessment (Table 3.8)	BE8	Exploit vulnerabilities corresponding to AT2	2.20E-01	Estimated by vulnerability assessment (Table 3.8)
BE9	Exploit vulnerabilities corresponding to AT3	2.10E-01	Estimated by vulnerability assessment (Table 3.8)	BE10	V3 safety failure	4.00E-02	(Taylor, 2010)
BE11	WP safety failure	3.125E-02	(OREDA, 2002)	BE12	External fire	5.52E-02	(Debray et al., 2004)
BE13	Operator fails to shutdown	1.00E-02	(Andersen et al., 2004)	BE14	Exploit vulnerabilities corresponding to AT9	4.50E-01	Estimated by vulnerability assessment (Table 3.8)
BE15	PLC2 safety failure	2.19E-03	(Hauge & Onshus, 2010)	BE16	Exploit vulnerabilities corresponding to AT5	2.10E-01	Estimated by vulnerability assessment (Table 3.8)
BE17	Exploit vulnerabilities corresponding to AT6	4.50E-01	Estimated by vulnerability assessment (Table 3.8)	BE18	P safety failure	3.29E-04	(Hauge & Onshus, 2010)
BE19	Exploit vulnerabilities corresponding to AT7	2.10E-01	Estimated by vulnerability assessment (Table 3.8)	BE20	Exploit vulnerabilities corresponding to AT8	4.50E-01	Estimated by vulnerability assessment (Table 3.8)

BE21	V2 safety failure	3.50E-06	(Hauge & Onshus, 2010)	BE22	SV safety failure	2.40E-03	(Hauge & Onshus, 2010)
BE23	External physical attacks	3.30E-02	Estimated with Table 3.2 (1/Λ, Λ=30 years).	BE24	Exploit vulnerabilities of PPSs	4.20E-01	Estimated by vulnerability assessment (Table 3.6)
EF1	Immediate ignition	7.00E-01	(Vílchez et al., 2011)	EF2	Fireball (BLEVE)	7.00E-01	(Vílchez et al., 2011)
EF3	Flame front acceleration	4.00E-01	(Vílchez et al., 2011)	/	/	/	/

Table 3.10 Explanations of the leaf node and intermediate nodes.

Symbols	**Node names**	**Symbols**	**Node names**	**Symbols**	**Node names**	**Symbols**	**Node names**
IE1	AT4 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	AT9 success
IE13	AT5 success	IE14	AT6 success	IE15	AT7 success	IE16	AT8 success
IE17	PLC2 failure	IE18	ESD control failure	IE19	P failure	IE20	V2 failure
IE21	ESD failure	IE22	Shell rupture due to overpressure	IE23	AT10 success	CE	Liquid release
CON	Consequences	/	/	/	/	/	/

3.4.3 Probabilistic risk assessment results

A Bayes net toolbox developed based on MATLAB (Murphy, 2001) was used to perform the risk assessment by using the prior probabilities presented in Table 3.9. Additionally, the risks purely caused by safety causes are assessed by configuring the prior probabilities of security attacks into zeros. We used a risk matrix to visualize the major adverse risks induced by safety causes and the combination of safety causes and security threats, as shown in Figure 3.15. It is observed that the risks of safety-associated major adverse events (fireball, explosion, cloud fire, and toxic dispersion) are all within the green region, which means the safety risks are acceptable. However, the risks estimated by integrating safety-associated scenarios and security-related attack scenarios become unacceptable and are much higher than the pure-safety risks. The estimated annual frequencies of the occurrence of major adverse events (fireball, explosion, cloud fire, and toxic dispersion) considering both safety hazards and security threats are between 10^{-3} to 10^{-2}. This result indicates that it is necessary to integrate security attack scenarios into the safety risk assessment of ICPSs, otherwise, major adverse risks may be underestimated.

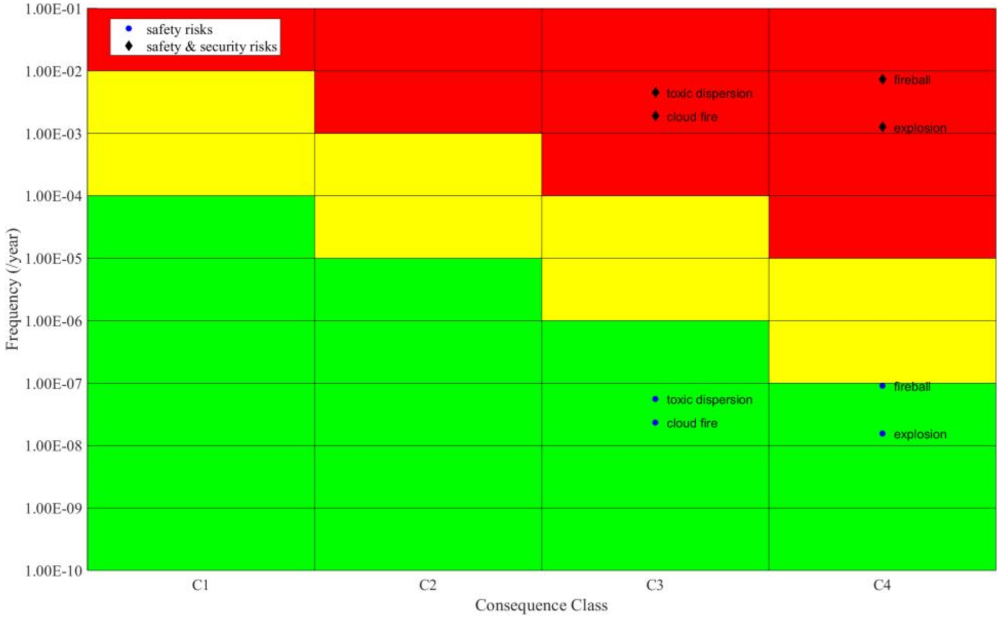
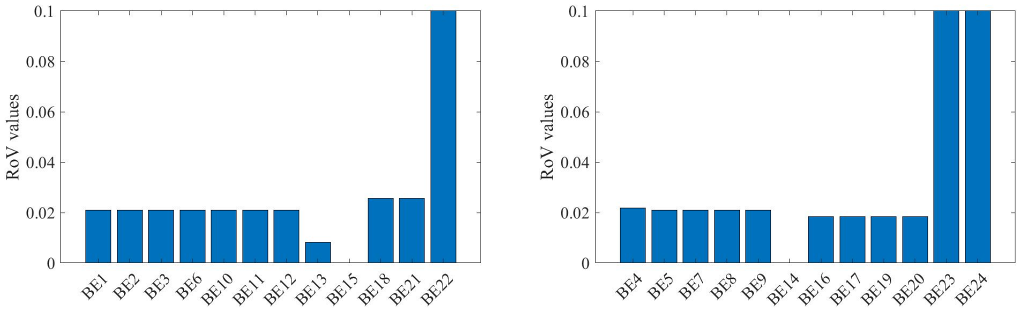


Figure 3.15. A risk matrix presenting major adverse risks induced by safety causes and the integration of safety causes and security attacks.

3.5 Sensitivity analysis of root nodes

To identify critical causes leading to major adverse events, this section uses the RoV (ratio of variance) measure to perform a sensitivity analysis of the basic events (root nodes). Concerning the occurrence of major adverse events (fireball, explosion, cloud fire, and toxic dispersion), the RoV value of each root node is calculated. The calculated RoV values of the safety-associated basic events and security-associated basic events are presented in Figure 3.16 (a) and (b) respectively. Figure 3.16 (a) shows that BE22 (safety relief valve safety failure) has a sensitivity significantly ahead of others, followed by BE18 (pressure sensor safety failure) and BE21 (shutdown valve safety failure). Those three events are related to the failure of safety barriers (emergency shutdown system and safety relief valve), which means safety barriers play important roles in preventing major adverse events. Particularly, as a passive safety barrier, the safety relief valve (SV) is the most critical equipment. BE15 (PLC2 safety failure) has the smallest sensitivity, and it is followed by BE13 (operator fails to shutdown). Because BE15 and BE13 take the same responsibility to activate the shutdown valve (V2) based on the received measurement signals from pressure sensor (P), they may reduce each other's criticality/sensitivity to certain extents. The remaining safety-associated basic events (BE1, BE2, BE3, BE6, BE10, BE11, and BE12) have nearly the same sensitivity, which means they have similar importance.

Regarding security-associated basic events, BE23 (External physical attacks) and BE24 (Exploit vulnerabilities of PPSs) have dominant sensitivities. It means that physical attack is the most threatening attack mode and the vulnerability subject to the physical attack is the most critical security vulnerability. One reason for this result is that physical attacks can overpass the protection of some safety barriers (ESD system and safety relief valve in this study) and induce the loss of contaminant by damaging equipment directly. Regarding C2P attacks, BE14 (Exploit vulnerabilities corresponding to AT9) has the smallest sensitivity, while BE4, BE5, BE7, BE8, BE9, BE16, BE17, BE19, and BE20 have almost the same sensitivity. This result demonstrates that the vulnerabilities subject to each C2P attack mode have similar sensitivities, except the vulnerabilities subject to AT9 (setpoint manipulation of the overpressure threshold of PLC2). The reason may be that manual emergency shutdown in case of PLC2 failures reduces the danger of cyberattacks against PLC2. The sensitivity analysis results indicate the importance of protecting digitalized safety barriers from cyberattacks and the necessity of deploying physical/passive barriers and human barriers to prevent major adverse events.



(a) RoV values of safety-associated basic events. (b) RoV values of security-associated basic events.

Figure 3.16. Sensitivity analysis of BN root nodes.

3.6 Conclusions

This study proposes an approach for integrated safety and security risk analysis of industrial cyber-physical systems (ICPSs) with respect to major adverse event scenarios in chemical plants. According to the risk assessment results, major safety-related accident risks may increase to a large extent with the involvement of potential security attack scenarios. The assessment of only safety-associated scenarios or security attack scenarios may lead to a risk underestimation. A sensitivity analysis was performed to identify critical basic events. The results indicate that the vulnerabilities of ICPSs to cyberattacks should be given enough attention, particularly considering the possible C2P attacks on digitalized safety barriers. Moreover, it is found that physical/passive barriers and human barriers play an important role in preventing the happening of disastrous consequences. Because physical/passive barriers and human barriers are not subjected to cyberattacks, they can be considered critical measures to prevent the occurrence of C2P attack-induced disastrous scenarios.

Due to the lack of data, attack likelihood estimation regarding both physical attacks and C2P attacks is somewhat dubious in practice. Although some reference data in the chemical process industries or other similar sectors is helpful, the estimation of attack likelihood may be still subjective. Additionally, some conservative assumptions were made in the vulnerability analysis, for instance, the shortest global TTC is used with the ignorance of attackers' intrusion path selection and attackers are assumed with high knowledge levels. The uncertainties associated with the rough or conservative assumptions in threat analysis and vulnerability analysis need to be evaluated and properly handled in future studies.

Chapter 4 Integrated safety and security barrier management regarding emerging cyber-physical risks under uncertainties

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

This chapter is drafted with modifications based on the following publication:

- Yuan, S., Reniers, G., Yang, M. (2024). Integrated management of safety and security barriers in chemical plants to cope with emerging cyber-physical attack risks under uncertainties. Reliability Engineering & System Safety, (Accepted).

4.1 Introduction

With the automation and digitization of chemical process facilities, industrial cyber-physical systems (ICPSs), also called industrial control systems (ICSs), are widely applied to chemical plants to integrate the operation of the physical process with computing and communication infrastructures (Derler et al., 2011; Ji et al., 2016). Due to the vulnerabilities of ICSs to cyberattacks, it is evidenced that dangerous failure of industrial facilities may be induced by either safety causes or cyberattacks (Flaus, 2019). The latter is known as cyber-physical (C2P) attacks. For instance, the malware "Stuxnet" is regarded as the world's first publically known digital weapon, which can target programmable logic controllers (PLCs) and induce physical damage (Kushner, 2013).

Some attempts have already been made by researchers to quantify the vulnerabilities or security risks of ICSs regarding C2P attacks. Remarkably, a time-to-compromise (TTC) based approach was proposed to quantify the vulnerability of ICSs to cyberattacks (McQueen et al., 2006a; 2006b). Then, the TTC approach has been adapted for quantitative risk/reliability assessment of industrial cyber-physical systems regarding C2P attacks (Zhang et al., 2015; Semertzis et al., 2022). By contrast, some researchers used the exploitability subscores from the CVSS (Common Vulnerability Scoring System) approach to quantify the vulnerability of ICSs regarding cyberattacks (Poolsappasit et al., 2011; Huang et al., 2018). More recently, the TTC approach was modified and combined with CVSS scores for the vulnerability assessment of ICSs (Ling & Ekstedt, 2022), in which the exploitabilities of known vulnerabilities are also considered. Additionally, Markov decision processes (MDP) (Chen et al., 2016), game-theoretic methods (Orojloo & Azgomi, 2017) and the combination of a stochastic game with Markov processes (Lalropuia & Gupta, 2019) were employed to model cyber-physical attacks and support security assessment of ICSs.

The consequence/impact analysis of C2P attacks against ICSs has also been investigated by previous studies. For instance, mean-time-to-failure (MTTF) (Lalropuia & Gupta, 2019) and mean-time-to-shutdown (MTTSD) (Orojloo & Azgomi 2017; Huang et al., 2018) were used to describe the physical impact caused by C2P attacks on industrial systems. The potential economic consequences associated with production loss, operating cost, and loss of incidents caused by C2P attacks were also investigated by previous studies (Huang et al., 2009; Li et al., 2017).

However, the uncertainties associated with the attackers' knowledge levels were usually ignored in previous studies. Because security attacks are more difficult to predict compared to accidental events and the complexity of the ICSs makes the security analysis harder, the uncertainties involved in the integrated safety and security risks may be significant. The importance of considering probability distributions rather than using fix-point probability values or expected values of probability distributions in risk modeling and the decision-making process was highly emphasized by researchers (Bier & Lin, 2013). Appropriate treatment of probability distributions and uncertainty propagation in risk assessment helps decision-makers understand the strength of guidance for decision-making. Additionally, the effects of safety barriers (such as safety instrumented systems and manual shutdown) on C2P attack protection and the corresponding damage mitigation were rarely considered in previous studies. A systematic approach is needed to manage safety and security barriers based on the assessment of the integrated safety and security risks under significant uncertainties.

Therefore, targeting the gaps in uncertainty treatment and barrier management with respect to C2P attack scenarios, this study develops a systematic approach for integrated safety and security barrier management and achieving risk-based decision-making on barrier optimization by integrating C2P attack risks with conventional safety risks. The remainder of this chapter is organized as follows. Section 4.2 elaborates on the overall structure and operating procedures of each step of the proposed approach. An illustrative case study is employed to demonstrate the application of the proposed approach to cost-effective barrier optimization in section 4.3. Discussions are presented in section 4.4 before conclusions are given in section 4.5.

4.2 Methodology

4.2.1 Overview of the proposed approach

This study proposes a systematic approach for integrated safety and security barrier management based on integrated risks considering uncertainties. An overview of the proposed methodology is given in Figure 4.1. The proposed methodology consists of three parts, which aim to address accident scenario integration considering both safety causes and C2P attacks, uncertainty handling and risk assessment, and risk-based decision-making, respectively. A detailed illustration of each part of the methodology is presented in the following sub-sections.

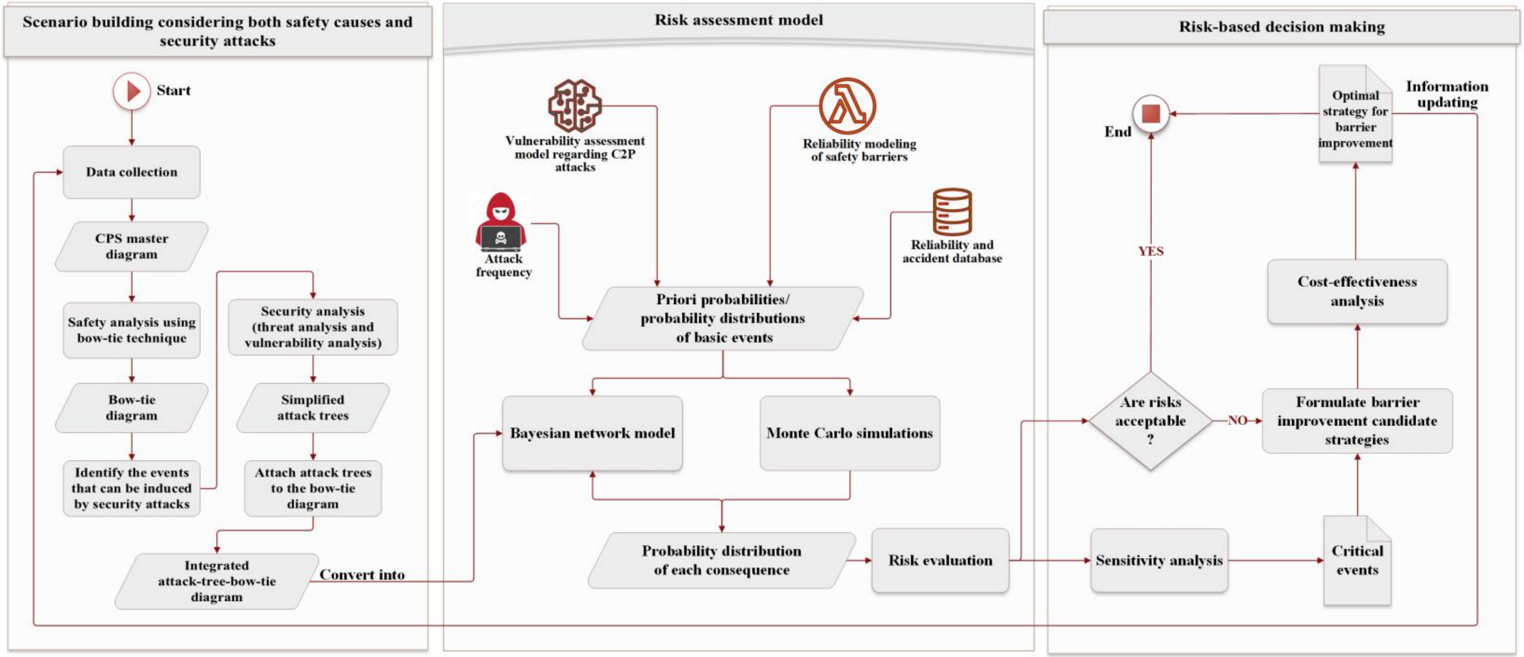

Scenario building considering both safety causes and security attacks
Start
Data collection
CPS master diagram
Safety analysis using bow-tie technique
Bow-tie diagram
Identify the events that can be induced by security attacks
Security analysis (threat analysis and vulnerability analysis)
Simplified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

Figure 4.1. Overview of the proposed methodology.

4.2.2 Scenario identification and integration

Bow-tie diagrams (CCPS/EI, 2018) and attack trees (Gribaudo et al., 2015) are widely used for accident scenario building in terms of safety and security, respectively. ICSs are usually complex engineered systems with the integration of IT (information technology) and OT (operational technology) infrastructures. A tool named CPS master diagram (Guzman et al., 2021), which is capable of representing ICSs in a multi-layered manner with the demonstration of energy flows and information flows, can serve as a basis for accident scenario building. Chapter 3 (section 3.3) presents a systematic approach for accident scenario building for ICSs, in which the CPS master diagrams, bow-tie analysis, and attack trees are combined to identify and integrate safety-associated and security-associated adverse scenarios (Yuan et al., 2024). The same procedures are used here for adverse scenario identification and integration. Basically, the integration can be done by checking each event in the bow-tie diagram by asking if this event can also be induced by security attacks. If the answer is yes, the possible security-associated scenarios should be analyzed, and the developed attack trees need to be attached to the appropriate places of the bow-tie diagram. The workflow of this approach is presented in the left block of Figure 4.1. An example of the CPS master diagram can be found in Figure 4.6(b) and the integrated attack-tree-bow-tie diagram can be found in Figure 4.7. It is worth mentioning that the role of OT safety barriers in preventing the success of C2P attacks is also considered during scenario building. These OT barriers are identified in the bow-tie diagram and are expected to activate in response to specific adverse events, whether triggered by safety-related causes or C2P attacks.

4.2.3 An integrated risk assessment model

A risk assessment model is developed based on a Bayesian network (BN) and with the help of a vulnerability assessment model for C2P attacks, reliability modeling of safety barriers, and Monte Carlo simulations.

4.2.3.1 Bayesian network and uncertainty propagation handling

1) Convert the attack-tree-bow-tie diagram into a BN model

The obtained attack-tree-bow-tie diagram incorporates safety-associated and security-associated scenarios. However, a quantitative risk analysis is still challenging.

Due to the complexity of the identified scenarios, multiple occurring events (MOE) are usually unavoidable. While those events are not allowed in fault/attack tree analysis when performing probability calculations (Ericson, 2005). To solve this problem, minimal cut sets must be determined and the tree should be translated into an equivalent set of Boolean equations for probability calculations. Alternatively, the fault/attack trees can be converted into BN models for probability modeling, which is capable of handling dependent basic events (Khakzad et al., 2011). BNs are probabilistic graphical models and are widely used for risk assessments due to their advantage of representing random variables with their interdependencies (Tong et al., 2018). In a BN, a joint probability distribution $P(X)$ of variables $X = \{X_1, ..., X_n\}$ is presented as follows (Jensen & Nielsen, 2007):

$$P(X) = \prod_{i=1}^{n} P(X_i | Pa(X_i)) \tag{4.1}$$

$$P(X|E) = \frac{P(E|X) \cdot P(X)}{P(E)} = \frac{P(E, X)}{\sum_X P(E, X)} \tag{4.2}$$

where $Pa(X_i)$ is the parent node set of X_i. When evidence E becomes available, the posterior probabilities $P(X|E)$ can be derived based on Bayes theorem using Eq (4.2). The topology and conditional probability tables (CPTs) of the BN can be derived from the integrated attack-tree-bow-tie diagrams by following dedicated mapping algorithms. Previous studies have already given guidance on transforming attack trees (Gribaudo et al., 2015) and bow-tie diagrams (Khakzad et al., 2013) into BNs. We avoid repeating illustrations here. After the converting process, the root nodes of the BN model can be divided into safety-related and security-related nodes, which correspond to the basic events in the bow-tie diagram and in the attack trees, respectively. Each state of the leaf node presents each outcome event (consequence) in the attack-tree-bow-tie diagram.

2) Determination of priori probabilities/probability distributions

After the topology and CPTs of the BN model are determined, prior probabilities of the root nodes are required to perform the risk assessment. This study uses probability distributions or fixed-point probability values for the root nodes. Four ways are used to determine prior probabilities or probability distributions for different types of root nodes, as illustrated below.

- ■ Regarding safety-related initiating events, for instance, critical failure of a technical component, human failure, external fires, etc., reliability databases

(Hauge & Onshus, 2010), human reliability data (Kirwan, 2017), accident databases (Debray et al., 2004), or data available in the literature may be used to derive the occurrence probabilities. When a probability distribution of the occurrence of an event is available, it may be used instead of a fixed-point probability value.

- Regarding safety barriers, the probability of failure on demand (PFD) is used to quantify the reliability of safety barriers because they usually follow a low-demand mode (IEC, 2010; Yuan et al., 2022b). For human components, human reliability data may be used to obtain the PFDs of human actions. The approach for reliability modeling of technical components of safety barriers considering barrier maintenance strategies is illustrated in section 4.2.3.2.
- Regarding security-associated basic events, attack likelihood/frequency and the conditional probabilities of successful attacks given attack attempts are needed. Attack likelihood of C2P attacks may be estimated according to incident statistics of the investigated chemical plants or comparable companies with the help of expert judgment. In case of lacking incident data, the estimation may mainly be performed based on expert judgment, which is one significant source of uncertainty.
- A vulnerability assessment is performed to estimate the conditional probabilities of successful attacks given attack attempts. This is achieved by assessing the vulnerabilities of both IT systems and OT systems and considering the uncertainties associated with attackers' knowledge levels. Details about the vulnerability assessment are elaborated on in section 4.2.3.3.

3) Uncertainty propagation handling

In this study, Monte Carlo simulations are used to handle uncertainty propagation in the risk assessment when probability distributions are used for root nodes. This is achieved by sampling point values from the probability distributions as inputs while accumulating the inferred probability of each leaf node state. Finally, probability distributions for each state (representing each consequence) of the leaf node can be obtained. Regarding the consequence assessment, a severity class for typical dangerous phenomena in chemical plants suggested by the ARAMIS project is used (Andersen et al., 2004). A risk matrix is used to visualize and evaluate risk profiles by mapping the expected values and ranges of the probability distributions of the potential consequences into the risk matrix. A flowchart is presented in Figure. 4.2 to demonstrate the main procedures in the risk assessment and uncertainty propagation handling.

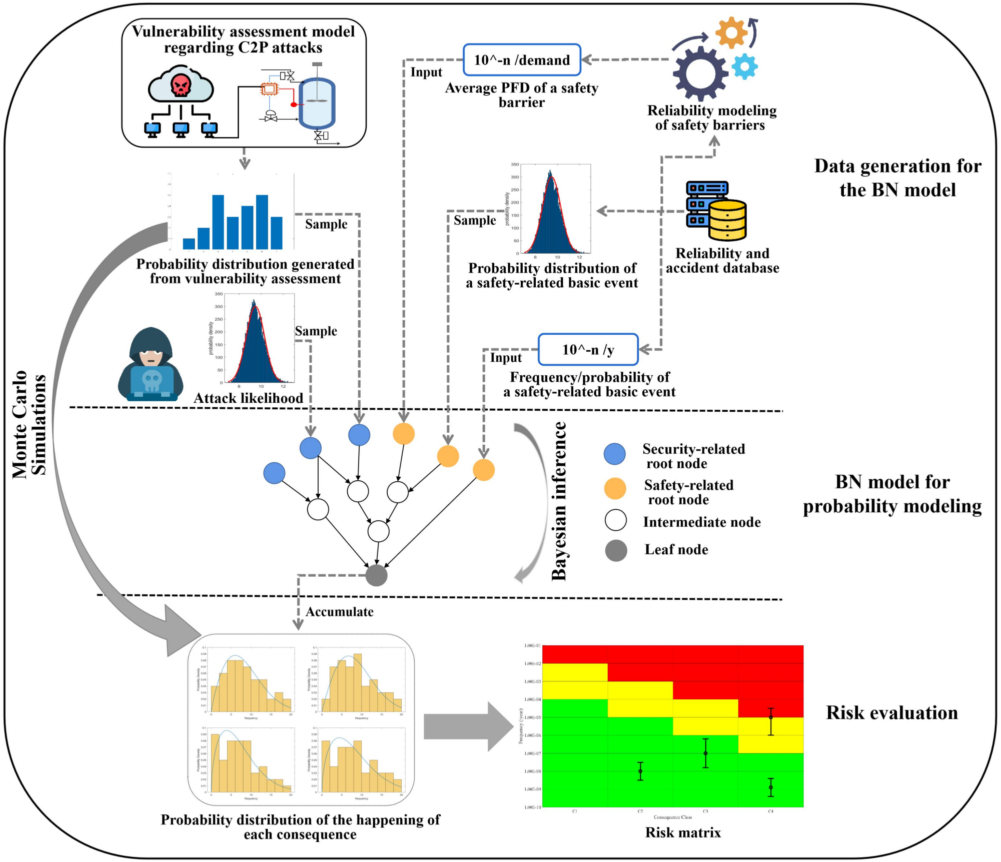

Vulnerability assessment model regarding C2P attacks
Sample
Probability distribution generated from vulnerability assessment
Sample
Attack likelihood
Monte Carlo Simulations
Input
10^-n /demand
Average PFD of a safety barrier
Reliability modeling of safety barriers
Data generation for the BN model
Sample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

Figure 4.2. Structure of the integrated risk assessment model.

4.2.3.2 PFD calculation of safety barriers under preventive maintenance

Safety barriers play important roles in protecting industrial systems from disastrous damage in case of dangerous failures/deviations. For a complex safety barrier system, fault tree analysis may be implemented to calculate the PFD of the whole barrier system. In practice, corrective maintenance and preventive maintenance are widely used for technical facilities in chemical plants. For the sack of safety, preventive maintenance is usually performed at specific intervals (e.g. once per year) for safety barriers, which is also known as periodic maintenance. Because some safety barriers are allocated under harsh environments, degradation inevitably happens and impacts the reliability of those safety barriers (Zhang et al., 2020). For the components subject to degradation, for instance, emergency shutdown valves, a Gamma degradation process is adapted to simulate the continuous aging degradation, as follows (Zhang et al., 2020):

$$X(t) \sim \Gamma(\alpha t,\beta) = f_{X(t)}(x) = \frac{\beta^{\alpha t}}{\Gamma(\alpha t)} x^{\alpha t-1} e^{-\beta x},\ \alpha,\beta > 0 \tag{4.3}$$

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

$$F_{X(t)}(x) = \int_0^x f_{X(t)}(x)dx \tag{4.4}$$

$$A(t) = Pr(X(t) < L) = F_{X(t)}(L) \tag{4.5}$$

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

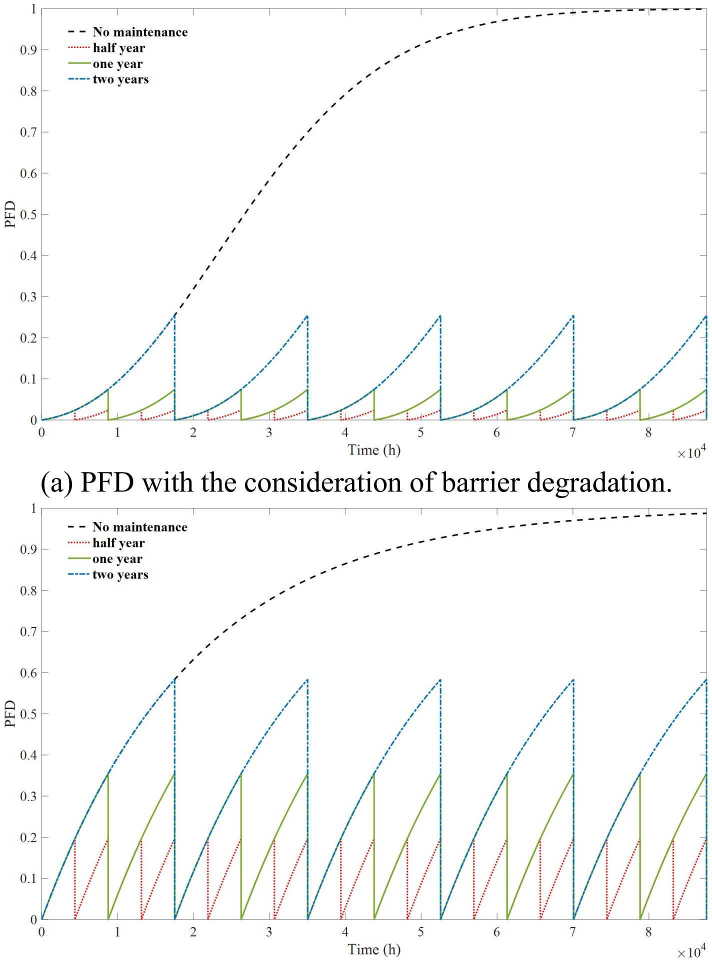
$$PFD_d(t) = 1 - A(t\%T) = 1 - F_{X(t\%T)}(L),\quad nT \le t < (n+1)T \tag{4.6}$$

where $t\%T$ means the remainder when dividing t by T. n is an integer from 0 to positive infinity ($n = 0, 1, 2, \dots, +\infty$). The calculated PFDs considering different maintenance intervals using Eq. (4.6) are compared in Figure 4.3 (a).

By contrast, for the components (e.g. programmable logic solvers) that are not obviously subjected to degradation, their PFDs are assumed to follow exponential distributions with constant failure rates (IEC, 2016; Schmitz et al., 2021). With the assumption that perfect barrier maintenance with a time interval, T, is implemented and ignoring the time spent on maintenance, PFD can be calculated as below.

$$PFD_{nd}(t) = 1 - e^{-\lambda*(t\%T)}, \quad nT \leq t < (n+1)T \tag{4.7}$$

where $PFD_{nd}(t)$ is the PFD without the consideration of barrier degradation. λ is failure rate. By using Eq. (4.7), the calculated PFDs considering different maintenance intervals are shown in Figure 4.3 (b). Finally, the average values of the PFDs over time are used as prior probabilities for the BN model.



(a) PFD with the consideration of barrier degradation.

(b) PFD without the consideration of barrier degradation.

Figure 4.3. PFD calculation of technical safety barrier components with different preventive maintenance strategies.

4.2.3.3 Vulnerability assessment of ICS to C2P attacks considering uncertainties

A C2P attack process can be divided into two phases, intrusion into the IT systems and manipulation of the OT systems. A vulnerability assessment model is developed to assess both two phases. A time-to-compromise (TTC) based approach was proposed to quantify the vulnerability of IT infrastructures to cyberattacks (McQueen et al., 2006a; 2006b). Ling & Ekstedt (2022) modified this TTC estimation approach and combined it with an ICS-specific vulnerability dataset (Thomas & Chothia, 2020) considering attackers' skill levels, the number of known vulnerabilities on a component, and the exploitability of the known vulnerabilities. This study combines the approach developed by Ling & Ekstedt (2022) with compromise graphs to estimate the global TTC of each attack path. A detailed explanation of the TTC estimation approach can be found in Appendix II.

In previous studies, two approaches were used to address the attack path selection issue of attackers. McQueen et al. (2006b) and Semertzis et al. (2022) assumed that all possible attack paths are executed by attackers in parallel, which is a conservative assumption. By contrast, Zhang et al. (2015) related attack path selection to the exploitability of each attack path and it is assumed that a more exploitable attack path is more likely to be selected by attackers. However, the exploitability of unknown vulnerabilities and the uncertainty in attackers' knowledge levels are not well considered in previous studies. To address the uncertainties associated with attackers' knowledge levels in attack path selection, this study considers two different attack path selection mechanisms, which are random attacks and strategic attacks (Bier & Gutfraind, 2019). Random attack presents that an attacker selects one attack path from all possible attack paths randomly, which is applicable to attackers with low knowledge levels. A strategic attack presents that an attacker selects the attack path based on the exploitability ranking of all possible attack paths (a more exploitable attack path is more likely to be selected), which is applicable to attackers with advanced knowledge levels of the targeted system. We assign different probabilities of executing random attacks and strategic attacks for the attackers with different knowledge levels, as shown in Table 4.1.

Table 4.1 Attack path selection mechanisms for attackers with different knowledge levels.

Attacker categories[8]	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:

$$Pr(i) = \frac{a}{n} + \frac{b}{GTTC_i} / \sum_{j=1}^{n} \frac{1}{GTTC_j} \tag{4.8}$$

Where $Pr(i)$ is the probability of attack path i being selected. a and b are the likelihoods of executing a random attack and executing a strategic attack respectively, which are determined in Table 1. $GTTC_i$ is the global time-to-compromise of attack path i, which is calculated by summing the TTC of each attack step along the attack path. The TTC estimation method can be found in Appendix II. The conditional probability of attack path i is executed successfully given an attack attempt (L^i) is estimated as follows (Yuan et al., 2024):

$$L^i = Pr(i) \times \frac{MTTD_i}{GTTC_i + MTTD_i} \times \beta_i \tag{4.9}$$

$$MTTD_i = \frac{\sum_{k=1}^{N} TTD_k}{N} \tag{4.10}$$

$$L_{at}^{i} = L^a + L^b, \ldots, L^n \tag{4.11}$$

where $MTTD_i$ presents the mean-time-to-detect regarding attack path i. MTTD (mean-time-to-detect) is a widely used performance indicator describing the average time needed by the security operations center (SOC) to detect a cyber intrusion successfully (Mughal, 2022). The MTTD regarding a specific intrusion type is calculated by averaging all incident detection times of this intrusion type, as presented in Eq. (4.10). For simplification, a reference value (14 days) from Semertzis et al. (2022) is used as the MTTD for remote cyber intrusions. In practice, it may be determined based on actual incident data collected by SOCs. In case multiple attack

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

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 presented in Eq. (4.11).

Because a successful cyber intrusion cannot always induce a physically dangerous scenario (Huang et al., 2009), coefficient β_i is defined to describe the likelihood that a successful intrusion of attack path i induces a dangerous phenomenon. β_i depends on the fault detection capability and deviation suppression/tolerance capability of the OT system regarding the specific attack mode, and it is calculated as $\beta_i = \beta_i^d \times \beta_i^r$. β_i^d presents the probability that the attack-induced deviations escape the fault/anomaly detection algorithm successfully. β_i^r presents the likelihood that the attack-induced deviations cause a dangerous phenomenon successfully. β_i^d and β_i^r are determined as below.

Regarding the manipulation of OT systems, five types of C2P attacks with representativeness are investigated, as illustrated in Figure 4.4. Some basic descriptions of how FDI (false data injection) attacks, DoS (denial-of-service) attacks, and setpoint manipulations compromise the industrial control system are given in Table 4.2.

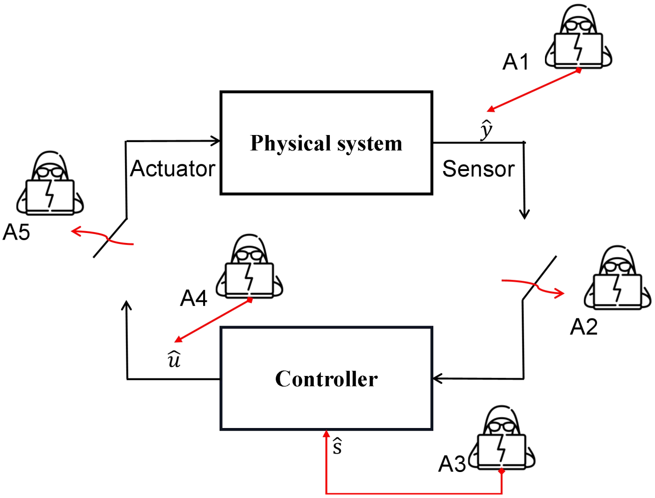


Figure 4.4. Typical C2P attacks against industrial control systems, adapted from (Huang et al., 2009).

Table 4.2 Explanations of the C2P attacks.

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

The value of β_i^d should be determined considering both the attack mode and the fault detection algorithm of the OT system. In this study, we assumed that predefined ranges were applied for sensors' and actuators' signals as the fault detection method (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. Referenced β_i^d values, 1, 0.8, 0.5, and 0.2, are used for FDI attacks executed by attackers with expert, intermediate, beginner, and novice knowledge levels, respectively. The values of β_i^d considering different attack modes and different attackers' knowledge levels are configured in Table 4.3. In practice, the β_i^d values may be modified considering the specific fault detection algorithms the OT system uses.

Table 4.3 Configurations of β_i^d for attackers with different knowledge levels.

Attacker's knowledge levels	$\boldsymbol{\beta_i^d}$ for FDI attacks	$\boldsymbol{\beta_i^d}$ for DoS attacks	$\boldsymbol{\beta_i^d}$ for Setpoint manipulations
expert	1	1	1
intermediate	0.8	1	1
beginner	0.5	1	1
novice	0.2	1	1

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

$$\begin{cases} X(k+1) = f(X(k), \widetilde{U}(k), w) \\ Y(k) = g(X(k), v) \\ U(k) = h(\tilde{S}(k), \widetilde{Y}(k)) \end{cases} \tag{4.12}$$

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

$$\beta_i^r = \begin{cases} 1, \ if \ X_i(k) \notin R \\ 0, \ if \ X_i(k) \in R \end{cases}, \ k \in K \tag{4.13}$$

where $K = \{k_s, \cdots, k_e\}$ represents the attack duration from the start time k_s to the stop time k_e. $X_i(k)$ is estimated by solving Eq. (4.12) regarding the attack mode of attack path i. A group of simulations (Monte Carlo simulations) is used to address the uncertainties in β_i^r associated with process noises and observation noises. To make the vulnerability assessment model structured, a pseudocode is presented in Figure 4.5 to demonstrate the procedures of the vulnerability assessment model.

Algorithm for vulnerability assessment of ICSs regarding C2P attacks

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

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

```
Establish a compromise graph with $n$ attack paths and $k$ attack modes (ending nodes);
Initiate a probability distribution of attackers' skill levels and Monte Carlo simulations with $m$ trials;
for $j = 1{:}m$ do
    Sample an attacker according to the probability distribution of attackers' skill levels;
    for $i = 1{:}n$ do
        Calculate $GTTC_i$ using the method in Appendix II;
        Calculated the probability of each attack path being selected, $Pr(i)$, using Eq. (4.8);
        Determine the values for $\beta_i^d$, $\beta_i^r$, and $\beta_i$;
        Calculate $L^i$ for each attack path using Eq. (4.9);
    end for
    for $i = 1{:}k$ do
        Calculate $L_{at}^i$ using Eq. (4.11);
    end for
    Accumulate the values of $L_{at}^i$ $(i = 1{:}k)$;
end for
return probability distributions of $L_{at}^i$ $(i = 1{:}k)$.
```

Figure 4.5. Pseudocode of the vulnerability assessment model.

4.2.4 Sensitivity analysis and risk-based decision making

Sensitivity analysis of BN root nodes can identify critical causal factors regarding undesired events and therefore supports risk treatment. Because the Birnbaum importance measure (van der Borst & Schoonakker, 2001) can be applied to integrated risk assessment models easily, this study uses Birnbaum importance measure for sensitivity analysis.

$$I_n = p_s(p_n = 1) - p_s(p_n = 0) \tag{4.14}$$

where I_n is the criticality of causal factor n. p_s is the probability of occurrence of the unwanted accident scenario. p_n is the probability of happening of causal factor n. With the identification of critical root nodes, corresponding risk treatment strategies may be proposed to reduce undesired risks.

In practice, the objectives of risk management usually include ensuring the risks are at acceptable levels, saving the costs used for risk reduction, reducing production losses resulting from downtime, meeting legislation requirements, etc. It is crucial to make decisions on risk reduction while considering the trade-offs between multiple objectives, for instance, the trade-off between safety and costs. Among risk-based decision analysis methods, risk matrix is one of the widely-used tools because it is straightforward and user-friendly. Particularly, the combination of cost-effectiveness analysis and a risk matrix helps to investigate the trade-off between safety and costs (Reniers & Van Erp, 2016). An optimization problem under constraints is formulated to characterize the decision analysis for barrier improvements, as follows:

$$\begin{cases} Min(C_i) \\ Risk_i \leq Risk_{threshold} \\ i \in \{1,2,3,\cdots,N\} \end{cases} \tag{4.15}$$

where C_i means the cost of strategy i regarding barrier improvements. $Risk_i$ is the risk estimation after implementing strategy i. $Risk_{threshold}$ is the risk threshold that could be the risk acceptable level in the risk matrix. In case probability distributions are used to represent risks considering uncertainties, thresholds may be used to constrain the expected values of probability distributions, and other constraints can also be applied to the probability distributions (for instance, setting up thresholds for the boundary values of the probability distributions). Eq (4.15) can be solved by using exhaustive search algorithms. In case of massive candidate strategies are proposed, evolutionary algorithms (for instance, genetic algorithms) may be used to obtain approximately optimal strategy while saving computation efforts.

4.3 Case study

4.3.1 System description and scenario building

In this case study, a continuous stirred tank reactor (CSTR) with its SCADA system is investigated, as demonstrated in Figure 4.6 (a). The CSTR model performs a hypothetical exothermic reaction A→B (Pilario & Cao, 2018). Product B is assumed to be a flammable liquid with toxicity. The reactor temperature is controlled using a jacketed cooling system with a water pump (WP), a control valve (V3), a temperature sensor (T), and a programmable logic controller (PLC1). Reactant A is fed at a fixed flow rate using a control valve (V1). An emergency shutdown system (ESD) with a programmable logic controller (PLC2), a block/shutdown valve (V2), and a pressure sensor (P) is deployed to block feeding in case of overpressure. Additionally, a safety relief valve (SV) is installed. Both PLCs are connected to the SCADA system and linked to the corporate network and the outside Internet/WAN. A CPS master diagram considering the multi-layered structure of the ICS is constructed, as shown in Figure 6 (b). Remote hackers (probably having different knowledge levels) are identified as threat agents, and ten specific attack modes are identified, as illustrated in Table 4.4. For the identified typical attack modes, OT safety barriers play a crucial role in preventing their success. For example, the emergency shutdown system (ESD), manual shutdown, and safety relief valve (SV) are expected to activate promptly in overpressure scenarios caused by either safety-related issues or C2P attacks. In the case of C2P attack modes (AT6-AT10) targeting digitized safety barriers, the safety relief valve (SV) will remain unaffected by cyberattacks since it is a passive safety barrier. In contrast, the automatic ESD system may be compromised by attack modes AT6-AT10, and the manual shutdown could fail under attack modes AT6-AT10 due to the compromise of the pressure sensor (P) or shutdown valve (V2). Based on the scenario integration approach presented in section 4.2.2, an attack-tree-bow-tie diagram was developed to integrate accidental scenarios considering both safety causes and C2P attacks, as shown in Figure 4.7.

Table 4.4 Descriptions of the identified attack modes.

