Integrating Security in Hazard Analysis Using STPA-Sec and GSPN: A Case Study of Automatic Emergency Braking System

xxx[a,b], xxx[a], xxx[a], xxx[a], xxx[a,*] and xxx[a]

[a] *School of Computer Engineering and Science, Shanghai University, Shanghai, 200444, China*
[b] *Purple Mountain Laboratories, Nanjing, 211111, China*

ARTICLE INFO

Keywords:
Hazard analysis
Safety
Security
Quantification
Intelligent connected vehicles

ABSTRACT

Hazard analysis constitutes a vital step in developing intelligent connected vehicles, aiming to eliminate or control hazards in the initial stages of system development and providing theoretical support for the system's safety design. However, conventional hazard analysis methods such as Failure Mode and Effects Analysis, Hazard and Operability Analysis suffer from two shortcomings: they do not account for the impact of cybersecurity factors on system safety and provide insufficient consideration for quantification. To this end, we propose a quantifiable hazard analysis method, which integrates System Theoretic Process Analysis for Security (STPA-Sec) and Generalized Stochastic Petri Net (GSPN), supporting the extraction, modeling, and quantification of hazards. Specifically, we consider cybersecurity factors on system safety, employing STPA-Sec for qualitative analysis to identify causal scenarios, safety and security requirements, and corresponding mitigations. Then, based on the identified causal scenarios, we establish a GSPN model to quantify system-level hazards. We take an automatic emergency braking system of an open-source test vehicle as a study to demonstrate this approach. Comparative assessments reveal that the proposed approach exhibits an advantage in both analysis processes (integrating security) and results (quantification) when contrasted with existing methods.

1. Introduction

The interconnection, autonomy, sharing, and electrification of vehicles are driving intelligent connected vehicles (ICVs) advancement. ICVs are expected to significantly reduce traffic accidents, transportation costs, and traffic congestion, and ensure passenger safety (Brenner and Herrmann, 2018). Meanwhile, ICVs face serious hazards caused by random functional failures (safety) and malicious cyberattacks (security). For example, from 1966 to 2019, the National Highway Transportation Safety Administration and its predecessor agencies managed recalls of more than 766 million cars, trucks, buses, recreational vehicles, motorcycles and mopeds due to safety defects (Xiaorui et al., 2022). Moreover, a cyber adversary could manipulate millimetre-wave radars, ultrasonic sensors, and forward-looking cameras to perform jamming and spoofing attacks on a vehicle, where the system is tricked into perceiving a fake obstacle in its path (Petit et al., 2015) (Yan et al., 2016). The misbehavior conduct resulting from failures of the vehicle's components or deliberate attacks threatens the communication and perception systems (Bouchouia et al., 2023). As representatives of complex cyber-physical systems (CPS), the safety and security of ICV are two crucial system attributes that need to be simultaneously considered during the hazard analysis process.

In order to identify, assess, and mitigate potential hazards, it is critical to employ hazard analysis methods in the early safety phase of system development. Meanwhile, hazard analysis is not a one-time activity but an ongoing process that aims at evaluating the design and determining the possible events that may result in hazards at the system level and forms a baseline for the numerous activities that are subsequently conducted (Zhou et al., 2020). Today, the widespread use of software and increased functional requirements have significantly increased the complexity of road vehicle systems. ISO 26262 is the functional safety standard for electrical/electronic systems in road vehicles. The standard establishes various procedures and processes to ensure functional safety. In its concept phase section, it outlines the requirement for conducting hazard analysis and risk assessment (HARA) but does not restrict the types of HARA methods. Furthermore, The standard does not provide a quantifiable measure for safety and security performance (Lala et al., 2020).

[*] Corresponding author
✉ xxx (xxx); xxx (xxx); xxx (xxx); xxx (xxx); xxx (xxx); xxx (xxx)

出现的参考文献太旧了。Traditional methods, such as Fault Tree Analysis (FTA) (Watson et al., 1961) and Failure Mode and Effects Analysis(FMEA) (Standard, 1980), concentrate on discerning the effects and probabilities associated with individual component
failures. Quantifiability of failure-based techniques is facilitated by considering the probabilities of component failures. However, they do not account for the impact of cybersecurity factors on system safety and not conside interactions between system components. The limitations of failure-based techniques and the increasing complexity of modern systems led to a new type of accident model. System Theoretic Process Analysis (STPA) (Leveson, 2004) is a relatively recent hazard analysis technique that views safety as a control problem. Accidents originate from inadequate control of component failures, misalignments in interactions between components, external disturbances, and similar factors. In recent years, the STPA method has been applied to the analysis of various automated and highly complex systems, encompassing aerospace exploration systems (Ishimatsu et al., 2014), autonomous vessels (Banda and Kannos, 2017), autonomous vehicles (Wróbel et al., 2018), aircraft systems (Scarinci et al., 2019), driving systems (Khastgir et al., 2021). System Theoretic Process Analysis for Security (STPA-Sec) (Young and Leveson, 2013) is an extension of the STPA, augmenting the safety analysis methodology by incorporating security considerations (Schmittner et al., 2016). While the analysis framework based on STPA possesses advantages over traditional methods, it still strictly emphasizes qualitative analysis, lacking guidance for further quantitative analysis of loss scenarios and system safety.

To address the limitations of current hazard analysis methods, we propose a quantifiable hazard analysis method with security consideration, which integrates STPA-Sec with the Generalized Stochastic Petri Net (GSPN) model. STPA-Sec is utilized to identify and analyze hazards within the system, generating necessary constraints and requirements to ensure secure system operation. GSPN is utilized to model system operational processes and the causal scenarios derived from STPA-Sec analysis for steady-state analysis. This approach enables the calculation of the probability of system-level hazard, facilitating the assessment of system availability. The main contributions of this paper are as follows:

- Employing an improved STPA-Sec methodology, incorporate safety and security considerations into the hazard analysis process. Discerning unintentional and intentional causes for unsafe control actions and their corresponding factors and constraints.
- Proposing a quantitative analysis approach based on GSPN. Building upon the causal scenarios identified in the STPA-Sec analysis, establish a GSPN model for system-level hazards and undertake a steady-state analysis to ascertain the probability.
- A case study of the automatic emergency braking (AEB) system on an open-source test vehicle is to validate the method's effectiveness. Compared with existing hazard analysis approaches, our method exhibits an advantage regarding analysis processes (integrating security) and results (quantification).

The structure of this paper is as follows. In Section 2, we present related work. In Section 3, we introduce our proposed method that combines STPA-Sec and GSPN. Section 4 demonstrates the application of our proposed method in a case study of the AEB system in an open-source test vehicle. Finally, in Section 5, we discuss the results of our case analysis and the limitations of our work. Section 6 concludes the paper.

2. Related Work

Based on the needs of system safety engineering, researchers have extensively explored methods for hazard analysis. From the perspective of modeling approaches, we review prior works by dividing relevant literature related to this paper into three categories: 1) event chain model-based, 2) system theory-based, and 3) state transition-based.

2.1. Event chain model-based

The event chain-based hazard analysis method models the cause-and-effect relationships of accidents as event chains, assuming that accidents result from a series of directly related events. By observing this event chain, a deeper understanding of the accident can be gained. Such hazard analysis methods are widely applied in the engineering field. Of the methods provided in ISO 26262:2018, FMEA, FTA, and Hazard and Operability (HAZOP) Analysis are the most commonly used hazard analyses in the automotive industry (Kölln et al., 2019).

(Alexander et al., 2009) have integrated techniques to identify hazards and establish safety requirements for autonomous systems. They initially employed Energy Trace and Barrier Analysis alongside checklists to offer a foundational hazard identification. Subsequently, a detailed analysis was conducted using Functional Failure Analysis (FFA) and the HAZOP method. (Silva and Lopes, 2013) tackled railway signaling systems scenarios and implemented an integrated safety and security approach. They integrated security-related failure modes into FMEA and failure events into FTA, identifying design constraints that wouldn't have been identified solely from safety concerns. (Schmittner et al., 2014) introduced an extended FMEA analysis technique incorporating security aspects. This technique employs FMEA as a template for a vulnerability cause-effect chain. In this approach, threats are quantified using threat agents representing attackers. Threat modes are identified using a STRIDE model, leading to threat effects and attack probabilities. When new threats or vulnerabilities are identified, previously obtained results can be reused, allowing for reanalysis. (Mohammadfam et al., 2022) devised an integrated model based on the Bowtie method and Bayesian networks (BNs) to evaluate safety and health risks associated with transporting hazardous materials in railway systems. The initial cause-consequence analysis was performed using the Bowtie method, closely connected to Event Tree Analysis (ETA) and FTA. BNs were then employed to establish a quantitative cause-consequences model starting from the root events.

However, with the widespread use of software, the autonomy of control systems, and the increasing interactions between components, the complexity of cyber-physical system continues to rise. This complexity leads to new types of hazards not caused by component failures. Traditional hazard analysis techniques based on reliability theory and functional decomposition, such as FMEA, FTA, and HAZOP, depict systems as several nearly independent subsystems (Scarinci et al., 2019). They are insufficient to represent the indirect interactions between system components and consider conditions leading to inappropriate behaviors.

2.2. System theory-based

Hazard analysis methods based on system theory overcome the limitations of event chain-based approaches. They can comprehensively describe the interaction patterns and coupling relationships in modern complex systems, addressing the challenges posed by the increasing complexity of systems. These methods enhance the comprehensiveness and systemic nature of accident analysis and prediction.

