

Collaborative Analysis Framework of Safety and Security for Autonomous Vehicles

Cui, J., Sabaliauskaite, G., Liew, L. S., Zhou, F. & Zhang, B.

Published PDF deposited in Coventry University's Repository

Original citation:

Cui, J, Sabaliauskaite, G, Liew, LS, Zhou, F & Zhang, B 2019, 'Collaborative Analysis Framework of Safety and Security for Autonomous Vehicles', IEEE Access, vol. 7, pp. 148672 - 148683.

<https://dx.doi.org/10.1109/ACCESS.2019.2946632>

DOI 10.1109/ACCESS.2019.2946632

ESSN 2169-3536

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see <http://creativecommons.org/licenses/by/4.0/>

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

Received September 9, 2019, accepted October 4, 2019, date of publication October 11, 2019, date of current version October 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2946632

Collaborative Analysis Framework of Safety and Security for Autonomous Vehicles

JIN CUI¹, GIEDRE SABALIAUSKAITE², LIN SHEN LIEW²,
FENGJUN ZHOU², AND BIAO ZHANG³

¹School of Information and Science Technology, Northwest University, Xi'an 710127, China

²iTrust, Singapore University of Technology and Design, Singapore 487372

³Institute of Flexible Electronics (IFE), Northwestern Polytechnical University, Xi'an 710129, China

Corresponding authors: Jin Cui (jin.cui@nwu.edu.cn) and Biao Zhang (iambzhang@nwpu.edu.cn)

ABSTRACT Human error has been statistically proven to be the primary cause of road accidents. This undoubtedly is a contributory cause of the rising popularity of autonomous vehicles as they are presumably able to maneuver appropriately/optimally on the roads while diminishing the likelihood of human error and its repercussion. However, autonomous vehicles are not ready for widespread adoption because their safety and security issues are yet to be thoroughly investigated/addressed. Little literature could be found on collaborative analysis of safety and security of autonomous vehicles. This paper proposes a framework for analyzing both safety and security issues, which includes an integrated safety and security method (S&S) with international vehicle safety and security standards ISO 26262 and SAE J3061. The applicability of the proposed framework is demonstrated using an example of typical autonomous vehicle model. Using this framework, one can clearly understand the vehicle functions, structure, the associated failures and attacks, and also see the vulnerabilities that are not yet addressed by countermeasures, which helps to improve the in-vehicle safety and security from researching and engineering perspectives.

INDEX TERMS Autonomous vehicle, safety, security, ISO 26262, SAE J3061, SAE J3016.

I. INTRODUCTION

An ever increasing number of vehicles on the roads worldwide has apparently increased the frequency of the traffic accidents, which is recognized as a major societal and public safety problem. In 2016 alone, more than thirty thousand people died in road accidents in United States, an increase of 5.6% over 2015 [1]. The economic cost of road traffic crashes was substantial, amounting to over 200 billion dollars a year [2]. Statistically, human error tops the list of factors of road accidents (causing 94% of road accidents), followed by vehicle malfunction, environmental factors and others [3]. The human error encompasses recognition error (e.g., driver's inattention and distraction), decision error (e.g., reckless driving and misjudging others' action), and performance error (e.g., overcompensation and poor driving skill) [3].

Driving automation is considered a solution to mitigate the human driving errors [4], [5]. A Driving Automation System (DAS) [6] usually makes use of a great variety of advanced

sensors and technologies such as Light Detection and Ranging (LiDAR), Global Positioning System (GPS), 3D mapping, path planning and Electronic Controlled Units (ECUs). Therefore, as compared to an average human driver, it should perform better with respect to recognition, decision-making as well as vehicle motion control. Vehicles equipped with DAS are the so-called Autonomous Vehicles (AVs). Another feature of AV is its V2X communication technology; V2X is a shortened form of Vehicle-to-Anything. In other words, an AV can communicate with other AVs, infrastructure and pedestrians. The AVs, once widely deployed, are expected to diminish human errors, optimize traffic flow, and ultimately enhance overall safety and experience of road users.

However, many issues concerning AVs' reliability and safety have to be tackled before AVs are indeed ready for wide adoption. Fatal crash of an AV including pedestrian has been reported in March 2018 [7]. This undoubtedly increase the emphasis on AV safety: keeping the AV safe is of paramount importance. AV is a safety-critical system, and any failures of AV may result in severe human injuries or even deaths.

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng Yan¹.

Meanwhile, AV consists of a myriad of heterogeneous components, both cyber and physical, which pose additional security challenges.

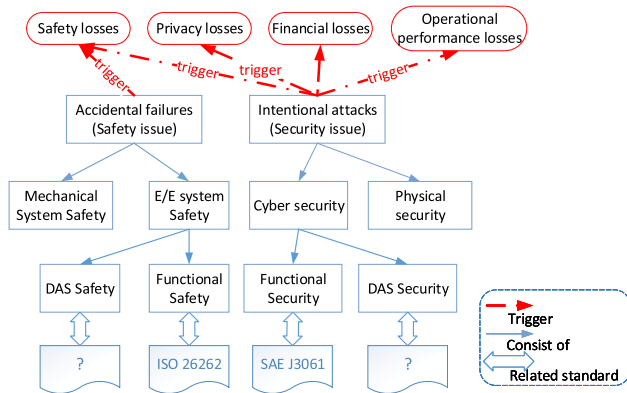


FIGURE 1. Safety and security composition for autonomous vehicles, and the related impacts.

AV safety aims at protecting the vehicle from accidental failures in order to avoid hazards, and security focuses on protecting the vehicle from intentional attacks [8]. Fig. 1 demonstrates safety and security composition for AVs and related impact. Safety of AV includes mechanical system safety and Electrical and Electronic (E/E) system safety, while E/E safety consists of functional safety and DAS safety. AV security includes physical security and cyber security, and cyber security contains DAS security and functional security. While accidental failures (safety issues) jeopardize AV's safety, intentional attacks (security issues) affect not only AV's safety but also its privacy, financial cost and operational performance. Thus, to ensure safety of AV, both accidental failures and intentional attacks have to be tackled; this implies the need of co-analyzing the safety and security issues of AV.

As shown in Fig. 1, analysis of functional safety and functional security of conventional road vehicles are addressed by the international standards ISO 26262 [9] and SAE J3061 [10], respectively. To the best of authors' knowledge, there are no standards or guidelines set particularly for DAS for ensuring safety and security of AVs. Generally, DAS can have either partial or full control of the AV. Different degree of control results in different degree of automation. As a result, both safety requirements and security requirements would vary accordingly. All this information should be taken into consideration during the analysis. With increasing level of driving automation, AVs should include more fail-operational mechanisms to be able to safety operate in case of failures or cyber-attacks. This should be reflected in safety and security requirements.

In this paper, a collaborative analysis of safety and security framework is proposed. By integrating safety and security engineering processes from automotive standards ISO 26262 and SAE J3061, an alignment is established between safety and security activities. Then, safety and

security integrated method (S&S) is employed to co-analyze AV functions, structure, failures (safety issues), attacks (security issues) and the associated countermeasures. The implementation of proposed framework is described, and an S&S model example is included. Researchers/engineers can clearly see the vehicle details, and the unaddressed vulnerabilities, thereby helping to improve the in-vehicle safety and security.

The rest of the paper is organized as follows: preliminary information is introduced in Section II, including driving automation system, related work, and S&S method introduction. Section III explains the proposed framework. Section IV describes the process of implementation the collaborative framework. Section V presents an implementation example of proposed framework. Finally, Section VI concludes the paper and with a glimpse of our future work.