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

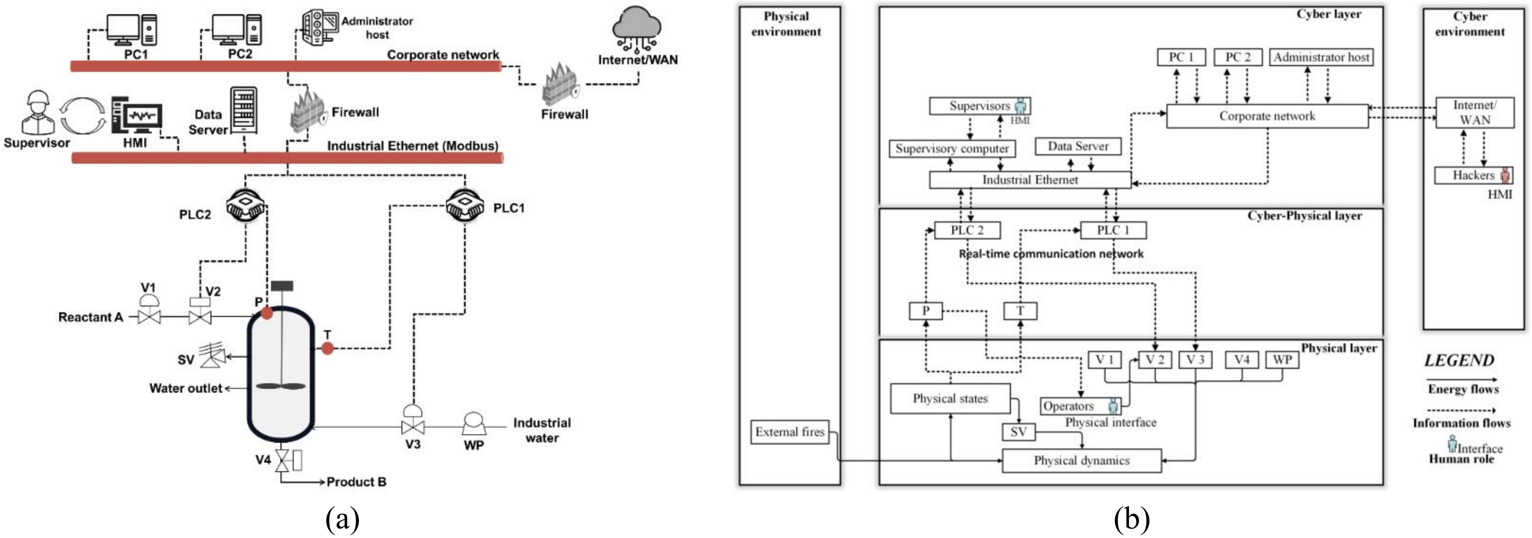


Figure 4.6. The investigated industrial control system (a) and its CSP master diagram (b), adapted from (Pilario & Cao, 2018) and (Yuan et al., 2024).

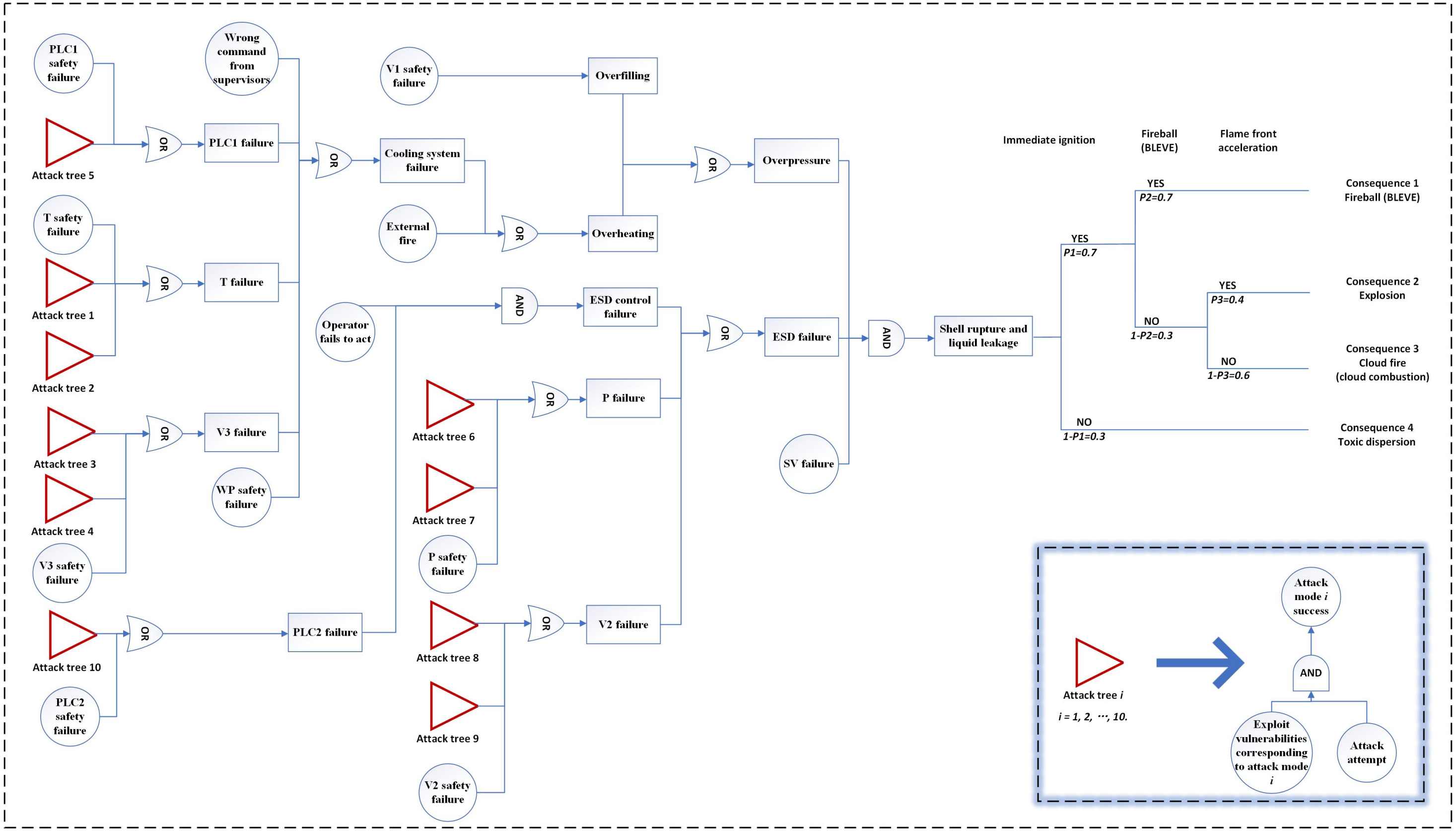


Figure 4.7. An integrated attack-tree-bow-tie diagram considering both safety causes and C2P attacks, adapted from (Yuan et al., 2024) and (Vílchez et al., 2011).

4.3.2 Security vulnerability assessment and risk assessment

Based on the developed attack-tree-bow-tie diagram, a BN model was developed following the mapping algorithm suggested by Khakzad et al. (2013) using the Bayes net MATLAB toolbox (Murphy, 2001). The BN nodes with pink color are derived from the bow-tie diagram, while the nodes with blue color are derived from attack trees, as shown in Figure 4.8. The explanation of the abbreviations of the BN nodes is given in Table 4.5. 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).

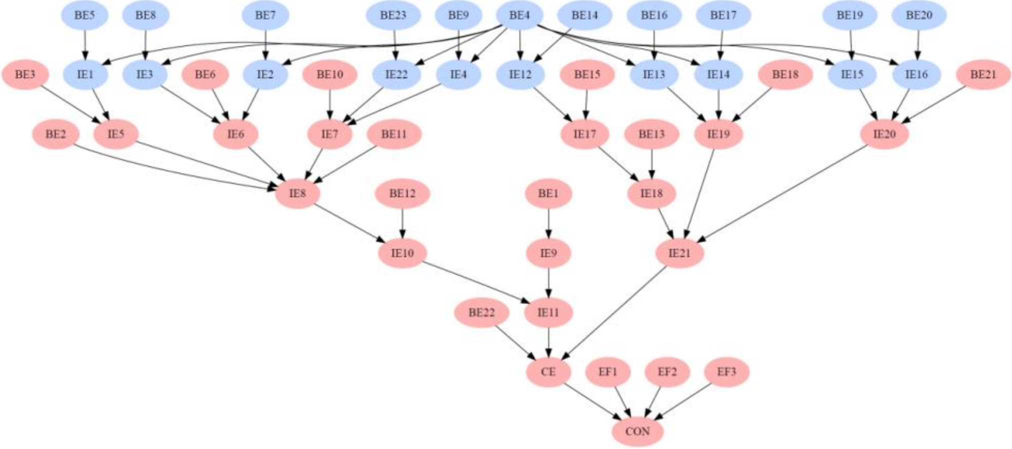


Figure 4.8. Bayesian network model for integrated risk assessment (nodes with pink and blue colors are derived from the bow-tie diagram and attack trees, respectively).

Table 4.5 Explanations of the BN nodes.

Symbols	Node names	Symbols	Node names	Symbols	Node names	Symbols	Node names
BE1	V1 safety failure	BE2	Human error in giving commands	BE3	PLC1 safety failure	BE4	C2P attack attempts
BE5	Exploit vulnerabilities corresponding to AT5	BE6	T safety failure	BE7	Exploit vulnerabilities corresponding to AT1[9]	BE8	Exploit vulnerabilities corresponding to AT2
BE9	Exploit vulnerabilities corresponding to AT3	BE10	V3 safety failure	BE11	WP safety failure	BE12	External fire
BE13	Operator fails to shutdown	BE14	Exploit vulnerabilities corresponding to AT10	BE15	PLC2 safety failure	BE16	Exploit vulnerabilities corresponding to AT6
BE17	Exploit vulnerabilities corresponding to AT7	BE18	P safety failure	BE19	Exploit vulnerabilities corresponding to AT8	BE20	Exploit vulnerabilities corresponding to AT9
BE21	V2 safety failure	BE22	SV safety failure	BE23	Exploit vulnerabilities corresponding to AT4	EF1	Immediate ignition
EF2	Fireball (BLEVE)	EF3	Flame front acceleration	CON	Consequences	CE	Central event (Liquid leakage)
IE1	AT5 success	IE2	AT1 success	IE3	AT2 success	IE4	AT3 success
IE5	PLC1 failure	IE6	T failure	IE7	V3 failure	IE8	Cooling system failure
IE9	Overfilling	IE10	Overheating	IE11	Overpressure	IE12	AT10 success
IE13	AT6 success	IE14	AT7 success	IE15	AT8 success	IE16	AT9 success
IE17	PLC2 failure	IE18	ESD control failure	IE19	P failure	IE20	V2 failure
IE21	ESD failure	IE22	AT4 success	/	/	/	/

[9] AT1 means attack mode 1, and the explanation of each attack mode can be found in Table 4.4.

According to a data analysis of cyber security incidents in a large American organization (Kuypers & Maillart, 2018), the recurrence intervals of severe cyber incidents remain overall stable, and the recurrence interval of C2P attacks may be estimated as approximately 150~465 days. Due to the lack of actual incident data, a Gamma distribution ($\Gamma(\alpha,\beta), \alpha = 6.08;\ \beta = 5$) with a mean value of 1.22 (corresponds to a recurrence interval of 300 days) is used to depict the C2P attack annual frequency in this case study. When new incident data becomes available, the Gamma distribution may be updated based on Bayes' theorem (Eide et al., 2007).

Regarding vulnerability assessment, the approach presented in section 4.2.3.3 was used to quantify the conditional probabilities of successfully executing each attack mode. A compromise graph was constructed to demonstrate all possible attack paths according to the identified attack modes and the IT structure of the ICS, as presented in Figure 4.9. The vulnerabilities at each attack step are presented in Table 4.6. Each ending node (AT1-AT9) of the compromise graph presents an attack mode, as illustrated in Table 4.4. For attack path 1, which is highlighted in red colour, the global time-to-compromise is calculated as: $GTTC_1 = TTC_1 + TTC_2 + TTC_4$.

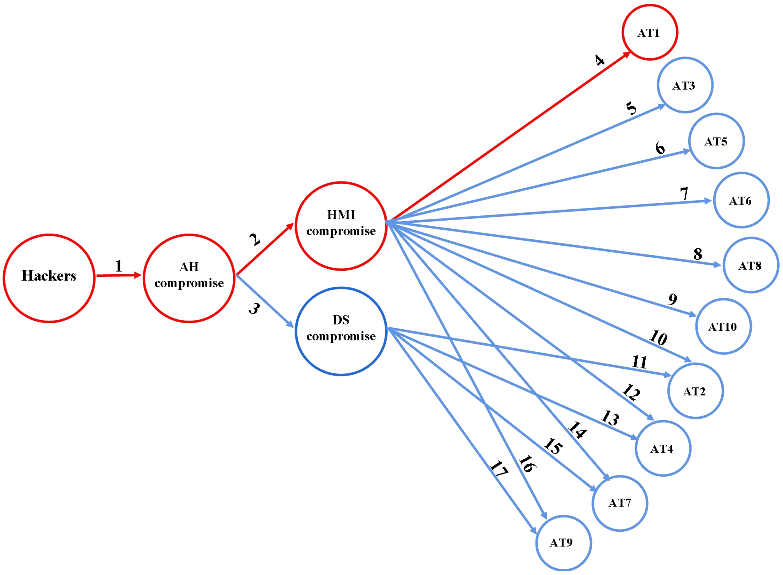


Figure 4.9. A compromise graph regarding the investigated ICS.

Table 4.6 Known vulnerabilities at each attack step.

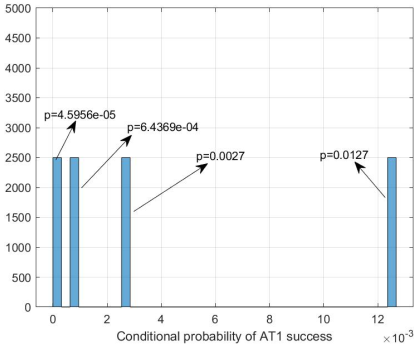
Attack step number	Known vulnerabilities	Attack step number	Known vulnerabilities (cve_id)

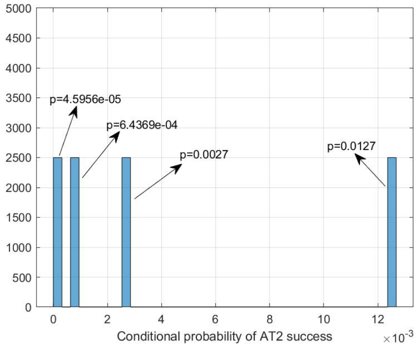
	(cve_id[10])		
1	CVE-2015-7871; CVE-2017-2683	2	CVE-2017-13997
3	CVE-2018-13799	4	no
5	no	6	CVE-2018-5459
7	no	8	no
9	CVE-2018-5459	10	CVE-2016-2200
11	CVE-2016-2200	12	CVE-2016-2200
13	CVE-2016-2200	14	CVE-2016-2200
15	CVE-2016-2200	16	CVE-2016-2200
17	CVE-2016-2200	/	/

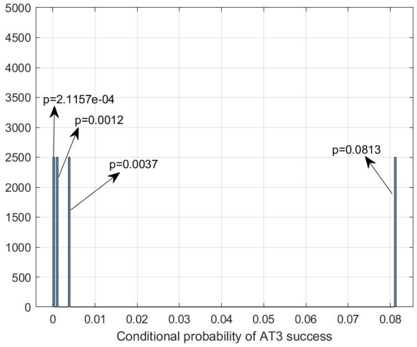
A MATLAB/Simulink model was developed based on the CSTR model from Pilario & Cao (2018) to assess the physical effects of different attack modes and to determine the value of β_i^r based on Eq (4.13). The developed MATLAB/Simulink model and selected simulation results are presented in Appendix I. According to the simulation results, the ICS's process and observation noises influence the vulnerability assessment results because they may decide if a DoS attack on PLC1 can succeed (if $\beta_7^r = 1$ and $\beta_8^r = 1$).

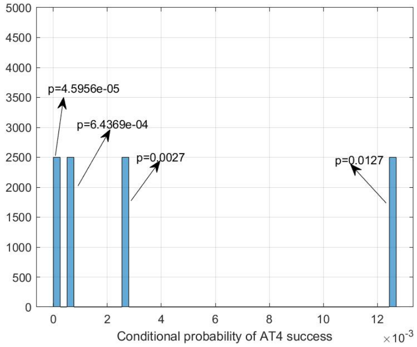
Considering the uncertainties associated with attackers' skill levels and process and observation noises, Monte Carlo simulations with 10000 trials were performed to obtain the probability distribution of successful execution of each attack mode following the algorithm presented in Figure 4.5. A uniform distribution (a ratio 1:1:1:1) was used for potential attackers with different skill levels (expert, intermediate, beginner, and novice). In practice, it may be configured based on expert judgment considering possible threat agents. The calculation results of the vulnerability assessment model are presented in Figure 4.10. Safety barrier maintenance time intervals were initialed as one year in PFD calculations by using the reliability models in section 4.2.3.2. Then, the prior probabilities/probability distributions for all root nodes are summarized in Table 4.7, based on which the Bayesian inference was performed to obtain the probability distributions of the occurrence of each consequence, as shown in Figure 4.11. Finally, the mean values and ranges of the obtained probability distributions are visualized in a risk matrix to present risk profiles considering parameter uncertainties, as shown in Figure 4.12.

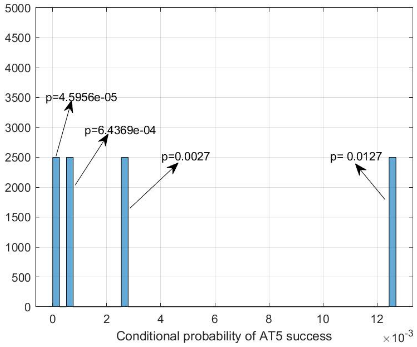
[10] A cve_id uniquely identifies one vulnerability from the Common Vulnerabilities and Exposures (CVE) database (NVD, 2023).

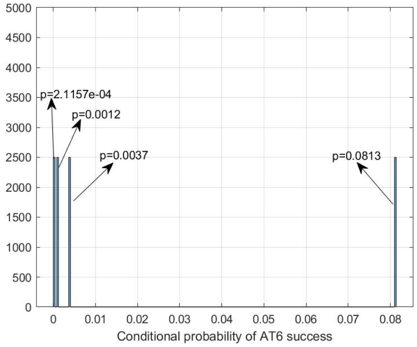

5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
12
Conditional probability of AT1 success
$\times 10^{-3}$
p=4.5956e-05
p=6.4369e-04
p=0.0027
p=0.0127


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
12
Conditional probability of AT2 success
$\times 10^{-3}$
p=4.5956e-05
p=6.4369e-04
p=0.0027
p=0.0127


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Conditional probability of AT3 success
p=2.1157e-04
p=0.0012
p=0.0037
p=0.0813


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
12
Conditional probability of AT4 success
$\times 10^{-3}$
p=4.5956e-05
p=6.4369e-04
p=0.0027
p=0.0127


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
12
Conditional probability of AT5 success
$\times 10^{-3}$
p=4.5956e-05
p=6.4369e-04
p=0.0027
p= 0.0127


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Conditional probability of AT6 success
p=2.1157e-04
p=0.0012
p=0.0037
p=0.0813

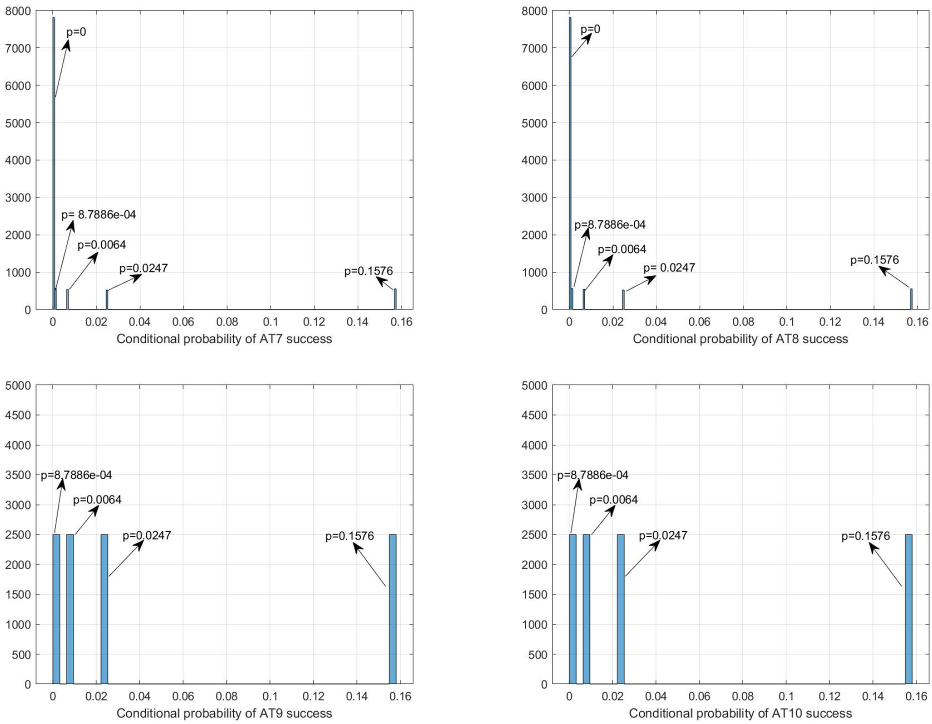
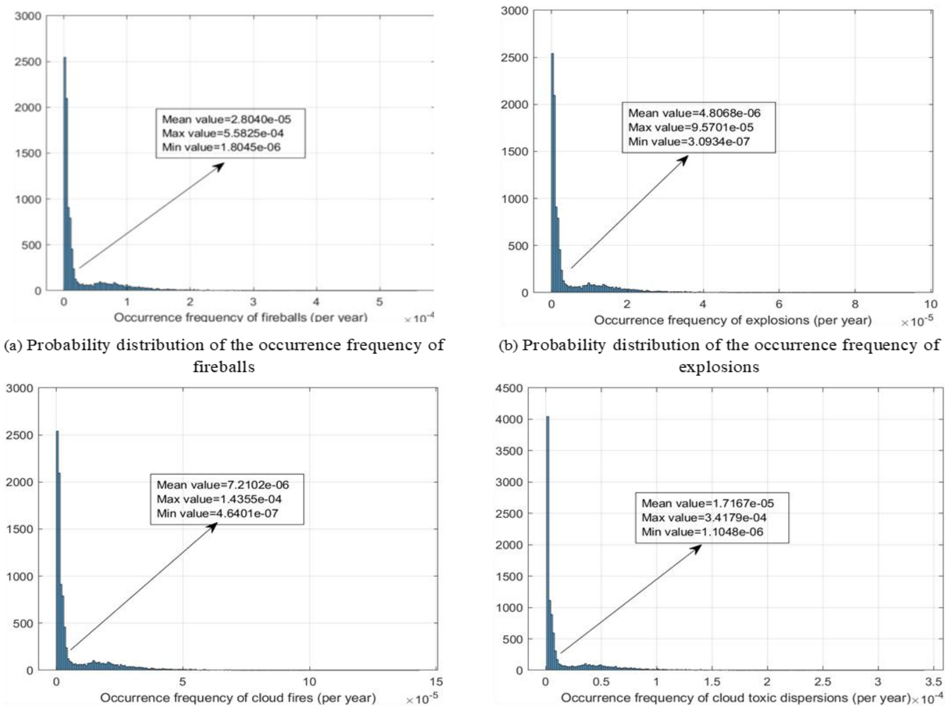


Figure 4.10. Conditional probabilities of the successful execution of each attack mode (*p* in the figures means conditional probability).

Table 4.7 Probabilities/probability distributions of the root BN nodes.

Nodes	**Prior probabilities (probability distributions)**	**Sources**	**Nodes**	**Prior probabilities (probability distributions)**	**Source**	**Nodes**	**Prior probabilities (probability distributions)**	**Source**
BE1	4.00E-02	(Taylor, 2010)	BE2	1.00E-02	(Andersen et al., 2004)	BE3	4.38E-02	(Hauge & Onshus, 2010)
BE4	Gamma distribution ($\Gamma(\alpha,\beta), \alpha = 6.08;\ \beta = 5$)	Assumed based on the data from (Kuypers & Maillart, 2018)	BE5	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE6	2.13E-02	(Hauge & Onshus, 2010)
BE7	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE8	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE9	As shown in Figure 4.10	Calculated from the vulnerability assessment model
BE10	4.00E-02	(Taylor, 2010)	BE11	3.125E-02	(OREDA, 2002)	BE12	5.52E-02	(Debray et al., 2004).
BE13	Beta distribution ($Beta(a,b)$, a=32.3; b=137.7)	(Roy et al., 2015)	BE14	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE15	average PFD=4.37E-03; λ=1.0E-06	λ is from Hauge & Onshus (2010); Eq (7) is used to calculate the PFD.
BE16	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE17	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE18	average PFD=6.57E-04; λ=1.5E-07	λ is from Hauge & Onshus (2010); Eq (7) is used to calculate the PFD.
BE19	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE20	As shown in Figure 4.10	Calculated from the vulnerability assessment model	BE21	average PFD=7.63E-03; α=1.02E-04, β=1.2E04, L=3E-04	α and β are from Zhang et al. (2020); Eq (6) is used to calculate the PFD.
BE22	average PFD=2.19E-03; λ=5E-07	λ is from (HSE, 2012); Eq (7) is used to calculate the PFD.	BE23	As shown in Figure 4.10	Calculated from the vulnerability assessment model	EF1	7.00E-01	(Vílchez et al., 2011)
EF2	7.00E-01	(Vílchez et al., 2011)	EF3	4.00E-01	(Vílchez et al., 2011)	/	/	/



(a) Probability distribution of the occurrence frequency of fireballs

(b) Probability distribution of the occurrence frequency of explosions

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

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

Figure 4.11. Probability distributions of the occurrence of each consequence

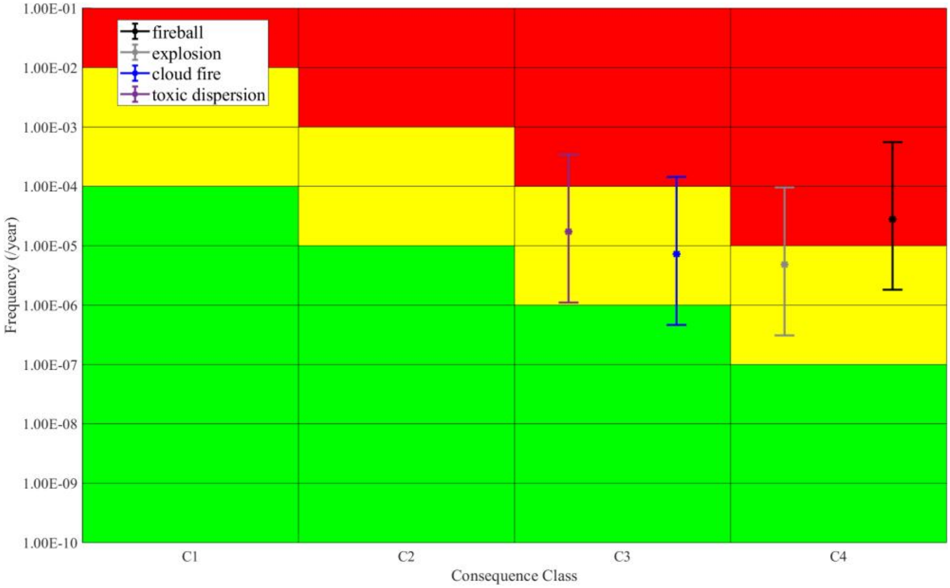


Figure 4.12. Estimated risks with uncertainties demonstrated in a risk matrix.

4.3.3 Decision making on barrier improvements

As shown in Figure 4.12, necessary improvements must be made to current safety and security barriers because the calculated “risk ranges” overlap with the red region in the risk matrix. Based on the developed BN model, a sensitivity analysis of the basic events (root nodes) regarding the happening of disastrous consequences was performed using Eq (4.14). The results from the sensitivity analysis are presented in Figure 4.13, where BE22 (SV safety failure) demonstrates the highest importance, with a Birnbaum importance measure value of 0.0207. BE16 to BE21, all related to the ESD system, also show relatively high sensitivities, with importance measure values around 4E-4. Other basic events have importance measure values below 2E-4. These results indicate that the safety relief valves (SV) and emergency shutdown (ESD) systems are critical components in mitigating undesired risks. While the safety relief valve is only subjected to safety failures, critical failures in the ESD system can arise from either safety failures or C2P attacks. Therefore, potential strategies can be proposed to enhance the safety and security protection of both the SV and the ESD systems.

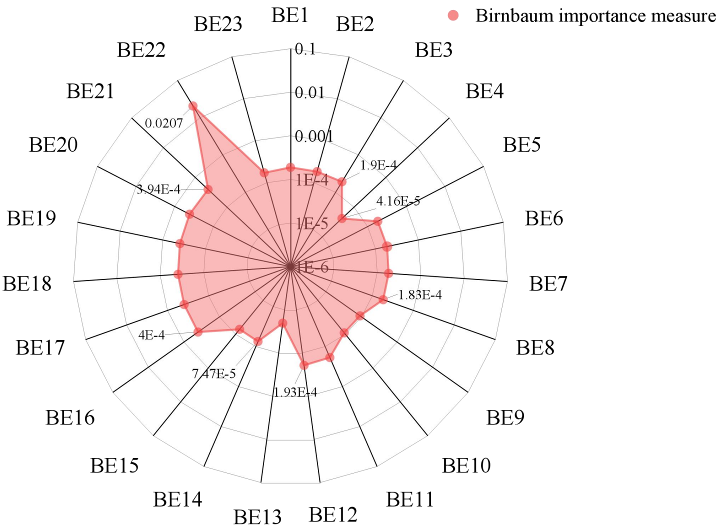


Figure 4.13. Sensitivity analysis of root nodes.

Vulnerability patching is a crucial way to remove vulnerabilities from an IT system by delivering security patches (Hong et al., 2014), and it is regarded a security barrier in this study. It is assumed that the security management team has the capability to patch vulnerabilities CVE-2016-2200, which is closely related to DoS attacks on PLCs.

Meanwhile, considering the maintenance scheduling of the ESD system and the safety relief valve, a cost analysis of the potential barrier improvement actions is given in Table 4.8. Considering also the technical constraints and practicability of the possible strategies, 18 schemes are proposed as candidate strategies for barrier improvements, as concluded in Table 4.9. The optimization problem is characterized as minimizing the cost of barrier improvements while ensuring the accident risks are within the acceptable thresholds. Because the threshold for fireball risks is the most difficult one to meet, the estimated mean values and maximum values of the occurrence frequency of fireballs after implementing each candidate strategy are compared in Figure 4.14. Because the proposed approach can present accident risks in the form of risk ranges, two thresholds (threshold A and threshold B) are used for the expected/mean value and maximum value of the estimated fireball risks respectively. Combined with the comparison of the cost of each strategy, the optimal strategy may be determined. As shown in Figure 4.14, the expected values of fireball risks after implementing strategies 7 to 9 and 16 to 18 are below threshold A, and strategy 7 has the lowest cost. Therefore, strategy 7 is most cost-effective when only the expected/mean value of fireball risks is considered decision-making criteria. By contrast, max values of fireball risks meet threshold B only after implementing strategies 16 to 18, and strategy 16 has the lowest cost. Therefore, strategy 16 is the optimal strategy when both the mean value and maximum value of the risk ranges are configured as decision-making criteria.

Table 4.8 Cost analysis of safety and security barrier improvement actions.

Improvement actions	**Cost analysis**[11]	**Total costs**
Maintenance of the ESD system	a1=10,000€ (one-time maintenance cost); a2=100,000€×2 days (downtime cost)	Total cost=a1+a2=210,000€
Maintenance of the safety relief valve	b1=2,000€ (one-time maintenance cost); b2=100,000€×1 day (downtime cost)	Total cost=b1+b2=102,000€
Patch vulnerability CVE-2016-2200	c1=20,000€ (patching cost); c2=100,000€×14 days (downtime cost)[12]	Total cost=c1+c2=1,420,000€

[11] Approximate costs are used for the perfect maintenance (replacement) because the cost varies for specific equipment.

[12] The time needed for vulnerability patching also varies. The downtime is estimated referring to an optimistic Mean Time To Patch (MTTP) for cost saving https://purplesec.us/learn/average-time-patch-vulneraiblity/.

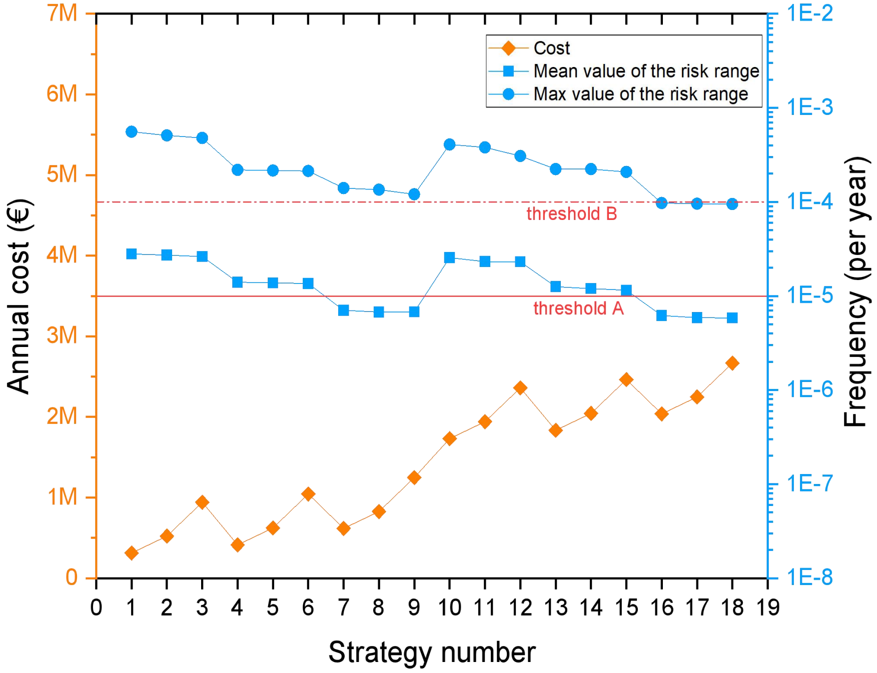


Figure 4.14. Risk profiles after implementing the candidate strategies (fireball risk as example).

Table 4.9 Proposed candidate strategies for improving the performance of safety and security barriers.

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

4.4 Discussions

4.4.1 Uncertainty treatment towards better decision-making

Considering the noticeable uncertainties in the integrated safety and security risks, multiple uncertain parameters are considered in this study. For instance, uncertainties associated with attackers' skill levels in cyber intrusion risks are assessed using Monte Carlo simulations. The combination of a Bayesian network and Monte Carlo simulations enables the use of probability distributions and handles uncertainty propagation in the risk assessment. As a result, the risk assessment results, which are in the form of probability distributions regarding the happening of each possible consequence, can provide more insights for the decision-making on barrier management. Using the proposed approach, risks can be evaluated according to the expected values and the risk ranges. Different thresholds may be set up for the expected/mean values and maximum values of the risk ranges to decide if the risks are acceptable. During the decision-making process, both the expected/mean values and the maximum values of the risk ranges can be used as criteria to determine appropriate strategies for safety and security barrier improvement. The maximum values of the risk ranges reflect worst-case risks considering parameter uncertainties, and they may be used as criteria when the worst cases are concerned by decision-makers. Therefore, with appropriate treatment and handling of parameter uncertainties, more criteria (optimization objectives) can be used to guide decision-making on safety and security barrier management according to the needs and interests of decision-makers.

4.4.2 Limitations and recommendations for future studies

The developed approach incorporates a relatively thorough list of parameters/factors in the risk assessment. Some approximate assumptions or referenced values are currently used for some variables due to the lack of data. For instance, a reference MTTD value was used for all C2P attack modes, and values of β_i^d were configured based on the assumption of a simple fault/anomaly detection algorithm. The configuration of those parameters may be improved according to the actual intrusion detection capability of the security operations center. The true positive rate and false alarm rate of the deployed fault/anomaly detection algorithm help to determine those parameters when related data is available.

Additionally, the configuration of the attack likelihood and the probability distribution of attackers with different skill levels depends on subjective judgment in the present study. The incorporation of more incident data and evidence in the configuration of those parameters helps to obtain more credible risk assessment results. Moreover, the possible dynamic variations were ignored in the risk assessment, and neither were in the barrier management. How to systematically address dynamic variations in the integrated risks assessment and further incorporate dynamic risks with the safety and security barrier management is worth investigating in future studies.

4.5 Conclusions

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

Chapter 5 Cost-effective maintenance of safety and security barriers via genetic algorithm

This chapter proposes a novel approach for optimizing a safety and security barrier maintenance strategy considering economic constraints. In the proposed approach, accident scenarios in terms of safety and physical security are constructed using an extended bow-tie diagram. After associated safety and security barriers are identified, a system simulation model is developed to conduct barrier modeling based on MATLAB/Simulink simulations, in which the barrier maintenance, the impacts of human and organizational barriers, and the correlations between barriers caused by shared components are considered. Finally, a combination of cost-effectiveness analysis (CEA) and genetic algorithm (GA) is employed to support the decision-making on barrier maintenance optimization. An illustrative case is employed in this study to validate the feasibility of the proposed approach.

This chapter is drafted with modifications based on the following publication:

- ◆ Yuan, S., Reniers, G., Yang, M., & Bai, Y. (2023). Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. *Process Safety and Environmental Protection, 170,* 356-371.

5.1 Introduction

Safety and security barriers are implemented in various forms (e.g., technical and non-technical) to protect chemical plants from undesired events in terms of prevention and mitigation of potentially catastrophic consequences (Zeng et al., 2020; Yuan et al., 2022a). The integration of safety and security in chemical plants has been emphasized and investigated in previous studies. For example, integrated safety and security risk assessments were recommended considering the interaction among safety and security-related causal factors through a dynamic risk assessment approach (Song et al., 2019a; 2019b). An approach based on dynamic graphs was proposed to integrate safety and security resources to reduce the risk of intentional domino effects (Chen et al., 2019). However, integrated management of safety and security barriers is still challenging, particularly in the use of quantitative risk assessment to support barrier management.

Typically, bow-tie metaphor/diagrams are widely used and recommended in the performance assessment and management of safety barriers due to their advantages in being capable of quantitative analysis and relatively understandable/straightforward. QRA (quantitative risk assessment) is highly suggested to support safety barrier management by researchers (Pitblado et al., 2016; Bucelli et al., 2017; Yuan et al., 2022a). The ARAMIS (Accidental Risk Assessment Methodology for Industries) project integrated add-on safety barriers into a QRA framework to facilitate safety barrier management with respect to major accident scenarios (Andersen et al., 2004). Additionally, an extension of bow-tie diagrams to the security risk analysis or safety and security (in one go) risk analysis was also investigated in previous studies (Abdo et al., 2018), demonstrating bow-tie diagrams have the potential to facilitate integrated safety and security risk management and barrier management. However, how to make good use of QRA approaches for barrier maintenance scheduling has rarely been investigated by previous studies.

Regarding barrier maintenance, barrier aging, degradation, and the influence of human and organizational factors should be considered (Fiorentini & Marmo, 2018; CCPS/EI, 2018). Generally, there are sorts of approaches widely-used for chemical process facility maintenance, for instance, reliability centered maintenance (RCM) (Eisinger & Rakowsky, 2001), condition-based maintenance (CBM) (Wang et al., 2022), preventive maintenance (PM) (Basri et al., 2017), risk based inspection (RBI) (Tan et al., 2011), or a combination of them. The objective of facility maintenance is to maximize the availability and efficiency of the facility and guarantee a safe and correct operation and minimize costs. Because the common goal of safety barriers and security barriers is to control risk, risk-based approaches are suitable for supporting barrier maintenance considering risk sources, including both safety hazards and malicious acts. By integrating barrier maintenance with QRA, barrier maintenance can be planned based on quantitative barrier importance to risk control in a manner similar to risk based inspection (RBI) (Pitblado et al., 2016). However, for a complex system with many safety and security barriers, it is difficult to determine specific optimal maintenance intervals for each barrier when the solution space is extremely large.

Targeting the challenges mentioned above, an approach is proposed to conduct risk assessments of accident scenarios considering the intervention of safety and security barriers and support cost-effective barrier maintenance in case of large-solution-space problems. The remainder of this chapter is organized as follows. Firstly, the proposed methodology is described in section 5.2. A system simulation tool for barrier modeling is introduced in section 5.3. Then, an illustrative case study is presented in section 5.4. Section 5.5 discusses the novelty of the proposed approach and suggests recommendations for future work. Conclusions are presented in section 5.6.

5.2 Methodology

5.2.1 Overall framework

To address the current gaps in cost-effective safety and security barrier maintenance, several main principles are proposed as follows. i) Both safety and security risk sources should be identified and depicted in the scenario building phase, meanwhile, the correlations/dependencies between safety and security barriers should be considered. ii) The effectiveness of barriers and also implementing barrier maintenance strategies should be measured by their risk-reduction performance regarding specific accident scenarios. iii) The barrier maintenance strategy should be optimized based on the synergistic effects of barriers on system risk reduction, rather than evaluating and optimizing each barrier according to its own probability of failure and consequence of failure. This is the main difference between our proposed approach and the conventional risk-based inspection (RBI) approaches (Tan et al., 2011). Based on the above principles, a novel approach with three steps (scenario building & barrier identification, barrier modeling, and optimization of barrier maintenance strategy) is proposed, as shown in Figure 5.1. A detailed illustration of the three steps is presented in the following sub-sections.

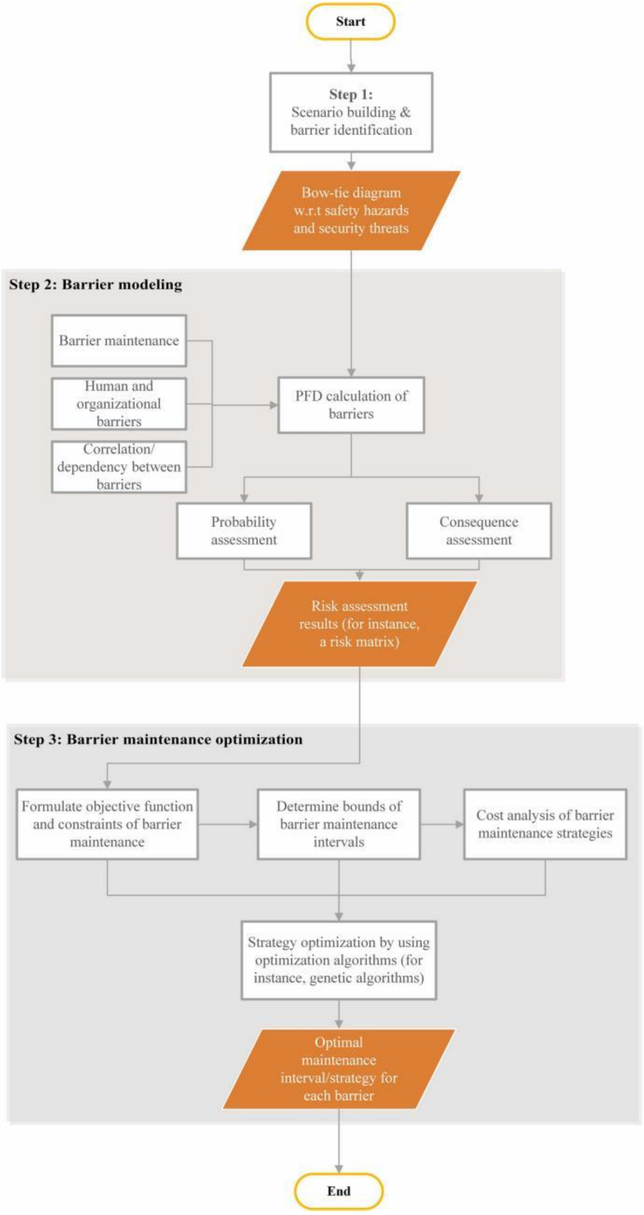


Figure 5.1. The framework of the proposed barrier maintenance approach.

5.2.2 Scenario building & barrier identification

Bow-tie identification techniques are widely used for HAZard IDentification (HAZID) and safety risk management (de Ruijter & Guldenmund, 2016), for instance, the MIMAH (methodology for identifying major accident hazards) (Andersen et al., 2004) and DyPAS (Dynamic Procedure for Atypical Scenarios Identification) (Paltrinieri et al., 2013). In this study, bow-tie diagrams are employed to identify and visualize accident scenarios in terms of both safety hazards and security threats. Safety and security barriers can be identified and located on the bow-tie diagrams with the help of existing documents or databases related to the investigated process control systems. For example, a database of checklists is available to support the barrier identification of industrial control systems (Guzman et al., 2021). Additionally, it is important to conduct a dedicated assessment of the adequacy of safety barriers in place to prevent security-attack scenarios during the scenario building and barrier identification process, as highlighted by Iaiani et al. (2023). Since certain safety barriers can be overridden by specific attack modes while others cannot, evaluating the adequacy of these barriers against specific attack scenarios is crucial. This helps to accurately determine where to place the attack events and safety barriers in the bow-tie diagram. A detailed example of this analysis is provided in Section 5.4.1.

Probability of failure on demand (PFD) is widely used to describe the unavailability of barriers. For a series of barriers following an AND logic gate, Eq (5.1) is used to calculate the output probability. For a series of barriers following an OR logic gate, formula (5.2) can be applied.

$$P_{OUT} = P_{IN} * (PFD_1 * PFD_2 \cdots PFD_n) \tag{5.1}$$

$$P_{OUT} = P_{IN} * (1 - (1 - PFD_1) * (1 - PFD_2) \cdots (1 - PFD_n)) \tag{5.2}$$

where P_{OUT} is the output probability and P_{IN} is the input probability of the branch. PFD_1 to PFD_n indicate the PFDs of barriers, and *n* is the number of barriers. For a barrier with two outlet branches, one branch presents the failing of the barrier with a probability (PFD). Another presents the barrier succeeding with a probability that is 1-PFD.

5.2.3 Barrier modeling

This section elaborates on how to determine PFDs of barriers and assess the performance of barriers based on their risk-reduction performance.