(Leveson, 2004) developed a method called System Theoretic Process Analysis (STPA) based on system theory, considering safety as a system's control problem rather than a component failure problem. It categorizes accident factors into three main classes: inadequate enforcement of safety constraints, inadequate execution of control actions, and inadequate or missing feedback. Based on STPA, (Young and Leveson, 2013) propose the STPA-Sec method, an extension of STPA focusing on security aspects. Similar to STPA, it presents hazards as control problems. Each control action is examined under various conditions and guidewords to identify potential loss scenarios. This approach enables a focus on vulnerable states to prevent threats from being exploited, thereby averting disruptions and potential losses. On the safety and security co-analysis, STPA-Sec has some limitations. (Schmittner et al., 2016) focus on improving an existing approach STPA-Sec and concept phase in the lifecycle. They aligned terminologies used in safety and security analyses, identified unclear guidelines in recognizing intentional causal scenarios using the STPA-Sec method, and suggested incorporating security-relevant elements into a control loop model for improved clarity. They identified and improved limitations in STPA-Sec during safety and security co-analysis. (Friedberg et al., 2017) presents a unified analysis method for safety and security called STPA-SafeSec based on STPA and STPA-Sec. The "general integrity threats" list and "general availability threats" list are introduced in this method. It unifies safety and security considerations while choosing suitable mitigation strategies and can also prioritize the most critical system components for an in-depth security analysis (e.g., penetration testing). (Khastgir et al., 2021) using the STPA method as a foundation and developed an extension to identify test scenarios. This approach was applied to a real-world case study involving an SAE Level 4 Low-Speed Automated Driving System (ADS). In contrast to random (or constrained random) scenario generation, the STPA-extension represents a targeted approach, uncovering actual weaknesses and flaws within the ADS.

In addition to STPA, there are systems theory-based methods such as the Functional Resonance Analysis Method (FRAM). (Guo et al., 2023) proposed a hybrid approach combining FRAM and Dynamic Bayesian Network (DBN). This approach incorporated new data from objective sources and subjective judgments to analyze the evolution characteristics of collision risks in ship pilotage operation processes. FRAM was employed to depict ship pilotage behavior, and the outcomes were integrated with Accident Causation Analysis and Taxonomy to scrutinize the factors contributing to collision accidents and the interrelationships between these factors, thereby developing the structure of a DBN.

Table 1
Relevant papers characterization

Category	Paper	Associated with a standard	Validation	Contribution origin	Application area
Event chain model-based	(Alexander et al., 2009)	No	Case Study	Academic	Generic
	(Silva and Lopes, 2013)	EN 5012x, IEEE 1474	Case Study	Industrial	Railway
	(Schmittner et al., 2014)	IEC 61508, ISO/IEC 27000	Example	Academic	Automotive
	(Mohammadfam et al., 2022)	No	Case Study	Academic	Railway
System theory-based	(Leveson, 2004)	No	Example	Academic	Generic
	(Young and Leveson, 2013)	No	Example	Academic	Nuclear
	(Schmittner et al., 2016)	ISO 26262, SAE J3061	Case Study	Academic	Automotive
	(Friedberg et al., 2017)	No	Example	Academic	Generic
	(Khastgir et al., 2021)	No	Case Study	Academic	Automotive
	(Guo et al., 2023)	No	Case Study	Academic	Ship
State transition-based	(Kriaa et al., 2014)	No	Case Study	Industrial	Control Systems
	(Gonçalves et al., 2017)	STANAG 4671	Case Study	Industrial	Air traffic
	(Zhang et al., 2020)	No	Case Study	Academic	Railway
	(Liang et al., 2022)	No	Case Study	Industrial	Nuclear
	(Mamdikar et al., 2023)	No	Case Study	Academic	Nuclear

2.3. State transition-based

The state transition-based hazard analysis method focuses on the system's behavior itself. By abstracting system states and the transition relationships between states, it constructs graphical or mathematical models of the system. Among these methods, Petri net analysis and Markov analysis are typical representatives.

(Kriaa et al., 2014) conducted a case study on an industrial control system using the Boolean logic-driven Markov processes formalism. This method was employed to identify and rank risks leading to a safety issue, regardless of whether these risks originated from accidents or malevolence. (Gonçalves et al., 2017) utilized Petri Net to demonstrate the safety and reliability of drones. The study aimed to illustrate the frequency of UAVs entering states identified as the most feared events. Additionally, it assessed the UAV's ability to react after encountering a fault situation, focusing on how it responds to inputs from the operating crew. This research was designed to enhance trust and simplify the operational authorization process in UAV operations. (Liang et al., 2022) introduces a method for simplification in Markov state-based models for the reliability assessment of complex safety systems. This method was applied to construct simplified models for a nuclear power plant's typical reactor protection system.

State transition-based hazard analysis methods have also been combined with other approaches. (Zhang et al., 2020) developed the FPN-FTA model by integrating dynamic weighting fuzzy Petri net and FTA. This integration aimed to enhance the accuracy and rationality of safety risk assessment for high-speed railway stations in China. The study employed the stampede accident at Shijiazhuang high-speed railway station as a case study example to assess the stampede risk level. Based on the assessment results, it further derived risk control schemes for this station. (Mamdikar et al., 2023) introduced a framework based on Unified Modeling Language and Petri net (PN). In this framework, UML is utilized to capture all safety-critical system requirements. At the same time, PN is employed to conduct a detailed analysis of the reliability aspects of a safety-critical system.

The pertinent information extracted from the related work has been summarized in Table 1, encapsulating the following characteristics of the papers: 1) association with relevant safety/security standards; 2) approach validation methods detailed in the paper; 3) origination from industry or academia; and 4) the specific application area demonstrated in the paper.

3. Methodology

This section first introduces the overview framework of our proposed method and then expands on the contents of the framework in detail.

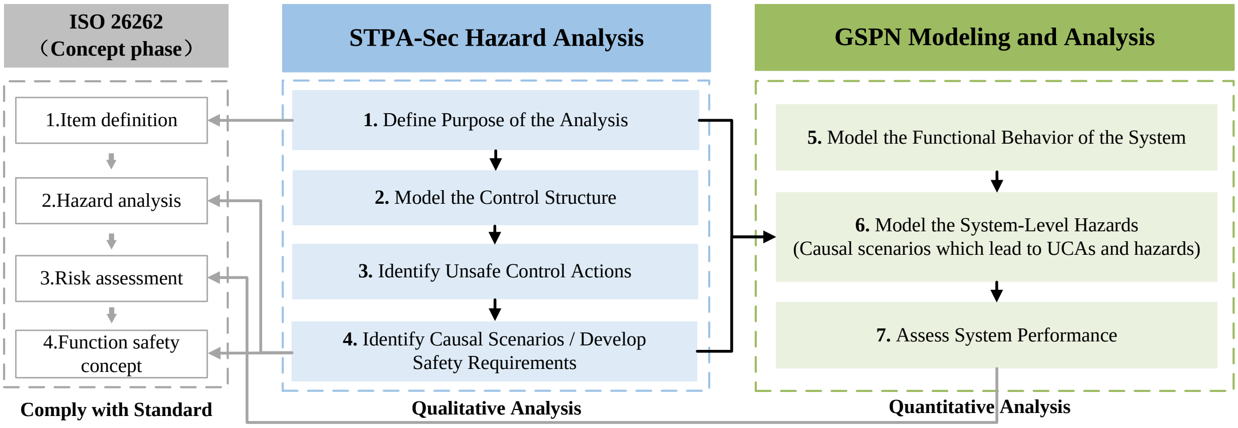
ISO 26262
（Concept phase）
1.Item definition
2.Hazard analysis
3.Risk assessment
4.Function safety concept
Comply with Standard
STPA-Sec Hazard Analysis
1. Define Purpose of the Analysis
2. Model the Control Structure
3. Identify Unsafe Control Actions
4. Identify Causal Scenarios / Develop Safety Requirements
Qualitative Analysis
GSPN Modeling and Analysis
5. Model the Functional Behavior of the System
6. Model the System-Level Hazards (Causal scenarios which lead to UCAs and hazards)
7. Assess System Performance
Quantitative Analysis

Figure 1: Overview of the proposed method

3.1. Overview of the Proposed Method

ISO 26262 is an international standard for functional safety in the automotive industry applying to electronic and electrical systems(Mallya et al., 2016). However, the hazard analysis techniques recommended by the standard may not be adequate to address the escalating complexity of modern software-intensive and safety-critical systems. We present an approach to hazard analysis that integrates STPA-Sec and GSPN, supporting the extraction, analysis, modeling, and quantification of hazard scenarios. The process of this method can be aligned with the HARA process outlined in the ISO 26262 conceptual phase illustrated in Fig.1.

We commence with a qualitative analysis employing STPA-Sec, followed by modeling and quantitative analysis using GSPN. STPA-Sec is employed to discern and scrutinize hazardous scenarios while concurrently considering functional safety and cyber security factors' influence on the system. The analysis of GSPN is conducted at the causal factor level obtained from STPA-Sec analysis. Firstly, a model is established for the system's normal operational scenarios. Subsequently, the causal scenarios derived from STPA-Sec analysis are incorporated to formulate a GSPN model for system-level hazards. Finally, the model undergoes steady-state analysis, evaluating the probability of system-level hazard occurrences. The detailed explanation of STPA-Sec Hazard Analysis and GSPN Modeling and Analysis is in section 3.2 and section 3.3.

3.2. STPA-Sec Hazard Analysis

System-theoretic Process Analysis (STPA) (Leveson, 2004) is a qualitative hazard analysis method based on the System Theoretic Accident Model and Processes, approaching safety concerns as control problems. It involves the creation of a layered control structure through analysis. Subsequently, identifying design flaws within the system's control structure and precisely defining potential faults and hazardous situations it seeks to unearth inappropriate control behaviors. Ultimately, this approach aims to prevent system accidents and address risks from interaction among components. The main steps encompass defining relevant items, defining system-level hazards, establishing a layered control structure for the system, and identifying unsafe control actions and causal scenarios.