II. PRELIMINARIES

A. DRIVING AUTOMATION SYSTEM

DAS is the hardware and software that are collectively capable of performing the entire Dynamic Driving Tasks (DDTs) on a sustained basis, which is the key property that can replace human driver in AVs [6]. The Driving Automation Levels (DALs) are classified based on the DDTs they could perform, and listed as follows:

- DAL 1: Driver Assistance, where DAS performs the longitudinal or the lateral vehicle motion control;
- DAL 2: Partial Driving Automation, where DAS performs the longitudinal and the lateral vehicle motion control;
- DAL 3: Conditional Driving Automation, where DAS also performs the Object and Event Detection and Response (OEDR);
- DAL 4: High Driving Automation, where DAS also performs DDT-fallback;
- DAL 5: Full Driving Automation, where DAS performs all the previous tasks, and is unlimited by Operational Design Domain (ODD).

ODD is a specific operating domain in which an automated function or system is designed to properly operate, including but not limited to roadway types, speed range, geography, traffic, environmental conditions (e.g., weather, daytime/nighttime), and other domain constraints [11]. For example, ODD can be designed like this: on expressway, the vehicle can hold a speed lower than 40km/h driving in the daytime only.

Fig. 2 outlines the five DDTs as well as the degree of human intervention needed by each DAL. In case of low driving automation (i.e., DAL 0, DAL 1, and DAL 2), a human driver is required to perform all/partial driving tasks. As for DAL 3, DAS can perform all DDTs but a fallback-ready user is required to control the vehicle when any DDT system failures occur or the DAS is about to leave its ODD. In case of high driving automation (i.e., DAL 4 and DAL 5), DAS alone can take full charge of the vehicle, and so human driver is not needed.

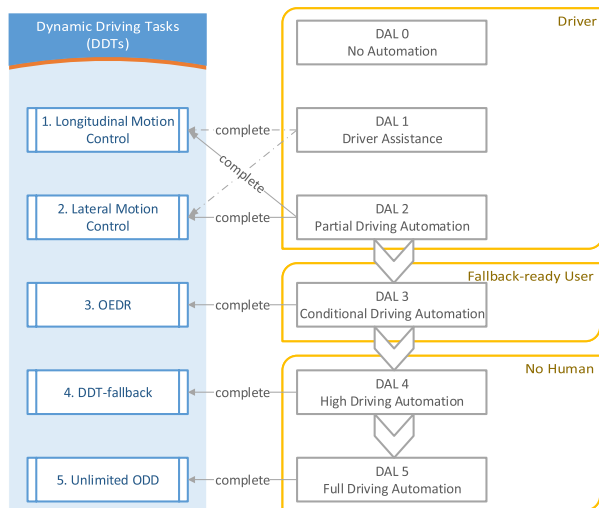


FIGURE 2. Driving automation levels' instructions.

B. RELATED WORK

Safety and security are two key properties of autonomous vehicles, and they share identical goals - protecting AVs from failing. An AV is considered safe when it is protected from accidental failures, and secure when it is unharmed by intentional attacks. Thus, safety and security co-analysis is required [12].

Standard SAE J3061 [10] proposes a way to integrate vehicle safety (ISO 26262) and security (SAE J3061) processes by establishing communication paths between safety and cybersecurity phases. Such communication can be considered as a type of method for integration. For example, SAE J3061 states that researchers should perform security threat analysis and safety hazard analysis simultaneously to ensure that no failure or attack has been missed. However, how to integrate the safety and security analysis is not proposed in this standard.

SAHARA [13] combines two well-known approaches, namely HARA [9] and STRIDE [14], to review system design in a methodical way. The safety analysis is done using HARA analysis of ISO 26262, while the security analysis is done based on the STRIDE method independently. Similar to SAHARA, US² [15] also performs safety and security co-analysis. For an attack, US² firstly quantifies its security level, and then determines if the attack introduces any safety hazards; if it does, then both security countermeasure and safety countermeasure are necessary; else, only the security countermeasure is needed.

Ponsard et al. [16] present a methodology that utilizes existing techniques, such as goal-oriented requirements engineering (GORE), to co-engineer safety and security. The approach takes results from safety and security analysis to build a goal tree connecting requirements with the related hazards/vulnerabilities where each object can be marked as safety or security relevant. The analysis of safety and security requirements is performed jointly, although the input to

this technique from hazard/threat identification activities may come from different sources.

STPA-SafeSec is presented in [17], which is based on STPA [18] and STPA-Sec [19], and used to choose the most effective mitigation strategies to ensure system safety and security. The strength of the approach is unified safety and security consideration while choosing suitable mitigation strategies, a possibility to prioritize the most critical system components for an in-depth security analysis (e.g., penetration testing). The analysis identifies potential system losses, caused by a specific security or safety vulnerability, and better mitigation strategies.

There are also lots of safety and security co-analysis methods [20], however they mainly focus on the risk analysis part and do not take driving automation level into consideration. Depending on DAL, an AV may vary in terms of functions, structure and vulnerabilities. In this paper, our collaborative framework is applicable to AV regardless of its DAL.

III. COLLABORATIVE ANALYSIS FRAMEWORK

A. ALIGNMENT OF AV SAFETY AND SECURITY

As any failures or attacks may lead to safety losses (as seen in Fig. 1), integration of safety and security is crucial for AVs. To save time and cost, the integration has to be considered in early development phase. In this section, we describe how to align the DAS with the safety and security standards.

SAE J3061 [10] is a cyber security guidebook for vehicle systems, which defines lifecycle process framework and provides guiding principles. In SAE J3061, the cyber security lifecycle can be divided into several phases: concept phase, product development phase (system level, hardware level and software level), production and operation phase. Concept phase is the first step for the whole lifecycle, which include the following activities: feature definition, Threat Analysis and Risk Assessment (TARA), functional security concept, security requirements, and security assessment. Feature definition describes the system being developed to which the cyber security process will be applied, i.e., it defines the boundary of the features. TARA identifies threats and assesses the risk, and the result of TARA drives all downstream activities. Security concept describes the high-level strategy for obtaining security from TARA phase, and once the concept is determined for satisfying the feature, the security requirement can be determined. Attack tree [21] is a popular method used for TARA; the processes of attacks are summarized into a graph comprising nodes (representing attack events), edges (denoting path of attacks through the system), and gates (e.g. logic AND and OR gates). Security assessment is performed to identify the current security posture of the cyber physical vehicle, and it is developed throughout the security lifecycle.

ISO 26262 [9] is an international standard for functional safety of E/E systems in production automobiles, defined by the International Organization for Standardization. It provides an automotive safety lifecycle (includes management,

development, production, operation, service, decommissioning) and supports tailoring the necessary activities during these phases. In the development part, the safety process is composed of item definition, Hazard Analysis and Risk Assessment (HARA), functional safety concept, safety requirement and safety assessment. Fault tree analysis [22] is often used for HARA. Fault trees are similar to attack trees, where the tree nodes represent failure events.

As ISO 26262 and SAE J3061 are not developed specifically for AVs, they do not take driving automation into consideration. In AVs, we need to consider safety, security, and DAS simultaneously.

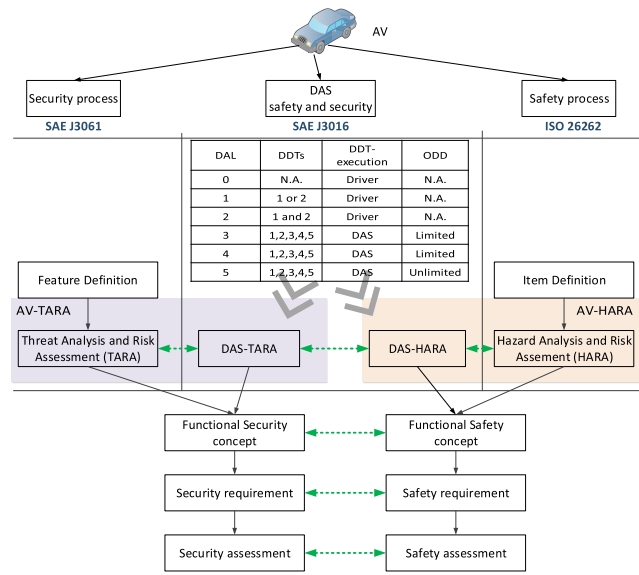


FIGURE 3. Concept phase alignment of ISO 26262, SAE J3061 and SAE J3016 standards. "N.A." denotes not applicable.