1) PFD calculation considering barrier maintenance

For a barrier constituted by multiple components, fault tree analysis is used to calculate the PFD of this barrier and then the calculated PFD is used for probabilistic risk assessment. The unavailability of a technical barrier was considered following the exponential distribution and can be expressed as a function of time (IEC, 2016; Redutskiy, 2017; Schmitz et al., 2021; Wu et al., 2022). For simplification purposes, PFD can be calculated according to Eq (5.3), in which a constant failure rate is assumed to calculate the cumulative failure probability.

$$PFD_{withoutBM}(t) = 1 - e^{-\lambda t} \tag{5.3}$$

where λ is the barrier failure rate and *t* denotes time. Some failure rate databases for safety barriers or the technical components of safety barriers are available and can be retrieved from (OREDA, 2002; Ottermo et al., 2021). It is assumed that the performance of a barrier can restore to its original state after the barrier maintenance, which can be called complete functional maintenance/test (Ottermo et al., 2021). We assume that the barrier failure rate will not change after complete functional maintenance, but it may not be equal to the original value in practice. We assume that the performance of a barrier follows a linear distribution during the maintenance period. If barrier maintenance with a time interval of *T* is conducted, the *PFD* of this barrier/barrier component is calculated according to Eq (5.4), which is a periodic piecewise function composed of exponential distributions and linear distributions. The starting times of barrier maintenance are the piecewise points.

$$PFD_{withBM}(t) = \begin{cases} 1 - e^{-\lambda * (t\%(T+h))}, & n(T+h) \le t < (n+1)T + nh \\ 1 - e^{-\lambda T} - (1 - e^{-\lambda T})/h * (t\%(T+h) - T), & (n+1)T + nh \le t < (n+1)(T+h) \end{cases} \tag{5.4}$$

where h is the required maintenance time. $t\%(T+h)$ means the remainder when dividing *t* by *T*+*h*. *n* is an integer from 0 to positive infinity. A comparison between the time-dependent PFD of a barrier using different maintenance intervals is shown in Figure 5.2.

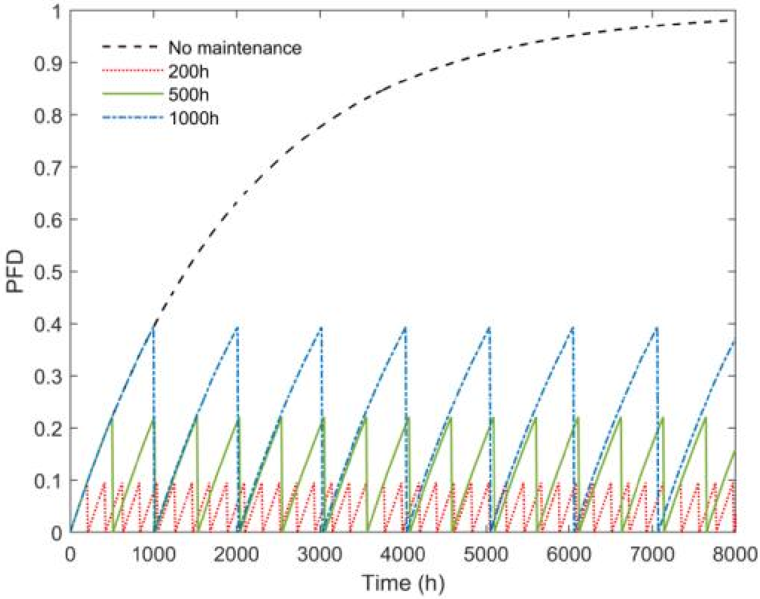


Figure 5.2. The PFD of a barrier using different maintenance intervals.

2) Human and organizational barriers

The need to involve human error probability (HEP) in the quantification of PFDs of the safety instruments executed by humans was suggested (Hauge et al., 2010). HEP can be estimated by Human Reliability Analysis (HRA) (Kirwan, 2017; Dimaio et al., 2021). Alternatively, there are some suggested rough PFD values for human actions and human barriers. For instance, the ARAMIS project provided the reference PFD values derived in an equivalent level of confidence (LC) for different types of human barriers, as shown in Table 5.1. Additionally, the quantification of the influence of the safety management system on QRA results through the audit of the safety management system quality/efficiency was suggested by both the ARAMIS project (Andersen et al., 2004) and the I-RISK project (Bellamy et al., 1999; Papazoglou et al., 2003). The ARAMIS project suggested evaluating the influence of safety management efficiency on safety barrier reliability by conducting site-specific questionnaires (Andersen et al., 2004).

Table 5.1 Reference PFDs for human barriers, adapted from (Andersen et al., 2004).

Human barrier/human action types	PFD (from literature and industry)	Level of confidence
Prevention	10^{-2}	LC 2

Normal operation	10^{-2}	LC 2
Intervention	10^{-1}	LC 1

3) Correlations between barriers

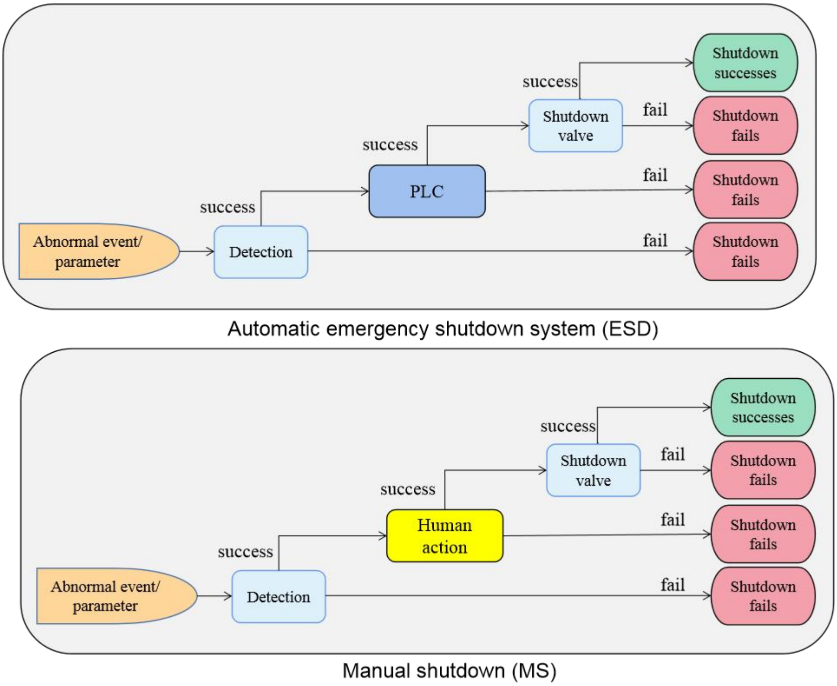
For the barriers designed to reduce the risks of the same accident scenarios, their criticalities to risk reduction are correlated inherently. In that case, the importance/criticality of one barrier in risk reduction is influenced by the reliability/availability of the other barriers because they have synergistic effects on risk reduction. Because both safety hazards and security threats can induce undesired accident scenarios, the assessment of safety and security barriers in a unified framework with the consideration of their synergistic effects on risk reduction is necessary.

Additionally, CCPS (USA) and the Energy Institute (UK) emphasized that active barriers should contain elements of 'detect-decide-act' and perform the complete intended function on its own when demanded (CCPS/EI, 2018). In real cases, it is possible that different barriers/barrier systems have some commonly used components responsible for completing specific tasks. For instance, an automatic emergency shutdown system (ESD) and a manual shutdown (MS) may use the same detector for monitoring the abnormal parameters/events and perform the shutdown by using the same valve, as shown in Figure 5.3. For two safety barriers with a shared component and located on the same branch, a conditional probability $P_2^{'}$ instead of PFD should be used for the second barrier. The conditional probability can be calculated as follows (Duijm, 2009):

$$P_{1,R} = \frac{P_1 - P_C}{1 - P_C} \tag{5.5}$$

$$\begin{aligned} P_2^{'} &= P(B_2\, fails \mid B_1\, has\, failed) \\ &= P_{2,R} + P(C\, fails \mid B_1\, has\, failed)\left[1 - P_{2,R}\right] \\ &= P_{2,R} + (P_C/P_1)\left[1 - P_{2,R}\right] \end{aligned} \tag{5.6}$$

where P_1 indicates the PFD of the barrier 1, which contains a common component C with a PFD P_C. $P_{1,R}$ is the PFD of the remaining components of barrier 1 in series with component C. It should be noticed that the above formulas can be adapted to calculate the conditional probabilities of multiple barriers with shared components as well, but those barriers have to be situated on the same branch of the bow-tie (Duijm, 2009).



Automatic emergency shutdown system (ESD)

Manual shutdown (MS)

Figure 5.3. A comparison between the components of an automatic emergency shutdown system (ESD) and manual shutdown (MS).

4) Probabilistic risk assessment

After determining the PFDs of the barriers, it is possible to conduct a probabilistic risk assessment of the undesired accident scenarios based on the bow-tie diagram obtained from step 1. By assigning the PFD values to the corresponding barriers located on the bow-tie diagram and following and calculation rules of the bow-tie, a probability assessment can be conducted. Alternatively, bow-tie diagrams can be transformed into Bayesian network models for probability assessment (Khakzad et al., 2013).

Consequence assessment is an important part of risk assessment. There are many methods or tools available for quantitative and qualitative consequence assessment of major accident scenarios in chemical plants. For instance, some software (PHAST, ALOHA, Ansys Fluent, FLACS, etc.) based on empirical models or CFD models can be used for physical effects modeling (Lewis, 2005). The combination of CFD simulations and probabilistic linear response models can be used for quantitative consequence assessment in terms of toxic leakage, fire, and explosion (Xie et al., 2022; Freeman, 1990). Alternatively, qualitative consequence assessment is also widely applied in the chemical process industries. For instance, a class of consequences was proposed by the ARAMIS project (Andersen et al., 2004), and the application of this class to typical dangerous phenomena was also presented, as shown in Table 5.2 and Table 5.3.

Table 5.2 Class of consequences, adapted from (Andersen et al., 2004).

Consequences			Class
Domino effect	**Effect on human targets**	**Effect on environment**	**Ranking**
To take into account domino effects, the class of consequence attributed to the studied dangerous phenomenon will be increased to the class of the secondary phenomenon that the first can bring about by domino effect.	No injury or slight injury with no stoppage of work	No action is necessary; just watching	C_1
	Injury leading to a hospitalization > 24 hours	Severe effects on the environment, requiring local means of intervention	C_2
	Irreversible injuries or death inside the site, Reversible injuries outside the site	Effects on environment outside the site, requiring national means	C_3
	Irreversible injuries or death outside the site	Irreversible effects on the environment outside the site, requiring national means	C_4

Table 5.3 Class of consequence of typical “fully developed” dangerous phenomena, adapted from (Andersen et al., 2004).

Dangerous phenomena	Consequence class
Pool fire	C_2
Tank fire	C_1
Jet fire	C_2 (increasing to C_3 or C_4 in case of domino effects)
VCE (Vapor Cloud Explosion)	C_3 or C_4 (according to the released quantity)
Flash fire	C_3
Toxic cloud	C_3 or C_4 (according to the risk phrases – C_4 for very toxic substances)
Fire	C_2
Missiles ejection	C_3
Overpressure generation	C_3

Fireball	C_4
Environmental damage	To judge on site
Dust explosion	C_2 or C_3 (according to the substance and the quantity)
Boilover and resulting poolfire	C_3

5.2.4 Barrier maintenance optimization

To make decisions on maintaining existing protection systems, which consist of a set of safety barriers and/or security barriers, an economic analysis is recommendable since the budget of a company for safety and security purposes is always limited (Chen & Reniers, 2021). This section illustrates the combination of cost-effectiveness analysis (CEA) and genetic algorithm (GA) for barrier maintenance optimization.

1) Cost-effectiveness analysis under constraints

The trade-off between safety and economy is a practical problem faced by chemical companies. There are a couple of methods that are useful to address the trade-off between safety and economy (Reniers & Van Erp, 2016). One of them is cost-effectiveness analysis (CEA), which has the advantages in conducting comparative studies and no need to monetize accident costs. The effectiveness of a strategy in CEA can be any safety indicator based on the preferences of decision-makers. In QRA-based CEA, the effectiveness of a candidate strategy is measured by risk-associated indicators (e.g. risk reduction of a specific accident scenario).

Two typical practices for conducting CEA with constraints are i) a minimum acceptable level of effectiveness (Eff_{min}) and ii) a maximum acceptable use of the safety budget (Bu_{max}). The first constraint applies to situations where a company has to reduce the risks below certain levels, corresponding to making the effectiveness of safety investment above certain levels. The second constraint applies to a company that only has a limited budget that can be used for safety investment. Those two constraints usually need to be matched with different objective functions. The two kinds of optimization problems are presented as follows (Reniers & Van Erp, 2016):

$$\begin{cases} Min(C_i) \\ Eff_i \geq Eff_{min} \\ i \in \{1,2,3,\cdots,N\} \end{cases} \tag{5.7}$$

or:

$$\begin{cases} Max(Eff_i) \\ C_i \leq Bu_{max} \\ i \in \{1,2,3,\cdots,N\} \end{cases} \tag{5.8}$$

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 of the implementation of strategy i. The effectiveness can be an indicator associated with safety and/or security risks. In this study, the effectiveness of implementing a barrier maintenance strategy is measured by the corresponding risk reduction regarding specific accident scenarios.

2) Optimization algorithm

Regarding barrier maintenance optimization, a series of candidate strategies should be formulated at first. If only a limited number of candidate strategies are formulated, the best strategy can be obtained through an exhaustive search. Otherwise, evolutionary algorithms (for instance genetic algorithms) help to solve the optimization problem with a large solution space. Because there are usually thousands or even millions of strategies concerning the variations in maintenance interval of each barrier, the application of evolutionary algorithms becomes necessary. For instance, the maintenance interval of a barrier can vary from the shortest time step (1 h in this study) to the maximum maintenance interval defined by users. It becomes unreasonable to assess all the maintenance strategies by an exhaustive method.

By contrast, genetic algorithms (GA) have proven to be able to solve multivariable, nonlinear, and combinatorial optimization problems where the solution space can be huge and too vast to search exhaustively in a reasonable amount of time (Caputo et al., 2011). Generally, GAs have five steps: i) initial population, ii) fitness function, iii) selection, iv) crossover, and v) mutation. A detailed illustration of GA can be found in Caputo et al. (2011). Regarding the safety optimization problem, the GA minimizes the objective function with respect to all the constraints and determines the optimal strategy within the entire space of possible solutions. The procedures of employing GA for solving the above-mentioned two kinds of CEA optimization problems are shown in Figure 5.4. This study uses a genetic algorithm toolbox based on MATLAB to address the barrier maintenance optimization.

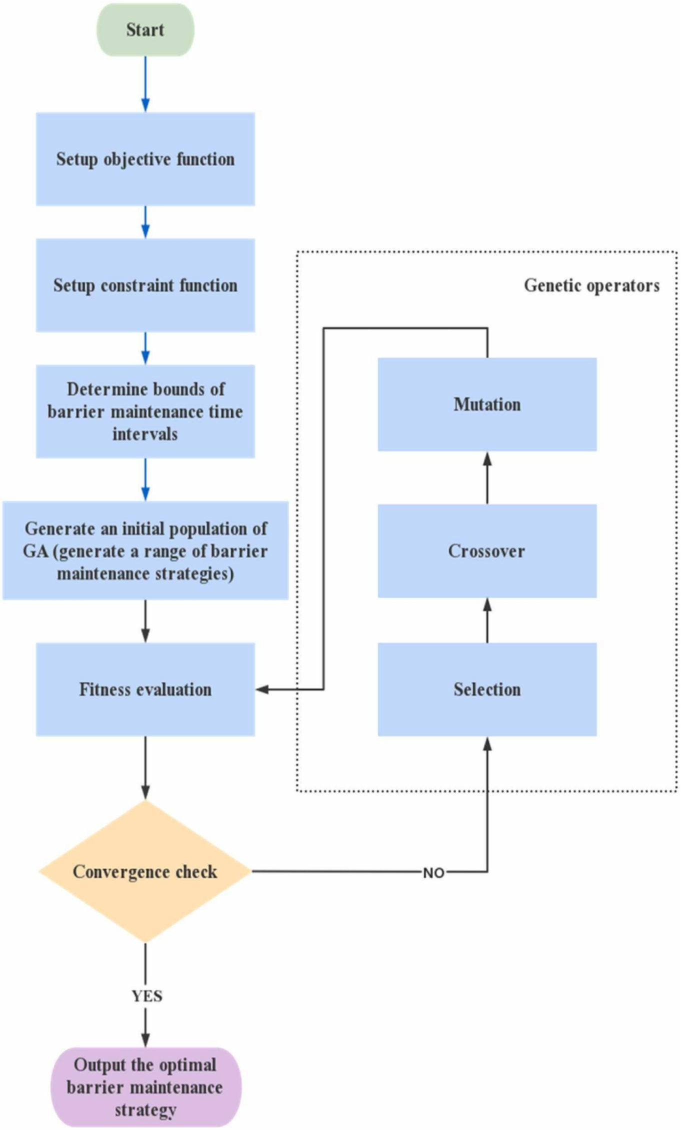

Start
Setup objective function
Setup constraint function
Determine bounds of barrier maintenance time intervals
Generate an initial population of GA (generate a range of barrier maintenance strategies)
Fitness evaluation
Convergence check
YES
Output the optimal barrier maintenance strategy
NO
Genetic operators
Selection
Crossover
Mutation

Figure 5.4. Genetic algorithm developed for safety and security barrier maintenance optimization.

5.3 A system simulation tool for barrier modeling

To implement the proposed approach in practice, a system simulation approach based on the MATLAB Simulink platform (Chaturvedi, 2017) is developed to conduct barrier modeling and risk assessment. The Simulink-based barrier modeling can be developed based on the obtained bow-tie diagram from step 1 (presented in section 5.2.2). The inputs for the barrier modeling are failure data of the barrier components and occurrence probabilities of the initiating events. The output of the simulation is a risk matrix with respect to major adverse events (for instance, VCE, flashfire, toxic cloud, etc.). A mapping algorithm for converting a bow-tie diagram into a Simulink-based barrier model is given in Figure 5.5. By following this mapping algorithm, the obtained accident scenarios presented by a bow-tie diagram can be transformed into a system simulation model, as shown in Figure 5.6.

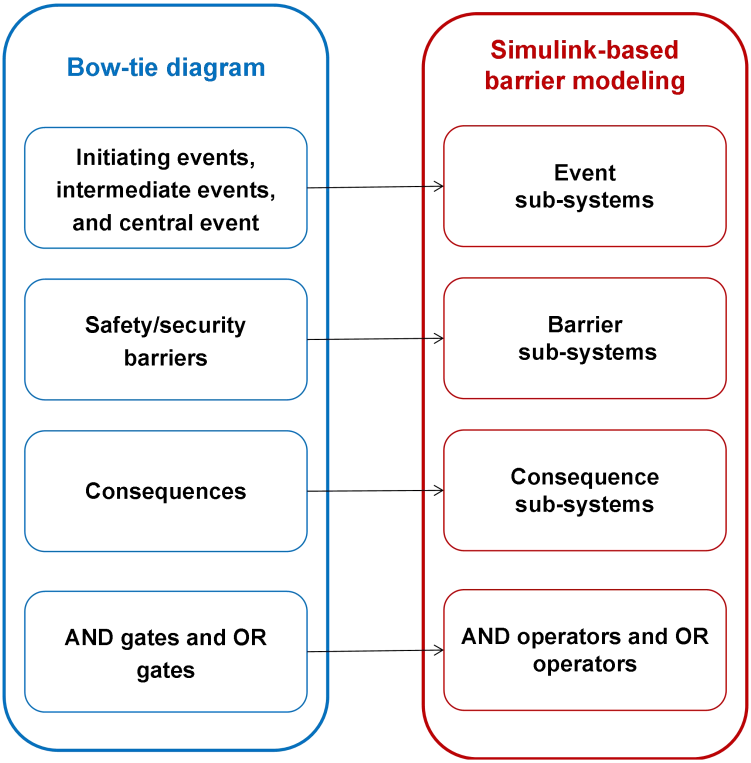


Figure 5.5. A mapping algorithm for converting a bow-tie diagram into a Simulink-based barrier model.

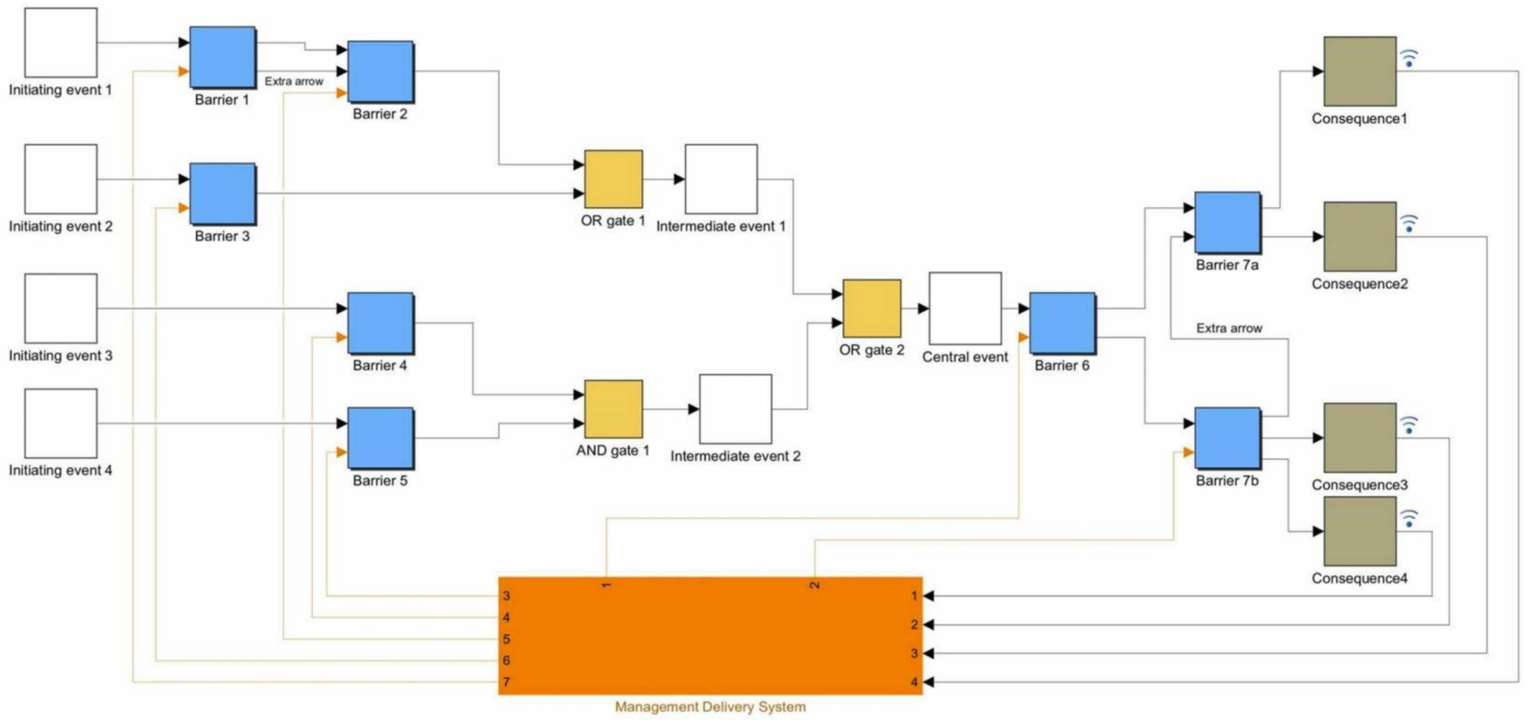


Figure 5.6. A Simulink model for barrier modeling.

All events, barriers, and consequences in the bow-tie diagram become sub-systems in the barrier modeling approach. “Event” sub-systems contain the frequencies/probabilities of such events happening. “Consequence” sub-systems contain information associated with the frequencies/probabilities and the severities of such consequences. In this study, the consequence assessment method proposed by the ARAMIS project is incorporated into the “consequence” sub-systems. “Barrier” sub-systems aim to calculate the time-dependent PFDs of such barriers. The arrows in the barrier model mainly transport probabilities, thus achieving a quantitative probability assessment. The basic rules for the probability calculation are adapted from the fault tree (Haasl et al., 1981) and event tree (Andrews & Dunnett, 2000), including the logical operators: AND gate and OR gate. PFDs of barriers can be determined according to the methods illustrated in section 5.2.3. The fault tree analysis of barriers can be performed based on the Simulink platform and incorporated into the “barrier” sub-system. For instance, the fault tree analysis of a schematic barrier system is shown in Figure 5.7. If multiple barriers use shared components, extra arrows should be used to link such barriers and transport necessary parameters (PFDs in this study) to ensure the correlations among barriers is considered (as mentioned in section 5.2.3). A barrier can also be placed on multiple branches on the right-hand side of the bow-tie. For instance, Barrier 7a and Barrier 7b in Figure 5.6 demonstrate the same barrier. In that case, extra arrows should link those barriers and transport time-dependent PFD values to ensure consistent PFDs are used.

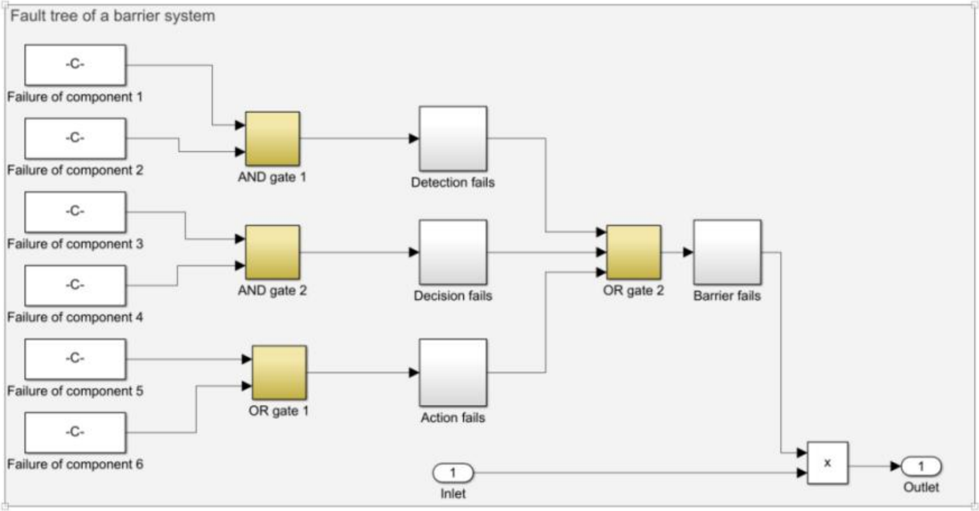


Figure 5.7. Fault tree analysis of a barrier system performed inside the “barrier” sub-system.

Our previous study defined management delivery systems (MDS) as a set of organizational and management factors that can prevent or mitigate undesired events indirectly by enhancing/maintaining the performance of scenario-specific barriers (Yuan et al., 2022a). In this study, a sub-system named “management delivery system” is used to tackle several tasks: i) collect risk assessment results, including both the probabilities/frequencies and severities of the consequences; ii) determine time intervals for barrier maintenance and give instructions to barrier sub-systems; iii) input PFDs for human barriers or human components of a barrier system. A summary of the features/tasks of each kind of sub-system in the Simulink-based barrier modeling is presented in Table 5.4. Due to the flexibility and compatibility of the MATLAB/Simulink simulations, various optimization algorithms (exhaustive search algorithms, evolutionary algorithms, etc.) can be integrated with the Simulink-based barrier modeling for barrier maintenance optimization. This study uses a genetic algorithm toolbox based on MATLAB (Mathworks, 2022a).

Table 5.4 Features/tasks of each sub-system in the Simulink-based barrier modeling.

Sub-systems	Features/tasks	Sub-systems	Features/tasks
Event	Event sub-systems contain and transport the frequencies or probabilities (or probability distributions in case of handling uncertainty propagation) of such events happening.	Barrier	Barrier sub-systems aim to determine the PFDs of such safety barriers and calculate and output the probabilities/frequencies for outlet branches. For complex safety barrier systems, the PFDs can be calculated with the help of fault tree analysis or reliability block diagrams.
Consequence	Consequence sub-systems contain and transport information associated with both the frequencies/probabilities and the severities of such consequences.	MDS	MDS sub-systems collect information, including both the probabilities/frequencies and severities of the consequences, from the consequence sub-systems and transport necessary parameters to barrier sub-systems for safety barrier configurations and PFD calculation. Decision-making modules can be incorporated into the MDS sub-systems.
AND operator	AND operators receive frequencies/probabilities from the inlets and calculate and output the frequencies/probabilities for the outlet by following an "AND" logic.	OR operator	OR operators receive frequencies/probabilities from the inlets and calculate and output the frequencies/probabilities for the outlet by following an "OR" logic.

5.4 Case study

5.4.1 Scenario building

A typical chemical reactor with its SCADA (supervisory control and data acquisition) system is investigated in this case study. The basic process control system of this reactor is adapted from Abdo et al. (2018), while an ESD system is considered as a system independent of the basic process control system, as shown in Figure 5.8. This reactor is used to run a chemical reaction in order to produce product C from two reactants A and B. We assumed that this reactor is used to produce a flammable liquid with toxicity, for instance, propylene oxide. The ESD system controls the block/shutdown valves (XV33012 and XV33013) in case of over-pressure based on the monitored pressure inside the reactor. The basic process control system includes a feeding system and a cooling system, which are controlled by PLC1 and PLC2, respectively. The temperature of the reaction is regulated with industrial water, and the temperatures of the water inside the cooling jacket and at the inlet are measured by the sensor TI and T2, respectively. The data collected by these two sensors is sent to PLC2, which regulates the water flow rate by controlling pumps (P1 and P2, P2 is a standby pump) and valves (CV33063 and XYSV33027). The physical components (valves, pumps, etc.) of the basic process control system are controlled by PLCs and supervised by a SCADA system. Site managers can access the information collected by the SCADA system and control the reaction process remotely inside the control center.

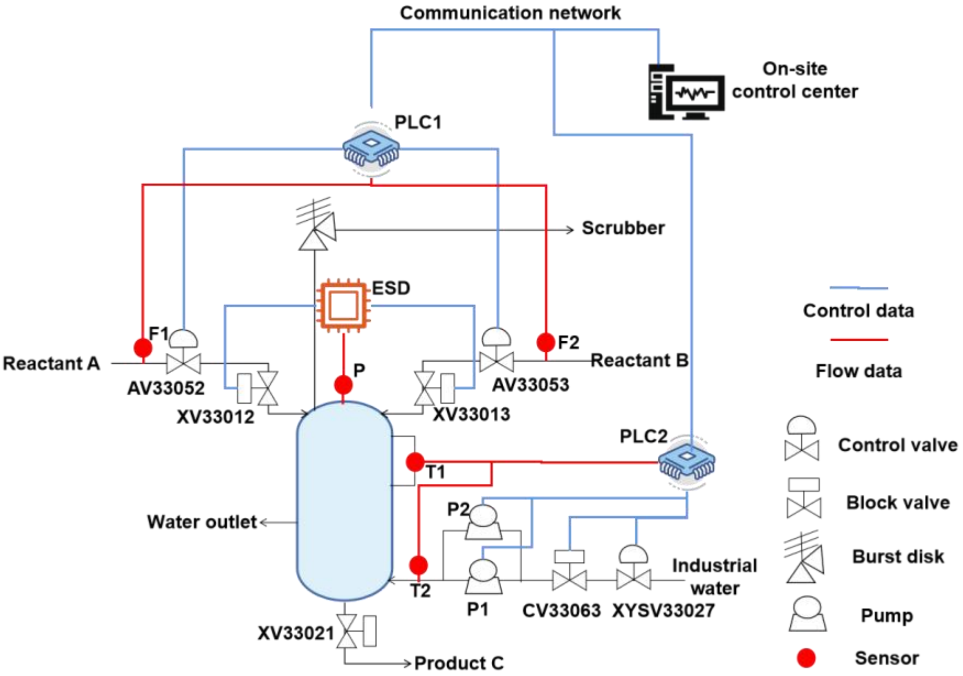


Figure 5.8. A chemical reactor with its SCADA system, adapted from (Abdo et al., 2018).

The liquid leakage-associated scenarios in terms of safety failures and malicious acts were built using a bow-tie diagram, which consists of a fault tree and an event tree. Associated safety and security barriers were identified according to the database/checklists from (Andersen et al., 2004), (Argenti et al., 2017), and (Guzman et al., 2021). A dedicated assessment of the adequacy of barriers to prevent each attack event was conducted to determine the appropriate placement of attack events and their associated barriers. For attacks executed by malicious insiders within the control center, security barrier B1 (entrance control system), along with OT safety barriers B9 (ESD), B10 (manual shutdown), and B11 (burst disk), can function to prevent the intrusion of malicious insiders and mitigate the risk of shell rupture caused by overpressure from insiders' manipulations. Consequently, this attack branch is placed before the OT safety barriers (B9 to B11). In contrast, OT safety barriers B9 to B11 cannot prevent direct physical sabotage or attacks on the reactor shell by external attackers. Therefore, these OT safety barriers are considered overridden by external physical attacks, and only security barriers B6 (entrance control) and B7 (guard response) are considered relevant in such cases. As a result of this assessment, a bow-tie diagram was constructed to illustrate both safety-related and security-related scenarios, identifying the locations of attack events and barriers. The resulting bow-tie diagram is presented in Figure 5.9 (fault tree) and Figure 5.10 (event tree). The explanations of those barriers are given in Table 5.5.

Table 5.5 Explanations of barriers in the bow-tie diagram.

Marks	Barriers	Marks	Barriers
B1	Entrance control system (unsupervised automatic credentials check)	B2	Training and authorization before work
B3	Inspection of supervisory computers	B4	Inspection of feeding system
B5	Inspection of cooling system	B6	Entrance control (unsupervised automatic biometrics check)
B7	Guard response	B8	Fire protection system
B9	Emergency shutdown system (ESD)	B10	Manual shutdown (MD)
B11	Burst disk	B12	Foam injection

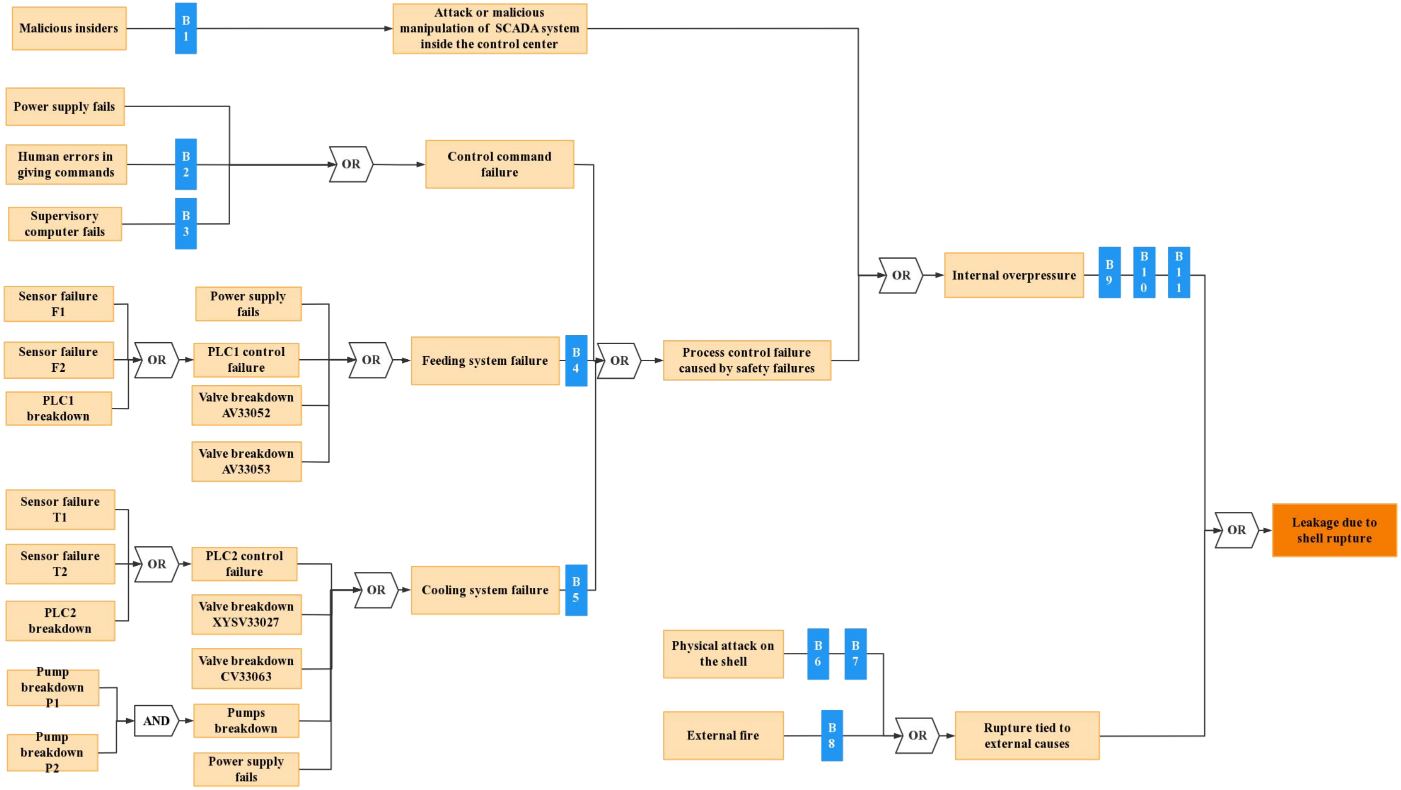


Figure 5.9. Fault tree of the chemical reactor with propylene oxide leakage as the top event.

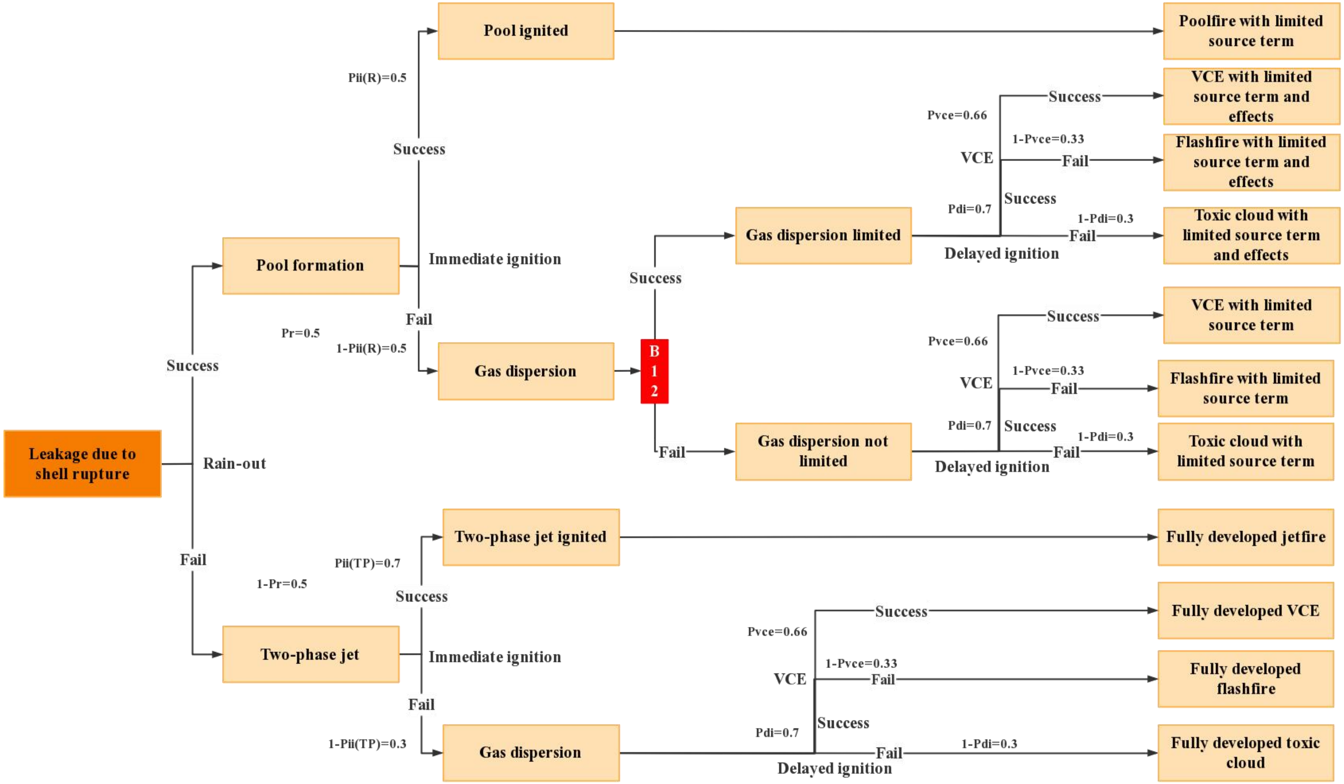


Figure 5.10. Event tree of the chemical reactor with propylene oxide leakage as the initiating event, adapted from (Andersen et al., 2004).

5.4.2 Barrier modeling configurations and results

In this case study, frequencies of the basic events are retrieved from other studies or datasets and are given in Table 5.6. The frequency of adversary attacks is adapted from (Landucci et al., 2017), in which the annual attack probability for chemical facilities in Italy was investigated. Configurations of the associated safety and security barriers and barrier components, including their failure rates/PFDs, maintenance time, and initial maintenance intervals, are given in Table 5.7. For simplification purposes, the maintenance times of all technical barrier components are set as 8 h. In practice, the barrier maintenance time can be configured according to the practical experience of workers. Constant PFDs are used to describe the performance of security barriers and human barriers/human actions due to the lack of historical data and the difficulties in formulating the time-varied PFDs for such barriers. Failure probabilities of the security barriers are mainly retrieved from (Argenti et al., 2017). The maintenance of all technical components of safety barriers is considered in the barrier modeling. With more data and studies on the evaluation of the time-dependent performance/PFDs of security barriers becoming available, the integrated optimization of safety and security barrier maintenance can also be achieved by employing the proposed methodology.

Table 5.6 Frequencies of basic events in the barrier modeling.

Events	**Frequencies (y^{-1})**	**Events**	**Frequencies (y^{-1})**
Malicious insiders	3.3E−02 (Landucci et al., 2017)	Physical attacks	3.3E−02 (Landucci et al., 2017)
Power supply fails	1.00E-01 (Çetinkaya, 2001)	Human error in giving commands	1.00E-02 (Andersen et al., 2004)
Supervisory computer fails	5.00E-04 (Çetinkaya, 2001)	PLC1 breakdown	4.38E-02 (Hauge & Onshus, 2010)
Sensor failure F1	3.50E-01 (Debray et al., 2004)	Sensor failure F2	3.50E−01 (Debray et al., 2004)
Valve breakdown AV33052	4.00E-02 (Taylor, 2010)	Valve breakdown AV33053	4.00E-02 (Taylor, 2010)
PLC2 breakdown	4.38E-02 (Hauge & Onshus, 2010)	Pump breakdown P1	3.125E-02 (OREDA, 2002)
Pump breakdown P2	3.125E-02 (OREDA, 2002)	Valve breakdown XYSV33027	4.00E-02 (Taylor, 2010)
Valve breakdown CV33063	4.00E-02 (Taylor, 2010)	Sensor failure T1	2.13E-02 (Hauge & Onshus, 2010)

Sensor failure T2	2.13E-02 (Hauge & Onshus, 2010)	External fire	5.52E-02 (Debray et al., 2004)

The developed barrier model is shown in Figure 5.11. In order to simplify the barrier modeling, the basic events for calculating feeding system failure frequency and cooling system failure frequency are not presented in the barrier model. The frequencies of feeding system failure and cooling system failure are calculated using fault tree analysis considering the frequencies of their associated basic events. With the combination of the consequences class proposed by the ARAMIS project, the calculated yearly-average frequency and severity of each consequence are presented in a risk matrix, as shown in Figure 5.12. A list of the consequences in the risk matrix is illustrated in Table 5.8.

Table 5.7 Configurations of safety and security barriers.

NO.	Barriers	Barrier components	Failure rates (/h)	PFDs	Maintenance time	Initial maintenance intervals	One-time maintenance cost[13]
1	Entrance control system (unsupervised automatic credentials check)	/	/	1.0E-02 (Argenti et al., 2017)	/	/	/
2	Training and authorization before work	/	/	1.0E-02 (Andersen et al., 2004)	/	/	/
3	Inspection of supervisory computers	/	/	1.0E-01 (Andersen et al., 2004)	/	/	/
4	Inspection of feeding system	/	/	1.0E-01 (Andersen et al., 2004)	/	/	/
5	Inspection of cooling system	/	/	1.0E-01 (Andersen et al., 2004)	/	/	/
6	Entrance control (unsupervised automatic biometrics check)	/	/	1.0E-02 (Argenti et al., 2017)	/	/	/
7	Guard response	Alarm assessment through CCTV system	/	3.0E-02 (Argenti et al., 2017)	/	/	/
		Communication to response force	/	5.0E-02 (Argenti et al., 2017)			
		Guard force response	/	1.62E-02 (Song et al., 2019a)			
8	Fire protection system	Smoke/combustion detector	4.12E-06 (OREDA, 2002)	/	8 h	500 h	150 €
		Programmable logic solver	1.0E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	300 €
		Fire pump	7.2E-05 (Gravestock, 2008)	/	8 h	500 h	300 €
		Deluge Valve	5.8E-06 (Gravestock, 2008)	/	8 h	500 h	200 €
9	Emergency shutdown system (ESD)	Pressure sensor*[14]	1.5E-07 (Hauge & Onshus, 2010)	/	8 h	500 h	150 €
		Programmable safety system	1.0E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	300 €
		Shutdown valve XV33012*	3.5E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	200 €

[13] One-time maintenance costs exclude downtime costs.

[14] A barrier component with * means it is a shared component.