System-theoretic Process Analysis for Security (STPA-Sec) (Young and Leveson, 2013) extends the safety-centered STPA approach for security analysis and posits that security is intricately linked to its impact on safety. The application of STPA-Sec to safety and security co-analysis has some limitations, and improvements have been proposed in terms of standardizing terminology and guiding intentional scenarios (Schmittner et al., 2016). The STPA-Sec analysis is carried out in five main steps, as shown in the following Fig.2.

STPA-Sec has the same four basic process steps as STPA:

Step 1: Define the Purpose of the Analysis. STPA-Sec employs a top-down approach with a primary focus on identifying unacceptable losses and vulnerable states. This aids in pinpointing critical system services and functions that require protection and control. The first step identifies such unacceptable losses, accidents, and high-level hazards of the system.

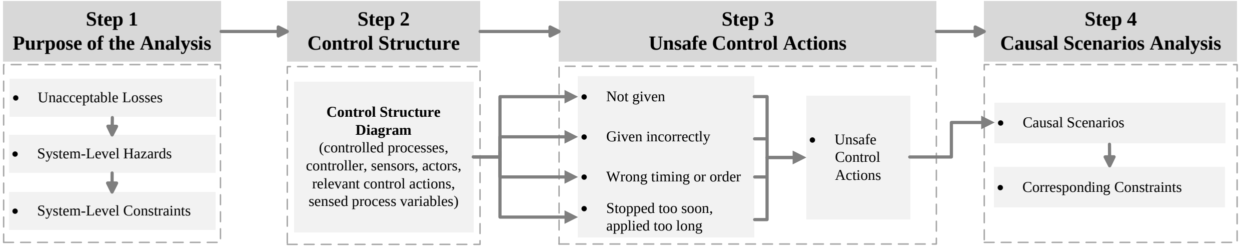

Figure 2: STPA-Sec steps and their outputs

Step 2: Model the Control Structure. In this phase, the control model of the system is formulated. A control model encompasses controlled processes, controllers, sensors, actors, pertinent control actions, and sensed process variables. The generalized control loop is shown in Fig 3.

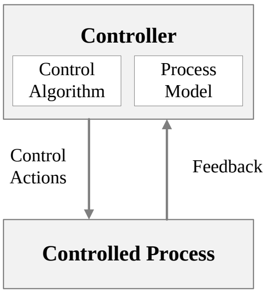

Figure 3: Generic control loop

Step 3: Identify Unsafe Control Actions. All control actions from the control model are reviewed. Unsafe or unsecured control actions or control actions leading to hazards (or vulnerable system states) are analyzed using predefined guidewords by verifying:

- control action not given
- control action given incorrectly
- wrong timing or order of control action
- control action stopped too soon or applied too long

Step 4: Identify Causal Scenarios. Causal scenario descriptions outline the potential triggers for unsafe control behaviors and hazards. In this phase, unsafe control actions are expanded into unsafe scenarios to identify any missing high-level safety constraints. Concurrently, the causal factors containing functional safety and cybersecurity attributes that could lead to unsafe control actions at the system level are extracted. As shown in Fig 4, annotated control loop graphs facilitate the identification of potential causes for unsafe control actions. Unintentional (safety) and intentional (security) causes of unsafe control operations are shown in blue and red, respectively.

3.3. GSPN Modeling and Analysis

The concept of Petri nets (PNs) was first introduced in 1962 in the doctoral thesis titled "Kommunikation mit Automaten" (Communication with Automata) by Carl Adam Petri, then at the University of Bonn in West Germany (Petri, 1962). PNs constitute a formal graphical and mathematical modeling tool tailored for depicting and analyzing the behavior of intricate, distributed, and concurrent systems (Kabir and Papadopoulos, 2019). PNs serve as a graphical and mathematical modeling tool with broad applicability across multiple systems. It functions as a means for delineating and comprehending information processing systems characterized by features such as concurrency, asynchrony, distribution, parallelism, nondeterminism, and/or stochasticity.(Signoret et al., 2013) Tokens are employed within these networks to emulate the dynamic and concurrent actions of the systems. Classical Petri Nets are binary directed graphs, typically represented as sextuples, with a formal representation as $\Sigma = \left(P, T; F, K, W, M_0\right)$,where:

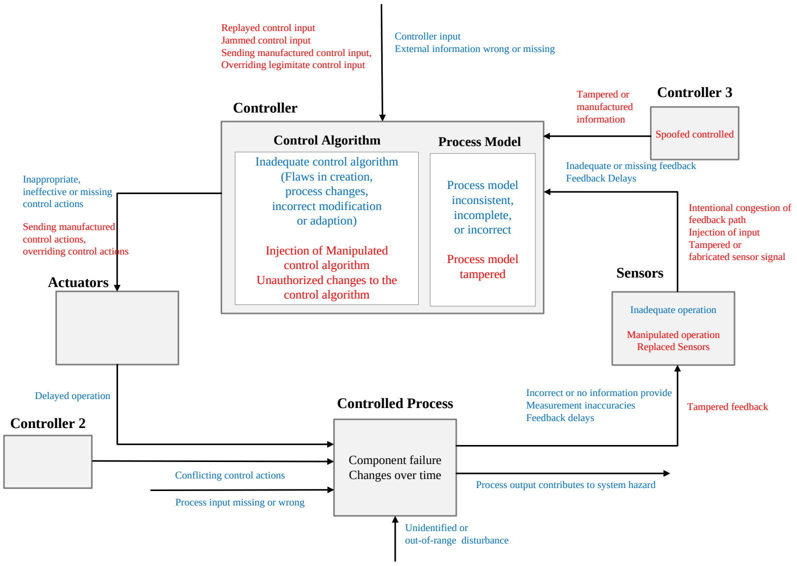

Figure 4: Control loop annotated with potential starting points for the identification of unintentional and intentional causes for unsafe control actions (Schmittner et al., 2016)

- $P = (p_1, p_2, \cdots p_n)$ is a finite set of places, where n is the number of places.
- $T = (t_1, t_2, \cdots t_m)$ is a finite set of transitions, where m is the number of transitions.
- $F \subseteq (P \times T) \cup (T \times P)$ represents the flow relationship, which is the set of arcs.
- $K : P \rightarrow N^+ \cup \{\infty\}$ is the place capacity function, it signifies the maximum quantity of resources that each place can hold.
- $W : F \rightarrow N^+$ is the arc weight function, indicating the actual number of resources consumed or produced by the arcs.
- $M_0 : P \rightarrow N$ represents the initial marking of a PN model, which reflects the number of tokens in each place within the model.

When simulating the behavior of moderately complex systems using classical PNs, it may lead to the issue of state explosion. Moreover, employing classical Petri Nets to model behaviors that change over time can be problematic (Kabir and Papadopoulos, 2019). To address the aforementioned limitations, traditional Petri Nets have undergone various modifications. The introduction of fixed-time parameters into the network model of Petri Nets is referred to as Timed Petri Net (Zuberek, 1991). Another approach to introducing time parameters in Petri Nets is by associating a random delay time between the enabling and firing of each transition, allowing for transitions to have delays following an exponential distribution. This type of Petri Net is referred to as a Stochastic Petri Net(SPN) (Molloy, 1982).

In an $SPN = (P, T; F, W, M_0, \lambda)$, where $(P, T; F, W, M_0)$ remains consistent with classical PN, $\lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_m\}$ represents the set of average firing rates for transitions. λ_i denotes the average firing rate for transition $t_i \in T$, signifying the average number of executions per unit time when it is enabled. GSPN (Marsan and Chiola, 1987) is an extension of SPN, characterized by the division of the transition set T into two subsets: $T = T_t \cup T_i$, $T_t \cap T_i = \emptyset$, the set of timed transitions $T_t = \left\{t_1, t_2, \cdots, t_k\right\}$, the set of immediate transitions $T_i = \left\{t_{k+1}, t_{k+2}, \cdots, t_n\right\}$, the set of average firing rates for timed transitions $\lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_t\}$. In GSPN, the set F allows for the inclusion of inhibitor arcs. Table 2 shows the graphical representation of GSPN modeling elements.

Most stochastic PNs' performance analysis is fundamentally based on their isomorphism with Markov chains in state space. Stochastic PNs offer a robust means of describing the system's performance model, while stochastic Markov processes provide a solid mathematical foundation for model evaluation. When transitions within the model exhibit both deterministic and stochastic delay characteristics, PNs become challenging to analyze and solve. In such scenarios, Monte Carlo simulations are typically coupled with SPN models to facilitate the assessment of system performance (Omeiri et al., 2021).

Table 2
The graphical representation of GSPN modeling elements

Element	Symbol	Meaning
Place	○	States, resources, or conditions of the system
Immediate transition	▮	Events that alter system states with zero delay
Exponential transition	▯	Events that alter system states with exponential distribution delays
Token	●	Number of resources owned by a place
Arc	↗	Bidirectional causal relationship between states and events
Inhibitor arc	⊸	Inhibition of transition firing when the number of tokens in places connected to inhibitory arcs satisfies the transition's activation condition.

The modeling and analysis in GSPN are based on the causal factors identified during the qualitative analysis in STPA-Sec. Specifically, an initial GSPN model is established for the system's normal operation. Then, the main causes of unsafe control actions are extracted from the causal scenarios identified in STPA-Sec analysis to create a model for system-level hazards. Finally, by calculating steady-state probabilities, the probability of system-level hazards occurring. This serves as the foundation for risk assessment.

4. Case Study

In this section, we conduct a case study of a critical system in a real open-source test vehicle. The vehicle is a project developed by Purple Mountain Laboratories for Network Communications and Security, as shown in Fig. 5. We selected the AEB system for analysis because it is crucial in enhancing vehicle and occupant safety. Firstly, we provide a concise description of the AEB system. Subsequently, we outline the methodology's application on this vehicle's AEB system and elucidate the outcomes derived from its implementation.