Fig. 3 depicts the proposed alignment of safety and security standards for AVs, where dotted line with arrowheads at both ends denotes simultaneous activities. Due to the automation levels of DAS, TARA and HARA should correspond with each DAL, i.e., TARA and HARA must consider particular properties of each DAL. The table shown in Fig. 3 shows the differences between the six DALs in terms of their DDTs, execution of all DDTs, and ODD (described in Section II). All the properties should be analyzed in DAS-TARA and DAS-HARA activities. After completion of AV-TARA, security concept phase is performed, which integrates the results of AV-TARA, followed by security requirement, and security assessment. In parallel, functional safety concept is performed after completion of AV-HARA, followed by safety requirement and safety assessment, as shown in Fig. 3. The results of each activity on the security side and the corresponding activity on the safety side have to be co-analyzed to assure their completeness and consistency, as shown in Fig. 3.

B. SAFETY AND SECURITY INTEGRATED METHOD (S&S)

S&S is an integrated safety and security analysis method. As the result of applying the method, an S&S model is

created, which incorporates six hierarchies of a system (functions, structure, failures, attacks, safety countermeasures, and security countermeasures), connected by relationship matrices. S&S model is an extension of the Six-Step Model, proposed in our earlier work [23]. S&S refines the relationship matrices, which analyze relationship between two elements, like elements' connection, countermeasure coverage, and interdependence between countermeasures. The structure of S&S model is depicted in Fig. 4. The sequential steps used to construct the S&S are as follows:

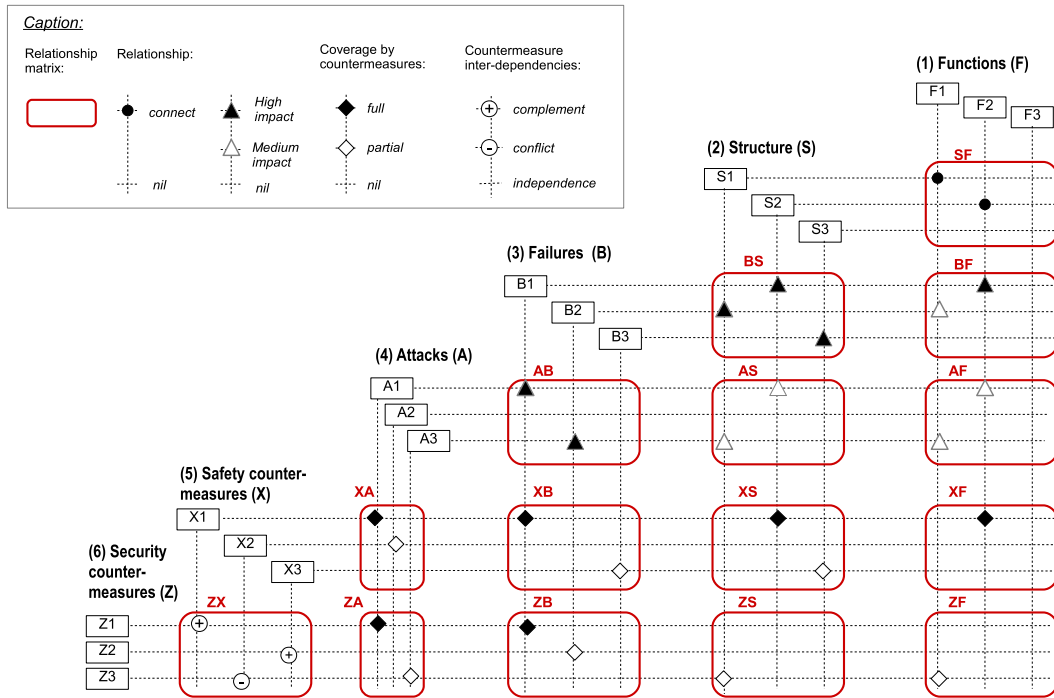
1. The construction of S&S model begins with identifying the functions of the system. \mathcal{F} is used to denote the function in Fig. 4.
2. The components that form the system's hierarchical structure are to be defined and appended to the model, marked as \mathcal{S} . Relationship matrix \mathcal{SF} defines the relationships between structure's component and the function components, which determines the connection between structure and function, i.e., whether the structure is used to implement the associated function. Matrix \mathcal{SF} can be obtained from the system directly, and the elements in \mathcal{SF} is 0 or 1, where 0 means no relationship, and 1 means the connection existing (the black circle is used to show the connection, as shown in Fig. 4).
3. The failures of the system (marked as \mathcal{B}) are to be identified and appended to the model. Matrix \mathcal{BS} defines the impact between failures and structure, which means whether the failure affects the corresponding structural component. For example, GPS failure impacts the GPS; camera failure impacts the camera. The relationship between failures and functions is demonstrated by matrix \mathcal{BF} , which can be obtained by the product of matrix \mathcal{BS} and \mathcal{SF} :

$$\mathcal{BF} = \mathcal{BS} \cdot \mathcal{SF} \quad (1)$$

To quantify the impact degree (marked as D_{impact}) and to correspond the high impact (marked as black triangle), medium impact (marked as white triangle) in Fig. 4, we define: in the relationship matrix, assuming the maximum value is $a_{u,v}$, and the minimum value is $a'_{u',v'}$, for any element $a_{i,j}$,

$$D_{\text{impact}}(i, j) = \begin{cases} \text{high} & \text{if } \frac{a_{i,j} - a'_{u',v'}}{a_{u,v} - a'_{u',v'}} \geq 0.5 \\ \text{medium} & \text{if } 0 < \frac{a_{i,j} - a'_{u',v'}}{a_{u,v} - a'_{u',v'}} < 0.5 \\ \text{nil} & \text{if } a_{i,j} - a'_{u',v'} = 0 \end{cases} \quad (2)$$

4. Attacks that lead to system's failure are to be identified and appended to the model, marked as \mathcal{A} . Relationship matrix \mathcal{AB} (Attacks - Failures) is used to determine which failures could be triggered by a successful attack. Matrix \mathcal{AS} (relationship of Attacks impact on Structure) and \mathcal{AF} (relationship of Attacks impact on Functions)



defined in [24]. However, in some situations, it might be difficult to distinguish between reinforcement and conditional dependency. Thus in S&S model, we consider following three types: complement (one countermeasure complement or support another), conflict (one countermeasure conflict or diminish another), independence (two countermeasures are mutually independent).

The matrices in S&S demonstrate the relationships between the six hierarchies, which will help to ensure consistency between these hierarchies. The hierarchies and relationships have to be maintained and updated throughout the entire development phase to sustain the consistency.

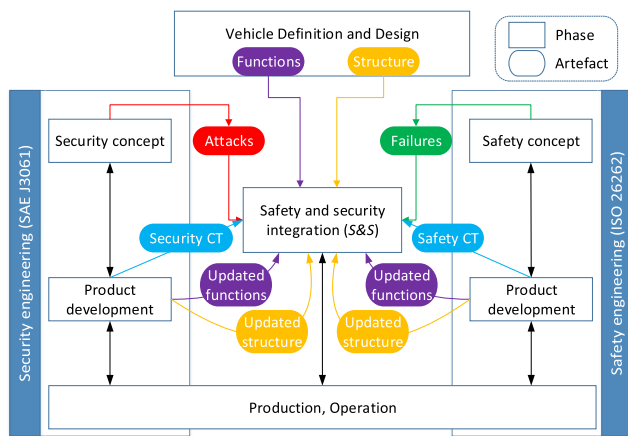


FIGURE 5. Collaborative analysis framework for autonomous vehicles. (CT denotes countermeasures).