		Shutdown valve XV33013*	3.5E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	200 €
10	Manual shutdown (MD)	Pressure sensor*	1.5E-07 (Hauge & Onshus, 2010)	/	8 h	500 h	150 €
		Human action	/	1.0E-02 (Andersen et al., 2004)	/	/	/
		ESD Push Button	1.2E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	100 €
		Shutdown valve XV33012*	3.5E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	200 €
		Shutdown valve XV33013*	3.5E-06 (Hauge & Onshus, 2010)	/	8 h	500 h	200 €
11	Burst disk	/	2.3E-05 (Lees, 1980)	/	8 h	500 h	200 €
12	Foam injection	Human response/intervention	/	1.0E-01 (Andersen et al., 2004)	/	/	/
		Injection pump	2.31E-06 (OREDA, 2002)	/	8 h	500 h	300 €
		Injection valve	1.862E-05 (OREDA, 2002)	/	8 h	500 h	200 €

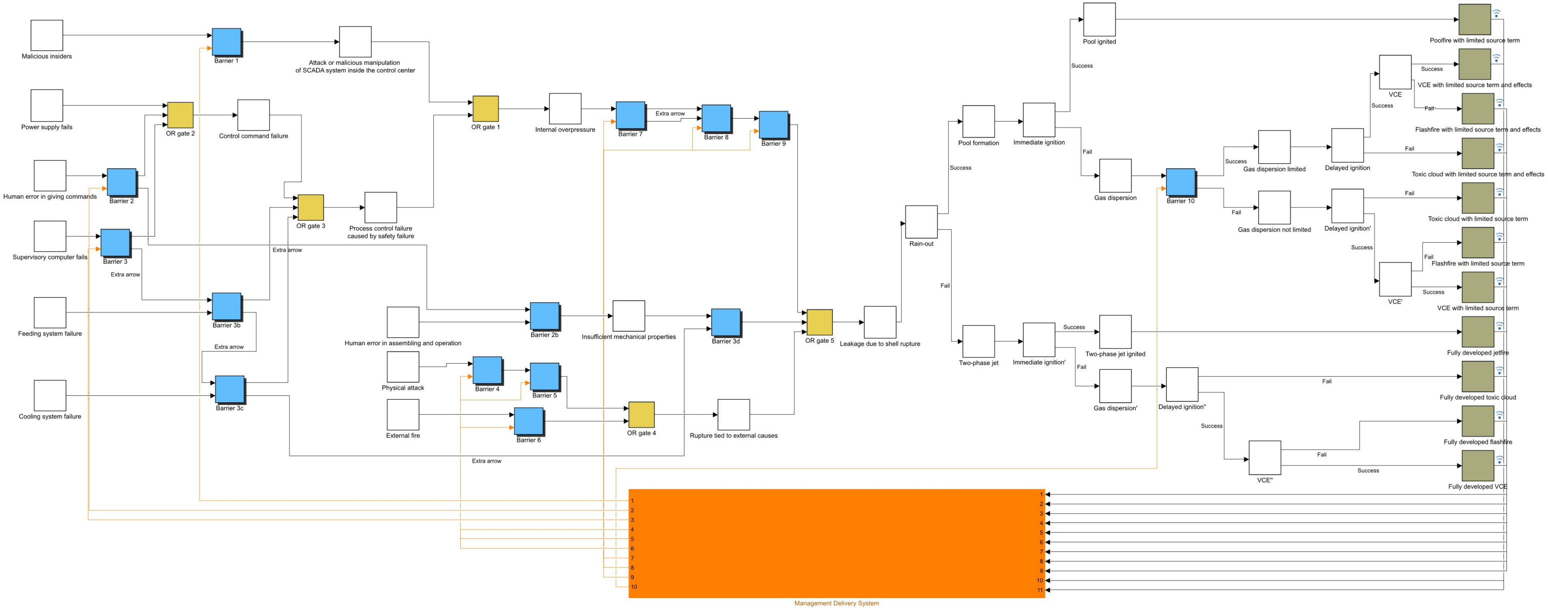


Figure 5.11. Barrier modeling with respect to flammable liquid leakage scenarios caused by chemical reactor shell rupture.

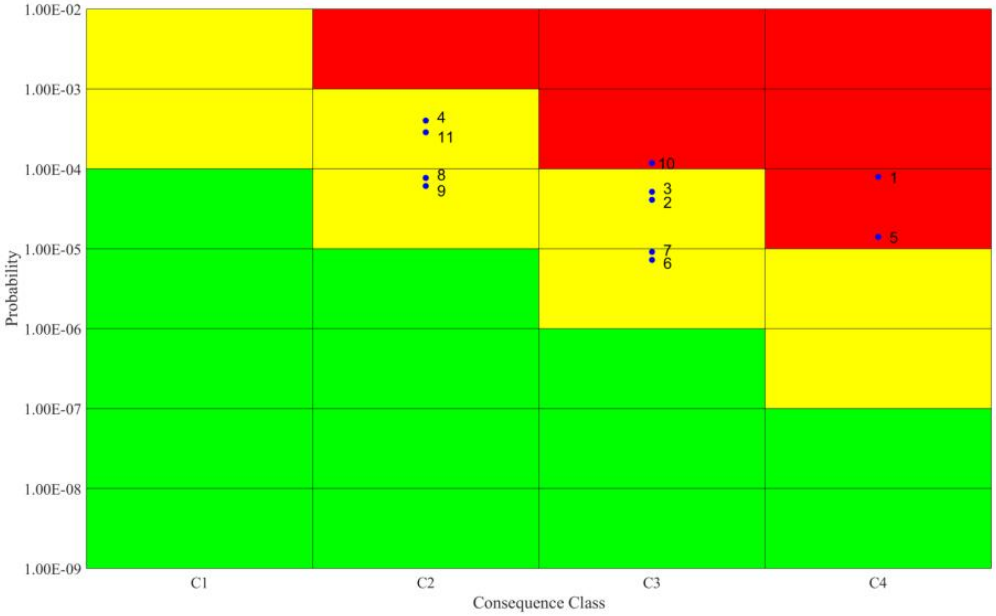


Figure 5.12. The obtained risk matrix (Note: numbers represent the consequence numbers, which are explained in Table 5.8).

Table 5.8 Table of consequences in the risk matrix.

Number	Consequence	Class	Number	Consequence	Class
1	Fully developed VCE	C_4	2	Fully developed flashfire	C_3
3	Fully developed toxic cloud	C_3	4	Fully developed jetfire	C_2
5	VCE with limited source term	C_4	6	Flashfire with limited source term	C_3
7	Toxic cloud with limited source term	C_3	8	Toxic cloud with limited source term and effects	C_2
9	Flashfire with limited source term and effects	C_2	10	VCE with limited source term and effects	C_3
11	Poolfire with limited source term	C_2	/	/	/

As shown in Figure 5.12, major adverse events are the dots with numbers 1, 5, and 10, corresponding to “fully developed VCE”, “VCE with limited source term”, and “VCE with limited source term and effects”. Since those dots are in the red region, which means they are unacceptable. Barrier maintenance strategy should be improved to ensure that all risks are situated in the yellow region (acceptable with mitigation) or green region (acceptable). The next section elaborates on how to achieve this optimization using the proposed GA-based method.

5.4.3 Barrier maintenance optimization

1) Cost analysis of barrier maintenance

To optimize the existing barrier maintenance strategy, a cost analysis of a series of candidate strategies should be conducted. Then, it is possible to determine the most cost-effective strategy through cost-effectiveness analysis (CEA). Reniers and Van Erp (2016) illustrated eight cost categories for protection measures, and the maintenance cost includes the costs for material, maintenance team, production loss, and start-up. Due to the difficulties in obtaining all the costs for barrier maintenance, we assumed one-time maintenance costs based on the product purchase prices (for small technical components, the cost of the maintenance team may take the main part). In some situations, the maintenance of safety barriers has to break off the production process (Wu et al., 2022). Therefore, a downtime cost should be considered. We assumed that the downtime cost per hour is 10,000 € and the downtime cost only applies to the ESD system in this case study. We list the maintenance costs (excluding downtime costs) for all the technical barrier components in Table 5.6. In practice, those costs should be configured according to the real expenses.

2) GA-based barrier maintenance strategy optimization

According to the obtained risk matrix, the risks of “fully developed VCE”, “VCE with limited source term”, and “VCE with limited source term and effects” are not acceptable. Therefore, the optimization objective is to minimize barrier maintenance costs with the constraints that ensure all risks are at least in the yellow region in the risk matrix. The objective function to be minimized is:

$$C = \sum_{i=1}^{n} U_i * N_i \tag{5.9}$$

where C is the annual total cost of barrier maintenance that can be calculated by summing the maintenance cost for each barrier. n is the number of barriers that need to be maintained. U_i is the one-time maintenance cost for barrier i and N_i is the number of maintenance of barrier i in one year. The nonlinear inequality constraints are:

$$\begin{cases} P_j \leq TS_j \\ j \in \{1,2,3,\cdots,N\} \end{cases} \tag{5.10}$$

where P_j is the probability of consequence j. TS_j is the threshold for consequence j. N is the number of consequences in the risk matrix. The thresholds were configured according to the boundaries of the yellow region in the risk matrix. Bounds of the barrier maintenance intervals were set as 1 h ~ 500 h. An integer constraint was applied to barrier maintenance intervals, which means that the maintenance intervals have to be integers.

A genetic algorithm toolbox based on MATLAB R2022a was used to solve this optimization problem. This toolbox is capable of solving smooth and non-smooth optimization problems with different types of constraints, including integer constraints. It searches the optimal strategy randomly by mutating and crossovering among a large number of population members. More instructions on how to use this toolbox can be found in (Mathworks, 2022a). The optimization results are shown in Figure 5.13. It can be observed that the individual penalty values (annual costs of barrier maintenance) are distributed relatively randomly after a few generations. The best penalty value (minimal annual maintenance cost) is 1479150€ after more than 250 generations, which is the annual cost corresponding to the optimal barrier maintenance strategy. Meanwhile, the mean penalty value at the end is 6572570€, which is the average cost of all individual strategies. The obtained optimal strategy for barrier maintenance is presented in Table 5.9.

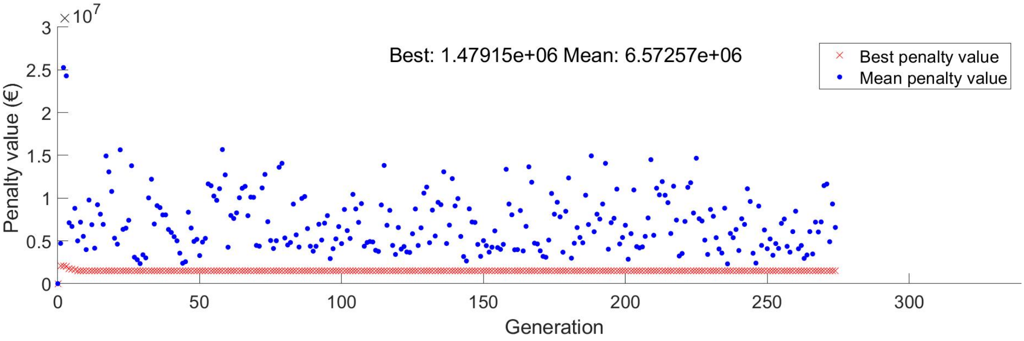


Figure 5.13. Calculation results of the genetic algorithm

Table 5.9 Optimal strategy for barrier maintenance.

Barriers	Technical components	Maintenance intervals (h)	Barriers	Technical components	Maintenance intervals (h)
Fire protection system	Smoke/combustion detector	137	Emergency shutdown system (ESD)	Pressure sensor	489
	Programmable logic solver	411		Logic solver	
	Fire pump	33		Shutdown valve XV33012	
	Deluge Valve	135		Shutdown valve XV33013	
Manual shutdown (MD)	ESD Push Button	479	Foam injection	Injection pump	485
				Injection valve	485
Burst disk	/	489	/	/	/

5.5 Discussion

5.5.1 Benefits of the GA-based barrier maintenance optimization

Compared to previous maintenance optimization approaches (such as RCM and RBI), the proposed approach has the advantage of integrating safety and security scenarios together for risk analysis and further optimizing the barrier maintenance strategy based on the synergistic effects of barriers on risk reduction, rather than evaluating each barrier according to its own probability of failure and consequence of failure. The implementation of GA addresses the large-solution-space optimization problems. In practice, barrier maintenance optimization is an optimization problem involving multivariable that may be unrealistic and unreasonable to search all strategies exhaustively. By using GA, the barrier maintenance optimization problems under economic constraints and technical constraints can be solved within affordable computation times. Compared with directly reducing the maintenance intervals of all barriers, implementing the proposed approach can achieve the same goal at a lower cost. For instance, the risks of all possible consequences are at acceptable levels by reducing the maintenance/inspection interval of all barriers to 45 h. The annual cost for barrier maintenance will be 13.63 M€ (including downtime cost). By contrast, the annual cost of barrier maintenance is 1.48 M€ (including downtime cost) using the proposed GA-based approach to ensure the risks are at acceptable levels. It shows that a large amount of the barrier maintenance cost may be saved by using the GA-based approach.

5.5.2 Recommendations for future work

i) In the proposed approach, barrier maintenance strategy is optimized based on the probabilistic risk assessment results, which also means uncertainty is inevitably involved in the approach. Selvik & Aven (2011) emphasized the importance of the identification and assessment of the uncertainty factors associated with the assumptions made in the reliability centered maintenance (RCM). Similarly, the obtained optimal strategy does not mean perfectly safe with saving costs due to the uncertainties involved in the proposed approach. An uncertainty analysis may be performed when applying the proposed approach in practice. The alleviation, assessment, and treatment of uncertainties in barrier maintenance optimizations may be focused on in future works.

ii) In this paper, exponential distributions were used to describe the time-varied PFDs of safety barriers. This is a relatively rough assumption and can be replaced by some more advanced models. For instance, the model with the consideration of multi-state transition of safety barriers (Wu et al., 2022), the model considering barrier degradation caused by aging degradation and damage caused by shocks (Pishro-Nik, 2016), and the model considering a series of intermediate factor (operation time, temperature, wind speed, pressure, and humidity) (Ouache et al., 2015). The integration of more sophisticated models for describing barrier degradation into the proposed approach helps to improve the accuracy of the results.

iii) Due to the lack of data related to security barriers, evaluating such barriers is challenging, and thus, the maintenance of such barriers is not considered in this study. With more data related to the performance of security barriers available, the quantitative assessment and maintenance/inspection optimization of security barriers will also be possible by employing the proposed approach.

5.6 Conclusions

This study investigates possible optimal barrier maintenance intervals concerning both safety hazards and security threats in chemical plants from a cost-effectiveness perspective. The results show that the combination of bow-tie diagrams and the Simulink-based barrier modeling is effective for risk assessment of accident scenarios considering the synergistic effects of barriers on risk reduction. Cost-effectiveness analysis and genetic algorithm can be combined to determine the optimal barrier maintenance intervals under economic constraints. It is possible to obtain acceptable risk levels with much lower costs for barrier maintenance using the proposed approach. The proposed barrier modeling approach has the potential to be implemented for quantitative and semi-quantitative risk assessment of various accident scenarios in terms of safety and security due to its flexibility and scalability.

Chapter 6 Dynamic safety barrier management based on data from multiple sources

A simulation approach is proposed in this chapter to conduct a dynamic risk assessment of chemical facilities considering the degradation of safety barriers and achieve dynamic management of safety barriers. In the proposed approach, multiple data (periodic proof test data, continuous condition-monitoring data, and accident precursor data) are combined to update barrier failure probabilities and risk profiles. Meanwhile, uncertainty propagation in probabilistic risk assessment is handled using Monte Carlo simulations. Cost-effectiveness analysis (CEA) is employed to support decision-making on safety barrier establishments and improvements based on dynamic risks. An illustrative case study is demonstrated to validate the feasibility of using the proposed approach in risk-based dynamic safety barrier management.

This chapter is drafted with modifications based on the following publication:

◆ Yuan, S., Reniers, G., & Yang, M. (2023). Dynamic-risk-informed safety barrier management: An application to cost-effective barrier optimization based on data from multiple sources. *Journal of Loss Prevention in the Process Industries, 83,* 105034.

6.1 Introduction

In terms of safety barrier management, Johansen & Rausand (2015) discussed the main principles related to barrier management in the offshore oil and gas industry. CCPS (USA) and Energy Institute (UK) developed guidance on employing bow-tie diagrams to facilitate safety barrier management through the proper depiction and allocation of safety barriers (CCPS/EI, 2018). Similarly, bow-tie diagrams were employed to support accident process monitoring and barrier alarm management based on the inspection of barrier status (Schmitz et al., 2020; 2021). By integrating safety barrier assessment into a QRA framework, the effectiveness of barriers can be reflected by how much risk can be reduced by implementing such barriers and further making decisions to achieve risk-based safety barrier management.

Moreover, the dynamic barrier management concept was introduced by Pitblado et al. (2016), who suggested the use of multiple data sources to determine near-real-time barrier status. In dynamic barrier management approaches, updating safety barrier status and risk profiles is necessary when new information becomes available over time. Bayes' theorem was introduced to achieve dynamic risk assessment of industrial systems by using near misses and incident data to update accident likelihood and the barrier failure probabilities (Kalantarnia et al., 2009; Khakzad et al., 2012). There are also some attempts to use condition-monitoring data for dynamic risk assessment (DRA). For instance, condition-monitoring data was employed to estimate the degradation states of chemical process systems with the help of Kalman filtering (Zadakbar et al., 2013a), particle filtering (Zadakbar et al., 2015), and principal component analysis (Zadakbar et al., 2013b). Zeng & Zio (2018) integrated statistical failure data and condition-monitoring data for dynamic risk assessment of an industrial system using a Bayesian updating algorithm. Additionally, probability distributions are widely used to interpret uncertainty in QRA and DRA (Yazdi et al., 2019). Thus, the handling of uncertainty propagation in DRA/QRA should be properly addressed to facilitate risk-based safety barrier management.

However, the decision-making methodologies for cost-effective safety barrier management are still challenging, particularly considering the utilization of multiple data for risk updating. Therefore, this chapter proposes a systematic approach for performance assessment and dynamic management of safety barriers. The remaining sections of this chapter are organized as follows. Firstly, the proposed methodology is introduced in section 6.2. Then, an illustrative case study is employed in section 6.3 to demonstrate the application of the proposed approach in dynamic risk assessment and cost-effective safety barrier optimization. Followed by the discussions are given in section 6.4, and conclusions are presented in section 6.5.

6.2 Methodology

6.2.1 Overview of the proposed approach

A dynamic-risk-based safety barrier management approach is proposed in this chapter, and the flowchart of the proposed approach is illustrated in Figure 6.1. Simulink is a MATLAB-based graphical programming environment for modeling, simulating, and analyzing multidomain dynamical systems (Chaturvedi, 2017). Due to several advantages of Simulink, such as its user-friendly primary interface, customizable block libraries, and good compatibility with the rest of the MATLAB environment, a Simulink-based safety barrier modeling (SSBM) tool is developed for implementing the proposed approach in practice. A detailed elaboration on the proposed approach is given in the following sub-sections.

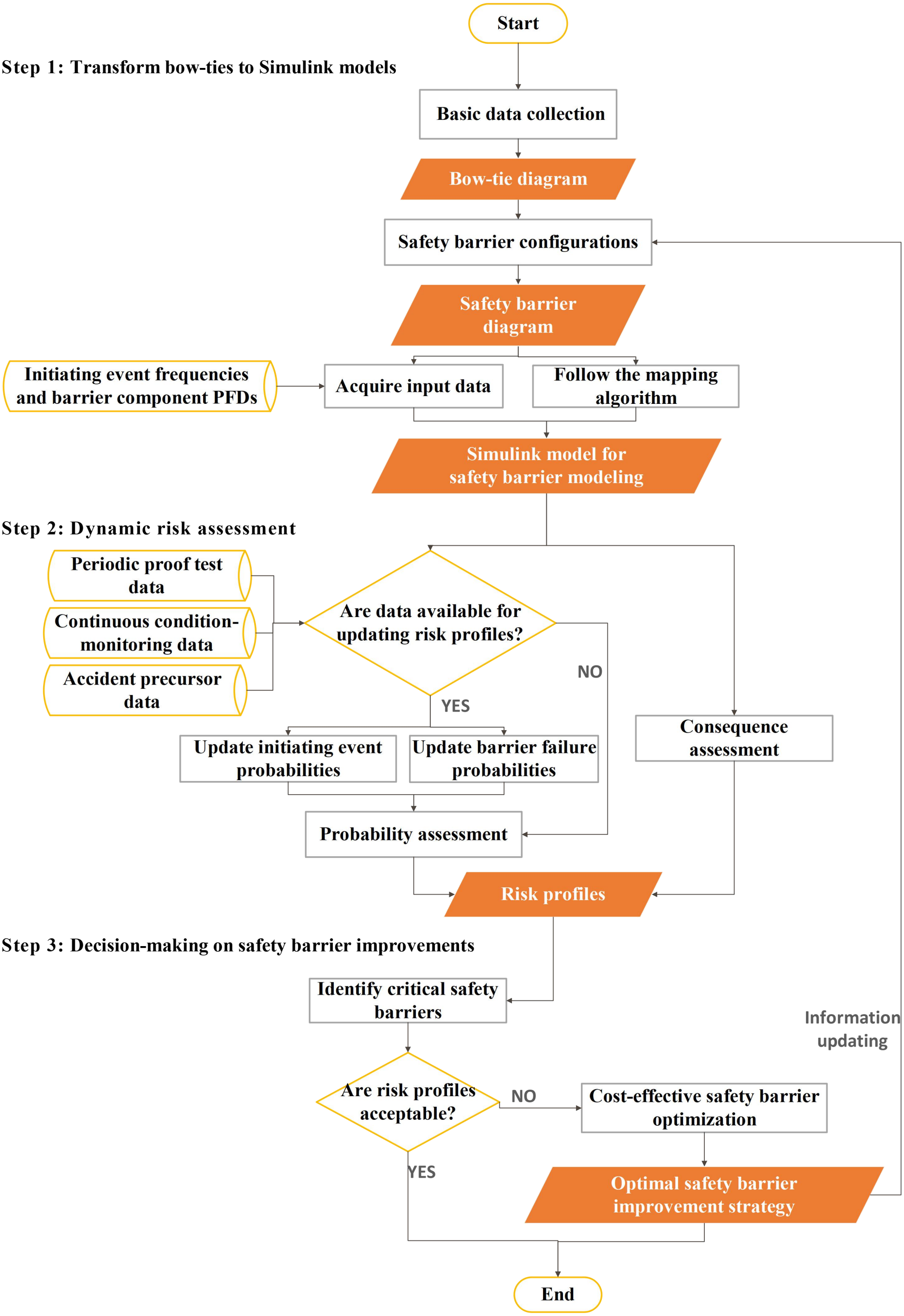

Start
Step 1: Transform bow-ties to Simulink models
Basic data collection
Bow-tie diagram
Safety barrier configurations
Safety barrier diagram
Initiating event frequencies and barrier component PFDs
Acquire input data
Follow the mapping algorithm
Simulink model for safety barrier modeling
Step 2: Dynamic risk assessment
Periodic proof test data
Continuous condition-monitoring data
Accident precursor data
Are data available for updating risk profiles?
NO
YES
Update initiating event probabilities
Update barrier failure probabilities
Consequence assessment
Probability assessment
Risk profiles
Step 3: Decision-making on safety barrier improvements
Identify critical safety barriers
Information updating
Are risk profiles acceptable?
NO
Cost-effective safety barrier optimization
YES
Optimal safety barrier improvement strategy
End

Figure 6.1. Flowchart of the dynamic-risk-based safety barrier management approach.

6.2.2 Transform bow-tie diagrams into Simulink models

This step aims to build accident scenarios with the consideration of the intervention of safety barriers. It is suggested to implement the bow-tie technique for HAZard IDentification (HAZID), then followed by constructing safety barrier diagrams (SBDs) by placing safety barriers on the paths of the bow-tie diagrams (Duijm, 2009).

1) Safety barrier diagrams

The safety barrier diagram (SBD) was introduced by Duijm (2009) as a safety management tool. It can be regarded as a modified version of the bow-tie diagram, and it has the advantages of i) relative simplicity that supports communication with non-expert stakeholders, ii) having deliberately inserted safety systems that support the management and maintenance of these systems, and iii) providing a useful framework for integrating information from risk analysis with operational safety management. A comparison between the conventional bow-tie diagram and the SBD is shown in Figure 6.2.

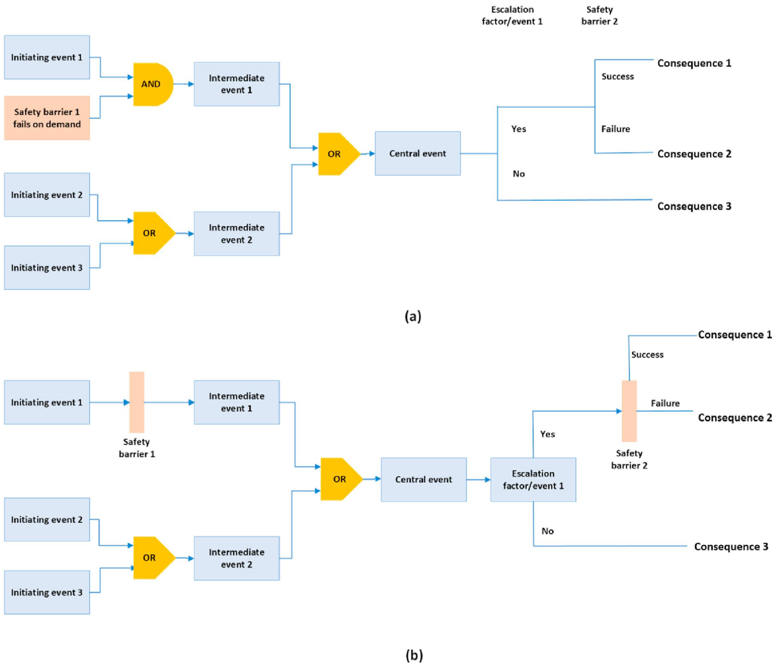


Figure 6.2. A comparison between the conventional bow-tie diagram (top) and the safety barrier diagram (bottom).

Among the intermediate events in an event tree, some can be presented as safety barriers because they are manageable. Although in some previous studies, all of those intermediate events were called safety barriers, we do not consider some (such as ignition and confined space nearby) as safety barriers. For those intermediate events that cannot be managed, we call them escalation factors/events. The calculation rules considering the failure of safety barriers are presented as follows:

$$P_{OUT1} = P_{IN} * PFD \tag{6.1}$$

$$P_{OUT2} = P_{IN} * (1 - PFD) \tag{6.2}$$

where P_{IN} is the input probability for the safety barrier. P_{OUT1} is the output probability of the branch with the condition that the safety barrier failed and P_{OUT2} is the output probability of the branch with the condition that the safety barrier functioned. PFD is the probability of failure on demand of this barrier. On the left-hand of the SBD, multiple barriers may be located before an undesired event on the same branch. In that case, the output probability of this branch can be calculated as follows:

$$P_{OUT} = P_{IN} * (PFD_1 * PFD_2 \cdots PFD_n) \tag{6.3}$$

where P_{OUT} is the output probability and P_{IN} is the input probability of this branch. PFD_1 to PFD_n denote the PFDs of safety barriers, and n is the number of barriers located on this branch. This equation is valid under the assumption that the occurrence of failure of each safety barrier in this branch is independent.

2) Mapping algorithm

This study employs a Simulink-based safety barrier modeling (SSBM) tool to conduct dynamic risk assessment and support safety barrier management. Previous studies already investigated the implementation of MATLAB/Simulink simulations for the calculation of fault trees (Latif-Shabgahi & Tajarrod, 2009; Papadopoulos & Maruhn, 2015) and the reliability analysis of safety instrumented systems (Ouache et al., 2015). To enhance the flexibility and adaptability of the SSBM, we suggest developing all the elements (exclude arrows/linkages) of the bow-tie diagram or SBD as sub-systems in the Simulink simulations. Then, the configuration of the elements can be made by developing specific simulation structures inside the sub-systems according to the needs of users. The basic rules for probability calculation are adapted from the fault tree (Haasl et al., 1981) and event tree (Andrews & Dunnett, 2000). A mapping algorithm for converting a bow-tie diagram (or a safety barrier diagram) into a

Simulink-based barrier model is illustrated in section 5.3 of this dissertation, and we avoid repeating illustrations here.

3) PFD calculation of safety barriers

The implementation of fault tree analysis (FTA) or reliability block diagram (RBD) helps to determine the PFD of a complex safety barrier system. In SSBM, a hierarchical structure for reliability analysis of safety barrier systems can be easily obtained by using hierarchical sub-systems due to the flexibility and adaptability of the Simulink platform. For instance, for a safety barrier system with the elements of 'detect-decide-act', the different functionalities should be achieved by using different components. The corresponding structure for this safety barrier system can be represented by a fault tree inside the "barrier sub-system", as shown in Figure 6.3. After the PFD of this safety barrier system was determined, the probabilities/frequencies for the outlet branches can be calculated according to Eqs (6.1) and (6.2).

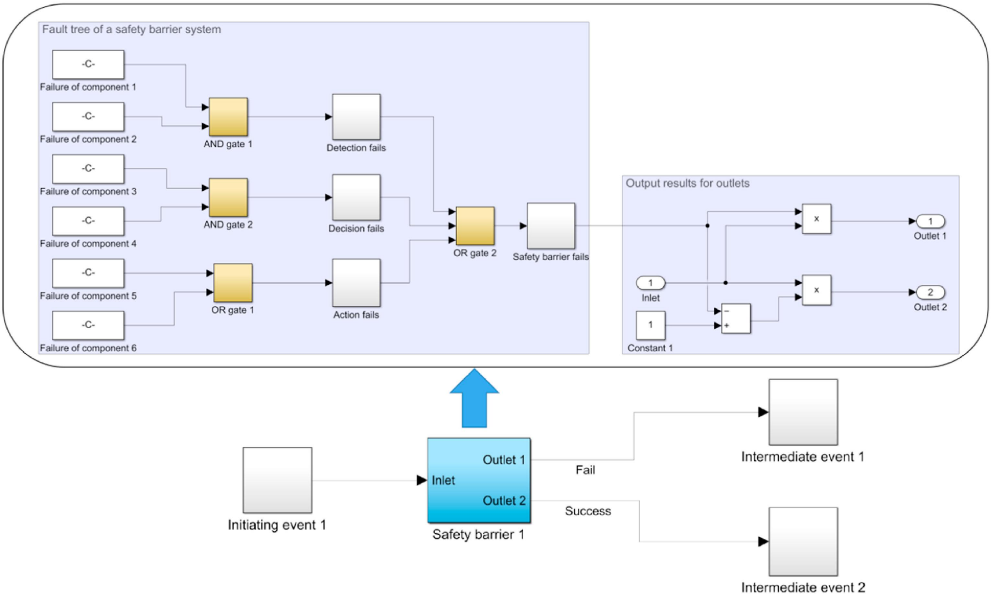


Figure 6.3. A safety barrier system illustrated using a barrier sub-system in the SSBM.

Additionally, in case several safety barriers have shared components, which are located on the same branch, the conditional probabilities should be used for the barriers excluding the first occurrence barrier (the barrier that may fail first). For instance, for two safety barriers with a shared component, a conditional probability $P_2^{'}$ should be used for barrier 2 given the failure of barrier 1. The conditional probability can be calculated as follows (Duijm, 2009):

$$P_{2,R} = \frac{P_2 - P_C}{1 - P_C} \tag{6.4}$$

$$\begin{aligned} P_2^{'} &= P(B_2\, fails \mid B_1\, has\, failed) \\ &= P_{2,R} + P(C\, fails \mid B_1\, has\, failed)\left[1 - P_{2,R}\right] \\ &= P_{2,R} + (P_C/P_1)\left[1 - P_{2,R}\right] \end{aligned} \tag{6.5}$$

where P_1 presents the PFD of the whole barrier 1, which contains a common component C with a PFD P_C. P_2 presents the PFD of the whole barrier 2, which also contains the component C. $P_{1,R}$ presents the PFD of the remaining components of the barrier 1 in series with component C. $P_{2,R}$ presents the PFD of the remaining components of the barrier 2 in series with component C. The above formulas can also be extended and adapted to calculate the conditional probabilities of multiple barriers with a shared component and located on the same branch of the bow-tie/SBD.

6.2.3 Dynamic risk assessment

A quantitative risk assessment (QRA) can be performed based on the developed Simulink-based model. Meanwhile, various data may be employed to update the failure probabilities of safety barriers and the probabilities of the initiating events and to achieve a dynamic risk assessment (DRA). With the DRA performed, updated risk profiles may be utilized to support safety barrier optimization.

6.2.3.1 Data sources for updating risks

In previous studies, statistical failure data, for instance, the counts of incidents or near-misses from the same or similar systems, were widely-used for risk updating based on Bayes's theorem (Meel & Seider, 2006; Kalantarnia et al., 2009; Khakzad et al., 2012). However, there are inherent challenges in collecting sufficient statistical failure data, and such data is often unable to provide timely insights into the health status of barriers (Zeng & Zio, 2018). Previous studies have shown that safety barriers can undergo significant degradation depending on their operating conditions and environmental factors, and this degraded performance can greatly influence the risk and severity of undesired events (Misuri et al., 2021; 2023). For example, while industrial PLCs are typically considered to follow a constant failure rate without accounting for degradation, shutdown valves are subject to degradation due to their operation in harsh environments (Zhang et al., 2020). Therefore, considering barrier degradation is essential in the context of dynamic barrier management.

Typically, safety barriers follow a low-demand mode. Periodic proof tests are implemented to evaluate and maintain the health status of safety barriers (Zhang et al., 2020). Health status indicators obtained from the periodic proof tests have the potential to update the failure probabilities of safety barriers. Additionally, continuous condition monitoring is gradually implemented to chemical process systems to facilitate the reliability analysis and maintenance management of critical facilities. The real-time information obtained from condition monitoring (temperature, pressure, vibration, etc.) may reveal the degradation states of the target systems and obtain more accurate failure probability prediction (Zeng & Zio, 2018). Therefore, periodic proof test data and continuous condition-monitoring data of process systems are combined with the accident precursor data to update risk profiles in this study. Moreover, in case probability distributions are used, Monte Carlo simulations are used to handle uncertainty propagations in the risk assessment (Hickman, 1983; Hauptmanns, 2002; Manno et al., 2012). Monte Carlo simulation is an alternative approach to mathematical analytic methods for the calculation of probability distributions. By implementing Monte Carlo simulations based on the SSBM, the proposed approach is able to perform dynamic risk assessment involving both deterministic probability point values and probability distributions. A detailed elaboration on the risk updating methods is presented in the following sub-sections.

6.2.3.2 Bayesian updating by using accident precursor data

Probability distributions are widely used to represent uncertainties in fault-tree-based risk assessment approaches (Yazdi et al., 2019). Beta distributions and gamma distributions were used to describe failure probabilities because they have advantages in serving as prior distributions in the Bayesian estimation of parameters (Eide et al., 2007; Khakzad et al., 2012). Beta-binomial model and Gamma-Poisson model have been used as the base for Bayesian updating of failure probabilities (Siu & Kelly, 1998). In this study, the Beta-binomial model is used, and the Beta distributions are suggested to describe prior PFDs of safety barriers or safety barrier components. Beta distribution can be presented as follows:

$$f(\mu) = \frac{\tau(a+b)}{\tau(a)\tau(b)}\mu^{a-1}(1-\mu)^{b-1} \propto \mu^{a-1}(1-\mu)^{b-1}, a > 0, b > 0 \quad (6.6)$$

where $f(\mu)$ is a Beta distribution of μ. a and b are distribution parameters. $\tau(a) = \int_0^\infty t^{a-1}e^{-t}\,dt$ is a gamma function. Generally, prior probability distributions are derived from the failure database or expert opinions to describe the uncertainty in failure probabilities. When new accident precursor data becomes available, the prior probability distributions can be updated using Bayes's theorem and obtaining posterior probability distributions, as follows:

$$f(x \mid Data) = \frac{g(Data \mid x)f(x)}{\int g(Data \mid x)f(x)\,dx} \propto g(Data \mid x)f(x) \quad (6.7)$$

where $f(x)$ is the prior distribution of x, $g(Data \mid x)$ is the likelihood function, and $f(x \mid Data)$ presents the posterior distribution. By using the binomial distribution, the conditional probability of observing *r* failures in *n* trials given a PFD, μ, can be presented as follows:

$$g(r\,failures\,in\,n\,trials \mid \mu) = \frac{n!}{r!(n-r)!}\mu^{r}(1-\mu)^{n-r} \quad (6.8)$$

By integrating Eq (6.8) and Eq (6.6) into Bayes's theorem, which is presented by Eq (6.7), the posterior distribution of μ can be obtained as follows:

$$f(\mu| \, r \text{ failures in } n \text{ trials}) = \frac{\tau(a'+b')}{\tau(a')\tau(b')}\mu^{a'-1}(1-\mu)^{b'-1} \quad (6.9)$$

where $a' = a + r$ and $b' = b + n - r$.

6.2.3.3 Updating risks by using periodic proof test data

This study assumes that the final safety barrier elements (such as shutdown valves) are subject to continuous aging degradation, and periodic proof tests are executed to evaluate and maintain the health status of the barrier components. The degradation process is modeled using a Gamma process and the degradation level $X(t)$ is presented as follows (Zhang et al., 2020):

$$X(t) \sim \Gamma(\alpha t, \beta) = f_{X(t)}(x) = \frac{\beta^{\alpha t}}{\Gamma(\alpha t)}x^{\alpha t-1}e^{-\beta x},\ \alpha, \beta > 0 \quad (6.10)$$

where $X(0) = 0$, the mean and variance of $X(t)$ are $\alpha t/\beta$ and $\alpha t/\beta^2$, respectively. The cumulative density function (CDF) of $X(t)$ for $t > 0$ is:

$$F_{X(t)}(x) = \int_0^x f_{X(t)}(x)dx \quad (6.11)$$

Proof tests are assumed as no harm tests in this study, which means the proof tests have no direct influence on the degradation process (Gamma process) and only observe/measure the barrier degradation levels. The time spent on tests is also ignored compared to the much longer test intervals. We also assume that the barrier component will fail to play its function when the degradation level reaches or overpasses a predefined failure threshold L. Under those assumptions, the availability of the barrier component in the i-th test interval given the observed degradation level from the *(i-1)*-th test is as follows:

$$A(t) = Pr\left(X(t) < L | X_{(i-1)\tau}\right) = F_{X(t-(i-1)\tau)}\left(L - X_{(i-1)\tau}\right),\ (i-1)\tau < t \leq i\tau \quad (6.12)$$

The average PFD of this barrier component in the i-th test interval is calculated as follows.

$$PFD_{avg}^{i} = 1 - \frac{\int_{(i-1)\tau}^{i\tau} F_{X(t-(i-1)\tau)}(L - X_{(i-1)\tau})dt}{\tau},\ (i-1)\tau < t \leq i\tau \quad (6.13)$$

where PFD_{avg}^{i} is the average PFD of the barrier component. τ is the time interval for proof tests. With the observed degradation level $X_{i\tau}$ becoming available from periodic proof tests continuously, the average PFD of this barrier component at the next time interval can be updated according to Eq (6.13).

6.2.3.4 Updating risks by continuous condition monitoring

Condition-monitoring data usually refers to the online-monitoring data that is related to the degradation of target systems of interest (Kim et al., 2015). Condition-monitoring data provides the opportunity to predict and anticipate the failures of target systems with reference to specific thresholds of the monitored variables (Zeng & Zio, 2018). Based on the assumption that the probability of a failure increases as the process moves away further from the normal operation, the probability of a failure can be calculated as follows (Zadakbar et al., 2013a):

$$for\, r > \mu,\ P = \varphi\left(\frac{r - (\mu + 3\sigma)}{\sigma}\right)$$

$$= \int_{-\infty}^{r} \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(r-(\mu+3\sigma))^2}{2\sigma^2}} dr \quad (6.14)$$

$$for\, r < \mu,\ P = 1 - \varphi\left(\frac{r - (\mu - 3\sigma)}{\sigma}\right)$$

$$= 1 - \int_{-\infty}^{r} \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(r-(\mu-3\sigma))^2}{2\sigma^2}} dr \quad (6.15)$$

where r is the residual value of the key variable with respect to the system performance, and P is the failure probability of the investigated system. $\mu + 3\sigma$ and $\mu - 3\sigma$ are the lower and upper threshold for the normal operation, respectively. Additionally, because condition-monitoring data are usually subject to process and observation noises, it is necessary to filter those noises and estimate the true degradation states of process systems with the help of filtering techniques, for instance, Kalman filter (Zadakbar et al., 2013a) and particle filtering (Zeng & Zio, 2018). Because particle filtering (PF) has the capability of being applied to nonlinear and non-Gaussian systems, this study implements the PF for generating residual values of the key variable and estimating the failure probabilities of basic process control systems. A detailed introduction to implementing the PF for residual value (r) generation and fault diagnosis can be found in (Zadakbar et al., 2015). We omit the repeated illustration here.

6.2.3.5 Consequence assessment

Consequence assessment is another important task of risk assessment. In the proposed approach, both quantitative and qualitative consequence assessments can be incorporated with the decision-making module for safety barrier optimization, as shown in Figure 6.4. "Consequence" sub-systems take responsibility for the consequence assessment. With respect to major adverse events, the calculation of disastrous physical effects (associated with fire, explosion, toxic leakage, etc.) and the assessment of their corresponding damages may be integrated to obtain quantitative consequence assessment results. Regarding physical effects modeling, some software (PHAST, ALOHA, Ansys Fluent, FLACS, etc.) based on empirical models or computational fluid dynamics (CFD) models can be employed (Lewis, 2005). Damage analysis models for heat radiation, explosion effects, acute intoxication, and fragments can be found in TNO Green Book (Van Den Bosh, 1989) and other studies (Gubinelli et al., 2004; Cozzani et al., 2005).

Due to the compatibility and scalability of the MATLAB/Simulink simulation platform, using blocks from the User-Defined Functions library (Mathworks, 2022b) helps to incorporate physical effect modeling results and damage analysis models. It is also possible to input the results from CFD simulations into the "consequence" sub-systems and combine them with the damage analysis models to perform a quantitative consequence assessment. Alternatively, qualitative consequence assessments can also be performed. For instance, a severity class for typical dangerous phenomena in chemical process industries suggested by the ARAMIS project (Andersen et al., 2004) can be used. If qualitative consequence assessment is performed, the “consequence” sub-systems take the responsibility to collect the occurrence probabilities and severity classes of the corresponding consequences to generate a risk matrix.

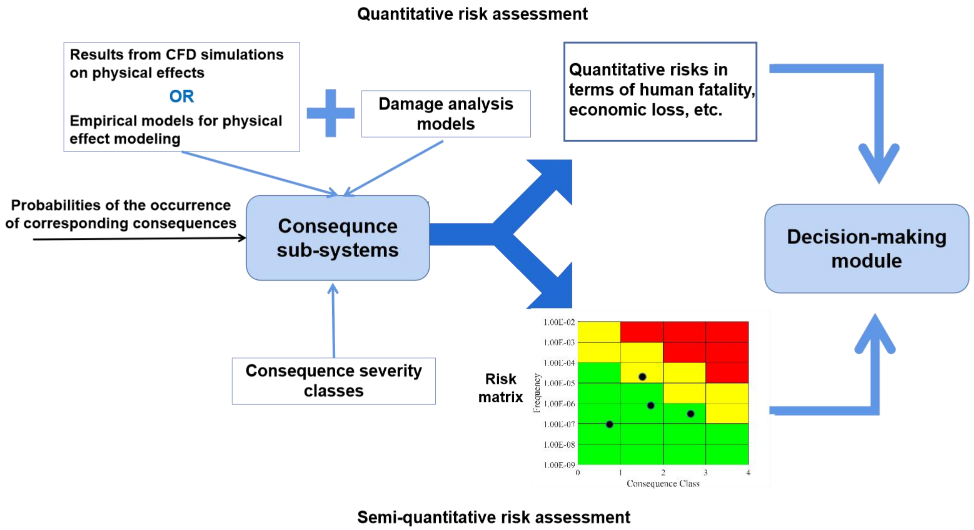

Quantitative risk assessment
Results from CFD simulations on physical effects
OR
Empirical models for physical effect modeling
Damage analysis models
Quantitative risks in terms of human fatality, economic loss, etc.
Probabilities of the occurrence of corresponding consequences
Consequnce sub-systems
Decision-making module
Consequence severity classes
Risk matrix
1.00E-02
1.00E-03
1.00E-04
1.00E-05
1.00E-06
1.00E-07
1.00E-08
1.00E-09
Frequency
0
1
2
3
4
Consequence Class
Semi-quantitative risk assessment

Figure 6.4. Conduct consequence assessment based on the “consequence” sub-systems.

6.2.4 Decision-making on safety barrier improvements

This step aims to provide a decision-making module for safety barrier management based on the risk assessment results. This decision-making module has several functionalities to support safety barrier management. Firstly, it helps to identify critical safety barriers or safety barrier components using sensitivity analysis. Secondly, the integration of cost-effectiveness analysis (CEA) and optimization algorithms helps decision makers to obtain the optimal cost-effective strategies for safety barrier improvements (such as allocating new barriers and optimizing barrier redundancy structure).

1) Identify critical safety barriers in terms of risk-reduction

After risk assessment results are obtained, sensitivity analysis may be performed to identify critical safety barriers/barrier components that object to unacceptable risks.

Generally, the Birnbaum importance measure (van der Borst & Schoonakker, 2001), risk reduction measure (Yazdi & Kabir, 2017), and ratio of variance (RoV) measure (Zarei et al., 2017) are used to rank the importance of initiating/intermediate events on the happening of the top event in a fault tree. Similarly, we use two measures for ranking the importance/sensitivity of safety barriers on the happening of accident scenarios with unacceptable risks. Using Birnbaum importance measure, the importance of a safety barrier in the occurrence of an unwanted accident scenario is presented as follows:

$$I_n = p_s(p_n = 1) - p_s(p_n = 0) \tag{6.16}$$

where I_n is the importance of safety barrier n. p_s is the probability of occurrence of the undesired accident scenario. p_n is the probability of failure on demand (PFD) of safety barrier n. Meanwhile, the risk reduction measure of a safety barrier with respect to the happening of an accident scenario is presented as follows:

$$RI_n = p_s - p_s(p_n = 0) \tag{6.17}$$

By ranking the criticalities of safety barriers based on the above measures, decision-makers may propose candidate strategies form barrier improvement and give more priority to critical barriers.

2) Cost-effective safety barrier optimization

After critical safety barriers are identified, we should investigate the optimal improvement strategies for those barriers. Generally, a series of measures can be implemented to improve the performance of safety barriers, including establishing and allocating new barriers, improving the redundancy structure of safety barriers, revising barrier maintenance intervals, training operators involved in the operation of safety barriers, etc. The specific ways for improving safety barrier performance may be proposed based on the real situations. In the safety science domain, cost-effectiveness analysis (CEA) is widely used to handle the trade-offs between cost and safety due to its advantages of conducting comparative studies and its flexibility in determining safety indicators based on the preferences of decision-makers (Reniers & Van Erp, 2016; Chen & Reniers, 2021; Chen et al., 2021). The implementation of CEA helps decision-makers obtain optimal strategies for barrier improvements with the consideration of both economic and technical constraints.