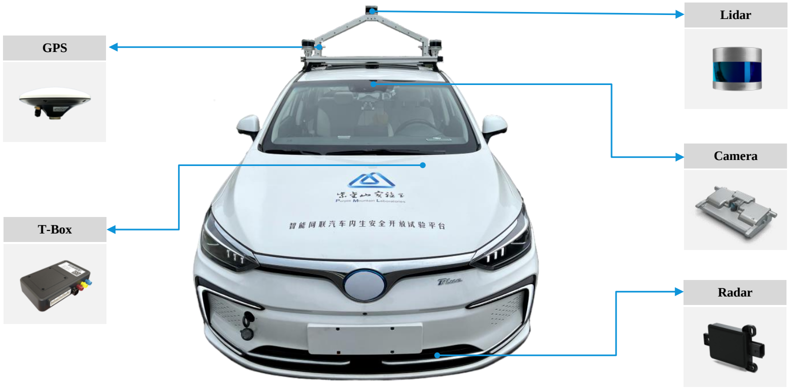

Figure 5: Open-source test vehicle for endogenous safety and security

4.1. AEB System description

AEB is an active safety technology, Fig.6 gives an overview of the AEB system workflow. It operates through a network of sensors, including cameras, radars, and LiDAR, constantly monitoring the vehicle's surroundings. Utilizing safety models, it calculates potential collision risks. When a collision danger emerges, it forewarns the driver to apply brakes preemptively. If necessary, it autonomously engages the braking system, thereby preventing collisions or minimizing the extent of the impact. From a functional perspective, the AEB system can categorize into subsystems, such as the human–machine interface (HMI), AEB control, communication and connection components, and other related systems. Under specific circumstances, based on sensor signals, other autonomous vehicle systems can request the suspension or deactivation of the AEB system. To limit the length of this paper, we analyzed the proper single-end AEB system without considering the advanced functions. Hazard analysis was implemented only on the critical functions of the AEB system.

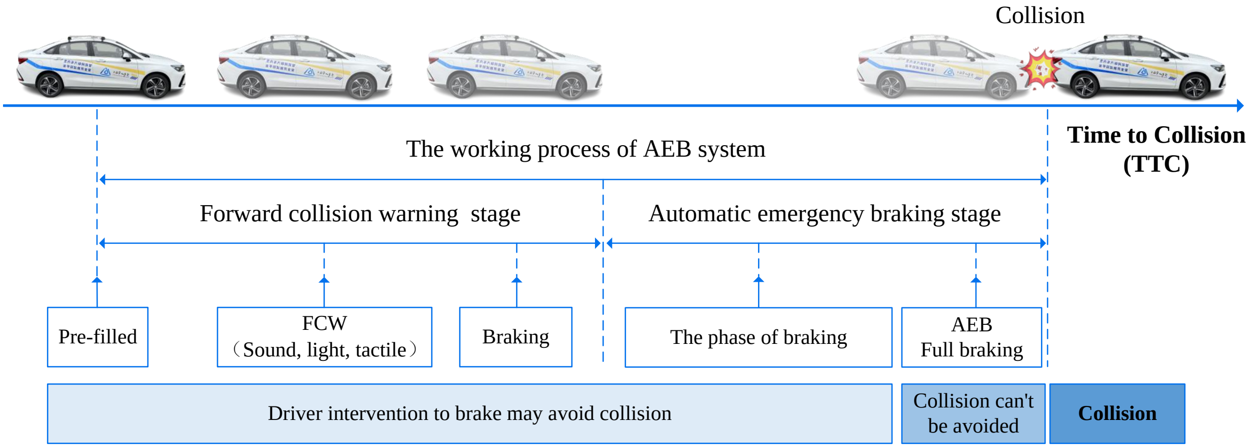
Collision
Time to Collision (TTC)
The working process of AEB system
Forward collision warning stage
Automatic emergency braking stage
Pre-filled
FCW
（Sound, light, tactile）
Braking
The phase of braking
AEB
Full braking
Driver intervention to brake may avoid collision
Collision can't be avoided
Collision

Figure 6: AEB system workflow

The inputs to the AEB system encompass obstacle data from the environmental perception module, self-vehicle status information from onboard sensors, and active operation data from the driver. Its outputs include warning signals such as optical, auditory, and vibrational cues and control commands for the braking system. As illustrated in Fig.7, the AEB system acquires pertinent information from the external environment through sensors. It interacts with the driver through the HMI system, conveying AEB status, sensor parameters, and other relevant details. Moreover, it transmits deceleration commands to the actuator module, affecting the vehicle accordingly.

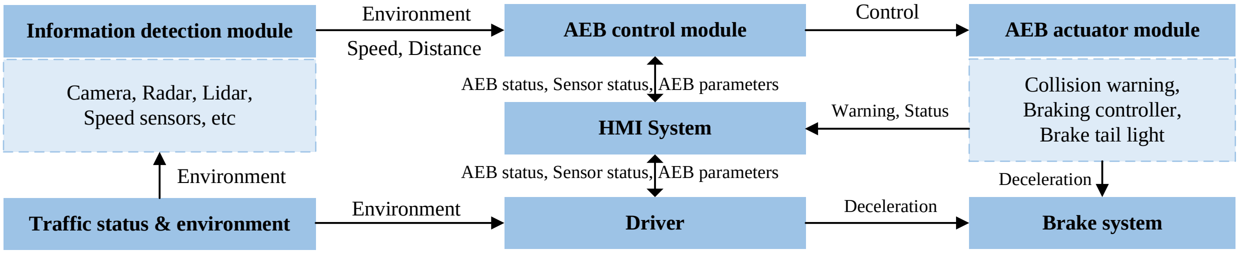

Figure 7: AEB System architecture

4.2. Hazard analysis with STPA-Sec

The STPA-Sec methodology is employed for the hazard analysis of the AEB system, considering both functional safety and security-affecting safety aspects. Presented herein is a systematic exposition of the application of the STPA-Sec analysis, elucidating each step along with the corresponding outcomes garnered.

Step 1: Define the purpose of the analysis (define system boundary). This step contains the following sub-steps:

It is imperative first to establish the system's losses to define the boundary of the analysis. According to the STPA Handbook, losses may encompass loss of life or bodily harm, property damage, environmental pollution, mission failure, damage to reputation, loss or leakage of sensitive information, or any other losses unacceptable to stakeholders. The losses defined in this case study are as follows:

L-1: Injury to individuals (including the self-vehicle driver, passengers, and other road users).

L-2: Vehicle collision losses (including self-vehicle and other road-involved vehicles).

Table 3
Table of identified hazards

Hazard	Comment	Constraints	Related Losses
H-1	Failure to maintain the minimum safe distance from the target vehicle in front of the self-vehicle	During the vehicle's operation, the self-vehicle must maintain a minimum safe distance from the target vehicle in front	L-1, L-2
H-2	The self-vehicle decelerates during braking, exceeding the comfort threshold for passengers or the vehicle's safety limit	The deceleration of the self-vehicle during braking should not exceed the safety threshold for passengers or the vehicle	L-1
H-3	Unexpected triggering of the AEB function leading to rear-end collision by the following vehicle	The vehicle should not activate the AEB function in situations where AEB activation is unnecessary	L-1, L-2
H-4	Failure to correctly execute AEB when it is necessary, resulting in a collision with the preceding vehicle or pedestrian	The vehicle should correctly activate the AEB function in scenarios where AEB activation is necessary	L-1, L-2

After identifying the system and its boundaries, the next step is to define system-level hazards by identifying system conditions or states that could lead to losses under the worst environmental conditions. System-level safety constraints specify the requirements or behaviours the system must meet to prevent hazards and mitigate losses. Each hazard corresponds to a safety constraint and its associated losses. Identified hazards as illustrated in Table 3.

Step 2: Model the control structure and identify unsafe control actions. Identifying Unsafe Control Actions (UCAs) based on the established control loop is instrumental for conducting causation scenario analysis and refining the system's safety requirements. The control structure captures functional relationships and interactions by modelling the system as feedback control loops. The hierarchical control structure utilized in STPA-Sec is a functional model. It can transmit, link, and display information such as commands and feedback, but it may not necessarily correspond to physical connections. Fig.8 shows the control model for the AEB System.

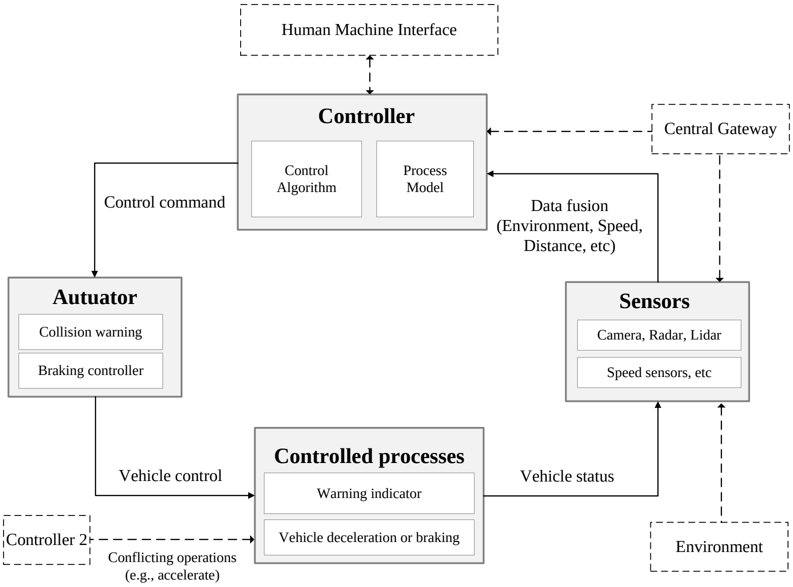

Figure 8: Control loop of the AEB system

Unsafe control actions refer to control behaviours that can lead to hazards in specific environments and under worst-case scenarios. Begin by identifying Control Actions that could potentially lead to the occurrence of hazards. Then, based on the four types of Unsafe Control Actions (UCAs), proceed with their identification. The UCAs were