C. COLLABORATIVE ANALYSIS FRAMEWORK

S&S method is used to combine artefacts from the safety and security processes of AVs in the framework. The collaborative framework is shown in Fig. 5; the rectangles denote the phases, and the rounded rectangles represent the artefacts (i.e., output from the associated phase). There are several phases in the framework, namely vehicle definition and design, safety/security concept, safety/security product development, integration and production and operation. The framework starts with vehicle definition and design, where AV functions and structure are defined. Functions and structure are important information for analyzing safety and security. When someone attacks certain component of AV (e.g., certain sensor, or certain ECU), knowing the structure and functions can be helpful to foresee the possible consequences and design the mitigation approach more efficiently.

The safety concept and security concept phases come from standards ISO 26262 and SAE J3061 respectively, as detailed in Fig. 3. During the AV-HARA and AV-TARA activities in the concept phase, AV's failures and attacks are analyzed. Subsequently, both safety countermeasures and security countermeasures are designed during development phase, to serve as detection and mitigation approaches. Thus, all the information, needed for constructing the S&S model and performing safety and security integration, can be obtained

from the concept and product development phases. During the product development phase, some of the initial AV functions and structure may be updated, which are also added to the S&S. The S&S assures the completeness and consistency of the AV safety and security countermeasures.

IV. IMPLEMENTATION OF COLLABORATIVE ANALYSIS FRAMEWORK

In this section, we describe the process of framework implementation.

A. AV FUNCTIONS

Fig. 6 shows the common functions of AVs. Autonomous driving includes three main functions: perception, decision & control, and vehicle platform manipulation [25], [26].

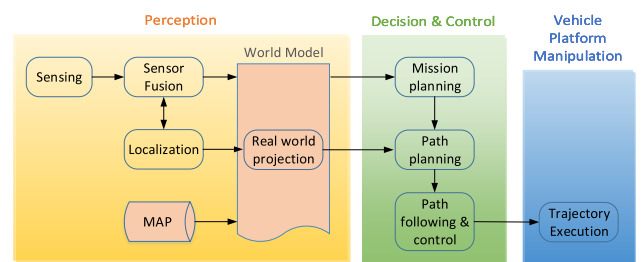


FIGURE 6. AV functions example [27].

In perception, sensors' measurements are collected by the sensing functional component, then sensor fusion component considers information of multiple sensors to construct a hypothesis about the state of the environment. In addition, to establish the confidence values for state variables, the sensing component may perform sensor data calibration. The localization component is responsible for determining the location of the vehicle with respect to its surroundings, with required accuracy. It may also aid the sensor fusion component to perform a task known as map matching, wherein physical locations of detected objects are referenced to the map's coordinate system. A map is usually pre-loaded to the on-board computer of AV. Once the AV is localized, the real-world projection component will be performed, which identifies the state of the external (and possibly, internal) environment, as perceived by the vehicle. Real-world projection includes object and traffic light detection, which may incorporate kinematic and dynamic models of the objects.

In decision & control part, the mission planning component repeatedly generates obstacle free trajectories in the world coordinate system and picks the optimal trajectory. The path planning considers the results of mission planning, localization and real-world projection to design the waypoints (waypoint is a data string with coordinate, direction and velocity information) at every control cycle. The path following and control component is responsible for transforming the path waypoints into control commands, which are then sent to vehicle platform manipulation to perform trajectory execution. This is achieved by a combination of

acceleration (e.g., powertrain and steering) and deceleration (e.g., braking). In case of manual driving, the driver has to handle partial perception, make all decisions based on his own judgments, and manipulate the vehicle accordingly.

In addition to the aforementioned AV functions, there are several functions needed for V2X communication [27], such as time synchronization and wireless communication. Time synchronization is needed because the data packets exchanged among V2X need to be timestamped. The timestamps should be synchronized across all participants in V2X (including vehicles, infrastructures and so on). Typically, a common clock source is needed to be a reference clock for all the system clocks. Periodic synchronization with the clock source is necessary due to the inevitable drift in clock mechanisms of any electronic device. Wireless communication is a basic requirement for V2X, which enables AV to communicate with other AVs, infrastructures, etc.

As mentioned in Sec. II, different DALs require different DDTs and hence different functions. In the case of DAL 4/5-enabled AV, all the functions mentioned in Fig. 6 should be included in the framework. The lower the level of driving automation, the more the tasks handled by a human driver, and hence the lesser the functions to be considered during the analysis.

B. AV STRUCTURE

In this paper, we consider a vehicle, capable to perform at any driving automation levels (as described in Sec. II), which includes both autonomous and manual driving systems. An example of such a vehicle is the ZMP Robocar [26]. The structure of AV includes Driving Automation System (i.e., DAS), manual driving system, and supporting systems, as shown in Fig. 7. DAS is in charge of autonomous driving, and can be decomposed into three systems, as shown in Fig. 7:

- Cognitive Driving Intelligence system (CDI system)
- Vehicle Platform Manipulation system (VPM system)
- Communication system

CDI system performs perception (perception of the external environment/context in which vehicle operates), and decision & control (decisions and control of vehicle motion with respect to the external environment/context that is perceived) functions [28]. It includes an on-board computer and several external sensors (e.g., LiDAR, GPS, camera), as shown in Fig. 7. VPM system deals mostly with sensing, control and actuation of the vehicle in order to achieve the desired motion. It includes ECUs, actuators (e.g., steering, and brake motors), and internal sensors (e.g., wheel encoders). Communication system enables communication between DAS elements. It can be further broken down into in-vehicle and V2X (vehicle to vehicle, infrastructure, and humans) communication systems (see Fig. 7). Regarding the in-vehicle communication, several technologies have been considered, such as CAN Bus, Ethernet, USB and so on.

Manual Driving System is the system that enables the driver to manually control the vehicle (manual driving mode).

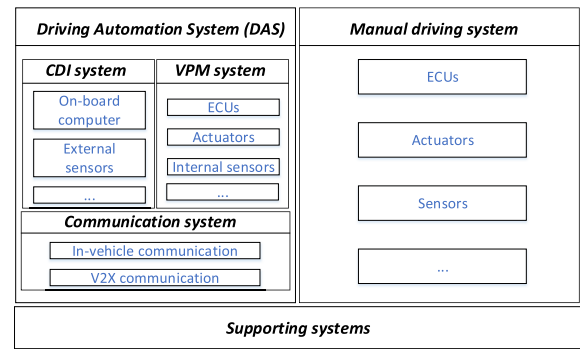


FIGURE 7. A reference of AV structure model.

It includes ECUs, sensors, actuators, and so on, as shown in Fig. 7. Note that the manual driving system may not be included in highly-automated AVs (AVs at DAL 4 or 5), since their DAS can perform all driving tasks and totally replace the manual driving system.

Finally, supporting systems are the systems, which support both driving modes (i.e., automated and manual), such as airbag and battery control systems.