A series of barrier improvement strategies should be formulated before conducting a cost-effectiveness analysis (CEA). Typically, there are two kinds of constraints imposed on decision-makers in terms of CEA. They are i) a minimum acceptable level of effectiveness (*Eff_{min}*) and ii) a maximum acceptable use of safety budget (Bu_{max}). The optimization problems considering the two kinds of constraints are presented as follows, respectively (Reniers & Van Erp, 2016):

$$\begin{cases} Min(C_i) \\ Eff_i \geq Eff_{min} \\ i \in \{1,2,3,\cdots,N\} \end{cases} \tag{6.18}$$

and

$$\begin{cases} Max(Eff_i) \\ C_i \leq Bu_{max} \\ i \in \{1,2,3,\cdots,N\} \end{cases} \tag{6.19}$$

where i means a strategy *i* from *N* possible strategies for improving safety barrier performance. C_i is the cost of implementing strategy *i*. Eff_i is the effectiveness (risk-reduction performance) of implementing strategy *i*. If only a limited number of strategies are formulated, the best strategy may be obtained through exhaustive search optimization. Otherwise, evolutionary algorithms (for instance, genetic algorithms) may be implemented to solve the optimization problem with a large solution space.

6.3 Case study

6.3.1 Scenario building and model configurations

In this study, a continuous stirred tank reactor (CSTR) for carrying out an exothermic first-order reaction A→B was investigated. This CSTR model is adapted from (Pilario & Cao, 2018), in which a jacketed tank is implemented and the reactor temperature T is maintained by manipulating the coolant flow rate Q_c. The dynamic process of the CSTR is simulated by following Eq (6.20) to Eq (6.22). The CSTR with its safety barrier systems is shown in Figure 6.5.

$$\frac{dC}{dt} = \frac{Q}{V}(C_i - C) - \alpha k C + v_1 \tag{6.20}$$

$$\frac{dT}{dt} = \frac{Q}{V}(T_i - T) - \alpha\frac{(\Delta H_r)kC}{\rho C_p} - b\frac{UA}{\rho C_p V}(T - T_c) + v_2 \tag{6.21}$$

$$\frac{dT_c}{dt} = \frac{Q_c}{V_c}(T_{ci} - T_c) + b\frac{UA}{\rho_c C_{pc} V_c}(T - T_c) + v_3 \tag{6.22}$$

where the inputs of this model are $u = [C_i\, T_i\, T_{ci}]^T$, the outputs are $y = \lfloor C\, T\, T_C\, Q_c \rfloor^T$, v_1, v_2, and v_3 are process noises, and k is an Arrhenius-type rate constant, $k = k_0 exp\,(\frac{-E}{RT})$. Table 6.1 shows the parameter values in Eq (6.20) to Eq (6.22). In the model, α and b are both equal to 1.00 at normal operating conditions. The CSTR model was developed based on the Simulink platform, and it is available online (Karl, 2022).

Table 6.1 Parameter configurations in the CSTR model, adapted from (Pilario & Cao, 2018).

Parameters	Descriptions	Values	Units
Q	Inlet flow rate	100.0	L/min
V	Tank volume	150.0	L
V_c	Jacket volume	10.0	L
ΔH_r	Heat of reaction	-2.0×10^{5}	cal/mol
UA	Heat transfer coefficient	7.0×10^{5}	cal/min/K
k_0	Pre-exponential factor to k	7.2×10^{10}	min^{-1}
E/R	Activation energy	1.0×10^{4}	K
ρ, ρ_c	Fluid density	1000	g/L
C_p, C_{pc}	Fluid heat capacity	1.0	cal/g/K

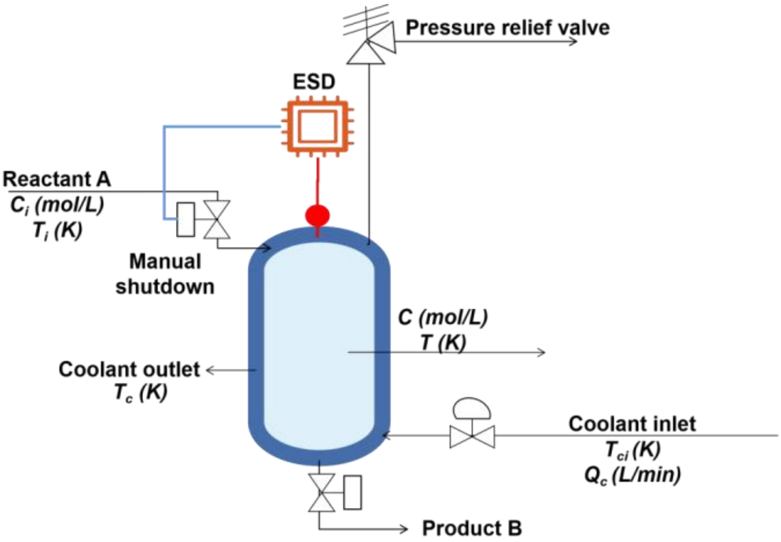


Figure 6.5. A continuous stirred tank reactor (CSTR) with its safety barrier systems, adapted from (Pilario & Cao, 2018).

By following the scenario-building procedures presented in section 6.2, the accident scenarios associated with the CSTR were identified by using the bow-tie technique. Then, the constructed bow-tie was transformed into a Simulink model, as shown in Figure 6.6. Meanwhile, the configurations of the initiating events and safety barriers in the Simulink model are illustrated in Table 6.2. It should be noted that the safety barriers with multiple components all follow an "OR" logic, which means the failure of any one component could lead to the failure of the whole safety barrier.

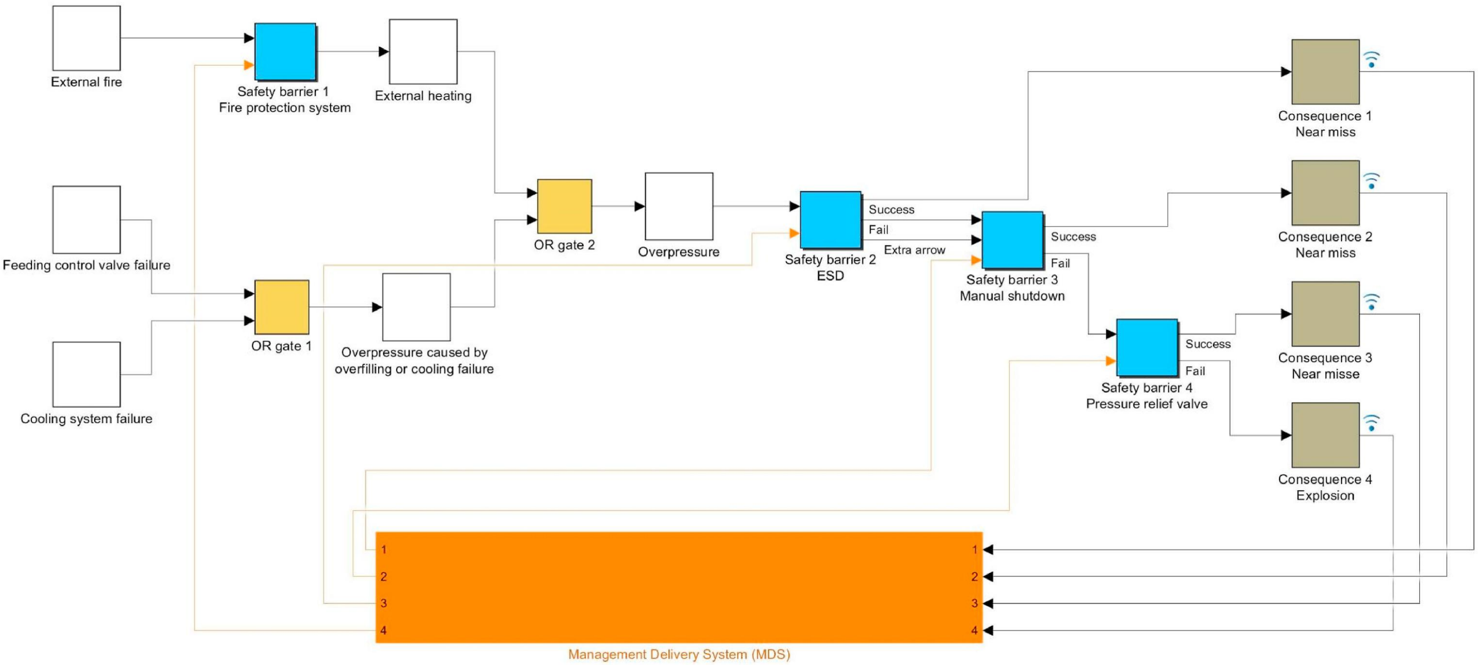


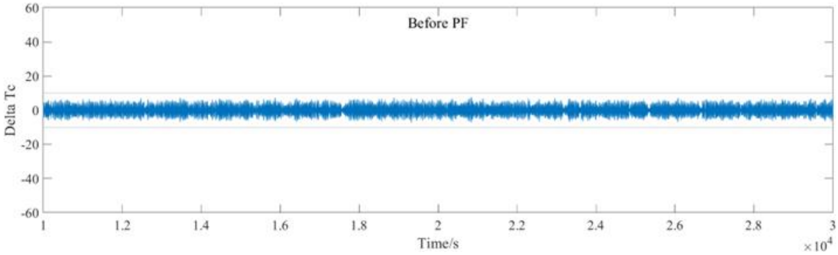
Figure 6.6. The developed Simulink model for safety barrier modeling.

Table 6.2 Configurations of initiating events and safety barriers.

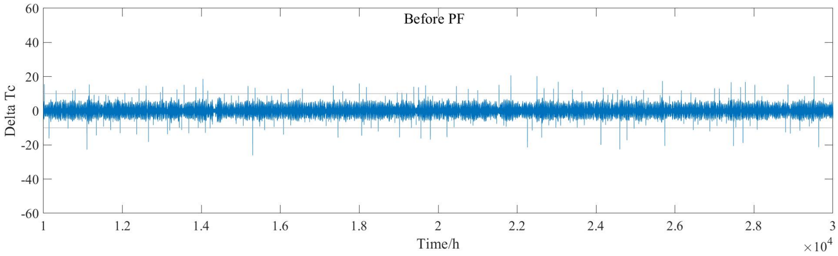
Initiating event OR safety barrier	**Descriptions (some safety barriers consist of multiple components)**		**Configurations of the probabilities/PFDs**
Initiating event	External fire		probability=5.52E-02 $\mathbf{y}^{-1}$ (Debray et al., 2004)
Initiating event	Feeding control valve failure		probability=4.00E-02 $\mathbf{y}^{-1}$ (Taylor, 2010)
Initiating event	Cooling system failure		The probability of cooling system failure is determined and updated by using continuous condition-monitoring data.
Safety barrier	Fire protection system	Smoke/combustion detector	PFD=9.02E-03, λ=4.12E-06 (OREDA, 2002)
		Programmable logic solver	PFD=2.19E-03, λ=1.0E-06 (Hauge & Onshus, 2010)
		Fire pump	PFD=1.58E-01, λ=7.2E-5 (Gravestock, 2008)
		Deluge Valve	PFD=1.27E-02, λ=5.8E-06 (Gravestock, 2008)
Safety barrier	ESD (emergency shutdown system)	Pressure sensor*[15]	PFD=3.29E-04, λ=1.5E-07 (Hauge & Onshus, 2010)
		Programmable safety system	PFD=2.19E-03, λ=1.0E-06 (Hauge & Onshus, 2010)
		Shutdown valve*	PFD of the shutdown valve is calculated and updated based on periodic proof test data.
Safety barrier	Manual shutdown	Pressure sensor*	The same as the pressure sensor in ESD.
		Fail to close valve manually	Beta distribution parameters: a=32.3, b=137.7, (Roy et al., 2015). The Beta distribution is updated by using accident precursor data.
		Shutdown valve*	The same as the shutdown valve in ESD.
Safety barrier	Pressure relief valve	/	PFD=2.4E-03, λ=1.1E-06 (Hauge & Onshus, 2010)

[15] A barrier component with * means it is a shared component.

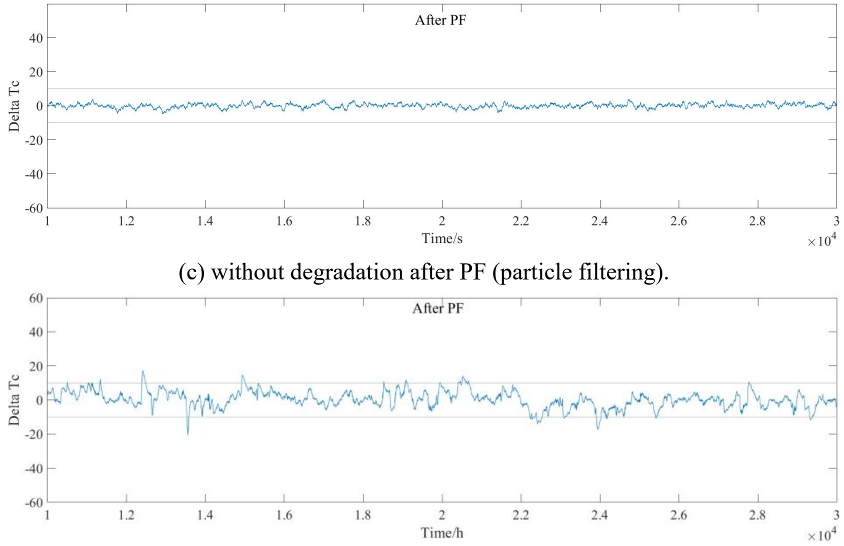
The probability of cooling system failure is determined and updated based on continuous condition-monitoring data by following the method illustrated in Section 6.2.3. The residual values of ΔT_C (demonstrates the temperature variation inside the cooling jacket) was selected as the variable for failure probability calculation based on Eq (6.14) and Eq (6.15). Under ideal operating situations, ΔT_C should be 0 because the temperature inside the cooling jacket remains stable. Therefore, the deviation of the ΔT_C values reflects the likelihood of the cooling system failure. It is assumed that the coolant inlet temperature, coolant outlet temperature, and the temperature inside the reactor (T_{ci}, T_c, T) were monitored. Based on that, particle filtering (PF) was integrated with Eq (6.22) to generate the residual values of ΔT_C, with the state vector $X = [T_{ci}, T_c, T, \Delta T_C]^T$ and measurement vector $Y = [T_{ci}, T_c, T]^T$. Then, the obtained residual values were integrated into Eq (6.14) and Eq (6.15) to estimate the failure probabilities of the cooling system. The normal operation threshold $\pm 3\sigma$ in Eq (6.14) and Eq (6.15) is set as ± 10 according to simulation results under normal operations (as shown in Figure 6.7). In real cases, this threshold may be determined based on the real monitored data under normal operating situations with the consultation with experts.



(a) without degradation before PF (particle filtering).



(b) with degradation before PF (particle filtering).



(c) without degradation after PF (particle filtering).

(d) with degradation after PF (particle filtering).

Figure 6.7. Residual values of ΔT_C with and without cooling system degradation.

In this study, three ways are used to calculate the PFDs of safety barriers or safety barrier components.

i) the PFDs can be calculated based on constant failure rates, as presented in Eq (6.23).

$$PFD = \frac{1}{2}\lambda T \tag{6.23}$$

where λ is the failure rate of the safety barrier/barrier component, which can be derived from existing databases, such as the OREDA database (OREDA, 2002) and PDS database (Hauge & Onshus, 2010). T is the periodic inspection interval, it is assumed as 4380 h in this study.

ii) In terms of the ESD (emergency shutdown system), constant failure rates were used for the pressure sensor and programmable safety system based on the assumption that constant failure rates is usually valid for electronic components (Zhang et al., 2020). By contrast, the degradation of the shutdown valve was considered because it is operated in harsh conditions. The PFD of the shutdown valve was calculated based on the approach presented in Section 6.2.3 and was updated by using periodic proof test data. We used the assumptions and configurations for shutdown valves from (Zhang et al., 2020), in which the designed closing time for a shutdown valve is 12 s. A predefined failure threshold L was set as 1.25×10^{-3}, and it is assumed that no maintenance action will be implemented unless the degradation level exceeds L. α and β used in Eq (6.10) were set as 1.02×10^{-4} and 1.2×10^{4}, respectively. The time interval for proof tests, $\tau,$ is set as 2190 h. By checking the closing time of the shutdown valve and evaluating the corresponding degradation level, the average PFD of the shutdown valve in the test interval can be updated according to Eq (6.13).

iii) A Beta distribution is used to describe the failure probability of human actions in the manual shutdown barrier. The Beta distribution parameters are set as: $a = 32.3$ and $b = 137.7$, adapted from (Roy et al., 2015). As new accident precursor data becomes available, the Beta distribution can be updated through Bayesian updating. Because the ESD barrier and the manual shutdown barrier have shared components (a pressure sensor and a shutdown valve), the conditional probability of manual shutdown failure given the failure of ESD is used, according to the method presented in Section 6.2.2.

6.3.2 Dynamic risk assessment results

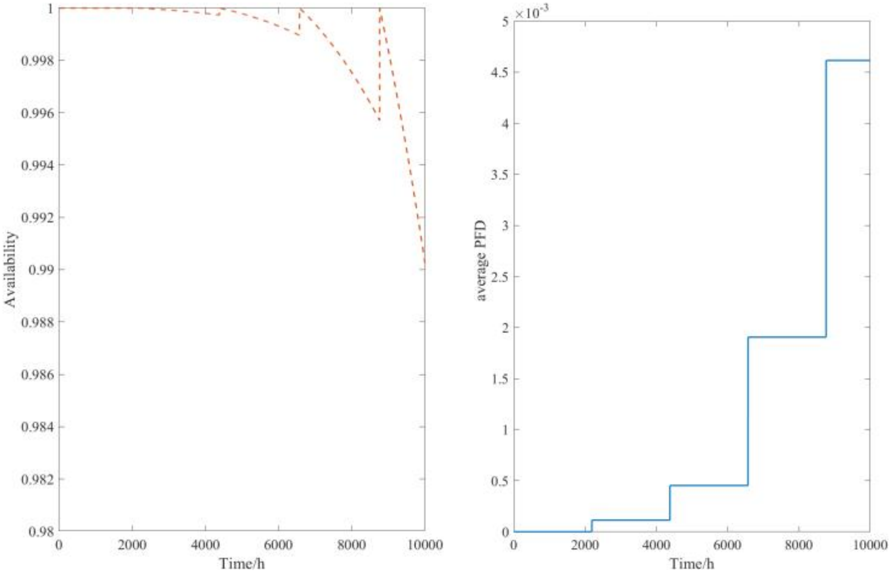
This section presents the dynamic risk assessment results regarding the investigated system. Hypothetical accident sequence precursor data of “fail to close valve manually” is given in Table 6.3. Hypothetical periodic proof test data for the emergency shutdown valve is presented in Table 6.4. Figure 6.8 shows the availability and average PFDs of the shutdown valve over time, which are calculated based on the hypothetical data in Table 6.4. It is assumed that a shock degradation happened to the cooling system at 5000 h. We simulated this degradation by adding a disturbance following a Gaussian distribution $N(0.01, 0.03)$ to the coolant inlet temperature T_{ci}. The residual values of ΔT_C with and without a certain degree of degradation are compared in Figure 6.7. The average residual values before and after the degradation were used to calculate the failure probabilities of the cooling system. By employing the periodic proof test data, continuous condition-monitoring data, and accident precursor data for risk profile updating, the dynamic risk profiles of the CSTR explosion accident are obtained and are shown in Figure 6.9. The probability distributions obtained from Monte Carlo simulations with 10000 trails are demonstrated in Figure 6.9 (a) and their corresponding mean values are demonstrated in Figure 6.9 (b). As we can see from Figure 6.9, the risk profile increases gradually and peaks at 8760 h, which results from the degradation of the safety barriers and the degradation of the cooling system.

Table 6.3 Hypothetical accident sequence precursor data of “fail to close valve manually”.

Time (h)	**Cumulative failure number**	**Cumulative trial number**
3000	6	30
5000	9	50
7000	11	70

Table 6.4 Hypothetical periodic proof test data of the shutdown valve.

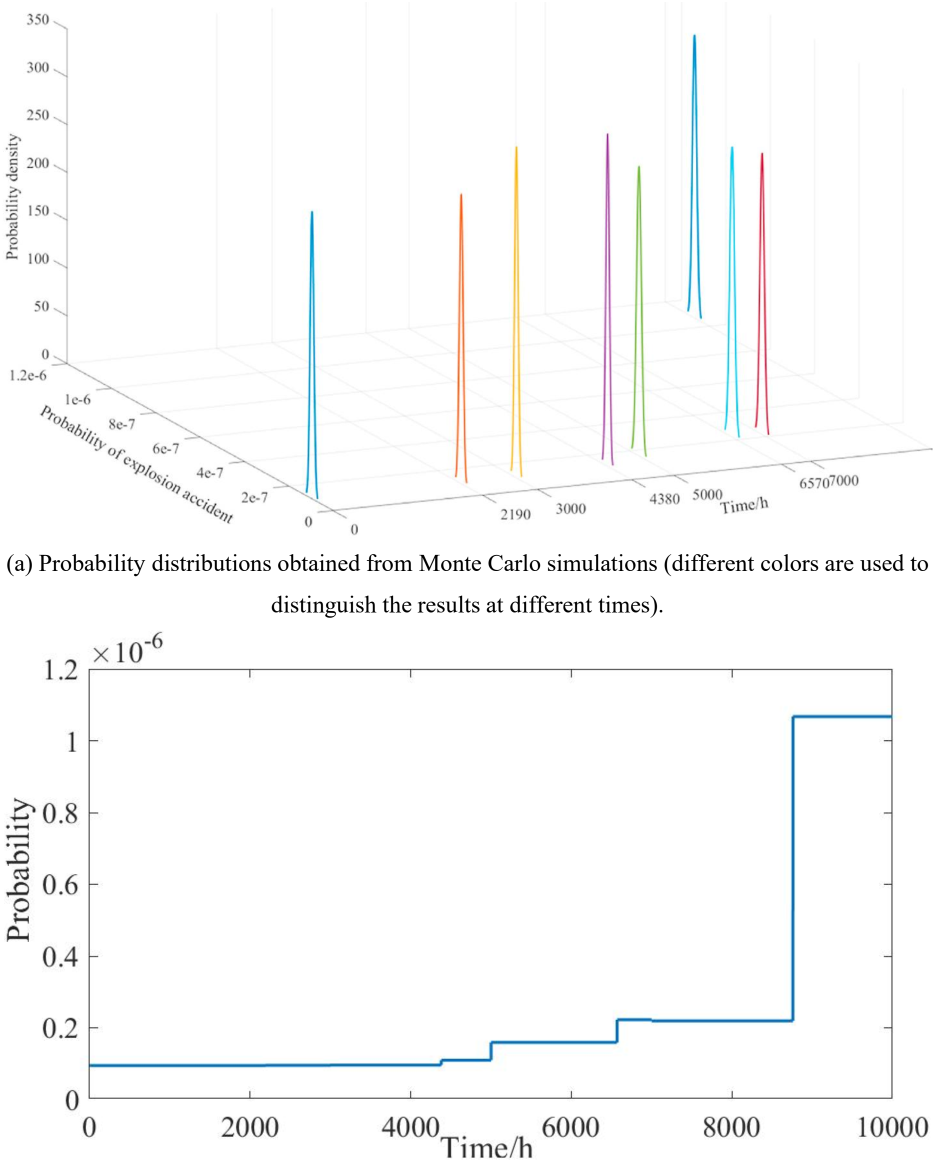
Test time (h)	**Degradation level**	**Test time (h)**	**Degradation level**
2190	8×10^{-4}	4380	9×10^{-4}
6570	1×10^{-3}	8760	$1.1 \quad \times 10^{-3}$



(a) Availability

(b) Average PFDs

Figure 6.8. Availability and average PFDs of the shutdown valve over time.



(a) Probability distributions obtained from Monte Carlo simulations (different colors are used to distinguish the results at different times).

(b) Mean values of the probability distributions.

Figure 6.9. Risk profiles of the CSTR explosion accident over time.

6.3.3 Cost-effective barrier optimization

Based on the obtained risk assessment results, we demonstrate the time-varied explosion risks in a risk matrix, as shown in Figure 6.10. Based on the risk profiles at the final stage (8760 h~10000 h), a sensitivity analysis of the safety barriers and initiating events was conducted. Two measures (Birnbaum importance measure and risk reduction measure) were used to rank the criticalities of safety barriers, as shown in Table 6.5.

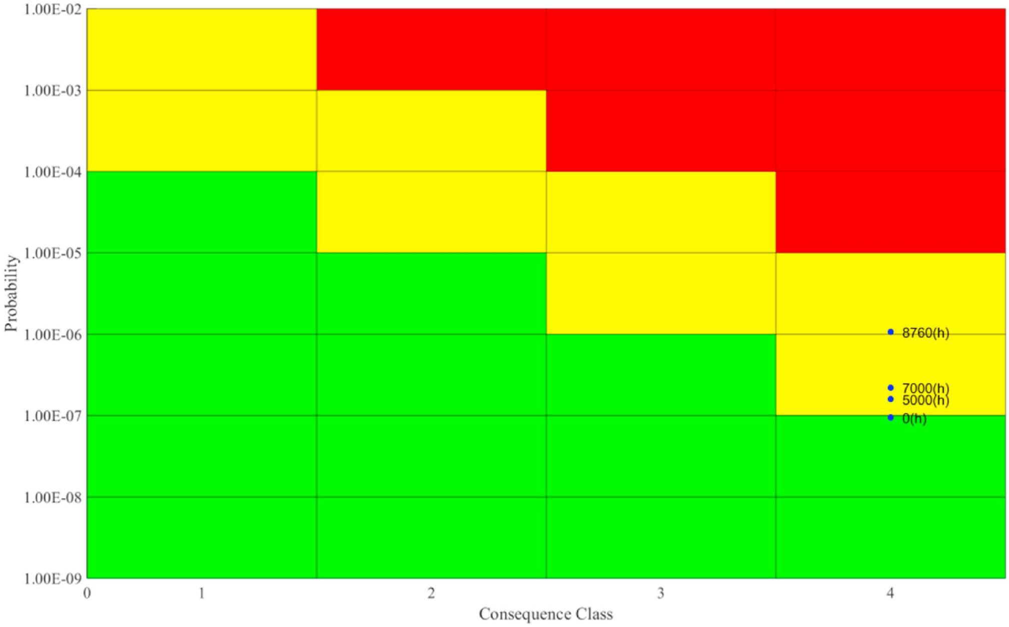


Figure 6.10. A risk matrix with respect to time-varied explosion risks.

Table 6.5 Sensitivity of safety barrier failures and initiating events in the explosion risk.

<table>
<tr><th>Safety barrier OR Initiating event</th><th colspan="2">Descriptions</th><th>Risk reduction measure</th><th>Birnbaum importance measure</th></tr>
<tr><td>Initiating event</td><td colspan="2">External fire</td><td>1.28E-07</td><td>2.32E-06</td></tr>
<tr><td>Initiating event</td><td colspan="2">Feeding control valve failure</td><td>5.38E-07</td><td>1.35E-05</td></tr>
<tr><td>Initiating event</td><td colspan="2">Cooling system failure</td><td>3.97E-07</td><td>1.33E-05</td></tr>
<tr><td>Safety barrier</td><td colspan="2">Fire protection system</td><td>1.28E-07</td><td>7.20E-07</td></tr>
<tr><td>Safety barrier</td><td rowspan="3">ESD</td><td>Pressure sensor</td><td>6.11E-08</td><td>1.86E-04</td></tr>
<tr><td>Safety barrier</td><td>Shutdown valve</td><td>9.32E-07</td><td>1.86E-04</td></tr>
<tr><td>Safety barrier</td><td>Programmable safety system</td><td>9.49E-08</td><td>4.33E-05</td></tr>
<tr><td>Safety barrier</td><td colspan="2">Fail to close valve manually</td><td>9.49E-08</td><td>4.07E-07</td></tr>
<tr><td>Safety barrier</td><td colspan="2">Pressure relief valve</td><td>1.09E-06</td><td>4.54E-04</td></tr>
</table>

Table 6.5 shows that the most important safety barriers/barrier components are the

pressure sensor, shutdown valve, and pressure relief valve, with importance measures as 1.86E-04, 1.86E-04, and 4.54E-04, respectively. Considering their current PFD estimations, the most sensitive safety barriers/barrier components are the pressure relief valve and shutdown valve, with risk reduction measures as 1.09E-06 and 9.32E-07, respectively. In this case study, three optimization principles are formulated as follows: i) When the explosion risk exceeds 1.00E-06 y-1, barrier improvements should be made. ii) The goal of safety barrier optimization is to limit the explosion risk below 1.00E-08 y-1. iii) When the degradation of safety barriers or basic process control systems is detected, the maintenance or replacement of the degraded components should be given more priority instead of allocating new safety barriers. In real cases, the safety barrier management/optimization principles may be determined according to the realistic needs.

Based on the proposed optimization principles, the optimization objective function and constraints are presented as follows:

$$\begin{cases} Min(C_i) \\ Risk_i \leq Risk_{threshold} \\ i \in \{1,2,3,\cdots,N\} \end{cases} \tag{6.24}$$

where C_i means the cost of strategy *i*. $Risk_i$ is the risk assessment outcome after implementing strategy *i*. $Risk_{threshold}$ is set as 1.00E-08 y^{-1} in this case study. This optimization aims to minimize the costs spent on safety barrier improvement and meanwhile to ensure the explosion risk is below 1.00E-08 y^{-1}. Because the cooling system degradation and shutdown valve degradation were detected, we considered the maintenance/replacement of such facilities. A series of barrier improvement strategies are proposed, as shown in Table 6.6, in which the cost analysis of each strategy is also given.

As shown in Table 6.6, strategy 4, strategy 6, and strategy 7 are the candidate strategies that meet the optimization constraint, which is to lower the explosion risk below the risk threshold. Among those three strategies, strategy 4 has the lowest total cost (2150€) for safety barrier system improvements, so it is selected as the most cost-effective strategy. In case more candidate strategies are proposed, a similar optimization may be done by comparing the results of those strategies and following exhaustive search optimization. If a large number of candidate strategies are formulated, it may become unreasonable to assess all the strategies using exhaustive optimizations. Alternatively, genetic algorithms may be implemented to solve optimization problems with a large solution space. Chapter 5 of this dissertation demonstrates the application of genetic algorithms in barrier optimizations.

Table 6.6 Candidate strategies for safety barrier optimization.

Strategy number	Description	Cost analysis	Explosion risk after barrier improvements (y^{-1})

1	a. Maintenance of the degraded cooling system (restore its performance to the initial);
	b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement).
		a= 1000€ (one-time maintenance cost);
		b= 400€ (replacement cost);
		Total cost=a+b=1400€.
			9.2453E-08

2	a. Maintenance of the degraded cooling system (restore its performance to the initial);
	b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement);
	c. Add a redundant pressure sensor.
		a= 1000€ (one-time maintenance cost);
		b= 400€ (replacement cost);
		c= 600€ (equipment and installation cost)+200€×2 (annual maintenance/inspection cost);
		Total cost=a+b+c=2600€.
			5.0599E-08

3	a. Maintenance of the degraded cooling system (restore its performance to the initial);
	b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement);
	c. Add a redundant shutdown valve.
		a= 1000€ (one-time maintenance cost);
		b= 400€ (replacement cost);
		c= 600€ (equipment and installation cost)+200€×4 (annual maintenance/inspection cost);
		Total cost=a+b+c=2800€.
			9.2452E-08

4	a. Maintenance of the degraded cooling system (restore its performance to the initial);
	b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement);
	c. Add a redundant pressure relief valve.
		a= 1000€ (one-time maintenance cost);
		b= 400€ (replacement cost);
		c= 450€ (equipment and installation cost)+150€×2 (annual maintenance/inspection cost);
		Total cost=a+b+c=2150€.
			2.2189E-10

5	a. Maintenance of the degraded cooling system (restore its performance to the initial); b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement); c. Add a redundant pressure sensor; d. Add a redundant shutdown valve.	a= 1000€ (one-time maintenance cost); b= 400€ (replacement cost); c= 600€ (equipment and installation cost)+200€×2 (annual maintenance/inspection cost); d= 600€ (equipment and installation cost)+200€×4 (annual maintenance/inspection cost); Total cost=a+b+c+d=3800€.	5.0599E-08
6	a. Maintenance of the degraded cooling system (restore its performance to the initial); b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement); c. Add a redundant pressure sensor; d. Add a redundant pressure relief valve.	a= 1000€ (one-time maintenance cost); b= 400€ (replacement cost); c= 600€ (equipment and installation cost)+200€×2 (annual maintenance/inspection cost); d= 450€ (equipment and installation cost)+150€×2 (annual maintenance/inspection cost); Total cost=a+b+c+d=3150€.	1.2144E-10
7	a. Maintenance of the degraded cooling system (restore its performance to the initial); b. Maintenance of the degraded shutdown valve (perfect maintenance/replacement); c. Add a redundant shutdown valve; d. Add a redundant pressure relief valve.	a= 1000€ (one-time maintenance cost); b=400€ (replacement cost); c= 600€ (equipment and installation cost)+200€×4 (annual maintenance/inspection cost); d= 450€ (equipment and installation cost)+150€×2 (annual maintenance/inspection cost); Total cost=a+b+c+d=3550€.	2.2189E-10

6.4 Discussions

6.4.1 A comparison of SSBM and BN

As a widely used tool for quantitative risk assessment (QRA) and dynamic risk assessment (DRA), Bayesian networks (BNs) are also employed to support safety barrier assessment and management. This section compares BN and the proposed SSBM approach, considering their characteristics and capabilities in QRA, DRA, and supporting decision-making. It helps practitioners get insights into the application prospects of the SSBM approach. The comparison results are presented in Table 6.7.

Table 6.7 A comparison of SSBM and BN with respect to safety barrier management.

Approaches	QRA capabilities	DRA capabilities	Decision-making capabilities
BN	BN has the advantage of representing the dependencies of events, incorporating multi-state variables, and updating probabilities in QRA.	Dynamic Bayesian networks (DBN) can be employed for conducting DRA. Hierarchical Bayesian networks can combine with Bayes's theorem to update the reliability of safety barriers and perform DRA (Khakzad et al., 2014).	The combination of BN and influence diagram can be employed to determine the optimal strategy for decision-making (Khakzad, 2021). However, it has difficulties in solving large solution space optimization problems.
SSBM	As a bow-tie-based approach, multiple occurring events (MOE) are not allowed on the left-hand side of the SSBM model. As a result, a simplification should be performed based on the minimal cut sets to determine the correct model structure. SSBM highlights deliberately inserted safety barrier. Its relative simplicity supports communication with non-expert stakeholders and facilitates safety barrier audition and management.	Due to the flexibility and compatibility of the Simulink simulations, SSBM is able to incorporate the data from various sources (periodic proof test data, continuous condition-monitoring data, accident precursor data, etc.) to update the failure probabilities of safety barriers and also update the happening probabilities of initiating events.	SSBM has the advantage of integrating with various optimization algorithms (exhaustive search algorithms, evolutionary algorithms, etc.) to solve large solution space optimization problems and support decision-making.

6.4.2 Recommendations for future work

In the proposed approach, the imperfection of the proof tests in revealing barrier health states was not considered. In real situations, the proof tests/inspections are usually subject to errors and defects (Zhang et al., 2021). The quantification and modeling of the proof test errors in revealing barrier health states may be further studied and incorporated into the SSBM approach. Additionally, the degradation process of safety barriers can vary depending on operating and environmental conditions. Further research could focus on how to integrate environmental factors into the degradation modeling of barriers and the prediction of their failure probabilities within the context of dynamic barrier management.

The proposed approach incorporates a model-based fault detection and diagnosis (FDD) approach, which is a preliminary exploration. More advanced FDD methods with more accuracy and adaptability in solving complex non-linear chemical process models may be integrated into the proposed approach. Additionally, dynamic risk assessment was performed by this study through the updating of probabilities while the model structure remains static. In some cases, the model structure may also need to be updated when new evidence becomes available. The case study in this paper only demonstrates the application of the SSBM in qualitative consequence assessment using a risk matrix. The feasibility of the SSBM in quantitative consequence assessment with the integration of physical effects modeling and damage analysis models may be further validated.

6.5 Conclusions

This chapter provides a simulation tool, named Simulink-based safety barrier modeling (SSBM), for safety barrier management based on dynamic risks. Periodic proof test data, continuous condition-monitoring data, and accident precursor data are combined for risk updating with the consideration of the degradation of safety barriers and the degradation of basic process control systems. The combination of cost-effectiveness analysis (CEA) and optimization algorithms is used to determine the optimal strategies for safety barrier establishments and improvements. A dynamic risk assessment of a continuous stirred tank reactor (CSTR) was demonstrated as the case study to validate the feasibility of the proposed approach in dynamic risk-based safety barrier management. The results show that the pressure relief valve and shutdown valve are the most critical safety barrier/barrier components regarding explosion risks. Apart from maintaining or replacing the degraded facilities, the allocation of a redundant pressure relief valve is the most cost-effective strategy. A comparison of the SSBM and Bayesian networks is also given to demonstrate the characteristics and capabilities of the SSBM in risk assessment and safety barrier management. The comparison shows that the SSBM has advantages in supporting communication with non-expert stakeholders, and various techniques/methods can be incorporated into the SSBM to facilitate safety barrier assessment and risk-based safety barrier management due to the flexibility and adaptability of the Matlab/Simulink platform.

Chapter 7 Dynamic and integrated safety and security barrier management

Previous chapters of this thesis are dedicated to addressing most of the proposed research sub-questions. However, assessing dynamic variations in the performance of security barriers and the integrated safety and security risks is still lacking in enabling dynamic barrier management. This chapter identifies and characterizes multiple data sources that have the potential to reveal risk variations associated with integrated safety and security risks and enables dynamic barrier management by proposing a systematic framework that consolidates the methodologies developed throughout the entire Ph.D. thesis. Meanwhile, The methodological foundations and procedural process for applying this framework in practice are shown, offering insights into its feasibility. Furthermore, the foundational principles and precautionary considerations pertinent to adopting this new framework in real-world contexts are discussed. It is emphasized that the implementation of this framework holds significant promise in fostering a dynamic and integrated management of safety and security barriers.

This chapter is drafted with modifications based on the following publication:

- ◆ Yuan, S., Reniers, G., Yang, M. (2024). Dynamic and integrated safety and security barrier management in chemical plants: a new paradigm to manage complex major adverse risks. *Process Safety and Environmental Protection,* (To be submitted).

7.1 Introduction

The necessity of integrating safety and security risks and dynamic barrier management is elaborated on in previous chapters (chapter 1 and chapter 2) of this thesis. Several approaches have been developed to foster integrated safety and security risk assessment (chapter 3), integrated safety and security barrier management (chapter 4 and chapter 5), and dynamic safety barrier management (chapter 6), respectively. However, the application of the dynamic barrier management (DBM) concept in the security barrier domain has not been achieved in previous chapters of this thesis or existing literature (Øien et al., 2022). Systematic integration of the developed methodologies/approaches is lacking to wrap up the entire endeavors of this PhD research and to boost dynamic and integrated safety and security (S&S) barrier management in real-world practices.

Therefore, this chapter explores the application of the DBM concept considering the integrated safety and security risks and further develops a systematic framework for dynamic and integrated S&S barrier management. Additionally, multiple-source data that are capable of revealing risk variations and updating risk profiles are identified and characterized to enable dynamic barrier management. A case study is provided in this chapter to show the efficacy and advantages of the proposed framework, and the foundational principles and practical notes are discussed to foster the adoption of the dynamic and integrated S&S barrier management paradigm in real-world practices.

The remaining sections of this chapter are organized as follows. Firstly, an introduction to the proposed framework is provided in section 7.2. Following this, the systematic integration of the methodologies underpinning this framework is elaborated upon, with a focus on the pertinent data sources and models facilitating dynamic barrier management, in section 7.3. Section 7.4 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 7.5, and conclusions are given in section 7.6.

7.2 An overview of the proposed framework

An overview of the proposed framework for dynamic and integrated S&S barrier management is shown in Figure 7.1. 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.

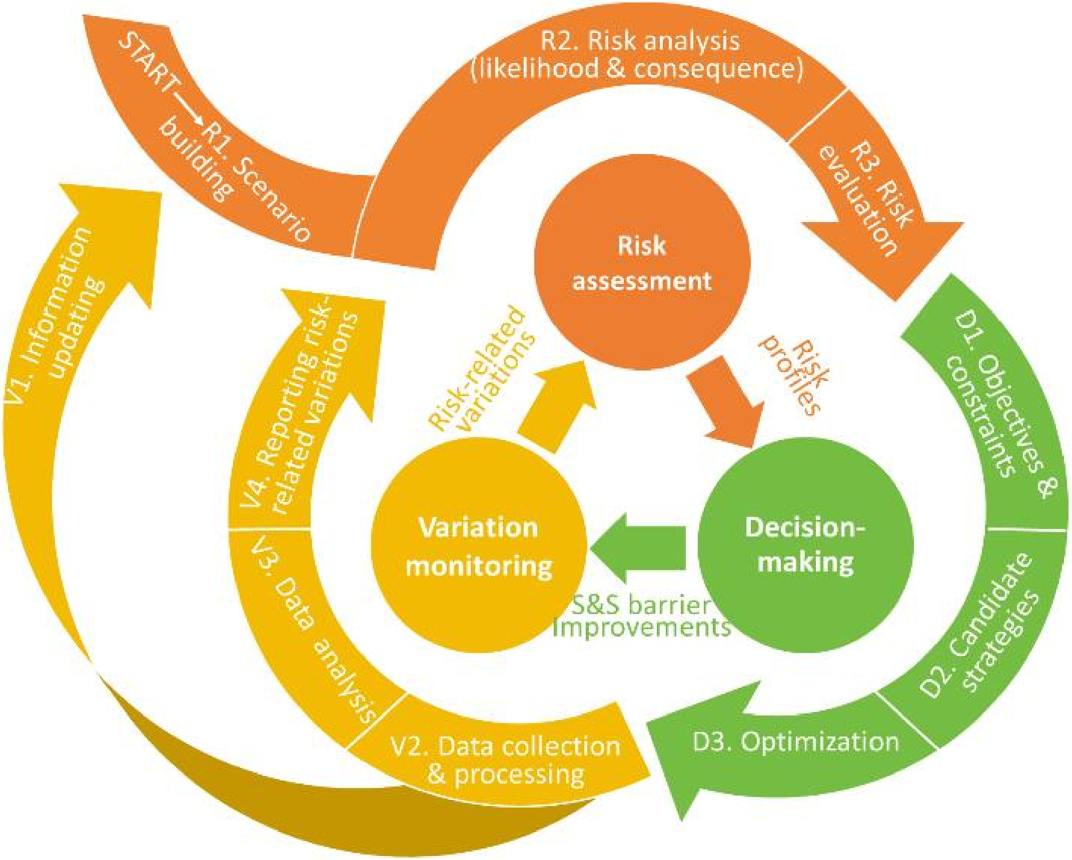


Figure 7.1. A framework for dynamic and integrated S&S barrier management.

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

- 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.

7.3 Methodologies

7.3.1 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 run this framework. Specific techniques/tools used in various phases of this framework are presented in Figure 7.2, in which the workflows are also given to integrate those techniques/tools systematically. Most techniques or approaches in Figure 7.2 have already been illustrated and applied in previous chapters of this thesis. For instance, the risk assessment approach has been applied to the case studies in Chapters 3 and 4. Cost-effectiveness analysis (CEA) and tailored optimization algorithms are combined for decision-making on barrier optimization, as demonstrated in Chapters 4, 5, and 6. The incorporation of safety-related multiple-source data for dynamic risk assessment and dynamic safety barrier management has been researched in Chapter 6. The methodology presented in Figure 7.2 is a systematic reframing of the endeavors of previous chapters and adds new elements to incorporating security-related data for dynamic S&S barrier management. Therefore, a brief introduction of each part of the proposed integrated methodology is given in the following sub-sections, with an emphasis on the newly added elements.