Table 4
Unsafe control actions

Control Action (CA)	Unsafe Control Action (UCA)			
	Not given	*Given incorrectly*	*Wrong timing or order*	*Stopped too soon or applied too long*
CA1: Driver controls the activation of the AEB function through the HMI system	UCA1: Failure to activate the AEB function when required (H-1/H-4)	UCA2: Activation of the AEB function when it is not required, for instance, when the vehicle is being towed (H-3)	UCA3: Activation of the AEB function is required, but the activation occurs too late (H-1/H-4)	——
CA2: Sensors gather information from the external environment and the self-vehicle, transmitting it to the AEB controller	UCA6: Sensors fail to transmit data to the controller when required (H-1/H-4)	UCA7: Sensors provide incorrect information (H-1/H-3/H-4)	UCA8: Information provided by sensors is not current, not real-time (H-1/H-3/H-4).	UCA9: Information provided by sensors has a too brief time duration (H-1/H-4)
CA3: Controller sends braking commands to the actuator	UCA10: Failure to send braking commands when required (H-1/H-4)	UCA11: Controller sends incorrect braking commands (H-1/H-2/H-3/H-4)	UCA12: Controller sends braking commands either too early or too late (H-1/H-3/H-4)	UCA13: Controller sends braking commands with a too brief duration (H-1/H-4)
CA4: Actuator executes the braking commands from the controller	UCA14: Failure to execute braking commands when required (H-1/H-4)	UCA15: Incorrect execution of braking commands when not needed (H-3)	UCA16: Executing braking commands either too early or too late (H-1/H-3/H-4)	UCA17: Braking commands executed with a too brief duration (H-1/H-4)
CA5: Driver's comfort level can influence the operation of the self-vehicle	——	UCA1: The driver, in an uncomfortable driving environment, may perform incorrect actions with the self-vehicle, such as accelerating during emergency braking (H-1/H-3/H-4)	——	——

identified based on the four guiding terms in Table 4. Each UCA can traced back to one or more system-level hazards. The potential system-level hazards associated with each UCA are listed in the table.

Step 3: Identify intentional and unintentional scenarios for unsafe control actions. Identifying causation scenarios follows the determination of UCAs. Causal scenarios describe factors that could lead to unsafe control actions and hazards. By analyzing components related to the control loop diagram, causal factors corresponding to each UCA are determined. Identify causation scenarios based on guiding terms such as signal errors/loss/delay, feedback errors/loss, connection failures, execution delays, and emergency interventions. As shown in Fig.9, based on the extended annotated control loop, we identified intentional and unintentional causal scenarios for unsafe control actions.

The attack surface of connected autonomous vehicles expands with the enlargement of networks. Vehicle-to-everything (V2X) network connections and in-vehicle communication networks are susceptible to attacks. Furthermore, connections to onboard computers, encompassing physical and wireless links, are increasingly vulnerable to

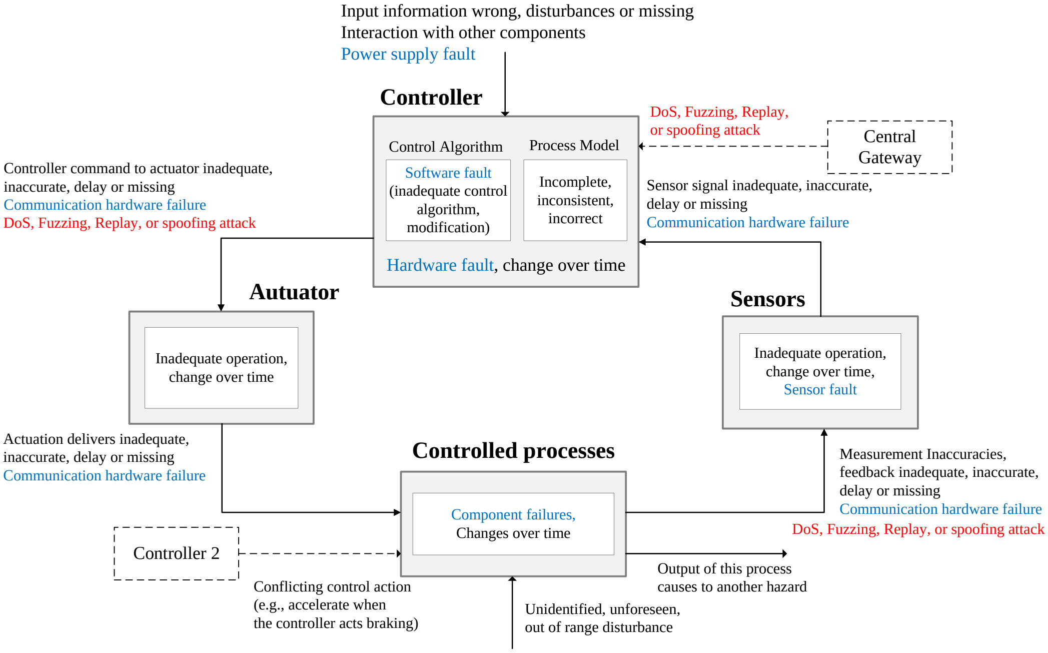
Input information wrong, disturbances or missing
Interaction with other components
Power supply fault
Controller
Control Algorithm
Software fault (inadequate control algorithm, modification)
Process Model
Incomplete, inconsistent, incorrect
Hardware fault, change over time
DoS, Fuzzing, Replay, or spoofing attack
Central Gateway
Sensor signal inadequate, inaccurate, delay or missing
Communication hardware failure
Controller command to actuator inadequate, inaccurate, delay or missing
Communication hardware failure
DoS, Fuzzing, Replay, or spoofing attack
Autuator
Inadequate operation, change over time
Sensors
Inadequate operation, change over time, Sensor fault
Actuation delivers inadequate, inaccurate, delay or missing
Communication hardware failure
Controlled processes
Component failures, Changes over time
Measurement Inaccuracies, feedback inadequate, inaccurate, delay or missing
Communication hardware failure
DoS, Fuzzing, Replay, or spoofing attack
Controller 2
Conflicting control action (e.g., accelerate when the controller acts braking)
Output of this process causes to another hazard
Unidentified, unforeseen, out of range disturbance

Figure 9: Control loop of the AEB system with examples of causal factors leading to hazards

security attacks (Cui et al., 2019). For the AEB system, there are potential network security risks in the feedback processes from sensors to the AEB controller, from the AEB controller to the actuator, and from the actuator back to the sensors (Berdich and Groza, 2023). The causation scenarios leading to UCAs encompass not only human factors, component failures, insufficiently robust control algorithms, unsafe control inputs, and incomplete control models as functional safety factors but also network security elements such as Denial of Service (DoS), fuzzing, replay, and spoofing attacks. Replay attacks involve the retransmission of data frames, which can potentially lead to delays in transmitting these frames. Fuzzing attacks are modification attacks that inject random values into data frames, causing errors in the transmitted data. DoS attacks typically involve writing high-priority frames on the bus, preventing the transmission of data frames on the bus and leading to the loss of correct signals. Spoofing attacks disrupt normal functionality by forging data. These intentional scenarios can lead to unsafe control actions in the AEB system. Table 5 lists the causal scenarios and main causes corresponding to each UCA.

As the final step in STPA-Sec, we have formulated a set of safety constraints and recommended safety requirements. These constraints are considered safety measures, specifying the system behaviours that must be fulfilled to prevent hazards. Based on the Causal Scenarios table (Table 5), the refined safety constraints and safety requirements for each elaborated UCA are outlined in Table 6.

4.3. Modeling and Analysis with GSPN

This section presents the integration part of the GSPN model for modeling and assessing system-level hazards of the AEB system based on causal scenarios obtained by STPA-Sec analysis. Firstly, a GSPN model is developed for the routine operational processes of the AEB system. Subsequently, the model is extended to incorporate causal scenarios to form a system-level hazard model. This methodology enables the quantification of AEB system availability.

4.3.1. Modeling of system-level hazards

The GSPN safety analysis is conducted based on the causal factor levels obtained from the STPA-Sec analysis, as illustrated in Table 5. TimeNET is a software for modelling and performability evaluation using stochastic Petri nets (Zimmermann, 2017). In this case study, the Time NET tool is used for modelling and analyzing the AEB system. Figure 10 illustrates the GSPN model depicting the system-level hazard. The analysis results represent the system's availability under specific circumstances. Based on the results obtained from the STPA-Sec method, four primary causes (sensor failures, controller failures, communication failures, and cyberattacks) leading to system-level hazards

Table 5
Causal scenarios derived from the STPA-Sec application

UCAs	Causal Scenarios	Main causes
UCA1/2/3	Misuse by the driver or passengers leading to unintended AEB activation	Human factor
UCA4/5	Sensor failure failing to detect pedestrians/target vehicles (leading to missed activation) or misidentifying objects such as manhole covers as targets (leading to false activation)	Sensor failure/inadequate algorithms
UCA4/5/6/7	Information obtained by sensors is lost/altered/delayed during transmission, leading to incorrect calculations by the controller	Communication failure/cyberattack (DoS attack/Fuzzing attack/Replay attack)
UCA8/9	Controller failure causing failure to send or sending incorrect braking commands	Controller failure
UCA10/11	Incomplete process model of the AEB controller, such as using maximum braking deceleration strategies or providing the driver with excessively short/long reaction times.	Inadequate/erroneous algorithms
UCA8	Controller subjected to DoS attack inundated with invalid data, resource depletion, leading to incorrect process model and failure to send braking commands	Cyberattack (DoS attack)
UCA9	Controller subjected to a spoofing attack, erroneously sending braking commands	Spoofing attack
UCA8/9/10/11	Braking commands sent from the controller to the actuator are lost/altered/delayed during transmission, leading to unintended braking of the vehicle	Communication failure/cyberattack (DoS attack/Fuzzing attack/Replay attack)
UCA12	Actuator failure resulting in the inability to execute braking commands correctly	Actuator failure
UCA13/14/15	Actuator executing incorrect braking commands from the controller	Incorrect braking commands
UCA16	driver's erroneous intervention in the self-vehicle due to panic or other reasons under unsafe conditions.	Human factor

were identified. After establishing the GSPN model, it enables the assessment of the AEB system's availability under specific circumstances through steady-state analysis.