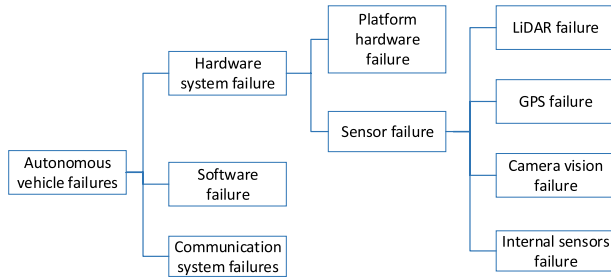
C. AV FAILURES AND RELATED SAFETY COUNTERMEASURES

Several failures related to AV and road surroundings are analyzed in [29], [30], as shown in Fig. 8 (a) and (b). AV related failures include hardware system failures, software failures and communication system failures. Hardware system failures include platform hardware failures and sensor failures, such as LiDAR failure, GPS failure, camera vision failure and internal sensors failure (e.g., wheel encoder failure). Road surroundings related failures include other road user (e.g., cyclist, pedestrian, other vehicles), weather impact, road conditions (e.g., improper lane marking, improper pavement conditions), construction zones, and traffic signals and signs (e.g., signal failure, sign failure). In this paper, as we focus on the AV itself safety and security co-analysis, only the AV-related failures are considered, i.e., failures shown in Fig. 8 (a).

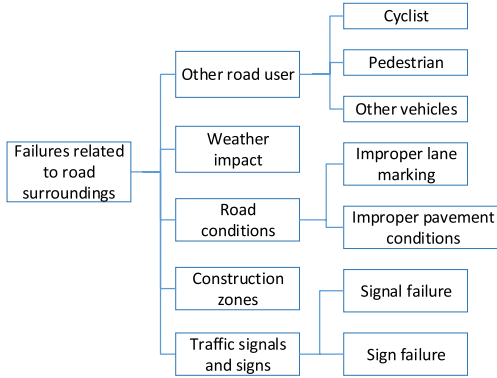
Regular inspection has been a common safety countermeasure [31]. Besides, multi-sensor fusion [32] could potentially be a countermeasure for sensor failure, when multiple sensors are collectively considered/fused to give a more reliable estimation; a faulty sensor could be complemented by other redundant sensors. Furthermore, fault detection and fault-tolerant control methods could be implemented in the on-board computer.

D. POTENTIAL ATTACKS ON AVS AND ASSOCIATED SECURITY COUNTERMEASURES

Attacks on AVs can be either physical or cyber, as shown in Fig. 1. The two main types of cyber-attacks are: deception attacks (e.g., spoofing, replay, and measurement substitution) and denial of service (DoS) attacks (e.g., jamming, network



(a) Failures related to AV.

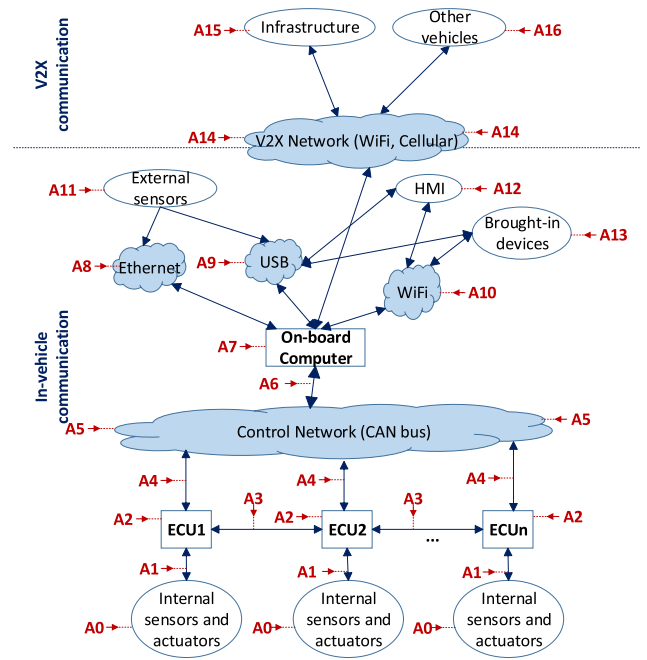


(b) Failures related to road surroundings.

FIGURE 8. Failures for autonomous vehicle.

flooding and increased communication latency) [33], [34]. Deception attack is a major challenge of AV security, where an attacker uses unauthenticated entity (e.g., ECUs, actuators and sensors) or fake information to deceive other entities. If such attack succeeds, AV might perform inappropriately or maliciously, thereby endangering not only the AV itself but also nearby road users and infrastructure. Therefore, to mitigate such attack, authentication of all existing entities in the AV should be performed before the access to available services is granted. DoS attack is highly related to the availability of information. For example, if jamming is successful, real-time information will be delayed, which will undoubtedly affect AV's performance. An AV should ensure that all the in-vehicle entities are functional, and useful information is available when needed.

Fig. 9 shows the potential attacks on AV [30], where several types of attacks are identified: A0 - direct physical attacks on internal sensors and actuators; A1 - deception and DoS attacks on internal sensor measurements and control actuation; A2 - direct physical attacks on ECUs; A3 - deception and DoS attacks on inter-ECU communication; A4 - deception and DoS attacks on CAN bus communication with ECUs; A5 - direct attacks on CAN bus, e.g., through diagnostic port; A6 - deception and DoS attacks on CAN bus communication with on-board computer; A7 - direct physical attacks on on-board computer; A8 - attacks on Ethernet; A9 - attacks on USB; A10 - attacks on WiFi; A11 - direct physical attacks on external sensors; A12 - direct physical attacks on Human Machine Interface (HMI); A13 - direct physical attacks on brought-in devices; A14 - attacks on V2X network;

**FIGURE 9. Potential attacks on AV.**

A15 - attacks on infrastructure; A16 - attacks on other vehicles.

There are numerous attack detection and mitigation techniques that can be used as security countermeasures in AVs. To mitigate deception attacks, authentication schemes are usually employed [35]. Authentication is an integral part of trust establishment between entities in AVs. With proper authentication schemes one can easily identify non-legitimate entities and fake messages, thereby providing security for autonomous vehicles. In order to avoid DoS attacks, cryptographic solutions are mainly used [36]. Cryptography uses methods like encryption/decryption algorithms and digital signatures (e.g., bit commitment and signature based mechanisms [37]) to provide confidential communication between legitimate entities, thereby securing the availability. Furthermore, various anomaly detection and mitigation methods can be implemented in the on-board computer.

V. FRAMEWORK IMPLEMENTATION EXAMPLE

The collaborative framework can be implemented in two stages:

Stage 1: The S&S model construction. In this stage, information from various phases of safety and security lifecycle on AV functions, structure, failures, attacks, and countermeasures is collected and added to the S&S model. Furthermore, the relationships among them are calculated and added to the relationship matrices.

Stage 2: In-depth safety and security analysis. In the second stage, further analysis of safety and security is performed. S&S model can be seen as a database of AV safety and security artefacts and their relationships. Thus, we extract the information from it related to particular artefacts and perform