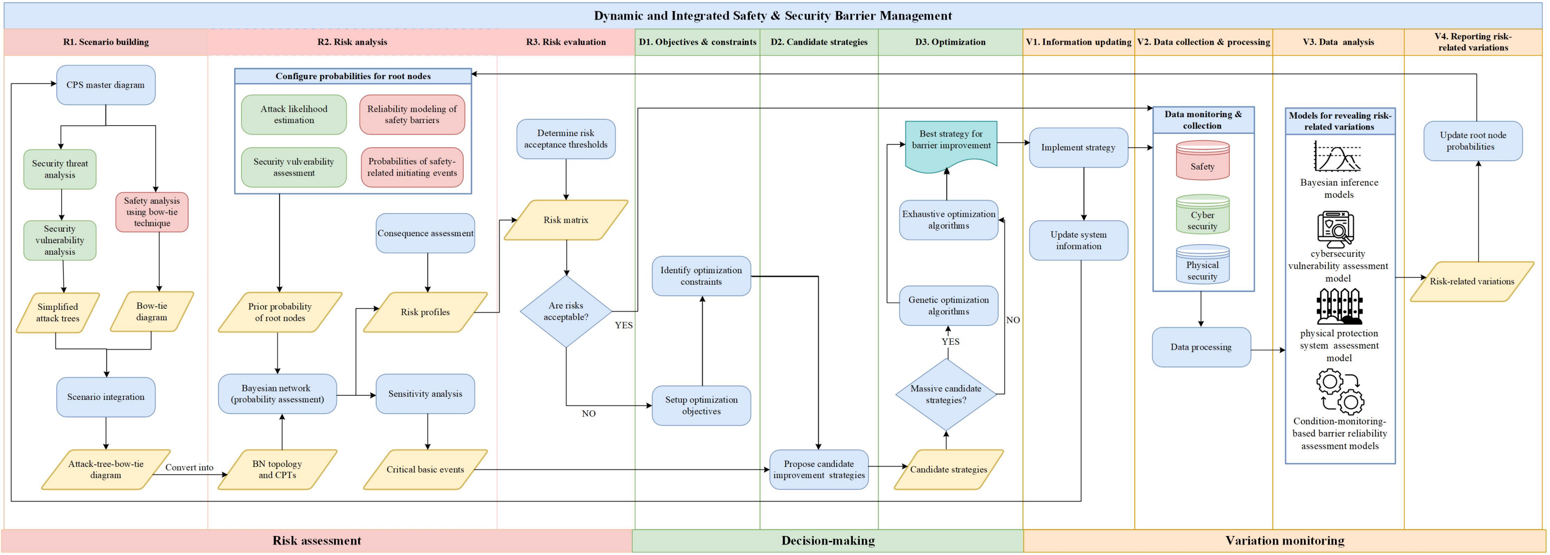

Dynamic and Integrated Safety & Security Barrier Management
R1. Scenario building
R2. Risk analysis
R3. Risk evaluation
D1. Objectives & constraints
D2. Candidate strategies
D3. Optimization
V1. Information updating
V2. Data collection & processing
V3. Data analysis
V4. Reporting risk-related variations
CPS master diagram
Security threat analysis
Security vulnerability analysis
Safety analysis using bow-tie technique
Simplified attack trees
Bow-tie diagram
Scenario integration
Attack-tree-bow-tie diagram
Convert into
Configure probabilities for root nodes
Attack likelihood estimation
Reliability modeling of safety barriers
Security vulverability assessment
Probabilities of safety-related initiating events
Prior probability of root nodes
Bayesian network (probability assessment)
BN topology and CPTs
Consequence assessment
Risk profiles
Sensitivity analysis
Critical basic events
Determine risk acceptance thresholds
Risk matrix
Are risks acceptable?
YES
NO
Identify optimization constraints
Setup optimization objectives
Propose candidate improvement strategies
Best strategy for barrier improvement
Exhaustive optimization algorithms
Genetic optimization algorithms
YES
NO
Massive candidate strategies?
Candidate strategies
Implement strategy
Update system information
Data monitoring & collection
Safety
Cyber security
Physical security
Data processing
Models for revealing risk-related variations
Bayesian inference models
cybersecurity vulnerability assessment model
physical protection system assessment model
Condition-monitoring-based barrier reliability assessment models
Update root node probabilities
Risk-related variations
Risk assessment
Decision-making
Variation monitoring

Figure 7.2. An integrated methodology for dynamic and integrated S&S barrier management.

7.3.2 Integrated safety and security risk assessment

7.3.2.1 BN models for integrated risk assessment

This section mainly adapts the approaches presented in Chapter 3 and Chapter 4 of this thesis for integrated risk assessment of industrial control systems (ICSs) considering safety causes, cyber-physical (C2P) attacks, and physical attacks. As shown in Figure 7.2, scenario building starts with establishing a CPS master diagram (Guzman et al., 2020), which demonstrates the ICS in a multi-layered manner and serves as a basis for 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 is employed to characterize potential threat agents using the methods presented in Section 3.3.4.1 of this dissertation. 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. Then, the integrated attack-tree-bow-tie diagram is converted into a Bayesian network (BN) model for quantitative risk assessment. More details on developing a BN model for integrated risk assessment can be found in Section 3.3 of this dissertation.

7.3.2.2 Prior probabilities/probability distributions

The BN model's topology and CPTs (conditional probability tables) are derived from the integrated attack-tree-bow-tie diagram. 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 4.2.3.1 of this dissertation. Details on determining the prior probabilities or probability distributions are given below.

1) For safety-related initiating events, reliability databases (Hauge & 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, 2010). Reliability databases, accident databases, and human reliability analysis are helpful for determining the

PFDs of safety barriers. The approaches presented in Section 4.2.3.2 are used to calculate PFDs for technical safety barriers.

2) Attack likelihood estimation relies more on expert judgment due to the lack of data found in both 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 3.2 in Section 3.3.4.1. 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 & Maillart (2018) may be used as a basis for attack likelihood estimation of C2P attacks, as presented in Table 3.3 in Section 3.3.4.1.

3) For security vulnerability assessment, the time-to-compromise (TTC) based approach presented in Section 4.2.3.3 is used to assess the vulnerability of industrial control systems to C2P attacks, considering the uncertainties associated with attackers' skill levels.

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), as illustrated in Section 3.3.4.2. A data repository built by Moreno et al. (2022) is used for typical PPSs in chemical plants, as shown in Table 7.1.

Table 7.1 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	$\mathrm{P}_{\mathrm{fail}} = PFD + (1-\eta) \times (1 - PFD)$ $P_{success} = (1 - PFD) \times \eta$
Entry control	Detection	0.40	0.80	
Fence	Delay	0.00	0.9968	
Closed Circuit TeleVision (CCTV)	Detection	0.205	0.97	
Intrusion detection by site personnel	Detection	dayshift: 0.233	0.248	
		nightshift: 0.4	0.248	

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 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):

$$P_S = P_D \times P_C \times P_T \tag{7.1}$$

$$P_T = \frac{1}{\sqrt{2\pi\sigma_t^2}} \int_0^{\infty} exp[-\frac{(t-\mu_t)^2}{2\sigma_t^2}]dt \tag{7.2}$$

$$t = ATT - RFT \tag{7.3}$$

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).

7.3.2.3 Consequence assessment and sensitivity analysis

Regarding consequence assessment, either qualitative or quantitative consequence assessment may be performed, and Section 6.2.3.5 gives a relatively thorough elaboration. In the methodology presented in Figure 7.2, 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. Additionally, a sensitivity analysis of basic events on the undesired risks can be performed using the BN model to get insights into the criticality of each event, as illustrated in Section 3.3.5.3.

7.3.3 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. Section 5.2.4 elaborates 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. The second applies to situations where a company only has a limited budget and aims to mitigate undesired risks as much as possible. Eqs. (5.7) and (5.8) characterize those two kinds of optimization problems, respectively. Apart from this, some additional constraints can also be added to the optimization problems according to actual needs.

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, the best strategy can be obtained through exhaustive search optimization, as demonstrated in the case study in Chapter 6. If a large amount of candidate strategies are proposed, it may 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 determine the approximately optimal strategy, as demonstrated in the case study in Chapter 5 of this dissertation.

7.3.4 Risk variation monitoring for dynamic barrier management

7.3.4.1 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 achieve dynamic barrier management. Various data in relation to the safety and security of the investigated system 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 7.2 characterizes the multi-source data and the models used to reveal and quantify risk-related variations.

Table 7.2 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).

Chapter 6 of this dissertation has already incorporated safety-related data for dynamic barrier management. Based on accident precursor data, Bayes's theorem (Beta-binomial model) 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 & Kelly, 1998). Additionally, condition-monitoring data, including periodic inspection data and continuous condition-monitoring data, has also 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 Chapter 6. The following sub-sections mainly introduce the incorporation of security-related data for revealing risk-related variations and dynamic barrier management.

7.3.4.2 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-based dynamic risk assessment may be applied to security risk analysis as it has been applied in the safety science domain. 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.

$$g(\lambda) = \frac{\beta^{\alpha}\lambda^{\alpha-1}}{\tau(\alpha)} e^{-\beta\lambda} \quad (7.4)$$

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).

$$P(r\ events\ in\ [0,t] \mid \lambda) = \frac{(\lambda t)^r}{r!} e^{-\lambda t} \quad (7.5)$$

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

$$g(\lambda \mid Data) = \frac{P(Data \mid \lambda) g(\lambda)}{\int P(Data \mid \lambda) g(\lambda)\, d\lambda} \propto P(Data \mid \lambda) g(\lambda) \quad (7.6)$$

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

$$g(\lambda \mid r\ events\ in\ [0,t]) = \frac{\beta'^{\alpha'} \lambda^{\alpha'-1}}{\tau(\alpha')} e^{-\beta'\lambda} \quad (7.7)$$

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

$$E(\lambda) = \frac{\alpha}{\beta} \quad (7.8)$$

$$E(\lambda') = \frac{\alpha'}{\beta'} = \frac{\alpha+r}{\beta+t} \quad (7.9)$$

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.

$$\frac{E(\lambda') - E(\lambda)}{E(\lambda)} = k\left(\frac{Pr' - Pr}{Pr}\right) \quad (7.10)$$

$$Pr' = Pr\left(\frac{E(\lambda') - E(\lambda) + E(\lambda) * k}{E(\lambda) * k}\right) \quad (7.11)$$

Where Pr is the prior probability of physical attacks against a chemical facility, which is estimated using the method presented in Table 3.2 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.

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 Figure 7.3. The prior annual frequency of facility/infrastructure attacks is initialized as a gamma distribution ($\Gamma(\alpha,\beta)$, α=58.18; β=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 is configured according to the method in Table 3.2. 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 Figure 7.4.

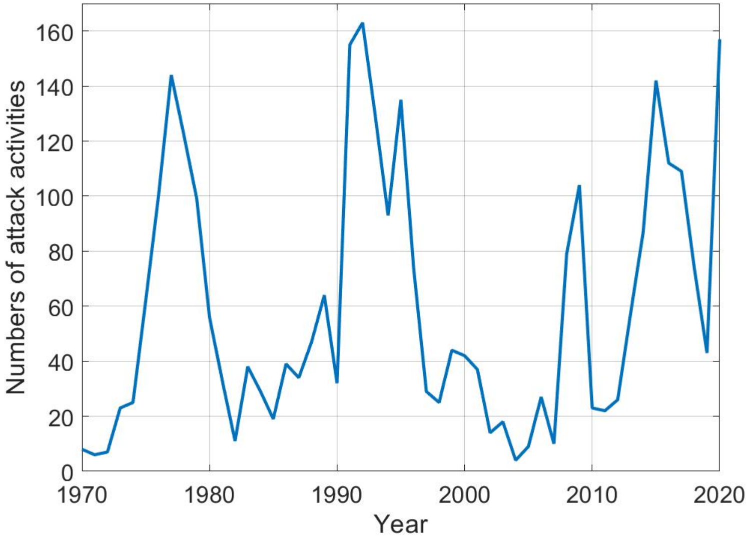
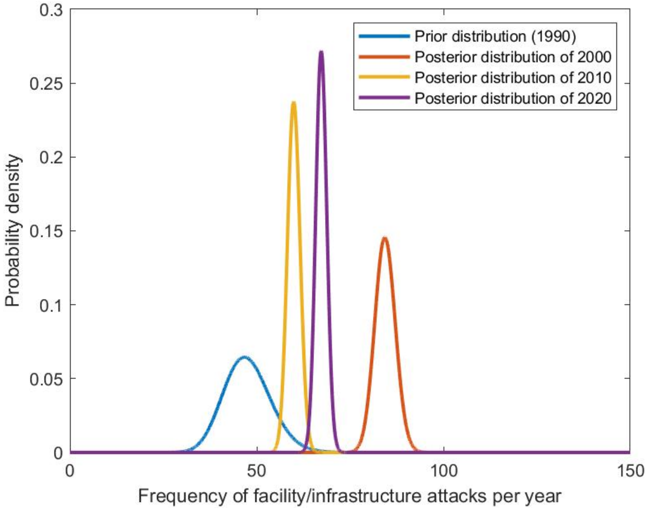
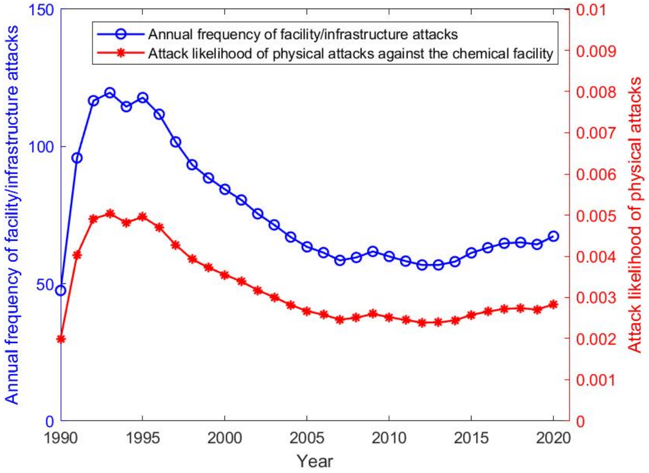


Figure 7.3. The numbers of facility/infrastructure attacks in Western Europe, adapted from (START, 2022).



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



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

Figure 7.4. Bayesian updating of attack likelihood estimations using precursor data.

Figure 7.4 (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 Eq. (7.4) to (7.7). Figure 7.4 (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. (7.11) ($k = 1$ is configured in Eq. 7.11).

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 Eq. (7.4) to (7.9). 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. (7.11) is used to update the plant-specific C2P attack likelihood based on indirectly related cyber incident data and expert judgment.

7.3.4.3 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 7.1 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. This study avoids repeated demonstration because Bayes's theorem has already been used in the vulnerability assessment of PPSs (van Staalduinen & Khan, 2015). 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 7.1). 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 usable 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 get acknowledged, for instance, disclosed by the CVE (Common Vulnerabilities and Exposures) data (NVD, 2023), the new vulnerabilities need to be accommodated to re-perform the vulnerability assessment based on the C2P attack vulnerability assessment model, as illustrated in Section 4.2.3.3. 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 (presented in Eq (4.10)). 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 variations mentioned above 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 7.2 and quantifying 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.

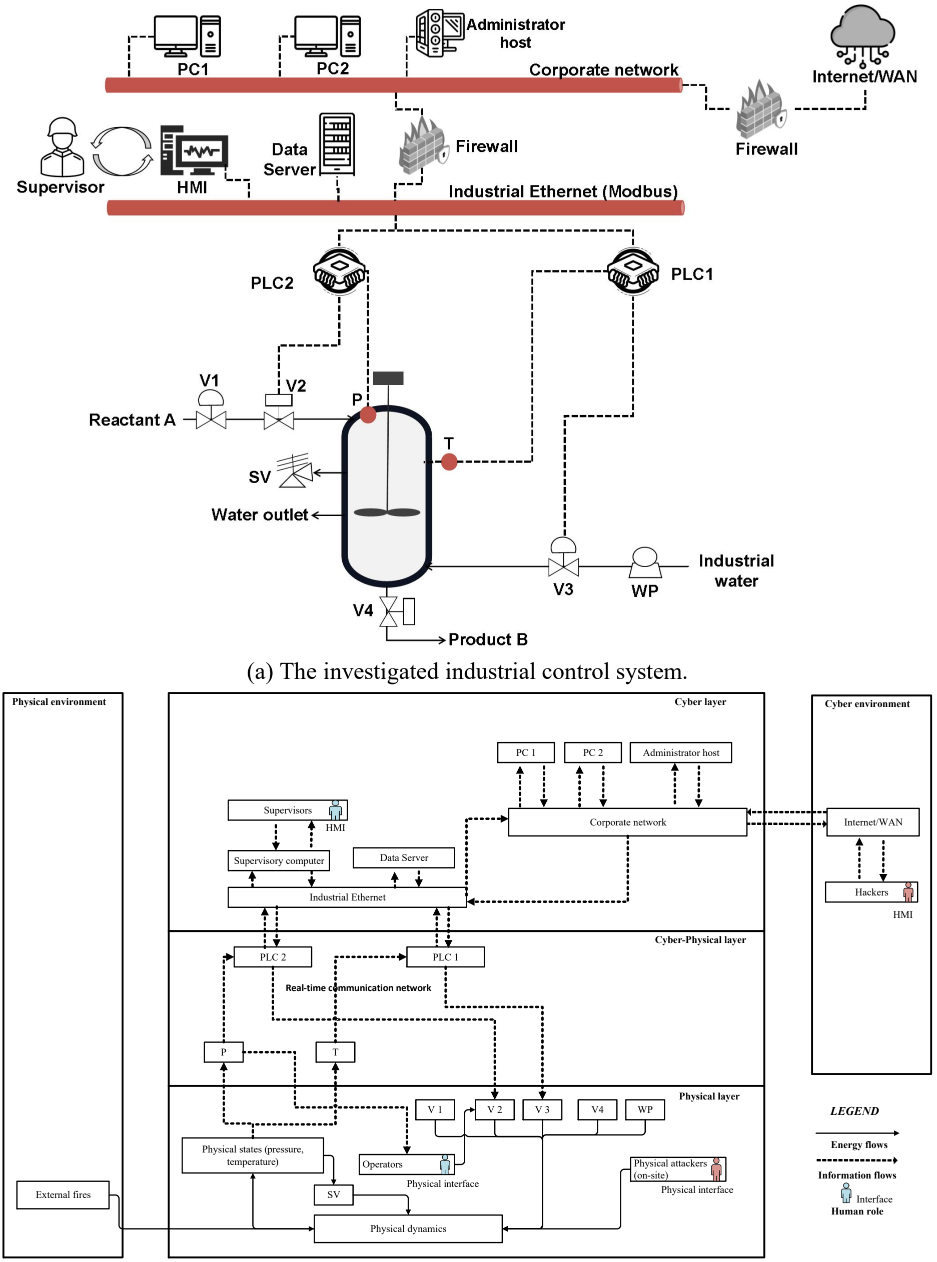
7.4 A case study of dynamic S&S barrier management

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.

7.4.1 Case study configurations

7.4.1.1 System description and BN model development

The industrial control system (ICS) demonstrated in Section 4.3 is adapted as the investigated system in this case study. Considering the potential safety failures, C2P attacks, and physical attacks, the proposed methodology is used to achieve dynamic and integrated S&S barrier management for this system. Figure 7.5 shows the basic information of the investigated ICS and its CPS master diagram. A detailed introduction to this system can be found in Section 4.3. We omit repeated elaboration here.

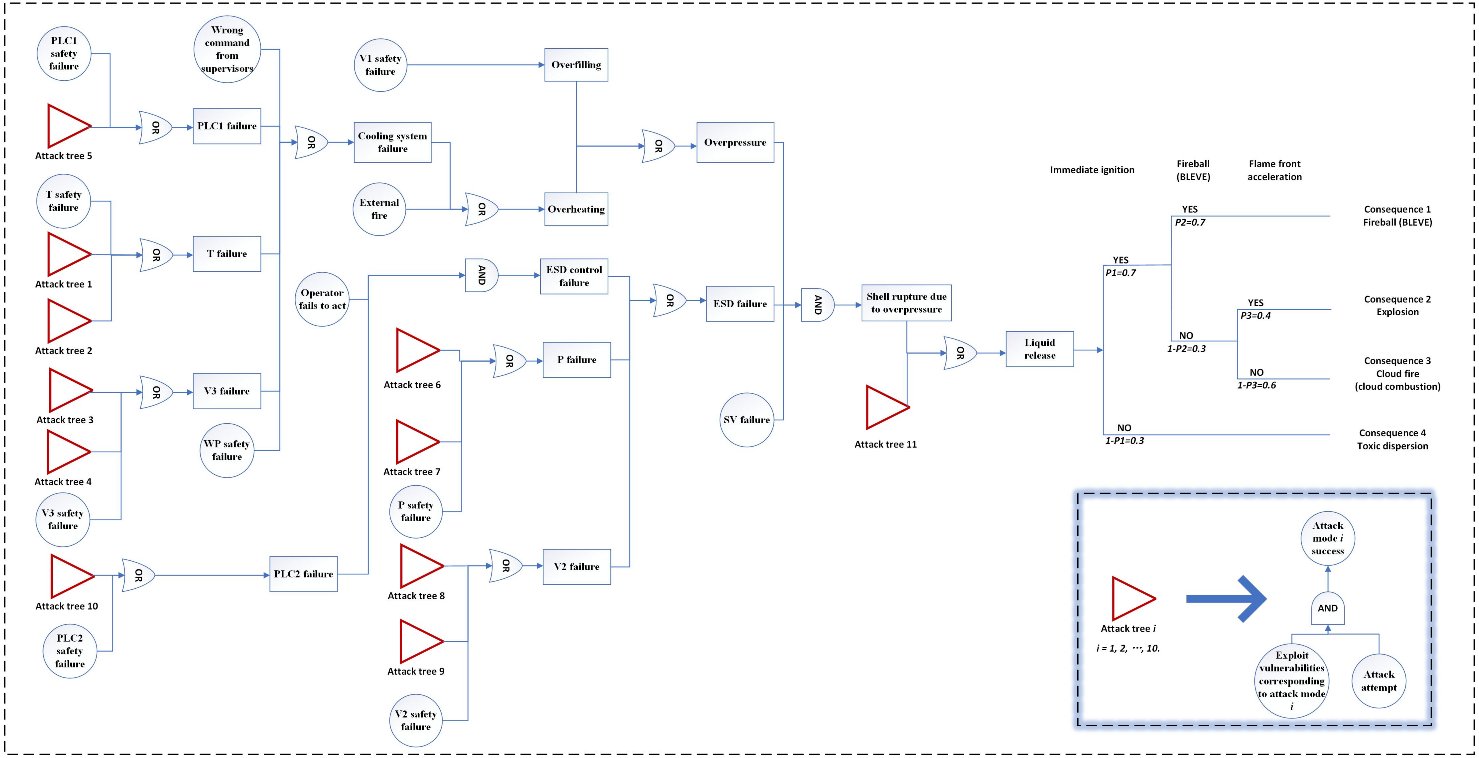


(a) The investigated industrial control system.

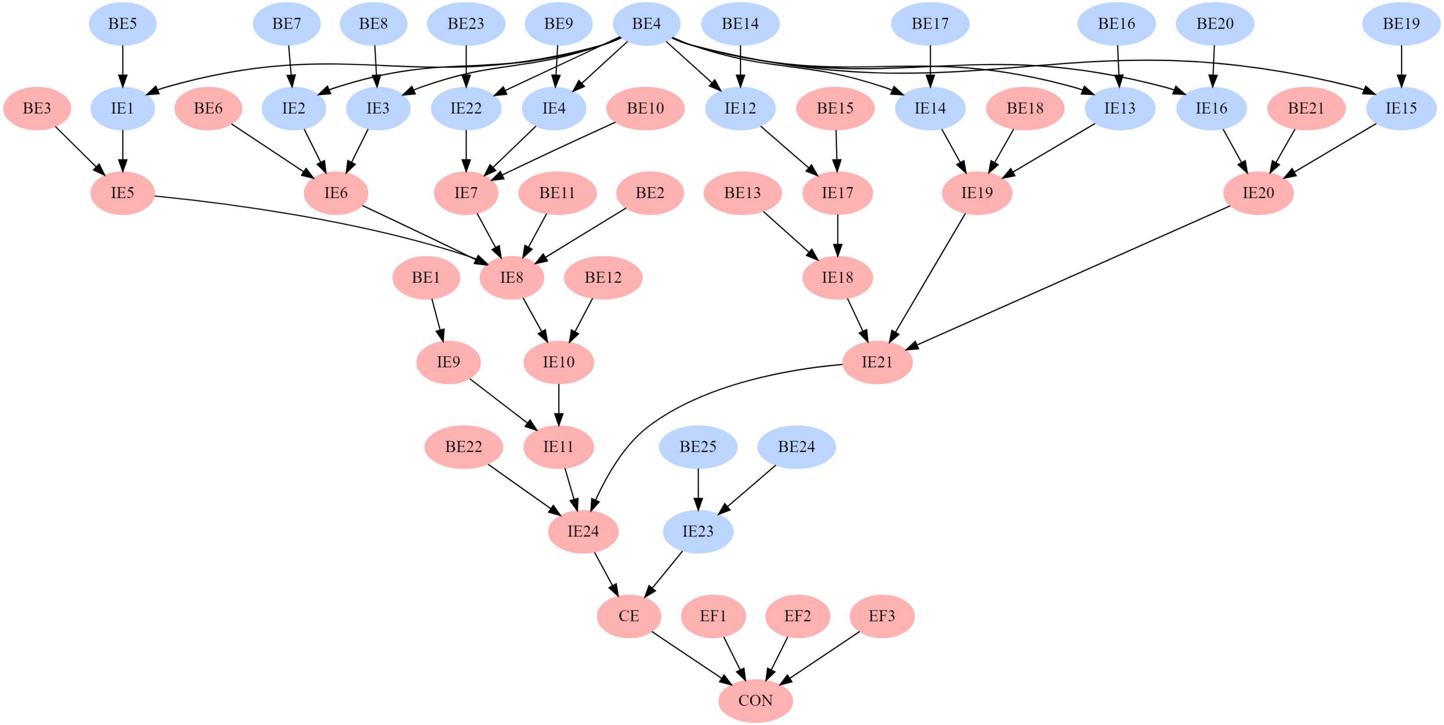
(b) The CPS master diagram.

Figure 7.5. The investigated industrial control system and its CPS master diagram.

Following the integrated safety and security risk assessment method, as presented in Section 7.3.2, an integrated attack-tree-bow-tie diagram is developed and then converted into a BN model (as shown in Figure 7.6). A list of the identified attack modes is given in Table 7.3. Explanations of the BN nodes are presented in Table 7.4. 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).



(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).

Figure 7.6. The developed attack-tree-bow-tie diagram and BN model.

Table 7.3 A summary of 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

Table 7.4 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[16]	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

[16] AT1 means attack mode 1, and the explanation of each attack mode can be found in Table 7.3.

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

7.4.1.2 Configurations of prior probabilities (probability distributions)

Configurations of the root nodes in the BN model are summarized in Table 7.5. 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.

$$PFD(t) = 1 - e^{-\lambda *(t\%T)},\ nT \leq t < (n+1)T \tag{7.12}$$

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 7.5.

2) The PFD calculation model from Zhang et al. (2020) is used for BE21, considering the degradation process of the shutdown valve. The configuration of the parameters in this model is also adapted from Zhang et al. (2020) and given in Table 7.5. The time interval for proof tests is configured as 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 6.2.3.3.

3) Regarding C2P attacks, the recurrence interval of attack attempts is estimated at approximately 150~465 days based on the data from Kuypers & Maillart (2018). A Gamma distribution ($\Gamma(\alpha,\ \beta)$, α=6.08; β=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 7.3.4.2.

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 in

the graph. The vulnerability assessment model presented in Section 4.2.3.3 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 Figure 7.8), are used as prior probability distributions for BN nodes: BE5, BE7, BE8, BE9, BE14, BE16, BE17, BE19, BE20, and BE23.

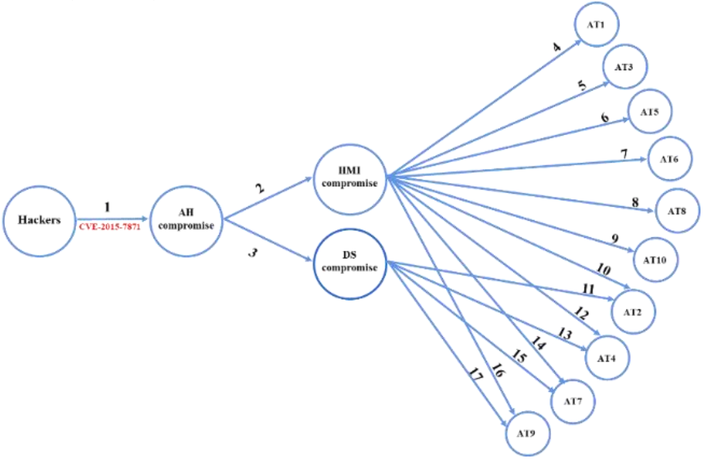
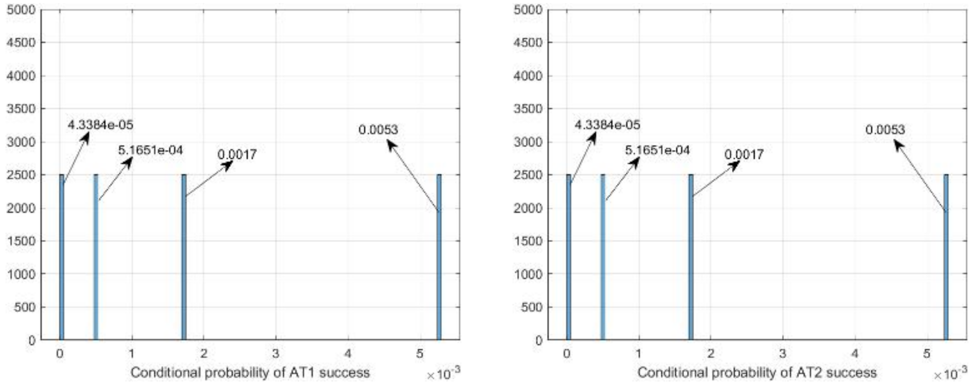
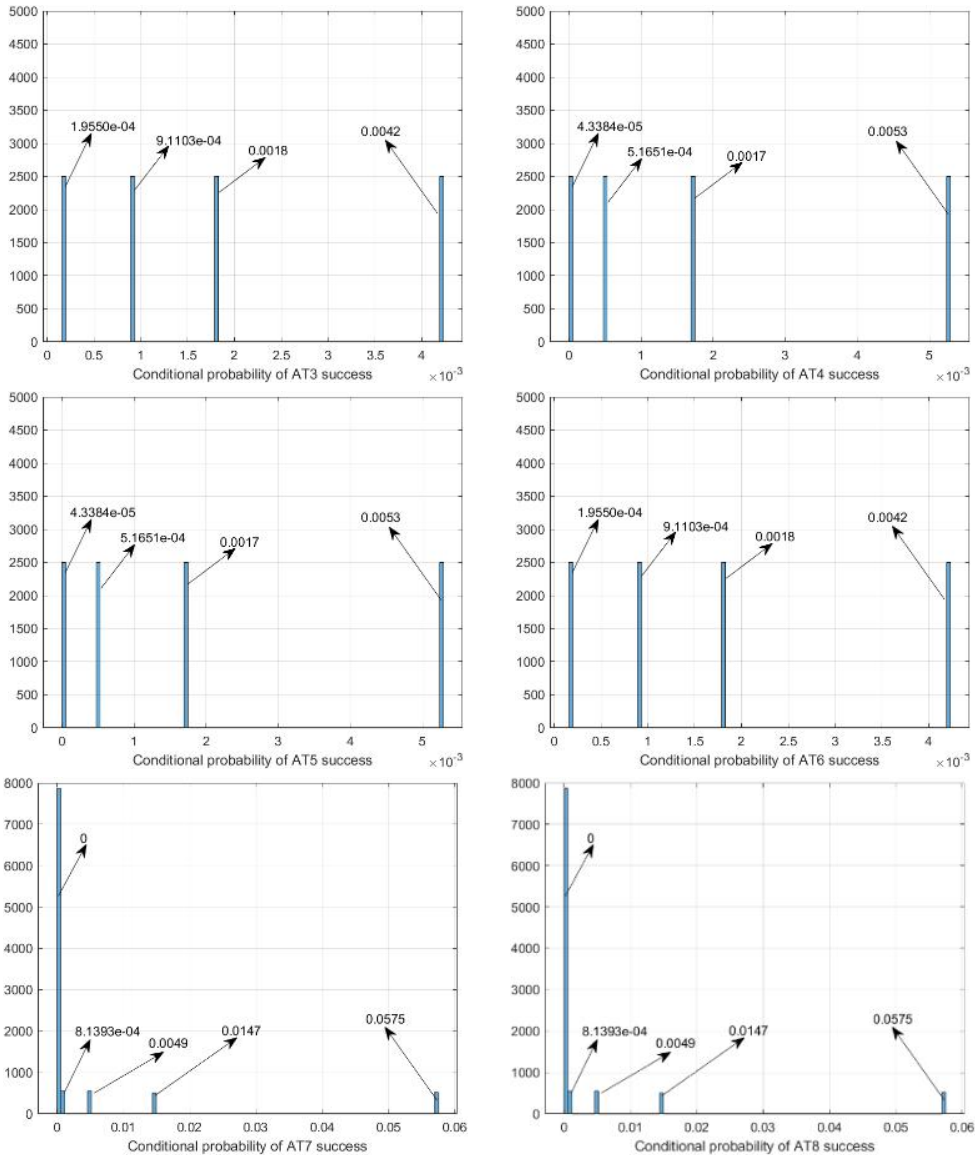


Figure 7.7. A compromise graph with the known CVEs along each attack step.




5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0 0.5 1 1.5 2 2.5 3 3.5 4
×10⁻³
1.9550e-04
9.1103e-04
0.0018
0.0042
Conditional probability of AT3 success
0 1 2 3 4 5
4.3384e-05
5.1651e-04
0.0017
0.0053
Conditional probability of AT4 success
4.3384e-05
5.1651e-04
0.0017
0.0053
Conditional probability of AT5 success
1.9550e-04
9.1103e-04
0.0018
0.0042
Conditional probability of AT6 success
8000
7000
6000
0
8.1393e-04
0.0049
0.0147
0.0575
0 0.01 0.02 0.03 0.04 0.05 0.06
Conditional probability of AT7 success
0
8.1393e-04
0.0049
0.0147
0.0575
Conditional probability of AT8 success

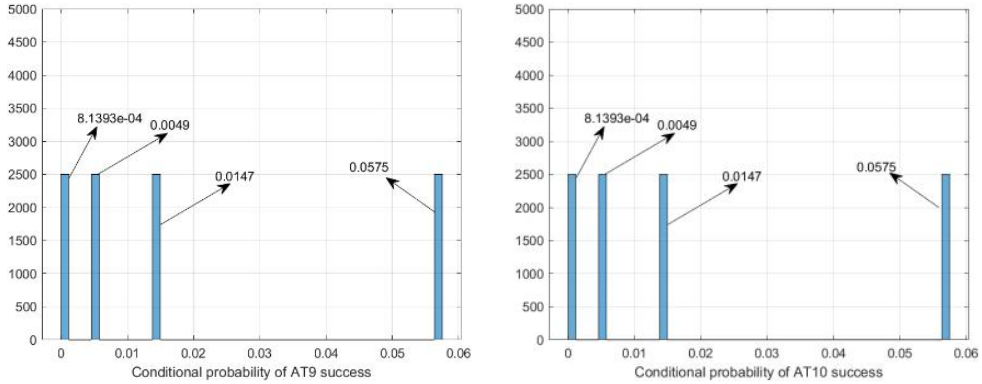
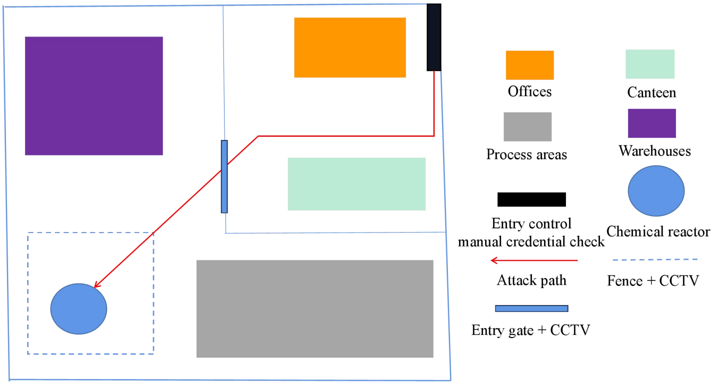


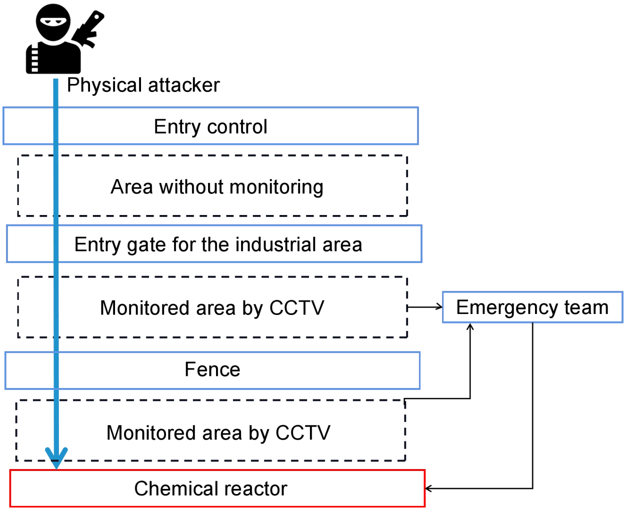
Figure 7.8. Probability distributions of successful implementation of each attack mode.

5) The annual frequency of physical attacks is initialized as 3.30E-03/year based on the method presented in Table 3.2, 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 7.3.4.2. 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 Figure 7.4. (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 7.3.2.2. A layout of the investigated chemical plant and its adversary sequence diagram are demonstrated in Figure 7.9 (a) and (b), respectively.



(a) Chemical plant layout.



(b) Adversary sequence diagram.

Figure 7.9. 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 7.1. The EASI model presented in Eqs (7.1) to (7.3) 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 Figure 7.10. According to Figure 7.10, the conditional probability of a successful physical attack given an attack attempt is calculated as 0.056. This result is used as prior probability for root node BE25.

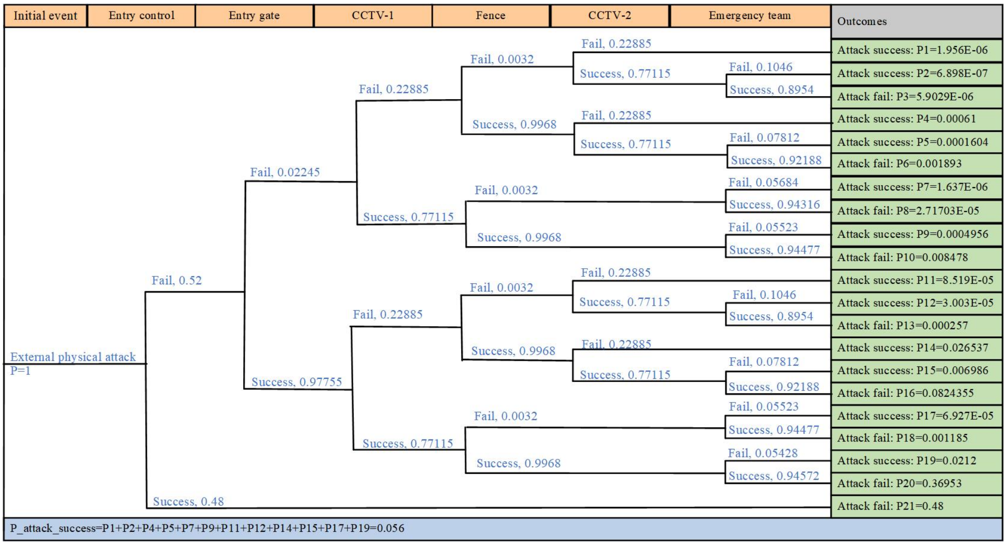


Figure 7.10. Vulnerability assessment of PPSs regarding external physical attacks using an event tree.

Table 7.5 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 & Onshus, 2010)	BE4	Gamma distribution ($\Gamma(\alpha, \beta)$, α=6.08; β=5)
BE5	Probability distribution for AT5, as shown in Figure 7.8.	BE6	2.13E-02 (Hauge & Onshus, 2010)
BE7	Probability distribution for AT1, as shown in Figure 7.8.	BE8	Probability distribution for AT2, as shown in Figure 7.8.
BE9	Probability distribution for AT3, as shown in Figure 7.8.	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 Figure 7.8.
BE15	average PFD=4.37E-03, λ=1.0E-06 (Hauge & Onshus, 2010); T= 1 year.	BE16	Probability distribution for AT6, as shown in Figure 7.8.
BE17	Probability distribution for AT7, as shown in Figure 7.8.	BE18	average PFD=6.57E-04, λ=1.5E-07 (Hauge & Onshus, 2010); T= 1 year.
BE19	Probability distribution for AT8, as shown in Figure 7.8.	BE20	Probability distribution for AT9, as shown in Figure 7.8.
BE21	Initial average PFD=3.5E-09; PFD is calculated and updated using Eq. (6.13), (α=1.02E-04, β=1.2E04, L=1.25E-03, τ=4380 h).	BE22	average PFD=5.47E-04, λ=5E-07 (HSE, 2012); T= 3 months.
BE23	Probability distribution for AT4, as shown in Figure 7.8.	BE24	Attack frequency estimations of physical attacks from 2016 to 2020 (as shown in Figure 7.4 (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)

7.4.2 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 Figure 7.11. 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.

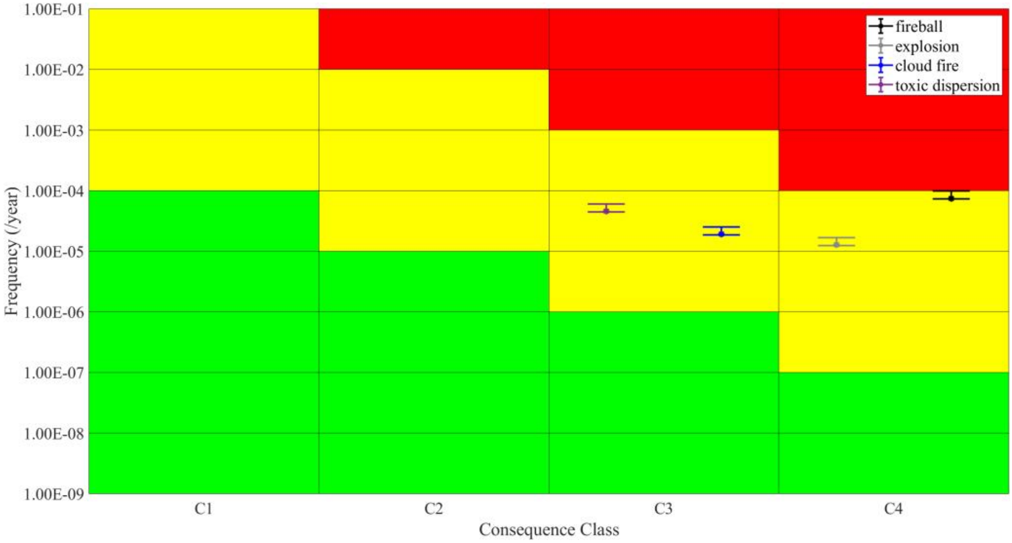


Figure 7.11. Initial risk profiles demonstrated in a risk matrix.

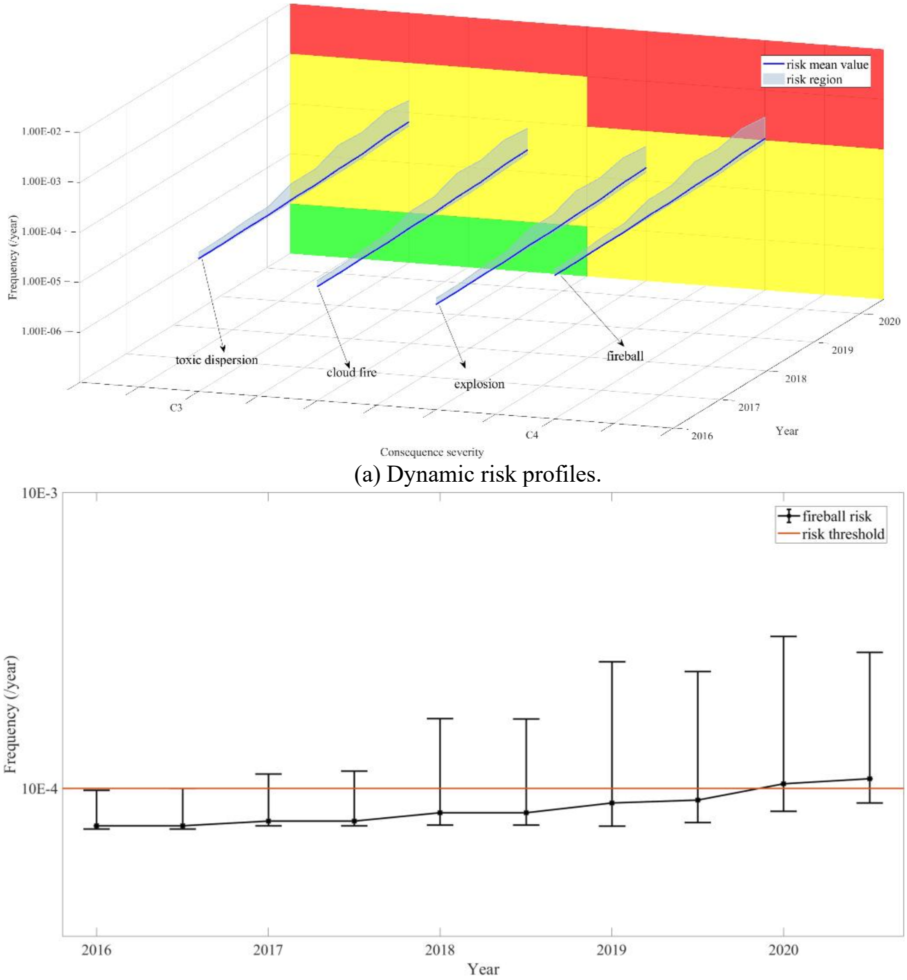
To demonstrate the capability of the proposed methodology in dynamic risk assessment and dynamic barrier management, we use some hypothetical data in relation to risk variations, as explained in Table 7.6. The risk profiles are updated based on the hypothetical data, as shown in Figure 7.12

Table 7.6 Hypothetical data for updating risks.

Data types	Descriptions										
Periodic proof test data on the shutdown valve	Time	2016		2017		2018		2019		2020	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
	Degradation level[17]	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[18]	Time	2016		2017		2018		2019		2020	
	Physical attack likelihood	0.002656		0.002724		0.002738		0.002707		0.002832	

[17] The failure threshold for degradation level: L=1.25E-03.

[18] Precursor data presented in Figure 7.3 is used to update the physical attack likelihood.



(a) Dynamic risk profiles.

(b) Fireball risk over time.

Figure 7.12. Dynamic risk profiles without timely S&S barrier improvement.

Figure 7.12(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 Figure 7.12(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. 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 Figure 7.13, 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 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 7.7. 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. 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 Figure 7.14.