Places P0-7 and transition T0-6 constitute the operational process model of the AEB system. In this model, the AEB system detects collision risks through sensors and controllers. When a collision risk is detected, the system issues a collision warning. If the driver does not intervene, the system autonomously engages emergency braking, ensuring the vehicle stops safely. Subsequently, the following vehicles can continue their operation. Transitions T0-T5 are considered immediate transitions, and T6 is an exponential transition. Their firing rates are set to 1, indicating that stochasticity does not influence these transitions. Places P8-13 and transitions T7-14 are extensions added to the AEB system based on normal operation, incorporating causal scenarios identified through STPA-Sec analysis. They form the GSPN model representing the system-level hazards. Transitions T18, T10, T12, and T14 are immediate transitions, while transitions T7, T9, T11, T13, and T15 are exponential transitions. The signification of places and transitions, shown in Fig. 10, is summarized in Table 7.

4.3.2. *Performance analysis*

The general disturbances consist of functional failures and cyberattacks. Here are the assumptions made for the system simulated using GSPN: (1) The components forming the system and the system itself can exist in only two states: normal or failure. (2) Each component is independent of the others, meaning there is no consideration for inter-component correlation. (3) The arrival intervals for failure and recoveries follow an exponential distribution.

Table 6
Safety constraints and safety requirements proposed from STPA-Sec results

UCA	Safety constraints	Recommended safety requirements
UCA1/2/3	Drivers should activate or deactivate the AEB function appropriately to prevent misuse	Drivers should maintain a vigilant state and exercise caution while driving
UCA4/5/6/7	The information obtained by sensors (such as cameras, radars, and LiDAR) must ensure accuracy and real-time capability, allowing for correct recognition	Enhancing the hardware reliability of sensors and the robustness of recognition algorithms
UCA4/5/6/7/8/9/10/11	Information exchange between all modules needs to be safeguarded	Measures such as authentication, data encryption, intrusion detection, and others can be employed as defence strategies
UCA10	Require a comprehensive process model ensuring emergency braking stays within the safety or comfort thresholds of both occupants and the vehicle	Optimize the control algorithm; the parameters need to be more precise and rational
UCA4/5/8/9/12	The components of the AEB system require stability and reliability to be maintained	Enhance the hardware reliability of all components within the AEB system, redundancy designs can be employed

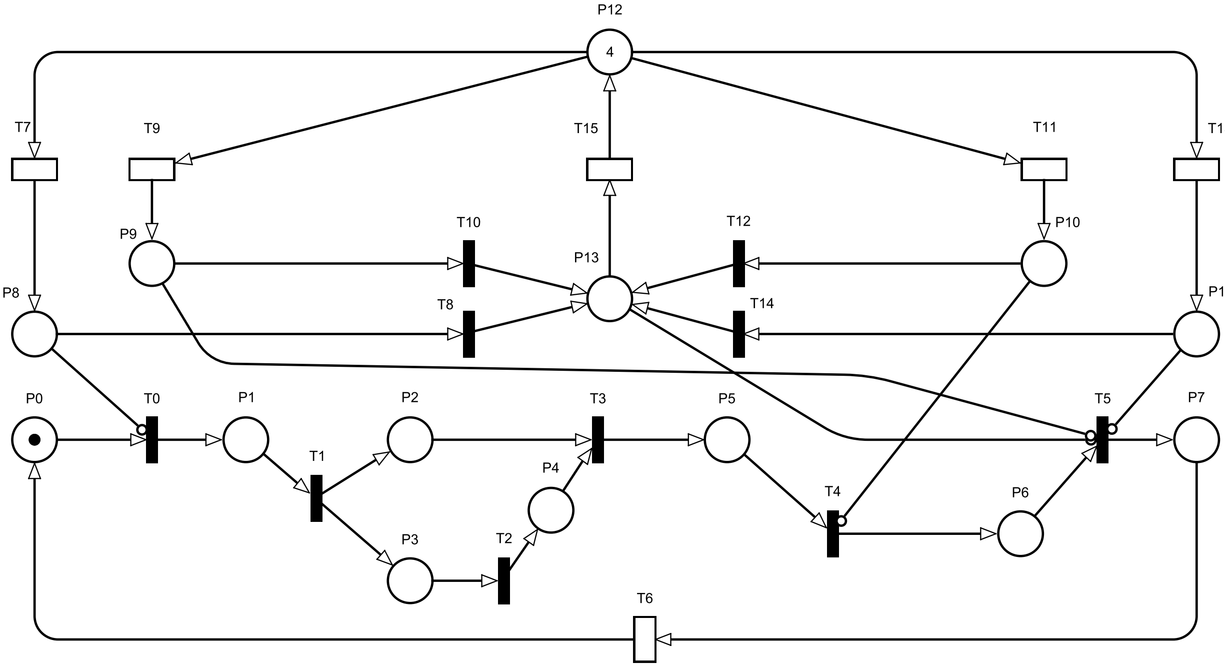

Figure 10: Main causal scenarios leading to system-level hazard

Functional failure probability data are retrieved from references based on what met the operating conditions. Common network attacks in vehicular networks, such as flooding, spoofing, and fault attacks, typically require around 10 seconds to cause a malfunction or gain control over the vehicle (Li et al., 2023). The average interarrival time of general disturbances is taken as 10 h under weak cyberattack conditions. The critical hardware circuits of the AEB system employ redundant design. The recovery rate of the function is determined by the damaged hardware and is set based on practical development experience. Table 8 gives the parameter values of the GSPN model. The established GSPN

Table 7
The signification of places and transitions of GSPN in Fig.10

Place	Signification	Transition	Signification
P0	Vehicle movement	T0	Enable AEB functionality
P1	Collision risk detection	T1	Detect collision risk
P2	Time-to-Collision (TTC) countdown	T2	Stop collision warning
P3	Collision warning	T3	Activate AEB functionality
P4	Driver non-intervention	T4	Send brake command
P5	Activation of AEB functionality	T5	Vehicle brakes
P6	Actuator receiving commands	T6	Vehicle resumes operation
P7	Vehicle engaging in safe braking	T7	Sensor failure occurs
P8	Sensor failure	T8	Sensor failure led to hazards
P9	Communication failure	T9	Communication failure occurs
P10	Actuator failure	T10	Communication failure led to hazards
P11	Subjected to cyberattacks	T11	Actuator failure occurs
P12	Number of potential disturbances	T12	Actuator failure led to hazards
P13	Occurrence of system-level hazards	T13	Cyberattacks
		T14	Cyberattacks led to hazards
		T15	Restore the system to normal function

Table 8
Firing transition rates of the GSPN model in Fig.10

Firing rate	Value h^{-1}	Signification	Data source
λ_7	$2.35 * 10^{-8}$	Communication line failure rate	(Knight et al., 2001)
λ_9	$1.57 * 10^{-5}$	Controller failure rate	(Knight et al., 2001)
λ_{11}	$3.31 * 10^{-5}$	Sensor failure rate	(Lu and Chen, 2019)
λ_{13}	0.1	Common cyberattacks rate	(Li et al., 2023)
λ_{15}	20	System recovery rate	(Li et al., 2023)

model comprises 62 states, including 53 transient states and 9 stable states. Fig.11 illustrates the reachability graph of the model. It demonstrates that the model is deadlock-free and bounded.

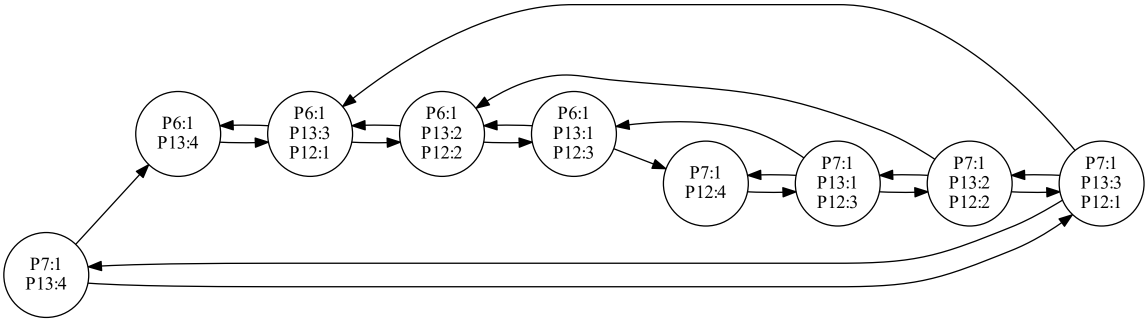

Figure 11: Reachability graph of the GSPN model

Using the Markov Chain Homomorphism method to calculate the steady-state probabilities of the GSPN. As shown in Table 9 results, under weak cyberattack disturbances, the AEB system can maintain a steady-state probability

Table 9
Stable probability of AEB System under weak cyberattack disturbances

Stable state	Probability	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13
M0	0.99499705	0	0	0	0	0	0	0	1	0	0	0	0	4	0
M1	0.0047398	0	0	0	0	0	0	0	1	0	0	0	0	3	1
M2	0.00023812	0	0	0	0	0	0	1	0	0	0	0	0	3	1
M3	0.00002258	0	0	0	0	0	0	0	1	0	0	0	0	2	2
M4	0.00000233	0	0	0	0	0	0	1	0	0	0	0	0	2	2
M5	0.00000011	0	0	0	0	0	0	0	1	0	0	0	0	1	3
M6	0.00000002	0	0	0	0	0	0	1	0	0	0	0	0	1	3
M7	0	0	0	0	0	0	0	0	1	0	0	0	0	0	4
M8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	4

of 0.99499705 of normal operational status. Due to the redundant design, components that fail are promptly replaced, preventing the occurrence of system-level hazards. The probability of system-level hazards occurring due to simultaneous communication line failure rate, actuator failure rate, sensor failure rate, and network attack is extremely low.