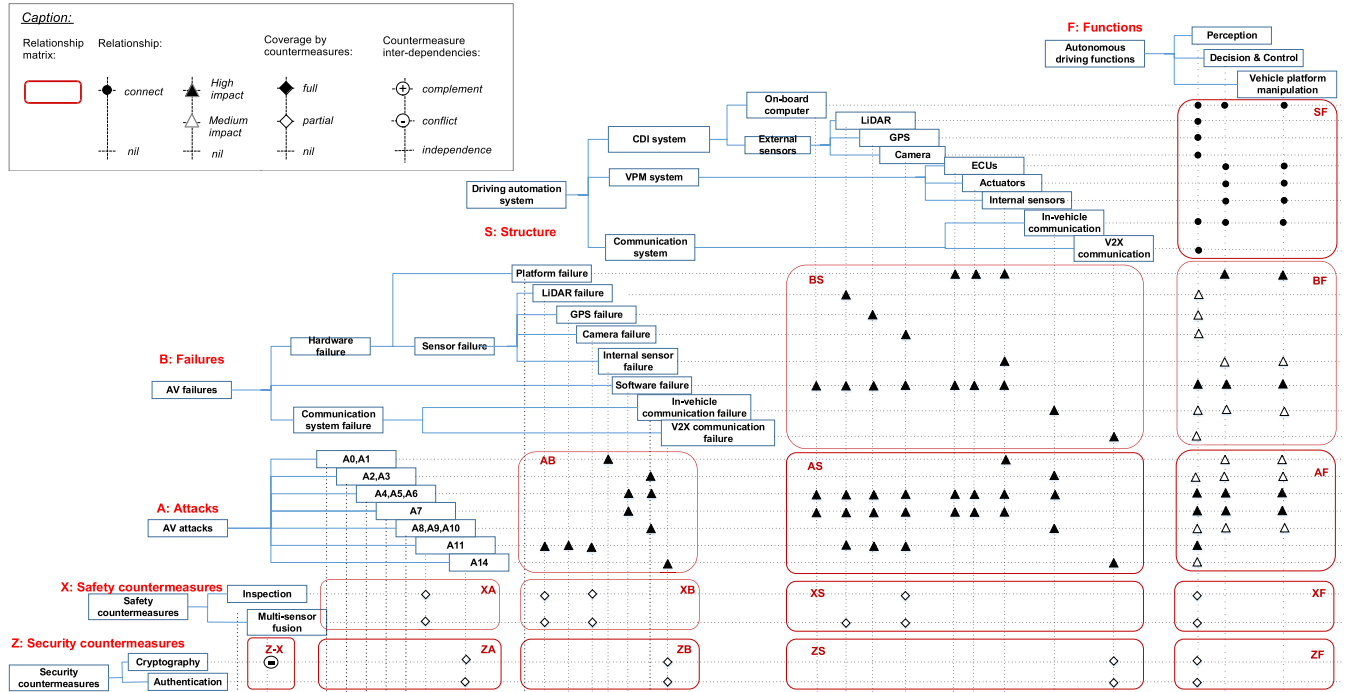


FIGURE 10. S&S model example.

an in-depth analysis of different aspects of safety and security. Then, the S&S model is updated based on the results of this analysis.

A. IMPLEMENTATION OF STAGE 1 – S&S MODEL CONSTRUCTION

Fig. 10 shows an example of S&S model of the autonomous vehicle, which includes AV Functions \mathcal{F} , Structure \mathcal{S} , Failures \mathcal{B} , Safety countermeasures \mathcal{X} , Attacks \mathcal{A} , and Security countermeasures \mathcal{Z} , and their relationship matrices as described in Section IV. As we can see in Fig. 10, \mathcal{SF} illustrates the relationship between structure and functions, for example, on-board computer is used to perform all the autonomous driving functions; external sensors (e.g., LiDAR, GPS, and Camera) as well as the in-vehicle communication (e.g., WiFi, Ethernet, USB, and CAN Bus) are connected to perception; the sub-structures of vehicle platform (e.g., ECUs, actuators and internal sensor) are used to perform decision & control and vehicle platform manipulation; V2X communication is connected to perception, since the information obtained from other vehicles or infrastructure via V2X is helpful for the vehicle to achieve precise perception [28].

\mathcal{BS} (Failures - Structure) can be obtained directly from the previous analysis (as shown in Section IV-B and IV-C). For example, platform failure impacts the vehicle platform, such as ECUs, actuators, internal sensors as shown in Fig. 10. GPS failure impacts the GPS. Software failure impacts the related software, which includes the software of on board computer, LiDAR, GPS, internal sensor etc. When the \mathcal{BS} is

analyzed, we can calculate the relationship between Failures and Functions (\mathcal{BF}) via Equation 1, and the degree of relationship is also given by Equation 2. Similarly, when matrix \mathcal{AB} (Attacks - Failures) is analyzed, \mathcal{AS} (Attacks - Structure) and \mathcal{AF} (Attacks - Functions) are computed by Equation 3 and Equation 4 respectively.

As we can see from matrices \mathcal{XB} and \mathcal{ZA} in Fig. 10, countermeasures cover certain failures or attacks. For example, **inspection**, a safety countermeasure, can partially cover LiDAR failure and camera failure (to inspect whether the devices are physical complete). **Multi-sensor fusion**, another safety countermeasure, can partially cover one-type sensor failures like LiDAR failure and camera failure. The security countermeasures (i.e., **cryptography** and **authentication** approaches) partially mitigate the attack on V2X network. The matrices \mathcal{XA} (coverage of Safety countermeasures on Attacks), \mathcal{XS} (coverage of Safety countermeasures on Structure) and \mathcal{XF} (coverage of Safety countermeasures on Functions) are obtained by Equation 6, 7 and 8 respectively. The corresponding coverage degrees are got by Equation 5. Note here, we use γ ($\gamma \ll 1$) to represent the partial coverage (white rhombus in Fig. 10) to compute the matrix. Similarly, matrices \mathcal{ZB} (Security countermeasures - Failures), \mathcal{ZS} (Security countermeasures - Structures) and \mathcal{ZF} (Security countermeasures - Functions) are calculated by Equation 9, 10 and 11.

The interdependency between safety and security countermeasure is illustrated by matrix \mathcal{ZX} in Fig. 10: there is a ‘conflict’ relationship between **cryptography** and **multi-sensor fusion**, which has to be further analyzed. To enable

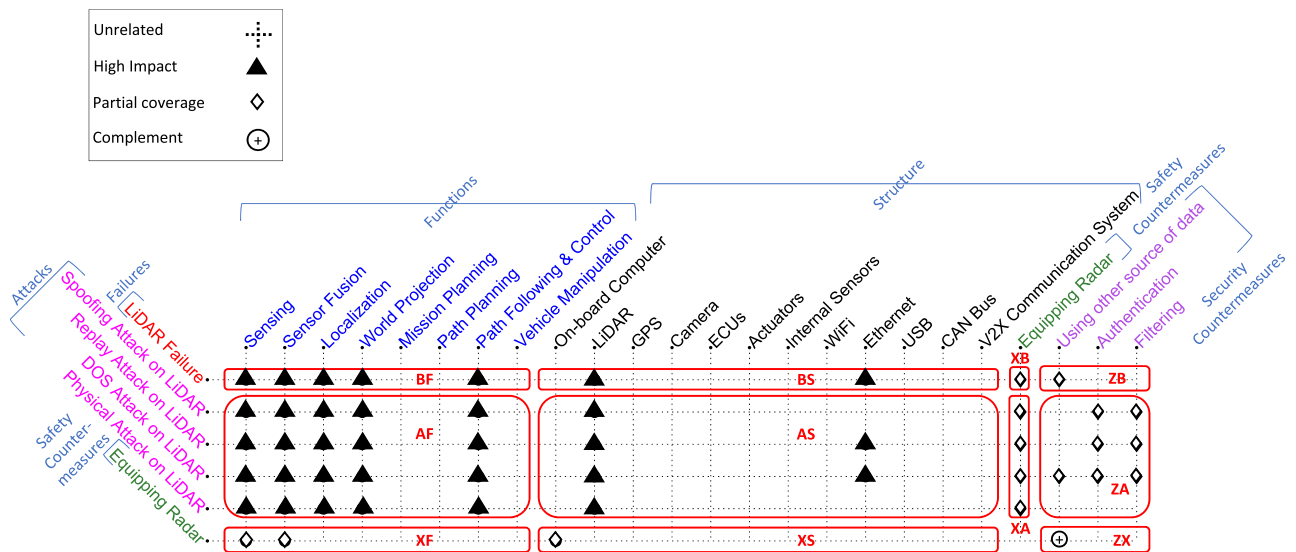


FIGURE 11. Relationships on LiDAR failures and attacks with AV functions, structure, and safety and security countermeasures.

multi-sensor fusion, the access to multiple types of sensor data must be granted; the access, however, might be restricted owing to the cryptography solution (which aims to keep the data private), thereby affecting the efficiency/effectiveness of multi-sensor fusion.