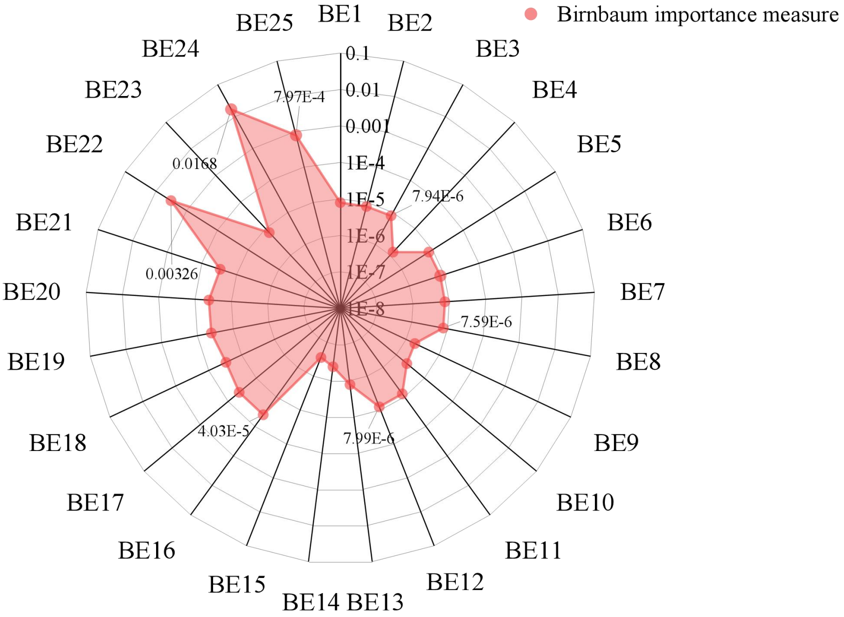


Figure 7.13. Sensitivity analysis of BN root nodes.

Table 7.7 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)[19]+ 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	/

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

	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	205,000€ (P maintenance cost)+102,000€ (V2	No	/

		● Maintain V2 immediately ● Patch CVE-2017-2683	maintenance cost)+1,420,000€ (patching CVE-2017-2683) = 1,727,000€		
	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	/
	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	/

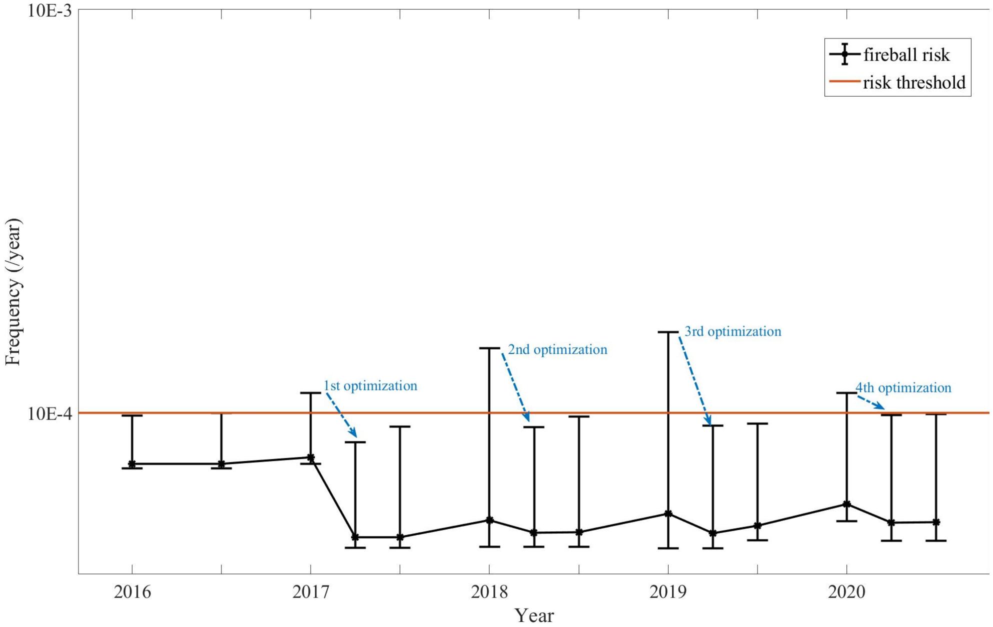


Figure 7.14. Dynamic risk profiles with timely S&S barrier improvement.

As shown in Figure 7.14, 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 7.7. Compared to Figure 7.12, the fireball risk is effectively mitigated by S&S barrier improvement when its maximum value is beyond the threshold in Figure 7.14, which demonstrates the advantage of dynamic barrier management in continuous and timely risk control.

7.5 Discussions

7.5.1 Notes when applying the proposed framework

The proposed framework is applied to a hypothetical case study by consolidating and implementing a set of approaches/methods developed for different purposes throughout this Ph.D. study. Although the case study demonstrates the feasibility and potential advantages of the proposed framework, some notes should be paid attention when applying the framework in practice.

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, 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 adverse events 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.

7.5.2 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.

7.6 Conclusions

This chapter proposed a systematic framework for dynamic and integrated S&S barrier management and demonstrated the application of the framework using a case study. The results show that multiple data in relation to the integrated safety and security risks may have the capability to reveal risk-related variations and further drive dynamic barrier management. The dynamic and integrated S&S barrier management has the advantages of making timely adaptations according to the new evidence, and ensuring the integrated S&S risks at acceptable levels continuously. Some practical notes are worthy of attention, and the establishment of a unified S&S risk management system is suggested for applying the proposed framework in practice.

Chapter 8 Conclusions and future research

This chapter wraps up the outcomes of this dissertation and answers the proposed main research question and research sub-questions. Limitations and future research directions are also discussed in this chapter.

8.1 Answer to research questions

This dissertation is dedicated to answering the proposed main research question by addressing the divided sub-questions point to point. The answers to the research questions are concluded as follows.

Sub-question 1: What are safety/security barriers?

To answer this question, a systematic literature review of the definition, classification, performance assessment, and management of safety and security barriers is conducted, as presented in Chapter 2. Based on the review of existing definitions of safety barriers, we define a safety barrier as a physical or non-physical tool planned to prevent, control, or mitigate undesired events or accidents. The means of safety barriers can vary from a technical facility or human action to a complex socio-technical system. The purposes of the safety barrier are to reduce the risk of an undesired event by reducing the occurrence probability of this event, limiting the extent and/or duration of the undesired event from escalation, or mitigating the impacts of the undesired event. Similarly, 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) or mitigate the corresponding consequences.

Sub-question 2. How to perform an integrated and quantitative risk assessment of chemical facilities considering both safety hazards/causes and security threats?

To answer this question, a quantitative risk assessment (QRA) approach is developed to assess the integrated safety and security risks of chemical facilities, considering the interactions between safety-associated and security-associated events. Due to the complexity of the industrial control systems (ICSs) in chemical plants, CPS master diagrams are employed to represent the ICSs in a multi-layered manner and help identify accident scenarios. Then, safety-associated and security-associated scenarios are integrated into an attack-tree-bow-tie diagram, based on which a BN model can be developed to perform quantitative risk assessment. Chapter 3 illustrates the developed approach and applies the approach to a case study.

Sub-question 3. How to handle uncertainties in safety and security risks properly and make decisions on barrier optimizations based on uncertain risks?

To handle the uncertainties in safety and security risks, a dedicated vulnerability assessment is developed to assess the vulnerability of ICS to cyber-physical attacks, considering the uncertainties associated with attackers' skill levels. Additionally, the combination of Monte Carlo simulations and BN modeling handles uncertainty propagation in the risk assessment, which enables the use of probability distributions for root nodes. Cost-effective analysis is employed to make decisions on barrier improvement based on a risk matrix demonstrating risk ranges. Chapter 4 elaborates on this approach and answers this research question. Using the developed approach, both the expected/mean values and the maximum values of the estimated risk ranges can be used as criteria to determine appropriate strategies for safety and security barrier improvement.

Sub-question 4. How to make decisions on safety and security barrier improvements effectively and cost-efficiently when facing large-solution-space optimization problems?

Chapter 5 is dedicated to answering this question. When facing a barrier optimization problem with a large solution space, it is usually too vast to search for the optimal strategy exhaustively in a reasonable amount of time. By contrast, evolutionary algorithms (for instance, genetic algorithms) help solve the optimization problem and determine the approximately optimal strategy. Chapter 5 develops a model for barrier modeling based on MATLAB/Simulink simulations considering both safety hazards and physical attacks. Meanwhile, a combination of cost-effectiveness analysis and a genetic algorithm is employed to solve the large-solution-space optimization problem and determine the approximately optimal strategy for barrier maintenance.

Sub-question 5. How to assess dynamic variations in safety and security risks and achieve integrated safety and security barrier management dynamically?

Data from different sources, including condition-monitoring data, accident precursor data, security incident data, etc., are characterized in Chapter 7. Those data have the potential to reveal the dynamic variations in risk profiles and, therefore, enable dynamic barrier management. Different models are combined to analyze the multi-source data and assess the risk-related variations. For instance, Bayesian updating handles accident precursor data, reliability models tackle condition-monitoring data, and the C2P attack vulnerability assessment model addresses cyber incident data. Chapter 6 provides an approach for dynamic barrier management considering the degradation of safety barriers based on multiple data (periodic proof test data, continuous condition-monitoring data, and accident precursor data). Chapter 7 proposes a systematic framework for dynamic and integrated safety and security barrier management combining multiple safety-related or security-related data.

Main question: How to integrate and manage safety and security barriers effectively

and cost-efficiently with respect to major adverse events in the chemical process industries?

To integrate safety and security barriers, a systematic risk assessment approach is developed to integrate safety-associated and security-associated scenarios regarding major adverse events in chemical plants. Meanwhile, the performance assessment of barriers is incorporated with the risk assessment to evaluate the effectiveness of barriers in controlling risk. The developed risk assessment approach is provided in Chapter 3. Furthermore, multiple uncertainties associated with the integrated safety and security risks are considered to provide more usable information for decision-making, as illustrated in Chapter 4. The combination of cost-effective analysis and tailored optimization algorithms helps get insights into the trade-offs between barrier management costs and undesired risks and decide the optimal barrier improvement strategy in case risk profiles are unacceptable.

Additionally, dynamic barrier management is achieved by analyzing multi-source data to reveal dynamic variations in barrier performance, update risk profiles, and make adaptions to barrier management strategies. When new evidence becomes available over time, barrier performance is maintained continuously to ensure the undesired risks are at acceptable levels despite dynamic risk variations. The approaches for dynamic barrier management are illustrated in Chapters 6 and 7. A framework and an integrated approach incorporating all methodologies developed by this study are proposed and demonstrated in Chapter 7 to facilitate dynamic and integrated safety and security barrier management in chemical plants.

8.2 Recommendations for future research

8.2.1 Human reliability analysis for barrier management

This study used rough reference values for human failure probabilities in the risk assessment. For instance, the reference values for the probabilities of failure on demand (PFDs) of human barriers are adapted from the ARAMIS project (Andersen et al., 2004). However, the human failure probabilities may vary depending on many factors, such as organizational factors, environmental factors, and task contexts. Human Reliability Analysis (HRA) techniques (Zarei et al., 2021) help assess more credible and dedicated human failure probabilities in the context of safety and security barrier management. Additionally, implementing autonomous control systems in chemical process facilities reduces the reliance on human actions in the systems' operation and, therefore, reduces human-made accidents (El-Kady, 2023). However, human interventions (human-machine interfaces) are still integral parts of various processes of industrial control systems. The task contexts of human roles changed dramatically with the increasing machine intelligence and digitization in the Industry 4.0 era. Assessing the reliability of human actions in the daily operation and emergency response processes of industrial cyber-physical systems is suggested to provide valuable databases for integrated safety and security barrier management in the digitalization era.

8.2.2 Incorporation with condition-monitoring techniques

This study incorporates the data from multiple sources for barrier management, among the data sources, periodic inspection data and continuous condition-monitoring data has been tentatively used to predict the health status of safety instrumented systems and basic process control systems in Chapter 6. With the advancement of artificial intelligence technologies and data-driven approaches, condition-monitoring techniques have been gradually used in asset maintenance and risk assessment over different industrial sectors. For instance, condition-monitoring techniques are utilized to drive condition-based maintenance, which aims to prevent the asset from losing its functionality by estimating the asset degradation based on the monitoring data and planning maintenance activities (Wang et al., 2022). Condition-monitoring can also enable predictive maintenance, which is able to process the monitoring data for prognostics and allows more proactive maintenance planning (Pinciroli et al., 2023). The condition-monitoring data has also been leveraged to support dynamic risk assessment (Zeng & Zio, 2018; Xing et al., 2019). The advancement of condition-monitoring techniques has the potential to reveal and even predict the health status of technical components in the context of S&S barrier management. How to leverage advanced condition-monitoring techniques to risk assessment and further support barrier management is recommended to be researched in future studies.

8.2.3 Resilience-based S&S barrier management

This study investigated risk-based frameworks and approaches for dynamic and integrated safety and security barrier management. Resilience is defined as the ability of a system to adapt and absorb any internal and external negative impacts and maintain a normal state or achieve recovery to a normal state after disruptions (Bento et al., 2021). The resilience concept has attracted significant attention from both academia and industries over the past decade. Applying resilience engineering principles to industrial systems helps design optimization for resilience enhancement, derive efficient restoration strategies, and so on (Pawar et al., 2021). Considering the complexity of industrial cyber-physical systems, there is a great need to call for a resilience assessment, which reveals the system's ability to perform the required functions under expected and unexpected disruptions (Colabianchi et al., 2021). Assessing the resilience of industrial cyber-physical systems before and after disruptions supports decision-making in selecting resilience strategies (Cassottana et al., 2023). Applying the resilience concept and resilience engineering principles in the context of safety and security barrier management helps guide barrier allocation and optimization from a resilience perspective.

8.2.4 Assessment of new attack modes and attack scenarios

This study considered several typical cyber-physical attack modes (false data injection attacks, denial-of-service attacks, and setpoint manipulations) that may induce physically dangerous scenarios. The study aims to demonstrate the noticeable cyber-physical attack risks and how they can be incorporated with the conventional safety QRA (quantitative risk assessment) instead of providing a thorough assessment of security threats and their attack modes. As a result, only selected attack modes and their corresponding attack scenarios have been investigated in the case studies. Due to the complexity of the industrial cyber-physical system (ICPS) and various possible threat agents, more attack modes and their credible attack scenarios may exist. For instance, APTs (advanced persistent threats) may exploit the ICPSs by implementing stealthy attacks (Ghafir & Prenosil, 2014). Physical attacks against IT infrastructures, the combination of cyberattacks and physical attacks (cyberattacks against the entrance control system followed by a physical intrusion), and the exploitation of ICPSs using social engineering techniques or internal fraud are also threatening ICPSs (Flaus, 2019). Apart from forcing the chemical process system into an unsafe state, the attackers may only aim to cause economic consequences by inducing abnormal deviations in the system (Huang et al., 2009). The attack mechanisms of those different attack modes are worthy of investigation. Considering more potential threat agents and their possible attack modes, the incorporation and assessment of more attack scenarios may be investigated in further research.

8.2.5 Multi-objective optimizations in barrier management

This study combines cost-effective analysis and exhaustive searching algorithms or genetic algorithms for decision-making on S&S barrier optimization, considering the trade-offs between barrier improvement costs and undesired risks. Chapter 5 presents a method (presented in Eqs. 5.7 and 5.8) to solve the barrier optimization problems with one objective function and multiple constraints, which can address two typical practices (minimize investment costs or maximize barrier effectiveness) in the cost-effectiveness analysis. Sometimes, the decision-makers may face more complicated problems that cannot be handled using proposed methods. For instance, decision-makers may have multiple objectives that need to be considered in some cases, and the method used in Chapter 5 becomes unable to solve. Alternatively, a multi-objective optimization needs to be conducted to solve the optimization problem with multiple objectives (Di Maio et al., 2023). Therefore, it is possible that more complicated optimization problems exist in the S&S barrier management process. Investigating multi-objective and other optimization problems in barrier management is recommended for future research.

Appendix I

Method for cyber-physical attack modeling

For C2P attacks against sensors, let $y_i(k)$ denote the measurement by sensor i at time k, and the sensor measurement is always within its predefined range, $y_i(k) \in Q = \left[y_i^{min}, y_i^{max}\right]$. Let $\tilde{y}_i(k)$ denote the received measurement by the controller at time k. If this sensor is under attack, $\tilde{y}_i(k)$ may be different from the real measurement $y_i(k)$, as follows (Cárdenas et al., 2011):

$$\tilde{y}_i(k) = \begin{cases} y_i(k) & for\ k \notin K \\ a_i(k) & for\ k \in K \end{cases} \tag{A1.1}$$

where $K = \{k_s, \cdots, k_e\}$ represents the attack duration between the attack start time k_s and the attack stop time k_e. $a_i(k)$ is the manipulated data by the attack. Because the manipulated measurement outside Q can be easily detected as a fault by the fault-tolerant algorithms, it is assumed that $a_i(k)$ also lies within Q (presented by $a_i(k) \in Q$). In the case of a DoS attack, a lack of measurement occurs during the attack. Consequently, the last received measurement will be used by the controller until new measurements are received after the DoS attack, as follows (Cárdenas et al., 2011):

$$a_i(k) = y_i(k_s),\ for\ k \in K \tag{A1.2}$$

where $y_i(k_s)$ is the last received measurement before the DoS attack starts. In terms of FDI attacks, attackers can inject any arbitrary value to manipulate the measurement data received by the controller. Thus, $a_i(k)$ can be any arbitrary value within the measurement range of the sensor. For instance, Min and Max Attacks, Scaling Attacks, and Additive Attacks are the possible methods implemented by attackers for FDI attacks, illustrated as follows (Huang et al., 2009):

i) Min and Max Attacks

$$a_i(k) = y_i^{min},\ for\ k \in K;\ Min\ attacks \tag{A1.3}$$

and

$$a_i(k) = y_i^{max},\ for\ k \in K;\ Max\ attacks \tag{A1.4}$$

ii) Scaling Attacks

$$a_i(k) = \begin{cases} \beta(t)y_i(k),\ for\ k \in K\ and\ \beta(t)y_i(k) \in Y_i \\ y_i^{min},\ for\ k \in K\ and\ \beta(t)y_i(k) \le y_i^{min} \\ y_i^{max},\ for\ k \in K\ and\ \beta(t)y_i(k) \ge y_i^{max} \end{cases} \tag{A1.5}$$

where $\beta(t)$ is the scale factor, which is a function of time t.

iii) Additive Attacks

$$a_i(k) = \begin{cases} y_i(k) + \gamma(t),\ for\ k \in K\ and\ y_i(k) + \gamma(t) \in Y_i \\ y_i^{min},\ for\ k \in K\ and\ y_i(k) + \gamma(t) \le y_i^{min} \\ y_i^{max},\ for\ k \in K\ and\ y_i(k) + \gamma(t) \ge y_i^{max} \end{cases} \quad \text{(A1.6)}$$

where $\gamma(t)$ is the scale factor, which is a function of time t. Similarly, the approach for modeling C2P attacks against sensors can also be adapted for C2P attacks against actuators, as follows (Huang et al., 2009):

$$\tilde{u}_i(k) = \begin{cases} u_i(k) & for\ k \notin K \\ a_i(k) & for\ k \in K \end{cases} \quad \text{(A1.7)}$$

where $u_i(k)$ is the correct control data from the controller to actuator i at time k. $\tilde{u}_i(k)$ is the control data received by the actuator at time k. When the actuator is under C2P attacks ($k \in K$), the manipulated signal ($a_i(k)$) will be used by the actuator instead of the correct control data $u_i(k)$. $a_i(k)$ in Eq (A1.7) can also be calculated following the methods presented by Eq (A1.2) to Eq (A1.6) with the replacement of sensor measurement y to control data u. A detailed illustration of the approach for impact analysis of FDI attacks and DoS attacks can be found in (Huang et al., 2009). For setpoint manipulations, the setpoint used by the controller may be modified by attackers as any arbitrary value ($a_i(k)$), as follows (Wen et al., 2023):

$$\tilde{s}_i(k) = \begin{cases} s_i & for\ k \notin K \\ a_i(k) & for\ k \in K \end{cases} \quad \text{(A1.8)}$$

where s_i is the predefined setpoint value. $\tilde{s}_i(k)$ is the setpoint value used by the controller at time k. By integrating the above attack modeling into a system control model, the system state vector with n variables ($X = \{x_1, \ldots, x_n\}$) under the influence of C2P attacks can be evaluated, as demonstrated below.

$$\begin{cases} X(k+1) = f(X(k), \tilde{U}(k), w) \\ Y(k) = g(X(k), v) \\ U(k) = h(\tilde{S}(k), \tilde{Y}(k)) \end{cases} \quad \text{(A1.9)}$$

where $X(k+1)$ is the system state vector at time $k+1$, which depends on $X(k)$, the control actions of l actuators, $\tilde{U}(k)$={$\tilde{u}_1(k), \ldots, \tilde{u}_l(k)$}, and the process noise (w). $Y = \{y_1, \ldots, y_m\}$ is the observation vector composed of the observation data of m variables. $Y(k)$ depends on the system state vector, $X(k)$, and the observation noise (v). $U(k) = \{u_1(k), \ldots, u_l(k)\}$ is the control data for actuators, which depends on the j setpoint values, $\tilde{S}(k) = \{\tilde{s}_1(k), \ldots, \tilde{s}_j(k)\}$, and the observed data from sensors ($\tilde{Y}(k)$={$\tilde{y}_1(k), \ldots, \tilde{y}_m(k)$}). $\tilde{Y}(k)$, $\tilde{U}(k)$, and $\tilde{S}(k)$ are modeled using Eq (3) to Eq (10) according to the specific attack modes and are used to estimate the system state vector. If we define a safety range for each system state variable, $R = \begin{bmatrix} x_1^{min},\ x_1^{max} \\ \ldots \\ x_n^{min},\ x_n^{max} \end{bmatrix}$, a dangerous system state is induced by a C2P attack when $X(k) \notin R$ and $k \in K$.

Therefore, the attack modeling helps to decide if a dangerous phenomenon can be induced by certain C2P attack modes based on the estimation of the system state vector. Additionally, we introduce a coefficient, β_i, to depict the likelihood of a physically dangerous scenario that may be induced by a specific attack mode, i, as follows.

$$\beta_i = Pr\{X(k) \notin R\},\ k \in K \tag{A1.10}$$

where $Pr\{X(k) \notin R\}$ is the probability of $X(k) \notin R$ regarding a specific attack mode. $X(k)$ is estimated by using Eq. (A1.9). $X(k)$ depends on $\widetilde{U}(k)$, $\tilde{S}(k)$, $\widetilde{Y}(k)$, w and v. Attack modes impact the configuration of $\widetilde{U}(k)$, $\tilde{S}(k)$ and $\widetilde{Y}(k)$. As a result, attack modes, process noise, and observation noise may impact the value of β_i. The determination of β_i for a specific attack mode should be conducted based on attack modeling of this attack mode.

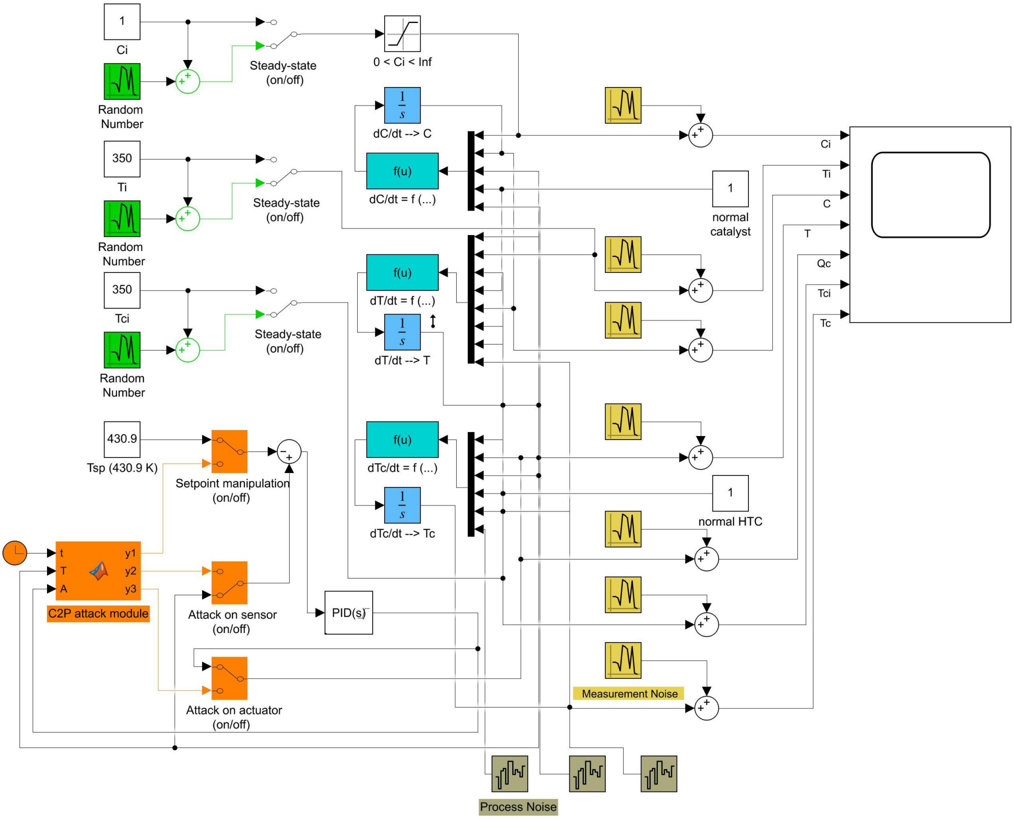
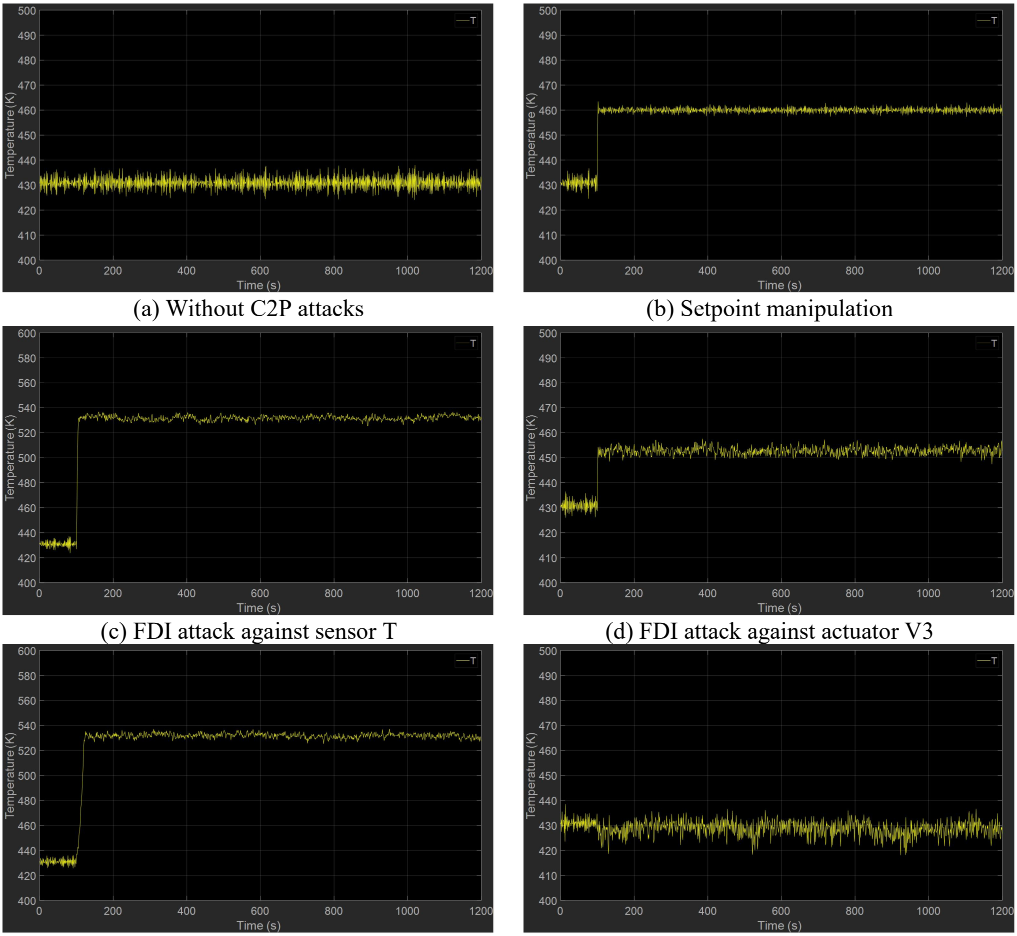


Figure A1. A C2P attack assessment model developed based on MATLAB/Simulink.



(a) Without C2P attacks

(b) Setpoint manipulation

(c) FDI attack against sensor T

(d) FDI attack against actuator V3

(e) DoS attack against sensor T

(f) DoS attack against actuator V3

Figure A2. C2P attacks' effects on the reactor temperature (attacks start from 100 s).

It is considered a dangerous overheating scenario when the reactor temperature overpasses 450 K ($X(k) \notin R$, when $T(k) > 450\ K$). Consequently, the β_i value for each C2P attack mode can be determined using Monte Carlo simulations ($\beta_i = Pr\{X(k) \notin R\}$). Regarding C2P attacks against ESD systems, $\beta_i = 1$ is always the case because a manipulation of the ESD system can stop the ESD system from performing its functionality on demand, no matter the process noise and observation noise.

Appendix II

Method for Time-to-compromise (TTC) estimation

In the TTC estimation approach, TTC is modeled as a random process composed of three subprocesses: i) at least one vulnerability is known, and the attacker has the exploit readily available that can be used to exploit the known vulnerability successfully, ii) at least one vulnerability is known, but the attacker must develop an exploit for it, and iii) the attacker must find and exploit new vulnerabilities because either no known vulnerabilities exist or the attacker is unable to exploit known vulnerabilities. The expected TTC of an attack step is estimated as follows (McQueen et al., 2006a; Ling & Ekstedt, 2022).

$$TTC = t_1P_1 + t_2(1-P_1)(1-u) + t_3u(1-P_1) \qquad \text{(A2.1)}$$

where t_i is the expected time used in subprocess i (i= *1, 2, 3*) in days and P_1 is the probability of being in subprocess 1. u is the probability that subprocess 2 is unsuccessful. The probabilities for an attacker to be in subprocess 1 and 2 are calculated as follows (Ling & Ekstedt, 2022).

$$P_1 = 1 - e^{-vm/k} \qquad \text{(A2.2)}$$

$$P_2 = e^{-vm/k} = 1 - P_1 \qquad \text{(A2.3)}$$

where v is the number of vulnerabilities on a specific component and m is the number of exploits readily available to the attacker. k is the total number of vulnerabilities in the database. The value of k is 2740 according to the ICS vulnerability dataset (Thomas & Chothia, 2020) available on October 5th, 2023. Subprocess 3 is considered running in parallel to subprocess 1 and 2, therefore, there is no need to estimate the probability of an attacker to be in subprocess 3. The time taken to complete each subprocess is estimated as below (Ling & Ekstedt, 2022).

$$t_1 = 1 * ((10/C_2 + 3.9/C_3)2) \qquad \text{(A2.4)}$$

$$t_2 = 37 \text{ (novice), } 27 \text{ (beginner), } 16 \text{ (intermediate), or } 6 \text{ (expert)} \qquad \text{(A2.5)}$$

$$t_3 = \left(f' - 0.5\right) * b + t_2 \qquad \text{(A2.6)}$$

where C_2 is the average base score of the vulnerabilities derived from CVSS v2.0[20] and C_3 is the average exploitability score of the vulnerabilities derived from CVSS v3.0[21]. In terms of t_2, 37 days, 27 days, 16 days, and 6 days are used for novice, beginner, intermediate, and expert attackers respectively. b is the MeanTime-Between-Vulnerabilities (MTBV) in days as calculated from the ICS advisory creation date (Thomas & Chothia, 2020). f is the fraction of vulnerabilities that are exploitable to the attacker, and it is determined based on Table A1. The

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

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

probability that subprocess 2 is unsuccessful (u) is calculated as $u = (1 - f)^v$. An Excel tool[22] developed by Thomas & Chothia (2020) was used to perform the TTC estimations.

Table A1 The number and fraction of exploitable vulnerabilities to attackers with different skill levels, adapted from (Ling & Ekstedt, 2022).

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

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

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. Computers & Security, 72, 175-195.

Alanen, J., Linnosmaa, J., Malm, T., Papakonstantinou, N., Ahonen, T., Heikkilä, E., & Tiusanen, R. (2022). Hybrid ontology for safety, security, and dependability risk assessments and Security Threat Analysis (STA) method for industrial control systems. *Reliability Engineering & System Safety, 220,* 108270.

Amin, M. T., Khan, F., Halim, S. Z., & Pistikopoulos, S. (2022). A holistic framework for process safety and security analysis. *Computers & Chemical Engineering, 165,* 107963.

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.

Andrews, J. D., & Dunnett, S. J. (2000). Event-tree analysis using binary decision diagrams. IEEE Transactions on Reliability, 49(2), 230-238.

Aven, T. (2007). A unified framework for risk and vulnerability analysis covering both safety and security. *Reliability Engineering & System Safety, 92(6),* 745-754.

Ayyub, B. M., McGill, W. L., & Kaminskiy, M. (2007). Critical asset and portfolio risk analysis: An all-hazards framework. *Risk Analysis: An International Journal, 27(4),* 789-801.

API. (2013). ANSI/API Standard 780 – Security risk assessment methodology for the petroleum and petrochemical industry. American Petroleum Institute, Washington, DC.

Argenti, F., Landucci, G., Cozzani, V., & Reniers, G. (2017). A study on the performance assessment of anti-terrorism physical protection systems in chemical plants. *Safety Science, 94,* 181-196.

Basri, E. I., Razak, I. H. A., Ab-Samat, H., & Kamaruddin, S. (2017). Preventive maintenance (PM) planning: a review. Journal of Quality in Maintenance Engineering, 23, 114-143.

Basheer, A., Tauseef, S. M., Abbasi, T., & Abbasi, S. A. (2019). Methodologies for assessing risks of accidents in chemical process industries. *Journal of Failure Analysis and Prevention, 19,* 623-648.

Badreddine, A., Romdhane, T. B., HajKacem, M. A. B., & Amor, N. B. (2014). A new multi-objectives approach to implement preventive and protective barriers in bow tie diagram. *Journal of Loss Prevention in the Process Industries, 32,* 238-253.

Bellamy, L., Oh, J. I. H., Hale, A. R., Papazoglou, I. A., Ale, B. J. M., Morris, M., Aneziris, O., Post, J. G., Walker, H., Brouwer, W. G. J. & Muyselaar, A. J. (1999). I-RISK development of an integrated technical and management risk control and monitoring methodology for managing and quantifying on-site and off-site risks. Final Project Report ENVA-CT96-0243.

Bento, F., Garotti, L., & Mercado, M. P. (2021). Organizational resilience in the oil and gas industry: A scoping review. *Safety Science, 133,* 105036.

Bier, V. M., & Lin, S. W. (2013). On the treatment of uncertainty and variability in making decisions about risk. *Risk Analysis, 33(10),* 1899-1907.

Bier, V., & Gutfraind, A. (2019). Risk analysis beyond vulnerability and resilience–characterizing the defensibility of critical systems. *European Journal of Operational Research, 276(2),* 626-636.

Bobbio, A., Portinale, L., Minichino, M., & Ciancamerla, E. (2001). Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering & System Safety, 71(3),* 249-260.

Bubbico, R., Lee, S., Moscati, D., & Paltrinieri, N. (2020). Dynamic assessment of safety barriers preventing escalation in offshore Oil&Gas. *Safety Science, 121,* 319-330.

Bucelli, M., Paltrinieri, N., Landucci, G., & Cozzani, V. (2017). Safety barrier management and risk assessment: integration for safer operations in the Oil&Gas industry. In Presented at HAZARDS 27, SYMPOSIUM SERIES NO 162. IChemE.

Bucelli, M., Landucci, G., Haugen, S., Paltrinieri, N., & Cozzani, V. (2018). Assessment of safety barriers for the prevention of cascading events in oil and gas offshore installations operating in harsh environment. *Ocean Engineering, 158,* 171-185.

Caputo, A. C., Pelagagge, P. M., & Palumbo, M. (2011). Economic optimization of industrial safety measures using genetic algorithms. Journal of Loss Prevention in the Process Industries, 24(5), 541-551.

Cai, B., Liu, Y., Liu, Z., Tian, X., Dong, X., & Yu, S. (2012). Using Bayesian networks in reliability evaluation for subsea blowout preventer control system. *Reliability Engineering & System Safety, 108,* 32-41.

Cárdenas, A. A., Amin, S., Lin, Z. S., Huang, Y. L., Huang, C. Y., & Sastry, S. (2011, March). Attacks against process control systems: risk assessment, detection, and response. In Proceedings of the 6th ACM symposium on information, computer and communications security (pp. 355-366).

Casciano, M., Khakzad, N., Reniers, G., & Cozzani, V. (2019). Ranking chemical industrial clusters with respect to safety and security using analytic network process. *Process Safety and Environmental Protection, 132,* 200-213.

Cassottana, B., Roomi, M. M., Mashima, D., & Sansavini, G. (2023). Resilience analysis of cyber-physical systems: A review of models and methods. *Risk Analysis, 43(11),* 2359-2379.

Çetinkaya, E. K. (2001). Reliability analysis of SCADA Systems used in the offshore oil and gas industry.

Chaturvedi, D. K. (2017). Modeling and simulation of systems using MATLAB® and Simulink®. CRC press.

CVSS v2.0 user guide. (n.d.). Retrieved July 19, 2023, from https://www.first.org/cvss/v2/guide.

CCPS. (1993). Guidelines for Safe Automation of Chemical Processes. Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York.

CCPS. (2001). Layers of protection analysis: simplified process risk assessment. American Institute of Chemical Engineers-Center of Chemical Process Safety, New York.

CCPS. (2003). Guidelines for Analyzing and Managing the Security Vulnerabilities of Fixed Chemical Sites (2nd ed.). Center for Chemical Process Safety, Wiley/AIChE, New York.

CCPS/EI. (2018). Bow Ties in Risk Management. Center for Chemical Process Safety and Energy Institute (UK), Wiley/AIChE, New York.

CIEHF. (2016). Human Factors in Barrier Management, White Paper, Chartered Institute of Ergonomics and Human Factors.

Chen, C., & Reniers, G. (2021). Economic model for tackling intentional domino effects in a chemical facility. In Dynamic Risk Assessment and Management of Domino Effects and Cascading Events in the Process Industry (pp. 193-222). Elsevier.

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. *Reliability Engineering & System Safety, 191,* 106470.

Chen, C., Reniers, G., & Khakzad, N. (2020). Cost-benefit management of intentional domino effects in chemical industrial areas. *Process Safety and Environmental Protection, 134,* 392-405.

Chen, C., Reniers, G., Khakzad, N., & Yang, M. (2021). Operational safety economics: Foundations, current approaches and paths for future research. *Safety Science, 141,* 105326.

Chen, Y., Hong, J., & Liu, C. C. (2016). Modeling of intrusion and defense for assessment of cyber security at power substations. *IEEE Transactions on Smart Grid, 9(4),* 2541-2552.

Cozzani, V., Gubinelli, G., Antonioni, G., Spadoni, G., & Zanelli, S. (2005). The assessment of risk caused by domino effect in quantitative area risk analysis. *Journal of Hazardous Materials, 127(1-3),* 14-30.

Colabianchi, S., Costantino, F., Di Gravio, G., Nonino, F., & Patriarca, R. (2021). Discussing resilience in the context of cyber physical systems. *Computers & Industrial Engineering, 160,* 107534.

Derler, P., Lee, E. A., & Vincentelli, A. S. (2011). Modeling cyber-physical systems. *Proceedings of the IEEE, 100(1),* 13-28.

Delvosalle, C., Fievez, C., Pipart, A., & Debray, B. (2006). ARAMIS project: A comprehensive methodology for the identification of reference accident scenarios in process industries. *Journal of Hazardous Materials, 130(3),* 200-219.

Debray, B., Piatyszek, E., Cauffet, F., & Londiche, H. (2004). Frequencies and probabilities data for the fault tree. Accidental risk assessment methodology for industries in the framework of SEVESO II directive (ARAMIS), Armines, École Nationale Supérieure de Mines de Saint Etienne, France, 100.

De Ruijter, A., & Guldenmund, F. (2016). The bowtie method: A review. *Safety Science, 88,* 211-218.

De Dianous, V., & Fiévez, C. (2006). ARAMIS project: A more explicit demonstration of risk control through the use of bow–tie diagrams and the evaluation of safety barrier performance. *Journal of Hazardous Materials, 130(3),* 220-233.

De Souza, J. A., Santos Fo, D. J., Squillante, R., Junqueira, F., Miyagi, P. E., & Silva, J. R. (2017). Safety active barriers considering different scenarios of faults in modern production systems. In Technological Innovation for Smart Systems: 8th IFIP WG 5.5/SOCOLNET Advanced Doctoral Conference on Computing, Electrical and Industrial Systems, Costa de Caparica, Portugal, May 3-5, 2017, Proceedings 8 (pp. 154-164). Springer International Publishing.

Di Maio, F., Scapinello, O., Zio, E., Ciarapica, C., Cincotta, S., Crivellari, A., ... & Larosa, L. (2021). Accounting for safety barriers degradation in the risk assessment of oil and gas systems by multistate Bayesian networks. *Reliability Engineering & System Safety, 216,* 107943.

Di Maio, F., Marchetti, S., & Zio, E. (2023a). A framework of sensitivity analysis for the performance assessment of safety barriers impacted by NaTech accidents. *Process Safety and Environmental Protection, 171,* 1022-1030.

Di Maio, F., Marchetti, S., & Zio, E. (2023b). Robust multi-objective optimization of safety barriers performance parameters for NaTech scenarios risk assessment and management. *Reliability Engineering & System Safety, 235,* 109245.

Ding, L., Ji, J., Khan, F., Li, X., & Wan, S. (2020). Quantitative fire risk assessment of cotton storage and a criticality analysis of risk control strategies. *Fire and Materials, 44(2),* 165-179.

Duijm, N. J., Andersen, H. B., Hale, A., Goossens, L., & Hourtolou, D. (2004). Evaluating and managing safety barriers in major hazard plants. In Probabilistic Safety Assessment and Management: PSAM 7—ESREL'04 June 14–18, 2004, Berlin, Germany, Volume 6 (pp. 110-115). Springer London.

Duijm, N. J. (2009). Safety-barrier diagrams as a safety management tool. *Reliability Engineering & System Safety, 94(2),* 332-341.

Eisinger, S., & Rakowsky, U. K. (2001). Modeling of uncertainties in reliability centered maintenance—a probabilistic approach. *Reliability Engineering & System Safety, 71(2),* 159-164.

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

Eide, S.A., Wierman, T.E., Gentillon, C.D., Rasmuson, D.M., & Atwood, C.L. Industry-Average Performance for Components and Initiating Events at US

Commercial Nuclear Power Plants; NUREG/CR-6928; Nuclear Regulatory Commission: Washington, DC, USA, 2007.

EC. (1996). Council directive 96/82/EC of 9 December 1996 on the control of major-accident hazards involving dangerous substances, 469–491.

EC. (1998). Directive 98/37/EC of the European Parliament and the Council of 22 June 1998 on the approximation of the laws of the Member States relating to machinery. *Official Journal of the European Community No. L, 207,* 1-46.

EI. (2020). Guidelines for the management of safety critical elements. Energy Institute, ISBN: 9780852934623, 3rd edition.

El-Kady, A. H., Halim, S., El-Halwagi, M. M., & Khan, F. (2023). Analysis of safety and security challenges and opportunities related to cyber-physical systems. *Process Safety and Environmental Protection, 173,* 384-413.

Flaus, J. M. (2019). *Cybersecurity of industrial systems.* John Wiley & Sons.

Fleming, K. N., & Silady, F. A. (2002). A risk informed defense-in-depth framework for existing and advanced reactors. *Reliability Engineering & System Safety, 78(3),* 205-225.

Fiorentini, L., & Marmo, L. (2018). Sound Barriers Management in Process Safety: Bow-tie Approach According to the First Official AIChE-CCPS Guidelines. *Chemical Engineering Transactions, 67,* 253-258.

Friedberg, I., McLaughlin, K., Smith, P., Laverty, D., & Sezer, S. (2017). STPA-SafeSec: Safety and security analysis for cyber-physical systems. *Journal of Information Security and Applications, 34,* 183-196.

Freeman, R. A. (1990). CCPS guidelines for chemical process quantitative risk analysis. *Plant/Operations Progress, 9(4),* 231-235.

Garcia, M. L. (2007). *Design and evaluation of physical protection systems (2nd edition).* Elsevier.

George, P. G., & Renjith, V. R. (2021). Evolution of safety and security risk assessment methodologies towards the use of bayesian networks in process industries. *Process Safety and Environmental Protection, 149,* 758-775.

Gibson, J. J. (1961). The contribution of experimental psychology to the formulation of the problem of safety–a brief for basic research. *Behavioral Approaches to Accident Research, 1(61),* 77-89.

Ghafir, I., & Prenosil, V. (2014). Advanced persistent threat attack detection: an overview. *Int J Adv Comput Netw Secur, 4(4),* 5054.

Gribaudo, M., Iacono, M., & Marrone, S. (2015). Exploiting Bayesian networks for the analysis of combined attack trees. *Electronic Notes in Theoretical Computer Science, 310,* 91-111.

Gravestock, N. (2008). Effectiveness of fire safety systems for use in quantitative risk assessments. New Zealand Fire Service Commission, Wellington, NZ.

Gubinelli, G., Zanelli, S., & Cozzani, V. (2004). A simplified model for the assessment of the impact probability of fragments. *Journal of Hazardous Materials, 116(3),* 175-187.

Guldenmund, F., Hale, A., Goossens, L., Betten, J., & Duijm, N. J. (2006). The development of an audit technique to assess the quality of safety barrier management. *Journal of Hazardous Materials, 130(3),* 234-241.

Guzman, N. H. C., Wied, M., Kozine, I., & Lundteigen, M. A. (2020). Conceptualizing the key features of cyber-physical systems in a multi-layered representation for safety and security analysis. *Systems Engineering, 23(2),* 189-210.

Guzman, N. H. C., Kozine, I., & Lundteigen, M. A. (2021). An integrated safety and security analysis for cyber-physical harm scenarios. *Safety Science, 144,* 105458.

Hale, A. (2003, September). Note on barriers and delivery systems. In PRISM conference, Athens.

Haasl, D. F., Roberts, N. H., Vesely, W. E., & Goldberg, F. F. (1981). Fault tree handbook (No. NUREG-0492). Nuclear Regulatory Commission, Washington, DC (USA). Office of Nuclear Regulatory Research.

Hauptmanns, U. (2002). Analytical propagation of uncertainties through fault trees. *Reliability Engineering & System Safety, 76(3),* 327-329.