The simulation experiment duration is 100,000 hours, with a confidence interval (CI) of 95%. The System unavailability was directly calculated by standard Monte Carlo style simulation approach using the TimeNET tool as depicted in Fig.12. The upper and lower bounds of CI define the confidence interval and indicate the convergence of the results. The system's unavailability sharply escalated within the first hour of operation under weak cyberattack disturbances. However, after 7 hours of runtime, it stabilized, reaching a level of 0.005.

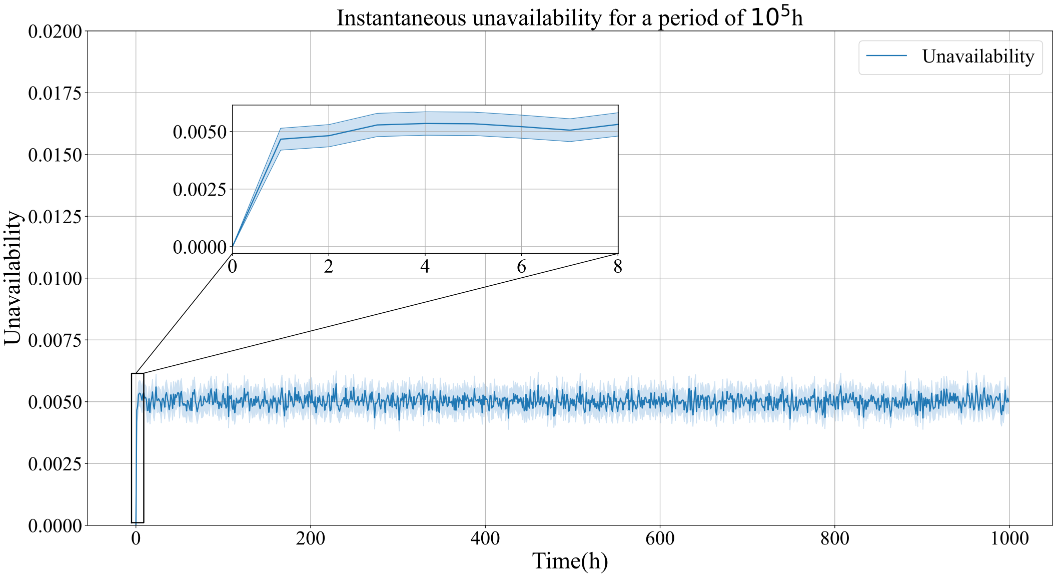

Figure 12: Unavailability of the AEB System under weak disturbance conditions

4.4. Methods Comparison Analysis

This section presents a comparative analysis of the HAZOP, FMEA, FRAM, STPA, and the method proposed in this paper for analyzing the hazards in the AEB system. HAZOP employs guiding words (such as "no," "more," and

Table 10
Analysis comparison (adapted from (Sun et al., 2022))

Attribute	HAZOP	FMEA	FRAM	STPA	STPA-Sec and GSPN
Inductive/ deductive	I(effect)/D(cause)	I	I(upstream)/ D(downstream)	I(control loop)/D(scenario)	I(control loop)/D(scenario)
Starts from	Function	Failure	Function/action	Loss/system-level hazard	Loss/system-level hazard
Abstract/specific	Specific	Specific	Specific/abstract	Specific/abstract	Specific/abstract
Identify hazards	Partially	Partially	Comprehensively	Comprehensively	Comprehensively
Identify root causes	Y	Y	Y	Y	Y
Qualitative/ quantitative	Qual. and quant. (deviation)	Qual. and quant. (priority)	Qual.	Qual.	Qual. and quant. (probability)
Safety/Security	Safety	Safety	Safety	Safety	Safety and security
Hazard causal factors	hardware (HW)/software (SW) failure	HW/SW failure	Resonance (HW/SW/ human/interface failures)	Interface (control command between HW/SW/human etc.)	Interface (control command between HW/SW/human/ cybersecurity etc.)
Level of detail	In depth	In depth	Adjustable	Adjustable	Adjustable

"less") in conjunction with process parameters (such as temperature, flow, and pressure) to discern how the process deviates from its intended design. After investigating the potential causes and consequences of these deviations, strategies are devised to prevent or mitigate the impacts of the identified deviations (Dunjó et al., 2010). FMEA, introduced in the aerospace industry in the 1960s, is a bottom-up analytical approach. It is employed to identify potential failure modes and the causes of failures in all components of a system to eliminate identified potential failures. The analysis commences at the lowest level of components and extends to encompass the entire system's failure impacts (Chiozza and Ponzetti, 2009). FRAM systematically analyzes six aspects of individual functions — input (I), output (O), time (T), resources (R), precondition (P), and control (C) — to identify functions and performance variabilities. Once these variabilities are aggregated, hazards and their corresponding uncontrollable performance variabilities can be pinpointed. Subsequent measures to eliminate or mitigate these hazards can then be implemented accordingly (Hollnagel, 2017).

The primary attributes of the comparative analysis include inductive or deductive reasoning, identified hazard on an abstract or specific level, comprehensiveness in identifying hazard sources, qualitative or quantitative analysis and assessment, the type of identified hazard causal factors, and the level of design detail that can be evaluated. Table 10 illustrates the comparative analysis of results obtained through different methodologies.

The five methods have categorized identified hazards into hardware failures, software failures, communication failures, and operational errors. Traditional and contemporary analytical methods can identify the hazards within the categories above. These methods share common identifications of faults/hazards, while some methods also reveal unique faults/hazards not identified by the others. Most common failures are hardware-related, whereas unique faults identified by FRAM and STPA are software-related. The improved STPA-Sec method also identified potential attack vectors within the AEB system. The causal scenarios analysis considered both functional safety and security factors.

As the software applications in the AEB system increase, HAZOP focuses on deviations from functionality or components, resulting in relatively straightforward analysis outcomes. FMEA evaluates the effectiveness of faults themselves, highlighting and determining countermeasures. FRAM relies heavily on qualitative analysis and expert insights (Yu et al., 2020), offering advantages in describing system dynamics. However, it does not classify accident causes and requires detailed expert analysis during application. STPA analysis results emphasize the correctness of control actions. For complex systems involving interactions among high-intelligence components, STPA can be employed to comprehend system behavior (Zhou et al., 2020). The first and fourth steps of these methods are similar. The first step involves defining the system boundaries, system functions, workflows, and interactions. The fourth step involves proposing relevant identified risk issues and analyzing them to minimize hazard risks to the maximum extent. Our proposed method simultaneously considers safety and security aspects. Building on qualitative analysis, it quantifies the probability of system-level hazard occurrences, providing more refined metrics for reducing hazard risks.

5. Discussion

The analysis results show that the primary issues impeding the secure operation of the AEB system are sensor failures, communication failures, controller failures, and cyberattacks. Through the GSPN steady-state calculation based on the provided data, under weak disturbance conditions, the unavailability of the AEB system is 0.005. It is noteworthy that due to the redundant control strategy implemented in the AEB system, occurrences of component failures or partial functional impairments resulting from cyberattacks have not led to system-level hazards. Redundant control designs can enhance the safety of ICV. The specific design process requires designers to make appropriate adjustments and optimizations based on the actual circumstances.

HARA is introduced in the concept phase of ISO 26262. However, STPA-Sec does not encompass risk assessment directly. It primarily focuses on the causes of accidents due to weak controls or insufficient execution. Building upon STPA-Sec, we combine the GSPN model, enabling the modeling and quantification of the system-level hazard identified. These quantifications serve as a basis for risk assessment. In this study, we select the system-level hazards caused by component failure and disturbance conditions while ignoring other factors such as human errors and abnormal control algorithms. Risk assessment is conducted by considering the combination of severity and occurrence probability. Our primary emphasis lies in hazard analysis and availability assessment. There is a need for a more comprehensive elaboration on risk severity segmentation.

6. Conclusion

In this study, addressing the limitations of traditional hazard analysis methods, we propose an integrated hazard analysis methodology that combines STPA-Sec and GSPN. A real-world case study is conducted on the AEB system in ICV to showcase the efficacy of our approach. The STPA-Sec methodology is employed to identify and analyze the causal scenarios with safety and security considerations. It generates necessary safety and security constraints, ensuring the safe operation of the AEB system within intricate network environments. Meanwhile, GSPN complements the limitation of STPA-Sec methodology, establishing a framework for modeling and quantifying system-level hazards.

In future work, we aim to undertake a more rigorous quantification of causal scenarios, thereby refining the risk assessment segment. Additionally, we plan to achieve a closer integration between STPA-Sec and GSPN. Beyond the amalgamation at the causal factor level, the control structures derived from STPA-Sec analysis can be mapped onto the GSPN model for event simulation. This approach enables a more precise assessment of the system's reliability and availability metrics.