B. IMPLEMENTATION OF STAGE 2 – IN-DEPTH ANALYSIS

Using the information recorded in S&S model, in-depth safety and security analysis can be performed. A tool has been developed using Matlab, which outputs graphical representation illustrating the relationships between artefacts (entities), selected for further analysis.

The following example includes in-depth analysis of LiDAR attacks and failures. LiDAR is a sensing system that uses rotating laser sensor to measure and map the surroundings into 3D [38]. For AVs, sensing and real world projection are all basic requirements for autonomous driving. Thus, LiDAR is widely used in AVs. The LiDAR failure or LiDAR attack will introduce perception fault, thereby affecting the autonomous driving.

Failures of LiDAR can be categorized into two groups: hardware failure (e.g., laser sensor fails) and software failure (e.g., algorithm on LiDAR fails). The safety countermeasures can be using additional LiDAR, or using Radar in addition to LiDAR. As discussed in [39], no single type of sensors works well for all driving functions and in all conditions, thus additional/redundant sensors/data would lead to more reliable estimation. [39] also deem that the fusion of data from both LiDAR and Radar could result in better object detection, distance estimation, etc.

Attacks on LiDAR comprise spoofing attack, replay attack, DoS attack, and physical attack. The associated security countermeasures include filtering, using other source of data [33] and authentication.

In order to further analyze possible LiDAR failures and attacks, we can extract the related information from the S&S model. Fig. 11 shows the relationships on LiDAR failures and attacks with AV functions, structure, and safety and security countermeasures. As LiDAR data is crucial for AV's perception as well as decision & control, any attacks on LiDAR or a faulty LiDAR would affect the AV functions considerably. For example, LiDAR data is used for sensing the surroundings, multi-sensor fusion, localization, world projection, and path following & control. So LiDAR failure/attacks affect all the above functions, as shown in Fig. 11. Besides, LiDAR, as a component of structure, is directly affected by failures or attacks, thus 'impact' (marked by black circle) is shown between failures, attacks and LiDAR. Moreover, replay attack and DoS attack can be executed via Ethernet (in ZMP vehicle, LiDAR and on-board computer is connected via Ethernet [26]), thus these two attacks also impact Ethernet.

Additional LiDAR is a safety solution for faulty or attacked LiDAR, which can fully cover the failure of original LiDAR. However, due to the high cost, **adding Radar** is considered an alternative safety countermeasure. Radar can be employed instead (along with camera) to enable multi-sensor fusion [40], which can partially cover LiDAR failures and attacks. **Using other source of data** (e.g., Radar or Camera) to recognize the surroundings can help to partially cover LiDAR failure and DoS attack on LiDAR. Besides, **filtering** data can partially cover both spoofing and replay attacks on LiDAR, by extracting the right and legitimate information from the distorted data, and can mitigate DoS attack partially. In addition, **authentication** is also a common security countermeasure that can partially cover spoofing, replay and DoS attacks on LiDAR.

Interdependence between safety countermeasure and security countermeasure is also illustrated in Fig. 11.

Adding Radar complements the security countermeasure of Using other source of data, and is independent of Filtering data and Authentication.

At the end of this stage, the S&S model is updated with the detailed failures, attacks, countermeasures, and relationships.

VI. CONCLUSION AND FUTURE WORK

The complex interactions between the cyber and physical components inside the AV introduce more potential safety- and security-related vulnerabilities. Identifying all potential vulnerabilities and applying appropriate countermeasures are challenges for researchers and engineers. In view of that, a collaborative analysis framework of safety and security is proposed in this paper. Combining safety engineering (ISO 26262) and security engineering (SAE J3061) processes, the framework analyzes AV functions, structure, failures, attacks, and the associated countermeasures simultaneously. An example is included to demonstrate the usefulness of the proposed framework based on a typical AV model. This framework can help researcher/engineer to address the safety failures and security attacks more intuitively, and to select appropriate safety and security countermeasures.

In the future, we will analyze the AV considering high-risk failures and attacks (e.g., attacks from V2X communication) using the framework. Moreover, we will implement the selected countermeasures on AV prototype, to test and validate the relationships between failures, attacks and countermeasures, and evaluate the performance of countermeasures.