Haddon Jr, W. (1973). Energy damage and the ten countermeasure strategies. *Human Factors, 15(4),* 355-366.

Hauge, S., Håbrekke, S., & Lundteigen, M. A. (2010). Reliability Prediction Method for Safety Instrumented Systems–PDS Example collection, 2010 Edition. SINTEF Report A, 17956, 42-50.

Hauge, S., & Onshus, T. (2010). Reliability Data for Safety Instrumented Systems PDS Data Handbook, 2010 Edition. SINTEF Report A, 13502.

Hauge, S., & Øien, K. (2016). Guidance for barrier management in the petroleum industry. *SINTEF report A,* 27623.

Hauge, S., Okstad, E., Paltrinieri, N., Edwin, N., Vatn, J., & Bodsberg, L. (2015). Handbook for monitoring of barrier status and associated risk in the operational phase. SINTEF F27045. *Center for Integrated Operations in the Petroleum Industry, Trondheim, Norway, Norway.*

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

Henry, M. H., & Haimes, Y. Y. (2009). A comprehensive network security risk model for process control networks. *Risk Analysis: An International Journal, 29(2),* 223-248.

Hickman, J. W. (1983). PRA procedures guide: a guide to the performance of probabilistic risk assessments for nuclear power plants. NUREG/CR-2300.

Hollnagel, E. (1999, September). Accidents and barriers. In *Proceedings of lex valenciennes* (Vol. 28, pp. 175-182). Presses Universitaires de Valenciennes.

Hollnagel, E. (2004). *Barriers and accident prevention*. Routledge.

Holland, P. (1997). *Offshore blowouts: causes and control*. Elsevier.

Hosseinniaa, B., Haskinsa, C., Reniersb, G., & Paltrinieria, N. (2019). A guideline for the dynamic barrier management framework based on system thinking. *Chemical Engineering Transactions, 77,* 103-108.

Hosseinnia Davatgar, B., Paltrinieri, N., & Bubbico, R. (2021). Safety barrier management: risk-based approach for the oil and gas sector. *Journal of Marine Science and Engineering, 9(7),* 722.

Hong, J. B., Kim, D. S., & Haqiq, A. (2014, June). What vulnerability do we need to patch first?. In 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (pp. 684-689). IEEE.

Hudson, P., & Hudson, T. (2015). Integrating cultural and regulatory factors in the bowtie: Moving from hand-waving to rigor. *Ontology modeling in physical asset integrity management,* 171-198.

Huang, K., Zhou, C., Tian, Y. C., Yang, S., & Qin, Y. (2018). Assessing the physical impact of cyberattacks on industrial cyber-physical systems. *IEEE Transactions on Industrial Electronics, 65(10),* 8153-8162.

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.

Hu, Y., Li, H., Yang, H., Sun, Y., Sun, L., & Wang, Z. (2019). Detecting stealthy attacks against industrial control systems based on residual skewness analysis. *EURASIP Journal on Wireless Communications and Networking, 2019(1),* 1-14.

Iaiani, M., Tugnoli, A., Bonvicini, S., & Cozzani, V. (2021). Analysis of cybersecurity-related incidents in the process industry. *Reliability Engineering & System Safety, 209,* 107485.

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 Safety and Environmental Protection, 172,* 69-82.

IEC:61508. (1998). Functional safety of electrical/electronic/programmable electronic safety-related systems. In: *International Electrotechnical Commission. Geneva, Switzerland.*

IEC: 61508. (2010). Functional safety of electrical/electronic/programmable electronic safety-related systems. IEC Standards Online.

IEC:61511. (2002). Functional safety: Safety instrumented systems for the process industry sector. In: *International Electrotechnical Commission* (Vol. 57, pp. 33-40). Geneva.

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.

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

IEC: 62443-3-2. (2020). Security for industrial automation and control systems–Part 3-2: Security risk assessment for system design. Geneva: International Electrotechnical Commission.

ISO:13702. (1999). Petroleum and natural gas industries-Control and mitigation of fires and explosions on offshore production installations-Requirements and guidelines. In: *International Organization for Standardization. Geneva, Switzerland.*

ISO:17776. (2000). Petroleum and Natural Gas Industries-Offshore Production Installations-Guidelines on Tools and Techniques for Hazard Identification and Risk Assessment. In: *International Organization for Standardization. Geneva, Switzerland.*

ISO:13702. (2015). Petroleum and Natural Gas Industries – Control and Mitigation of Fires and Explosions on Offshore Production Installations – Requirements and Guidelines. *International Standard Organization, Geneva, Switzerland.*

ISO:16530. (2017). Petroleum and Natural Gas Industries — Well Integrity. *International Standard Organization, Geneva, Switzerland.*

Innal, F., Cacheux, P. J., Collas, S., Dutuit, Y., Folleau, C., Signoret, J. P., & Thomas, P. (2014). Probability and frequency calculations related to protection layers revisited. *Journal of Loss Prevention in the Process Industries, 31,* 56-69.

Jaeger, C. D. (2002). Vulnerability assessment methodology for chemical facilities (VAM-CF). *Chemical Health & Safety, 9(6),* 15-19.

Janssens, J., Talarico, L., Reniers, G., & Sörensen, K. (2015). A decision model to allocate protective safety barriers and mitigate domino effects. *Reliability Engineering & System Safety, 143,* 44-52.

Jain, P., Pistikopoulos, E. N., & Mannan, M. S. (2019). Process resilience analysis based data-driven maintenance optimization: Application to cooling tower operations. *Computers & Chemical Engineering, 121,* 27-45.

Jensen, F. V., & Nielsen, T. D. (2007). Bayesian networks and decision graphs (Vol. 2). New York: Springer.

Johnson, C. W. (2003). Failure in Safety Critical Systems: A Handbook of Incident and Accident Reporting. *University of Glasgow Press, Glasgow, Scotland.*

Johnson, W. G. (1975). MORT: The Management Oversight and Risk Tree. *Journal of Safety Research, 7 (1),* 4-15.

Johnson, W. (1980). *MORT: Safety Assurance Systems*, NewYork: MarcelDekker. In: Inc.

Johansen, I. L., & Rausand, M. (2015). Barrier management in the offshore oil and gas industry. *Journal of Loss Prevention in the Process Industries, 34,* 49-55.

John, J. Fay., & David, Patterson. (2018). *Contemporary Security Management Fourth Edition.* Elsevier.

Ji, X., He, G., Xu, J., & Guo, Y. (2016). Study on the mode of intelligent chemical industry based on cyber-physical system and its implementation. *Advances in Engineering Software, 99,* 18-26.

Ji, Z., Yang, S. H., Cao, Y., Wang, Y., Zhou, C., Yue, L., & Zhang, Y. (2021). Harmonizing safety and security risk analysis and prevention in cyber-physical systems. *Process Safety and Environmental Protection, 148,* 1279-1291.

Kalantarnia, M., Khan, F., & Hawboldt, K. (2009). Dynamic risk assessment using failure assessment and Bayesian theory. *Journal of Loss Prevention in the Process Industries, 22(5),* 600-606.

Kang, J., Zhang, J., & Gao, J. (2016). Analysis of the safety barrier function: Accidents caused by the failure of safety barriers and quantitative evaluation of their performance. *Journal of Loss Prevention in the Process Industries, 43,* 361-371.

Karl Ezra Pilario (2022). Feedback-controlled CSTR Process for Fault Simulation (https://www.mathworks.com/matlabcentral/fileexchange/66189-feedback-controlled-cstr-process-for-fault-simulation), MATLAB Central File Exchange. Retrieved November 3, 2022.

Kim, H., Lee, S. H., Park, J. S., Kim, H., Chang, Y. S., & Heo, G. (2015). Reliability data update using condition monitoring and prognostics in probabilistic safety assessment. *Nuclear Engineering and Technology, 47(2),* 204-211.

Kirwan, B. (2017). A guide to practical human reliability assessment. CRC press.

Kosmowski, K. T., Śliwiński, M., & Piesik, E. (2015). Integrated safety and security analysis of hazardous plants and systems of critical infrastructure. *Journal of Polish Safety and Reliability Association, 6(2),* 31-45.

Kriaa, S., Pietre-Cambacedes, L., Bouissou, M., & Halgand, Y. (2015). A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering & System Safety, 139,* 156-178.

Kjellén, U. (2000). *Prevention of accidents through experience feedback.* CRC Press.

Khakzad, N., Khan, F., & Amyotte, P. (2011). Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering & System Safety, 96(8),* 925-932.

Khakzad, N., Khan, F., & Amyotte, P. (2012). Dynamic risk analysis using bow-tie approach. *Reliability Engineering & System Safety, 104,* 36-44.

Khakzad, N., Khan, F., & Amyotte, P. (2013). Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection, 91(1-2),* 46-53.

Khakzad, N., Khan, F., & Paltrinieri, N. (2014). On the application of near accident data to risk analysis of major accidents. *Reliability Engineering & System Safety, 126,* 116-125.

Khakzad, N., Landucci, G., & Reniers, G. (2017a). Application of dynamic Bayesian network to performance assessment of fire protection systems during domino effects. *Reliability Engineering & System Safety, 167,* 232-247.

Khakzad, N., Landucci, G., & Reniers, G. (2017b). Application of Graph Theory to Cost-Effective Fire Protection of Chemical Plants During Domino Effects. *Risk analysis, 37(9),* 1652-1667.

Khakzad, N., & Reniers, G. (2017). Cost-effective allocation of safety measures in chemical plants wrt land-use planning. *Safety Science, 97,* 2-9.

Khakzad, N., Landucci, G., Cozzani, V., Reniers, G., & Pasman, H. (2018a). Cost-effective fire protection of chemical plants against domino effects. *Reliability Engineering & System Safety, 169,* 412-421.

Khakzad, N., Martinez, I. S., Kwon, H. M., Stewart, C., Perera, R., & Reniers, G. (2018b). Security risk assessment and management in chemical plants: Challenges and new trends. *Process Safety Progress, 37(2),* 211-220.

Khakzad, N. (2021). Optimal fireFigurehting to prevent domino effects: Methodologies based on dynamic influence diagram and mathematical programming. *Reliability Engineering & System Safety, 212,* 107577.

Kornecki, A. J., & Zalewski, J. (2010). Safety and security in industrial control. In *Proceedings of the sixth annual workshop on cyber security and information intelligence research* (pp. 1-4).

Kushner, D. (2013). The real story of stuxnet. ieee Spectrum, 50(3), 48-53.

Kuypers, M., & Maillart, T. (2018, June). Designing organizations for cyber security resilience. In Proceedings of the 2018 The Workshop on the Economics of Information Security (WEIS), Innsbruck, Austria (pp. 18-19).

LawInsider (2024). *Security barrier Definition (N.d.).* Retrieved from https://www.lawinsider.com/dictionary/security-barrier.

Lalropuia, K. C., & Gupta, V. (2019). Modeling cyber-physical attacks based on stochastic game and Markov processes. *Reliability Engineering & System Safety, 181,* 28-37.

Latif-Shabgahi, G., & Tajarrod, F. (2009, March). A new approach for the construction of fault trees from system simulink. In 2009 International Conference on Availability, Reliability and Security (pp. 712-717). IEEE.

Landucci, G., Argenti, F., Tugnoli, A., & Cozzani, V. (2015). Quantitative assessment of safety barrier performance in the prevention of domino scenarios triggered by fire. *Reliability Engineering & System Safety, 143,* 30-43.

Landucci, G., Argenti, F., Spadoni, G., & Cozzani, V. (2016). Domino effect frequency assessment: The role of safety barriers. *Journal of Loss Prevention in the Process Industries, 44,* 706-717.

Landucci, G., Argenti, F., Cozzani, V., & Reniers, G. (2017). Assessment of attack likelihood to support security risk assessment studies for chemical facilities. *Process Safety and Environmental Protection, 110,* 102-114.

Landucci, G., Khakzad, N., & Reniers, G. (2020). Physical security in the process industry: Theory with applications. Elsevier.

Lees, F. P. (1980). Loss Prevention in the process Industries. Butterworth Hienemann Ltd, Oxford, ISBN 0-7506-1529-X, page 625.

Lewis, S. (2005). An overview of leading software tools for QRA. American Society of Safety Engineers–Middle East, 18-22.

Li, X., Zhou, C., Tian, Y. C., Xiong, N., & Qin, Y. (2017). Asset-based dynamic impact assessment of cyberattacks for risk analysis in industrial control systems. *IEEE Transactions on Industrial Informatics, 14(2),* 608-618.

Liu, Y., & Rausand, M. (2016). Proof-testing strategies induced by dangerous detected failures of safety-instrumented systems. *Reliability Engineering & System Safety, 145,* 366-372.

Liu, Y. (2020). Safety barriers: Research advances and new thoughts on theory, engineering and management. *Journal of Loss Prevention in the Process Industries, 67,* 104260.

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

Ling, E. R., & Ekstedt, M. (2023). Estimating Time-To-Compromise for Industrial Control System Attack Techniques Through Vulnerability Data. *SN Computer Science, 4(3),* 318.

Mathworks-Genetic algorithm. (n.d.). Retrieved September 28, 2022a, from https://nl.mathworks.com/help/gads/genetic-algorithm.html

Mathworks-User-defined functions. (n.d.). Retrieved October 14, 2022b, from https://nl.mathworks.com/help/simulink/user-defined-functions.html.

Manno, G., Chiacchio, F., Compagno, L., D'Urso, D., & Trapani, N. (2012). MatCarloRe: An integrated FT and Monte Carlo Simulink tool for the reliability assessment of dynamic fault tree. *Expert Systems with Applications, 39(12),* 10334-10342.

Markowski, A. S., & Kotynia, A. (2011). "Bow-tie" model in layer of protection analysis. *Process Safety and Environmental Protection, 89(4),* 205-213.

McQueen, M. A., Boyer, W. F., Flynn, M. A., & Beitel, G. A. (2006a). Time-to-compromise model for cyber risk reduction estimation. In Quality of Protection: Security Measurements and Metrics (pp. 49-64). Springer US.

McQueen, M. A., Boyer, W. F., Flynn, M. A., & Beitel, G. A. (2006b, January). Quantitative cyber risk reduction estimation methodology for a small SCADA control system. In Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06) (Vol. 9, pp. 226-226). IEEE.

McLeod, R. (2016). Issues in assuring human controls in layers-of-defences strategies. *Chemical Engineering Transactions, 48,* 925-930.

Meyer, T., & Reniers, G. (2022). Engineering risk management. Walter de Gruyter GmbH & Co KG.

Meel, A., & Seider, W. D. (2006). Plant-specific dynamic failure assessment using Bayesian theory. *Chemical Engineering Science, 61(21),* 7036-7056.

Myers, P. M. (2013). Layer of Protection Analysis–Quantifying human performance in initiating events and independent protection layers. *Journal of Loss Prevention in the Process Industries, 26(3),* 534-546.

Misuri, A., Landucci, G., & Cozzani, V. (2020). Assessment of safety barrier performance in Natech scenarios. *Reliability Engineering & System Safety, 193,* 106597.

Misuri, A., Landucci, G., & Cozzani, V. (2021). Assessment of risk modification due to safety barrier performance degradation in Natech events. *Reliability Engineering & System Safety, 212,* 107634.

Misuri, A., Ricci, F., Sorichetti, R., & Cozzani, V. (2023). The effect of safety barrier degradation on the severity of primary Natech scenarios. *Reliability Engineering & System Safety, 235,* 109272.

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. *Reliability Engineering & System Safety, 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. Boca Raton/London/New York: CRC press.

Norman, T. L. (2015). *Risk analysis and security countermeasure selection (2nd edition).* Boca Raton/London/New York: CRC press.

NIST. (2014). Security and privacy controls for federal information systems and organizations. Report, 800-53Ar4, *National Institute of Standards and Technology, Gaithersburg.*

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

Øien, K., Hauge, S., Jaatun, M. G., Flå, L., & Bodsberg, L. (2022). A Survey on Cybersecurity Barrier Management in Process Control Environments. In *2022 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)* (pp. 113-120). IEEE.

OREDA. (2002). Offshore Reliability Data Handbook. Trondheim, Norway: DNV.

Orojloo, H., & Azgomi, M. A. (2017). A game-theoretic approach to model and quantify the security of cyber-physical systems. *Computers in Industry, 88,* 44-57.

Ovidi, F., Zhang, L., Landucci, G., & Reniers, G. (2021). Agent-based model and simulation of mitigated domino scenarios in chemical tank farms. *Reliability Engineering & System Safety, 209,* 107476.

Ouache, R., Kabir, M. N., & Adham, A. A. (2015). A reliability model for safety instrumented system. *Safety Science, 80,* 264-273.

Ottermo, M; Hauge, S., & Håbrekke, S. (2021). Reliability Data for Safety Equipment: PDS Data Handbook. Trondheim: SINTEF Technology and Society.

Park, B., Kim, Y., Lee, K., Paik, S., & Kang, C. (2021). Risk assessment method combining independent protection layers (IPL) of layer of protection analysis (LOPA) and RISKCURVES software: Case study of hydrogen refueling stations in Urban Areas. *Energies, 14(13),* 4043.

Paltrinieri, N., Tugnoli, A., Buston, J., Wardman, M., & Cozzani, V. (2013). Dynamic procedure for atypical scenarios identification (DyPASI): a new systematic HAZID tool. *Journal of Loss Prevention in the Process Industries, 26(4),* 683-695.

Papazoglou, I. A., Bellamy, L. J., Hale, A. R., Aneziris, O. N., Ale, B. J. M., Post, J. G., & Oh, J. I. H. (2003). I-Risk: development of an integrated technical and management risk methodology for chemical installations. *Journal of Loss Prevention in the Process Industries, 16(6),* 575-591.

Papadopoulos, Y., & Maruhn, M. (2001, July). Model-based synthesis of fault trees from matlab-simulink models. In 2001 International Conference on Dependable Systems and Networks (pp. 77-82). IEEE.

Pawar, B., Park, S., Hu, P., & Wang, Q. (2021). Applications of resilience engineering principles in different fields with a focus on industrial systems: A literature review. *Journal of Loss Prevention in the Process Industries, 69,* 104366.

Pérez, S. I., Moral-Rubio, S., & Criado, R. (2021). A new approach to combine multiplex networks and time series attributes: Building intrusion detection systems (IDS) in cybersecurity. *Chaos, Solitons & Fractals, 150,* 111143.

Pishro-Nik, H. (2016). Introduction to probability, statistics, and random processes.

Pilario, K. E. S., & Cao, Y. (2018). Canonical variate dissimilarity analysis for process incipient fault detection. *IEEE Transactions on Industrial Informatics, 14(12),* 5308-5315.

Piètre-Cambacédès, L., & Bouissou, M. (2010). Modeling safety and security interdependencies with BDMP (Boolean logic Driven Markov Processes). In *2010 IEEE International Conference on Systems, Man and Cybernetics* (pp. 2852-2861). IEEE.

Pitblado, R., & Nelson, W. R. (2013). Advanced safety barrier management with inclusion of human and organizational aspects. *Chemical Engineering Transactions, 31,* 331-336.

Pitblado, R., Fisher, M., Nelson, B., Fløtaker, H., Molazemi, K., & Stokke, A. (2016). Concepts for dynamic barrier management. *Journal of Loss Prevention in the Process Industries, 43,* 741-746.

Pinciroli, L., Baraldi, P., & Zio, E. (2023). Maintenance optimization in industry 4.0. *Reliability Engineering & System Safety, 234,* 109204.

Poolsappasit, N., Dewri, R., & Ray, I. (2011). Dynamic security risk management using Bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing, 9(1),* 61-74.

PSA, N. (2013). *Principles for barrier management in the petroleum industry.* Retrieved October 2021, from Norway Petroleum Safety Authority: www.ptil.no.

Rathnayaka, S., Khan, F., & Amyotte, P. (2011a). SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection, 89(3),* 151-164.

Rathnayaka, S., Khan, F., & Amyotte, P. (2011b). SHIPP methodology: Predictive accident modeling approach. Part II. Validation with case study. *Process Safety and Environmental Protection, 89(2),* 75-88.

Ramzali, N., Lavasani, M. R. M., & Ghodousi, J. (2015). Safety barriers analysis of offshore drilling system by employing Fuzzy Event Tree Analysis. *Safety Science, 78,* 49-59.

Redutskiy, Y. (2017). Optimization of safety instrumented system design and maintenance frequency for oil and gas industry processes. *Management and Production Engineering Review, 8,* 46-59.

Reniers, G. L., Dullaert, W., Audenaert, A., Ale, B. J., & Soudan, K. (2008). Managing domino effect-related security of industrial areas. *Journal of Loss Prevention in the Process Industries, 21(3),* 336-343.

Reniers, G. L., & Van Erp, H. N. (2016). Operational safety economics: a practical approach focused on the chemical and process industries. John Wiley & Sons.

Reniers, G., & Khakzad, N. (2017). Revolutionizing safety and security in the chemical and process industry: applying the CHESS concept. *Journal of Integrated Security and Safety Science, 1(1),* 2-15.

Reniers, G., Khakzad, N., & van Gelder, P. (Eds.) (2017). Security Risk Assessment In the Chemical and Process Industry. (Integrated Security Science). Walter de Gruyter. https://www.degruyter.com/viewbooktoc/product/477259.

Roy, A., Kim, D. S., & Trivedi, K. S. (2010). Cyber security analysis using attack countermeasure trees. In *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research* (pp. 1-4).

Roy, A., Srivastava, P., & Sinha, S. (2015). Dynamic failure assessment of an ammonia storage unit: A case study. *Process Safety and Environmental Protection, 94,* 385-401.

Schmitz, P., Swuste, P., Reniers, G., & van Nunen, K. (2020). Mechanical integrity of process installations: Barrier alarm management based on bowties. *Process Safety and Environmental Protection, 138,* 139-147.

Schmitz, P., Swuste, P., Reniers, G., & van Nunen, K. (2021). Predicting major accidents in the process industry based on the barrier status at scenario level: A practical approach. *Journal of Loss Prevention in the Process Industries, 71,* 104519.

Schupp, B. A., Smith, S. P., Wright, P. C., & Goossens, L. H. (2004). Integrating human factors in the design of safety critical systems: A barrier based approach. In *Human Error, Safety and Systems Development: IFIP 18th World Computer Congress TC13/WC13. 5 7th Working Conference on Human Error, Safety and Systems Development 22–27 August 2004 Toulouse, France* (pp. 285-300). Springer US.

Semertzis, I., Rajkumar, V. S., Ştefanov, A., Fransen, F., & Palensky, P. (2022, May). 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) (pp. 1-6). IEEE.

Selvik, J. T., & Aven, T. (2011). A framework for reliability and risk centered maintenance. *Reliability Engineering & System Safety, 96(2),* 324-331.

Smith, E. (2002). Uncertainty analysis. Encyclopedia of environmetrics, 4, 2283-2297.

Störfall Kommission (SFK), 2002. SFK-GS-38 Report.

Svenson, O. (1991). The accident evolution and barrier function (AEB) model applied to incident analysis in the processing industries. *Risk Analysis, 11(3),* 499-507.

Sklet, S. (2006). Safety barriers: Definition, classification, and performance. *Journal of Loss Prevention in the Process Industries, 19(5),* 494-506.

Sultana, S., & Haugen, S. (2023). An extended FRAM method to check the adequacy of safety barriers and to assess the safety of a socio-technical system. *Safety Science, 157,* 105930.

Simon, C., Mechri, W., & Capizzi, G. (2019). Assessment of Safety Integrity Level by simulation of Dynamic Bayesian Networks considering test duration. *Journal of Loss Prevention in the Process Industries, 57,* 101-113.

Siu, N. O., & Kelly, D. L. (1998). Bayesian parameter estimation in probabilistic risk assessment. *Reliability Engineering & System Safety, 62(1-2),* 89-116.

Sobral, J., & Soares, C. G. (2019). Assessment of the adequacy of safety barriers to hazards. *Safety Science, 114,* 40-48.

Song, G., Khan, F., & Yang, M. (2019a). Probabilistic assessment of integrated safety and security related abnormal events: a case of chemical plants. *Safety Science, 113,* 115-125.

Song, G., Khan, F., & Yang, M. (2019b). Integrated risk management of hazardous processing facilities. *Process Safety Progress, 38(1),* 42-51.

Stluka, P., & Mařík, K. (2007). Data-driven decision support and its applications in the process industries. In *Computer Aided Chemical Engineering* (Vol. 24, pp. 273-278). Elsevier.

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

Taylor, J. R. (2010). The QRAQ Project Volume 4: Frequency of Releases and Accidents.https://www.academia.edu/35376294/The_QRAQ_Project_Volume_4_Frequency_of_Releases_and_Accidents. (accessible 2023, March)

Tan, Z., Li, J., Wu, Z., Zheng, J., & He, W. (2011). An evaluation of maintenance strategy using risk based inspection. Safety Science, 49(6), 852-860.

Tamascelli, N., Dal Pozzo, A., Scarponi, G. E., Paltrinieri, N., & Cozzani, V. (2024). Assessment of Safety Barrier Performance in Environmentally Critical Facilities: Bridging Conventional Risk Assessment Techniques with Data-Driven Modelling. *Process Safety and Environmental Protection, 181,* 294-311.

Tsunemi, K., Kihara, T., Kato, E., Kawamoto, A., & Saburi, T. (2019). Quantitative risk assessment of the interior of a hydrogen refueling station considering safety barrier systems. International *Journal of Hydrogen Energy, 44(41),* 23522-23531.

Thomas, H. W., & Day, J. (2015, April). Integrating Cybersecurity Risk Assessments Into the Process Safety Management Work Process. In *Poster Session, AiChe 2015 Spring Meeting & 11th Global Congress on Process Safety*.

Thomas, R. J. and Chothia, T. (2020). Learning from vulnerabilities - categorising, understanding and detecting weaknesses in industrial control systems. In Computer Security, Cham. Springer International Publishing.

Tong, X., Fang, W., Yuan, S., Ma, J., & Bai, Y. (2018). Application of Bayesian approach to the assessment of mine gas explosion. *Journal of Loss Prevention in the Process Industries, 54,* 238-245.

U.S. Congress, Office of Technology Assessment. (1992). *Technology Against Terrorism: Structuring Security. OTA-ISC-511.* Washington, DC: U.S. Government Printing Office.

Van Den Bosh, C. J. H., Merx, W. P. M., Jansen, C. M. A., De Weger, D., Reuzel, P. G. J., Leeuwen, D. V., & Blom-Bruggerman, J. M. (1989). Methods for the calculation of possible damage (Green Book). The Hague (NL): Committee for the Prevention of Disasters.

Van Staalduinen, M., & Khan, F. (2015, March). A barrier based methodology to assess site security risk. In SPE Health, Safety, Security, Environment, & Social Responsibility Conference-North America (pp. SPE-173561). SPE.

Van der Borst, M., & Schoonakker, H. (2001). An overview of PSA importance measures. *Reliability Engineering & System Safety, 72(3),* 241-245.

Van Nunen, K., Swuste, P., Reniers, G., & Schmitz, P. (2019). Developing leading safety indicators for occupational safety based on the bow-tie method. *Chemical Engineering Transactions, 77,* 49-54.

Vílchez, J. A., Espejo, V., & Casal, J. (2011). Generic event trees and probabilities for the release of different types of hazardous materials. *Journal of Loss Prevention in the Process Industries, 24(3),* 281-287.

Villa, V., & Cozzani, V. (2016). Application of Bayesian Networks to quantitative assessment of safety barriers' performance in the prevention of major accidents. *Chemical Engineering Transactions, 53,* 151-156.

Wahlstrom, B., & Gunsell, L. (1998). Reactor safety; A Description and Assessment of the Nordic safety work. *Risoforskningscenter: NKS-sekretariatet.*

Wang, Y., Cai, B., Zhang, Y., Liu, J., Khan, J. A., Liu, Y., ... & Liu, Y. (2022). Condition-based maintenance method for multicomponent system considering maintenance delay based on remaining useful life prediction: Subsea tree system as a case. *Ocean Engineering, 266,* 112616.

Whaley, A. M., Kelly, D. L., Boring, R. L., & Galyean, W. J. (2012). *SPAR-H step-by-step guidance* (No. INL/CON-12-24693). Idaho National Lab.(INL), Idaho Falls, ID (United States).

Wen, H., Khan, F., Ahmed, S., Imtiaz, S., & Pistikopoulos, S. (2023). Risk assessment of human-automation conflict under cyberattacks in process systems. *Computers & Chemical Engineering, 172,* 108175.

Wu, S., Li, B., Zhou, Y., Chen, M., Liu, Y., & Zhang, L. (2022). Hybrid Dynamic Bayesian network method for performance analysis of safety barriers considering multi-maintenance strategies. *Engineering Applications of Artificial Intelligence, 109,* 104624.

Xie, C., Huang, L., Wang, R., Deng, J., Shu, Y., & Jiang, D. (2022). Research on Quantitative Risk Assessment of Fuel Leak of LNG-fuelled Ship During Lock Transition Process. Reliability Engineering & System Safety, 108368.

Xing, J., Zeng, Z., & Zio, E. (2019). A framework for dynamic risk assessment with condition monitoring data and inspection data. *Reliability Engineering & System Safety, 191,* 106552.

Xue, L., Fan, J., Rausand, M., & Zhang, L. (2013). A safety barrier-based accident model for offshore drilling blowouts. *Journal of Loss Prevention in the Process Industries, 26(1),* 164-171.

Yan, F., Xu, K., Cui, Z., & Yao, X. (2017). An improved layer of protection analysis based on a cloud model: Methodology and case study. *Journal of Loss Prevention in the Process Industries, 48,* 41-47.

Yampolskiy, M., Horvath, P., Koutsoukos, X. D., Xue, Y., & Sztipanovits, J. (2013, April). Taxonomy for description of cross-domain attacks on CPS. In Proceedings of the 2nd ACM international conference on High confidence networked systems (pp. 135-142).

Yazdi, M., & Kabir, S. (2017). A fuzzy Bayesian network approach for risk analysis in process industries. *Process safety and environmental protection, 111,* 507-519.

Yazdi, M., Kabir, S., & Walker, M. (2019). Uncertainty handling in fault tree based risk assessment: state of the art and future perspectives. *Process Safety and Environmental Protection, 131,* 89-104.

Ylönen, M., Tugnoli, A., Oliva, G., Heikkilä, J., Nissilä, M., Iaiani, M., ... & Del Prete, E. (2022). Integrated management of safety and security in Seveso sites-sociotechnical perspectives. *Safety Science, 151,* 105741.

Yun, G., Rogers, W. J., & Mannan, M. S. (2009). Risk assessment of LNG importation terminals using the Bayesian–LOPA methodology. *Journal of Loss Prevention in the Process Industries, 22(1),* 91-96.

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. *Safety Science, 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. *Reliability Engineering & System Safety, 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 Safety and Environmental Protection, 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. *Journal of Loss Prevention in the Process Industries, 83,* 105034.

Yuan, S., Yang, M., & Reniers, G. (2024). Integrated process safety and process security risk assessment of industrial cyber-physical systems in chemical plants. *Computers in Industry, 155,* 104056.

Zadakbar, O., Imtiaz, S., & Khan, F. (2013a). Dynamic risk assessment and fault detection using a multivariate technique. *Process Safety Progress, 32(4),* 365-375.

Zadakbar, O., Imtiaz, S., & Khan, F. (2013b). Dynamic risk assessment and fault detection using principal component analysis. *Industrial & Engineering Chemistry Research, 52(2),* 809-816.

Zadakbar, O., Khan, F., & Imtiaz, S. (2015). Dynamic risk assessment of a nonlinear non-Gaussian system using a particle filter and detailed consequence analysis. *The Canadian Journal of Chemical Engineering, 93(7),* 1201-1211.

Zarei, E., Azadeh, A., Khakzad, N., Aliabadi, M. M., & Mohammadfam, I. (2017). Dynamic safety assessment of natural gas stations using Bayesian network. *Journal of Hazardous Materials, 321,* 830-840.

Zarei, E., Khan, F., & Abbassi, R. (2021). Importance of human reliability in process operation: A critical analysis. *Reliability Engineering & System Safety, 211,* 107607.

Zeng, T., Chen, G., Yang, Y., Chen, P., & Reniers, G. (2020). Developing an advanced dynamic risk analysis method for fire-related domino effects. *Process Safety and Environmental Protection, 134,* 149-160.

Zeng, Z., & Zio, E. (2018). Dynamic risk assessment based on statistical failure data and condition-monitoring degradation data. *IEEE Transactions on Reliability, 67(2),* 609-622.

Zhang, A., Barros, A., & Liu, Y. (2019). Performance analysis of redundant safety-instrumented systems subject to degradation and external demands. *Journal of Loss Prevention in the Process Industries, 62,* 103946.

Zhang, A., Zhang, T., Barros, A., & Liu, Y. (2020). Optimization of maintenances following proof tests for the final element of a safety-instrumented system. *Reliability Engineering & System Safety, 196,* 106779.

Zhang, A., Srivastav, H., Barros, A., & Liu, Y. (2021). Study of testing and maintenance strategies for redundant final elements in SIS with imperfect detection of degraded state. *Reliability Engineering & System Safety, 209,* 107393.

Zhang, A., Wu, S., Fan, D., Xie, M., Cai, B., & Liu, Y. (2022). Adaptive testing policy for multi-state systems with application to the degrading final elements in safety-instrumented systems. *Reliability Engineering & System Safety, 221,* 108360.

Zhang, Y., Wang, L., Xiang, Y., & Ten, C. W. (2015). Power system reliability evaluation with SCADA cybersecurity considerations. *IEEE Transactions on Smart Grid, 6(4),* 1707-1721.

Zhang, Q., Zhou, C., Tian, Y. C., Xiong, N., Qin, Y., & Hu, B. (2017). A fuzzy probability Bayesian network approach for dynamic cybersecurity risk assessment in industrial control systems. *IEEE Transactions on Industrial Informatics, 14(6),* 2497-2506.

Zhen, X., Han, Y., & Huang, Y. (2021). Optimization of preventive maintenance intervals integrating risk and cost for safety critical barriers on offshore petroleum installations. *Process Safety and Environmental Protection, 152,* 230-239.

Zhu, C., Qi, M., & Jiang, J. (2020). Quantifying human error probability in independent protection layers for a batch reactor system using dynamic simulations. *Process Safety and Environmental Protection, 133,* 243-258.

Summary

Concerning the accidental and intentional major adverse events in chemical process industries, particularly considering the emerging cyber-to-physical (C2P) attack risks affiliated with the automation and digitization process of industrial control systems, this study is dedicated to investigating the dynamic and integrated safety and security barrier management for ensuring the safety and security of chemical plant in the Industry 4.0 era. To achieve the objectives of this study, the following research steps have been made.

A systematic review has been conducted to understand the definitions and classifications of safety and security barriers and get insights into the fundamental aspects of safety and security barriers. Existing methodologies for the performance assessment and management of safety and security barriers have also been reviewed and discussed to identify research gaps, which provide valid foundations for the following steps.

With the identification of multi-dimensional risks (safety risks, physical attack risks, and C2P attack risks) threatening industrial control systems in chemical plants, an integrated approach is developed to construct accident scenarios concerning both safety hazards and security threats and quantitatively assess the risk of chemical facilities considering the interdependency between safety risks and security risks.

Considering the uncertainties associated with the integrated safety and security risks, particularly the uncertainties in attackers' knowledge levels, a vulnerability assessment model is developed to assess C2P attacks, and the combination of Monte Carlo simulations and a Bayesian network model is employed to handle uncertainty propagation in the risk assessment. Furthermore, combining cost-effectiveness analysis with a risk matrix yields the optimal strategy for safety and security barrier enhancements from a cost-effective perspective.

A novel approach for risk-based barrier maintenance is developed to tackle the challenges in solving barrier optimization problems with large-solution spaces. Accident scenarios regarding safety and physical security are constructed using an extended bow-tie diagram and then modeled based on MATLAB/Simulink simulations. A combination of cost-effectiveness analysis and genetic algorithms is employed to decide the approximately optimal strategy for barrier maintenance.

Multiple data (periodic proof test data, continuous condition-monitoring data, and accident precursor data) are combined to enable continuous safety barrier improvement by revealing the degradation of safety barriers and performing dynamic risk assessment. Furthermore, multi-source data capable of revealing risk variations are characterized and incorporated with the barrier management framework to empower dynamic and integrated safety and security barrier management. Dynamic and integrated S&S barrier management has the advantage of making timely

adaptations according to the new evidence and continuously ensuring the integrated safety and security risks at acceptable levels.

Finally, all methodologies developed in this study are structured into a systematic framework to foster the application of dynamic and integrated management of safety and security barriers in practices.

Curriculum Vitae

Shuaiqi Yuan (原帅琪)

Personal Information

Name: Shuaiqi Yuan (原帅琪) **Nationality:** Chinese

M/F: Male **Date of Birth:** 1994/10/11

E-mails: *S.Yuan-2@tudelft.nl* OR *cumtbyuanshuaiqi@163.com*

Research Interests

- Chemical process safety
- Chemical process security
- Barrier management
- Quantitative risk assessment
- Uncertainty quantification
- Risk-based decision-making

Educational Background

- Ph.D. candidate (Nov, 2020-Nov, 2024)

Safety and Security Science Group, Delft University of Technology, Delft, The Netherlands

Promoter and daily supervisor: Prof. Genserik Reniers and Dr. Ming Yang

- M.Eng. Degree in Safety Engineering (Sep, 2017-Jue, 2020)

Department of Safety Engineering, China University of Mining and Technology, Beijing, China

Supervisor: Prof. Jiansong Wu

- B.Eng. Degree in Mining Engineering (Sep, 2013-Jue, 2017)

Department of Mining Engineering, China University of Mining and Technology, Beijing, China

Journal Publications

1. **Yuan, S**., Reniers, G., Yang, M. (2024). Dynamic and integrated safety and security barrier management in chemical plants: a new paradigm to manage complex major adverse risks. Process Safety and Environmental Protection, (To be submitted).

2. **Yuan, S**., Reniers, G., & Yang, M. (2024). Integrated management of safety and security barriers in chemical plants to cope with emerging cyber-physical attack risks under uncertainties. Reliability Engineering & System Safety, 250, [110320].

3. **Yuan, S**., Yang, M., & Reniers, G. (2024). Integrated process safety and process security risk assessment of industrial cyber-physical systems in chemical plants. Computers in Industry, 155, [104056].

4. **Yuan, S**., Reniers, G., & Yang, M. (2023). Dynamic-risk-informed safety barrier management: An application to cost-effective barrier optimization based on data from multiple sources. Journal of Loss Prevention in the Process Industries, 83, [105034].

5. **Yuan, S**., Reniers, G., Yang, M., & Bai, Y. (2023). Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. Process Safety and Environmental Protection, 170, 356-371.

6. **Yuan, S**., Cai, J., Reniers, G., Yang, M., Chen, C., & Wu, J. (2022). Safety barrier performance assessment by integrating computational fluid dynamics and evacuation modeling for toxic gas leakage scenarios. Reliability Engineering and System Safety, 226, [108719].

7. **Yuan, S**., Yang, M., Reniers, G., Chen, C., & Wu, J. (2022). Safety barriers in the chemical process industries: A state-of-the-art review on their classification, assessment, and management. Safety Science, 148, [105647].

8. **Yuan, S.**, Wu, J., Zhang, X., & Liu, W. (2019). EnKF-based estimation of natural gas release and dispersion in an underground tunnel. Journal of Loss Prevention in the Process Industries, 62, [103931].

9. Wu, J., Cai, J., **Yuan, S***., Zhang, X., & Reniers, G. (2021). CFD and EnKF coupling estimation of LNG leakage and dispersion. Safety Science, 139, [105263].

10. **Yuan, S.,** Yang, M., Reniers, G., & Chen, C. (2022). An Approach for Identification of Integrated Safety and Security Barriers in the Chemical Process Industries. Chemical Engineering Transactions, 90, 571-576.

11. **Yuan, S**., Reniers, G., & Yang, M. (2022). The Necessity of Integrating Safety and Security Barriers in the Chemical Process Industries and its Potential Framework. Chemical Engineering Transactions, 91, 13-18.

12. Chen, X., Lin, W., Liu, C., Yang, F*., Guo, Y., Li, X., **Yuan, S***., & Reniers, G. (2023). An integrated EDIB model for probabilistic risk analysis of natural gas pipeline leakage accidents. Journal of Loss Prevention in the Process Industries, 83, [105027].

13. Yang, F*., Li, X., **Yuan, S*.**, & Reniers, G. (2023). Risk Analysis of Laboratory Fire Accidents in Chinese Universities by Combining Association Rule Learning and Fuzzy Bayesian Networks. Fire, 6(8), [306].

14. Cai, J., Wu, J., **Yuan, S**., Kong, D., & Zhang, X. (2022). Prediction of gas leakage and dispersion in utility tunnels based on CFD-EnKF coupling model: A 3D full-scale application. Sustainable Cities and Society, 80, [103789].

15. Cai, J., Wu, J., **Yuan, S**., Reniers, G., & Bai, Y. (2024). Risk-based optimization of emergency response systems for accidental gas leakage in utility tunnels. Reliability Engineering & System Safety,

244, [109947].

16. Wu, J., Cai, J., Liu, Z., **Yuan, S**., Bai, Y., & Zhou, R. (2023). BI-IEnKF coupling model for effective source term estimation of natural gas leakage in urban utility tunnels. Tunnelling and Underground Space Technology, 136, [105080].

17. Bai, Y., Wu, J., **Yuan, S**., Reniers, G., Yang, M., & Cai, J. (2022). Dynamic resilience assessment and emergency strategy optimization of natural gas compartments in utility tunnels. Process Safety and Environmental Protection, 165, 114-125.

18. Xu, Y., Reniers, G., Yang, M., **Yuan, S**., & Chen, C. (2023). Uncertainties and their treatment in the quantitative risk assessment of domino effects: Classification and review. Process Safety and Environmental Protection, 172, 971-985.

19. Wu, J., Xing, Y., Bai, Y., Hu, X., & **Yuan, S**. (2022). Risk assessment of large-scale winter sports sites in the context of a natural disaster. Journal of Safety Science and Resilience, 3(3), 263-276.

20. Hou, M., Hu, X., Cai, J., Han, X., & **Yuan, S**. (2022). An Integrated Graph Model for Spatial–Temporal Urban Crime Prediction Based on Attention Mechanism. ISPRS International Journal of Geo-Information, 11(5), [294].

21. Chen, X., Yang, F*., Cheng, S., & **Yuan, S***. (2023). Occupational Health and Safety in China: A Systematic Analysis of Research Trends and Future Perspectives. Sustainability, 15(19), [14061].

22. Xu, Y., Reniers, G., Yang, M., **Yuan, S**., & Chen, C. (2022). An Exploratory Study on Uncertainty Analysis in Quantitative Risk Assessment of Domino Effects. Chemical Engineering Transactions, 90, 565-570.

23. Cai, J., Wu, J., **Yuan, S.,** Liu, Z., & Kong, D. (2021). Numerical analysis of multi-factors effects on the leakage and gas diffusion of gas drainage pipeline in underground coal mines. Process Safety and Environmental Protection, 151, 166-181.

24. Wu, J., Liu, Z., **Yuan, S**., Cai, J., & Hu, X. (2020). Source term estimation of natural gas leakage in utility tunnel by combining CFD and Bayesian inference method. Journal of Loss Prevention in the Process Industries, 68, [104328].

25. Wu, J., Fang, W., Tong, X., **Yuan, S**., & Guo, W. (2019). Bayesian analysis of school bus accidents: a case study of China. Natural Hazards, 95, 463-483.

26. Wu, J., **Yuan, S.**, Zhang, C., & Zhang, X. (2018). Numerical estimation of gas release and dispersion in coal mine using Ensemble Kalman Filter. Journal of Loss Prevention in the Process Industries, 56, 57-67.

27. Tong, X., Fang, W., **Yuan, S**., Ma, J., & Bai, Y. (2018). Application of Bayesian approach to the assessment of mine gas explosion. Journal of Loss Prevention in the Process Industries, 54, 238-245.

28. Wu, J., **Yuan, S**., & Cai, J. (2020). Numerical simulation of gas leakage and dispersion in utility tunnel compartment based on OpenFOAM. Journal of Safety Science and Technology, 16(2), 168-173. (in Chinese)

29. Tong, X., **Yuan, S**., Fang, W., Ma, J. (2018). Classification and calculation of the best escape route of coal mine based on Dijkstra algorithm. Industry and Mine Automation, 44(04), 94-99. (in Chinese)

30. Ma, J., Wu, J., **Yuan, S**.. (2017). Simulation of Evacuation in Underground Coal Mines Based on Multi-Agent modeling. Science Technology and Engineering, 17(36), 152-157. (in Chinese)

Book Chapters

1. **Yuan, S**., Chen, C., Yang, M., & Reniers, G. (2021). Methods for domino effect risk management decision-making. In Methods in Chemical Process Safety (Vol. 5, pp. 461-494). Elsevier.

2. **Yuan, S**., Xu, Y., Bai, Y., Reniers, G., & Yang, M. (2023). Conceptual and practical applications of ISD. In Methods in Chemical Process Safety (Vol. 7, pp. 183-211). Elsevier.

3. Chen, C., Reniers, G., Yang, M., & **Yuan, S**. (2021). Domino effect security risk assessment. In Methods in Chemical Process Safety (Vol. 5, pp. 309-330). Elsevier.

Technical Reports

1. Yang, M., Chen, C., **Yuan, S**., Hermias, J., & Reniers, G. L. L. M. E. (2022). Value of Safety. TU Delft library.

2. van Nunen, K. L. L, Reniers, G. L. L. M. E, Yang, M., Chen, C., & **Yuan, S.** (2021). (Petro)chemical clusters and safety: A cluster-specific ranking of safety parameters. TU Delft library.

Conference Posters, Presentations, & Lectures

1. The 10th International Conference on Safety & Environment in Process & Power Industry, 2022, Firenze, Italy. Poster.

2. The 17th EFCE International Symposium on Loss Prevention and Safety Promotion in Process Industries, 2022, Prague, the Czech Republic. Poster.

3. The 31st Annual Conference of SRA Europe, 2023, Lund, Sweden. Oral presentation.

4. Invited lecture, Optimization and management of safety barriers based on quantitative risks. In College of Environment & Safety Engineering, FuZhou University. August 23, 2023.