Acknowledgements

This work was sponsored by Henan Science and Technology Major Project (221100240100), Shanghai Pujiang Program (23PJ1403000), the National Science Foundation of China (62002213), Shanghai Sailing Program (21YF1413800 and 20YF1413700), the Program of Industrial Internet Visualized Asset Management and Operation Technology and Products, Shanghai Science and Technology Innovation Action Plan (21511102500).

References

Alexander, R., Herbert, N., Kelly, T., 2009. Deriving safety requirements for autonomous systems, in: 4th SEAS DTC technical conference, Citeseer.

Banda, O.A.V., Kannos, S., 2017. Hazard analysis process for autonomous vessels, in: Technical Report.

Berdich, A., Groza, B., 2023. Secure by design autonomous emergency braking systems in accordance with iso 21434, in: Machine Learning and Optimization Techniques for Automotive Cyber-Physical Systems. Springer, pp. 155–187.

Bouchouia, M.L., Labiod, H., Jelassi, O., Monteuuis, J.P., Jaballah, W.B., Petit, J., Zhang, Z., 2023. A survey on misbehavior detection for connected and autonomous vehicles. Vehicular Communications , 100586.

Brenner, W., Herrmann, A., 2018. An overview of technology, benefits and impact of automated and autonomous driving on the automotive industry. Digital marketplaces unleashed , 427–442.

Chiozza, M.L., Ponzetti, C., 2009. Fmea: a model for reducing medical errors. Clinica chimica acta 404, 75–78.

Cui, J., Liew, L.S., Sabaliauskaite, G., Zhou, F., 2019. A review on safety failures, security attacks, and available countermeasures for autonomous vehicles. Ad Hoc Networks 90, 101823.

Dunjó, J., Fthenakis, V., Vílchez, J.A., Arnaldos, J., 2010. Hazard and operability (hazop) analysis. a literature review. Journal of hazardous materials 173, 19–32.

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.

Gonçalves, P., Sobral, J., Ferreira, L.A., 2017. Unmanned aerial vehicle safety assessment modelling through petri nets. Reliability Engineering & System Safety 167, 383–393.

Guo, Y., Jin, Y., Hu, S., Yang, Z., Xi, Y., Han, B., 2023. Risk evolution analysis of ship pilotage operation by an integrated model of fram and dbn. Reliability Engineering & System Safety 229, 108850.

Hollnagel, E., 2017. FRAM: the functional resonance analysis method: modelling complex socio-technical systems. Crc Press.

Ishimatsu, T., Leveson, N.G., Thomas, J.P., Fleming, C.H., Katahira, M., Miyamoto, Y., Ujiie, R., Nakao, H., Hoshino, N., 2014. Hazard analysis of complex spacecraft using systems-theoretic process analysis. Journal of spacecraft and rockets 51, 509–522.

Kabir, S., Papadopoulos, Y., 2019. Applications of bayesian networks and petri nets in safety, reliability, and risk assessments: A review. Safety science 115, 154–175.

Khastgir, S., Brewerton, S., Thomas, J., Jennings, P., 2021. Systems approach to creating test scenarios for automated driving systems. Reliability engineering & system safety 215, 107610.

Knight, I., Eaton, A., Whitehead, D., 2001. The reliability of electronicallly controlled systems on vehicles. Transport Research Laboratory (TRL) .

Kölln, G.C., Klicker, M., Schmidt, S., 2019. Comparison of hazard analysis methods with regard to the series development of autonomous vehicles, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE. pp. 2969–2975.

Kriaa, S., Bouissou, M., Colin, F., Halgand, Y., Pietre-Cambacedes, L., 2014. Safety and security interactions modeling using the bdmp formalism: case study of a pipeline, in: Computer Safety, Reliability, and Security: 33rd International Conference, SAFECOMP 2014, Florence, Italy, September 10-12, 2014. Proceedings 33, Springer. pp. 326–341.

Lala, J.H., Landwehr, C.E., Meyer, J.F., 2020. Autonomous vehicle safety: lessons from aviation. Communications of the ACM 63, 28–31.

Leveson, N., 2004. A new accident model for engineering safer systems. Safety science 42, 237–270.

Li, Y., Liu, Q., Zhuang, W., Zhou, Y., Cao, C., Wu, J., 2023. Dynamic heterogeneous redundancy-based joint safety and security for connected automated vehicles: Preliminary simulation and field test results. IEEE Vehicular Technology Magazine .

Liang, Q., Yang, Y., Zhang, H., Peng, C., Lu, J., 2022. Analysis of simplification in markov state-based models for reliability assessment of complex safety systems. Reliability Engineering & System Safety 221, 108373.

Lu, K.L., Chen, Y.Y., 2019. Iso 26262 asil-oriented hardware design framework for safety-critical automotive systems, in: 2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE), IEEE. pp. 1–6.

Mallya, A., Pantelic, V., Adedjouma, M., Lawford, M., Wassyng, A., 2016. Using stpa in an iso 26262 compliant process, in: Computer Safety, Reliability, and Security: 35th International Conference, SAFECOMP 2016, Trondheim, Norway, September 21-23, 2016, Proceedings 35, Springer. pp. 117–129.

Mamdikar, M.R., Kumar, V., Bharti, S., Singh, P., 2023. Reliability analysis of safety-critical systems using optimized petri nets. Progress in Nuclear Energy 164, 104841.

Marsan, M.A., Chiola, G., 1987. On petri nets with deterministic and exponentially distributed firing times, in: Advances in Petri Nets 1987 7, Springer. pp. 132–145.

Mohammadfam, I., Zarei, E., Yazdi, M., Gholamizadeh, K., et al., 2022. Quantitative risk analysis on rail transportation of hazardous materials. Mathematical Problems in Engineering 2022.

Molloy, 1982. Performance analysis using stochastic petri nets. IEEE Transactions on computers 100, 913–917.

Omeiri, H., Hamaidi, B., Innal, F., Liu, Y., 2021. Verification of the iec 61508 pfh formula for 2oo3 configuration using markov chains and petri nets. International Journal of Quality & Reliability Management 38, 581–601.

Petit, J., Stottelaar, B., Feiri, M., Kargl, F., 2015. Remote attacks on automated vehicles sensors: Experiments on camera and lidar. Black Hat Europe 11, 995.

Petri, C.A., 1962. Kommunikation mit automaten .

Scarinci, A., Quilici, A., Ribeiro, D., Oliveira, F., Patrick, D., Leveson, N.G., 2019. Requirement generation for highly integrated aircraft systems through stpa: An application. Journal of Aerospace Information Systems 16, 9–21.

Schmittner, C., Gruber, T., Puschner, P., Schoitsch, E., 2014. Security application of failure mode and effect analysis (fmea), in: Computer Safety, Reliability, and Security: 33rd International Conference, SAFECOMP 2014, Florence, Italy, September 10-12, 2014. Proceedings 33, Springer. pp. 310–325.

Schmittner, C., Ma, Z., Puschner, P., 2016. Limitation and improvement of stpa-sec for safety and security co-analysis, in: Computer Safety, Reliability, and Security: SAFECOMP 2016 Workshops, ASSURE, DECSoS, SASSUR, and TIPS, Trondheim, Norway, September 20, 2016, Proceedings 35, Springer. pp. 195–209.

Signoret, J.P., Dutuit, Y., Cacheux, P.J., Folleau, C., Collas, S., Thomas, P., 2013. Make your petri nets understandable: Reliability block diagrams driven petri nets. Reliability Engineering & System Safety 113, 61–75.

Silva, N., Lopes, R., 2013. Practical experiences with real-world systems: Security in the world of reliable and safe systems, in: 2013 43rd Annual IEEE/IFIP Conference on Dependable Systems and Networks Workshop (DSN-W), IEEE. pp. 1–5.

Standard, M., 1980. Procedures for performing a failure mode, effects and criticality analysis. Technical Report. MIL-STD-1629A.

Sun, L., Li, Y.F., Zio, E., 2022. Comparison of the hazop, fmea, fram, and stpa methods for the hazard analysis of automatic emergency brake systems. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering 8, 031104.

Watson, H.A., et al., 1961. Launch control safety study. Bell labs .

Wróbel, K., Montewka, J., Kujala, P., 2018. Towards the development of a system-theoretic model for safety assessment of autonomous merchant vessels. Reliability Engineering & System Safety 178, 209–224.

Xiaorui, F., Le Liu, Z., Li, W., et al., 2022. Analysis on the trends and characteristics of vehicle recalls in the united states, in: 2021 International Conference on Culture, Design and Social Development (CDSD 2021), Atlantis Press. pp. 1–5.

Yan, C., Xu, W., Liu, J., 2016. Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle. Def Con 24, 109.

Young, W., Leveson, N., 2013. Systems thinking for safety and security, in: Proceedings of the 29th Annual Computer Security Applications Conference, pp. 1–8.

Yu, M., Quddus, N., Kravaris, C., Mannan, M.S., 2020. Development of a fram-based framework to identify hazards in a complex system. Journal of Loss Prevention in the Process Industries 63, 103994.

Zhang, Q., Zhuang, Y., Wei, Y., Jiang, H., Yang, H., 2020. Railway safety risk assessment and control optimization method based on fta-fpn: A case study of chinese high-speed railway station. Journal of advanced transportation 2020, 1–11.

Zhou, X.Y., Liu, Z.J., Wang, F.W., Wu, Z.L., Cui, R.D., 2020. Towards applicability evaluation of hazard analysis methods for autonomous ships. Ocean Engineering 214, 107773.

Zimmermann, A., 2017. Modelling and performance evaluation with timenet 4.4, in: Quantitative Evaluation of Systems: 14th International Conference, QEST 2017, Berlin, Germany, September 5-7, 2017, Proceedings 14, Springer. pp. 300–303.

Zuberek, W.M., 1991. Timed petri nets definitions, properties, and applications. Microelectronics Reliability 31, 627–644.