REFERENCES

- [1] "2016 fatal motor vehicle crashes: Overview," Nat. Center Statist. Anal., U.S. Dept. Transp., Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS 812 456, 2017. [Online]. Available: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812456>
- [2] L. J. Blincoc, T. R. Miller, E. Zaloshnja, and B. A. Lawrence, "The economic and societal impact of motor vehicle crashes, 2010 (revised)," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS 812 013, May 2015. [Online]. Available: <http://www-nrd.nhtsa.dog.gov/pubs/812013.pdf>
- [3] "Traffic safety facts, a brief statistical summary: Critical reasons for crashes investigated in the national motor vehicle crash causation survey," Nat. Highway Traffic Saf. Admin., Nat. Center Statist. Anal., U.S. Dept. Transp., Washington, DC, USA, Tech. Rep. DOT HS 812 115, 2016. [Online]. Available: <http://www-nrd.nhtsa.dog.gov/pubs/812115.pdf>
- [4] J. M. Anderson, K. Nidhi, K. D. Stanley, P. Sorensen, C. Samaras, and O. A. Oluwatola, *Autonomous Vehicle Technology: A Guide for Policymakers*. Santa Monica, CA, USA: Rand Corporation, 2014. [Online]. Available: https://www.rand.org/pubs/research_reports/RR443-2.html
- [5] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations," *Transp. Res. A, Policy Pract.*, vol. 77, pp. 167–181, Jul. 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0965856415000804>
- [6] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Standard SAE-J3016, Society of Automotive Engineers, Sep 2016.
- [7] The Guardian. *Self-Driving Uber Kills Arizona Woman in First Fatal Crash Involving Pedestrian*. 2018. Accessed: Jun. 21, 2018. [Online]. Available: <https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe>
- [8] J. Cui and G. Sabaliauskaite, "On the alignment of safety and security for autonomous vehicles," in *Proc. IARIA CYBER*, Barcelona, Spain, Nov. 2017, pp. 1–6.
- [9] *Road Vehicles—Functional Safety*, Standard ISO-26262, Dec. 2016.
- [10] *Cybersecurity Guidebook for Cyber-Physical Vehicle Systems*, Standard SAE-J3061, Jan. 2016.
- [11] National Highway Traffic Safety Administration. (Sep. 2016). *Federal Automated Vehicles Policy-Accelerating the Next Revolution in Roadway Safety*. [Online]. Available: <https://www.transportation.gov/AV/federal-automated-vehicles-policy-september-2016>
- [12] E. Schoitsch, C. Schmittner, Z. Ma, and T. Gruber, "The need for safety and cyber-security co-engineering and standardization for highly automated automotive vehicles," in *Advanced Microsystems for Automotive Applications*. Cham, Switzerland: Springer, 2016, pp. 251–261.
- [13] G. Macher, H. Sporer, R. Berlach, E. Armengaud, and C. Kreiner, "SAHARA: A security-aware hazard and risk analysis method," in *Proc. IEEE DATE*, Grenoble, France, Mar. 2015, pp. 621–624.
- [14] *The Stride Threat Model*, Microsoft Corp., Redmond, WA, USA, 2005.
- [15] J. Cui and G. Sabaliauskaite, "US²: An unified safety and security analysis method for autonomous vehicles," in *Proc. Future Inf. Commun. Conf.* Singapore: Springer, Apr. 2018, pp. 600–611.
- [16] C. Ponsard, G. Dallons, and P. Massonet, "Goal-oriented co-engineering of security and safety requirements in cyber-physical systems," in *Proc. Int. Conf. Comput. Saf., Rel., Secur.* Trondheim, Norway: Springer, Mar. 2016, pp. 334–345.
- [17] I. Friedberg, K. McLaughlin, P. Smith, D. Lavery, and S. Sezer, "STPA-SafeSec: Safety and security analysis for cyber-physical systems," *J. Inf. Secur. Appl.*, vol. 34, pp. 183–196, Jun. 2017.
- [18] N. G. Leveson, *STPA: A New Hazard Analysis Technique*. Cambridge, MA, USA: MIT Press, 2012, pp. 211–249.
- [19] W. Young and N. Leveson, "Systems thinking for safety and security," in *Proc. 29th Annu. Comput. Secur. Appl. Conf.*, 2013, pp. 1–8.
- [20] E. Lisova, I. Šljivo, and A. Čaušević, "Safety and security co-analyses: A systematic literature review," *IEEE Syst. J.*, vol. 13, no. 3, pp. 2189–2200, Sep. 2019.
- [21] B. Schneier, "Attack trees," in *Secrets and Lies: Digital Security in a Networked World*. Hoboken, NJ, USA: Wiley, Oct. 2015, ch. 21, pp. 318–333.
- [22] E. Ruijters and M. Stoelinga, "Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools," *Comput. Sci. Rev.*, vol. 15, pp. 29–62, Feb./May 2015.
- [23] G. Sabaliauskaite, S. Adep, and A. Mathur, "A six-step model for safety and security analysis of cyber-physical systems," in *Critical Information Infrastructures Security*. Cham, Switzerland: Springer, 2017, pp. 189–200. doi: 10.1007/978-3-319-71368-7_16.
- [24] L. Piètre-Cambacède and M. Bouissou, "Cross-fertilization between safety and security engineering," *Rel. Eng. Syst. Saf.*, vol. 110, pp. 110–126, Feb. 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0951832012001913>
- [25] J. Becker, M. Helmle, and O. Pink, "System architecture and safety requirements for automated driving," in *Automated Driving*. Cham, Switzerland: Springer, 2017, pp. 265–283. doi: 10.1007/978-3-319-31895-0_11.
- [26] S. Kato, E. Takeuchi, Y. Ishiguro, Y. Ninomiya, K. Takeda, and T. Hamada, "An open approach to autonomous vehicles," *IEEE Micro*, vol. 35, no. 6, pp. 60–68, Nov./Dec. 2015.
- [27] S. Behere, M. Törngren, and D.-J. Chen, "A reference architecture for cooperative driving," *J. Syst. Archit.*, vol. 59, no. 10, pp. 1095–1112, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1383762113000957>
- [28] S. Behere and M. Törngren, "A functional reference architecture for autonomous driving," *Inf. Softw. Technol.*, vol. 73, pp. 136–150, May 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950584915002177>
- [29] P. Bhavsar, P. Das, M. Paugh, K. Dey, and M. Chowdhury, "Risk analysis of autonomous vehicles in mixed traffic streams," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2625, pp. 51–61, Jan. 2017.
- [30] J. Cui, L. S. Liew, G. Sabaliauskaite, and F. Zhou, "A review on safety failures, security attacks, and available countermeasures for autonomous vehicles," *Ad Hoc Netw.*, vol. 90, Jul. 2019, Art. no. 101823.
- [31] J. Cui and G. Sabaliauskaite, "US²: An unified safety and security analysis method for autonomous vehicles," in *Proc. FICC*, Singapore, Apr. 2018, pp. 600–611.
- [32] K. Chitnis, M. Mody, P. Swami, R. Sivaraj, C. Ghone, M. G. Biju, B. Narayanan, Y. Dutt, and A. Dubey, "Enabling functional safety ASIL compliance for autonomous driving software systems," *Auton. Vehicles Mach.*, vol. 2017, pp. 35–40, Jan. 2017.

- [33] J. Petit and S. E. Shladover, "Potential cyberattacks on automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 546–556, Apr. 2015.
- [34] R. G. Engoulou, M. Bellaïche, S. Pierre, and A. Quintero, "VANET security surveys," *Comput. Commun.*, vol. 44, pp. 1–13, May 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0140366414000863>
- [35] S. S. Manvi and S. Tangade, "A survey on authentication schemes in VANETs for secured communication," *Veh. Commun.*, vol. 9, pp. 19–30, Jul. 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2214209616300018>
- [36] M. N. Mejri, J. Ben-Othman, and M. Hamdi, "Survey on VANET security challenges and possible cryptographic solutions," *Veh. Commun.*, vol. 1, no. 2, pp. 53–66, Apr. 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2214209614000187>
- [37] A. M. Malla and R. K. Sahu, "Security attacks with an effective solution for dos attacks in VANET," *Int. J. Comput. Appl.*, vol. 66, no. 22, pp. 1–5, 2013.
- [38] B. Schwarz, "LIDAR: Mapping the world in 3D," *Nature Photon.*, vol. 4, no. 7, pp. 429–430, 2010.
- [39] Woodside Capital Partners. (2016). *Beyond the Headlights: Adas and Autonomous Sensing*. [Online]. Available: http://woodsidecap.com/wp-content/uploads/2016/12/20160927-Auto-Vision-Systems-Report_FINAL.pdf
- [40] S. Tuohy, M. Glavin, C. Hughes, E. Jones, M. Trivedi, and L. Kilmartin, "Intra-vehicle networks: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 534–545, Apr. 2015.



JIN CUI received the Ph.D. degree in computer science from the Institut National des Sciences Appliquées de Lyon, France, in 2016. She was a Postdoctoral Research Fellow with the Singapore University of Technology and Design, from 2017 to 2019. She is currently a Lecturer with Northwest University. Her current research interests include security of autonomous vehicle, VANETs and the IoT, and data aggregation.



GIEDRE SABALIAUSKAITE received the B.Sc. and M.Sc. degrees in information systems from the Kaunas University of Technology, Lithuania, and the Ph.D. degree in software engineering from Osaka University, Japan, in 2004.

She is currently a Research Scientist with the Singapore University of Technology and Design. She is interested in cross-disciplinary and emerging complex topics in relation to the organizations, the design and management of systems, the role of customers, and the strategies to deal with increasingly uncertain environments. Her current research interest includes safety and security of autonomous vehicles.

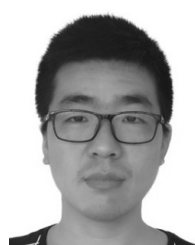


LIN SHEN LIEW received the B.Eng. degree in robotics and mechatronics engineering and the Ph.D. degree from the Swinburne University of Technology, Australia, in 2011 and 2017, respectively. He is currently a Postdoctoral Research Fellow with the Singapore University of Technology and Design. His current research interests include indoor tracking and localization systems and cybersecurity of autonomous systems.



FENGJUN ZHOU received the bachelor's degree from the Shandong University of Technology, China, and the Ph.D. degree in vehicle engineering from the Beijing Institute of Technology, China, in 2014. Then, he was an Engineer with the Chassis Department, Institute of Beijing Automotive Industry Company responsible for ABS and ESP development. He was with the China Automotive Engineering Research Institute Company Ltd. He is currently a Research Fellow with Singa-

pore University of Technology and Design (SUTD) and mainly focused on autonomous vehicle safety.



BIAO ZHANG received the bachelor's degree from Nanchang University, in 2009, the master's degree from Tianjin University, in 2012, and the Ph.D. degree from the Institut National des Sciences Appliquées de Lyon, France, in 2016. He was a Postdoctoral Research Associate with the Singapore University of Technology and Design (SUTD), from 2016 to 2018. He is currently an Associate Professor with the Institute of Flexible Electronics, Northwestern Polytechnical Univer-

sity (NPU), China. His current research interests include flexible electronics and safety materials.

...