On the Alignment of Safety and Security for Autonomous Vehicles

Jin Cui, Giedre Sabaliauskaite
Centre for Research in Cyber Security (iTrust)
Singapore University of Technology and Design, SUTD
Email:{jin_cui,giedre}@sutd.edu.sg

***Abstract*—Safety is the primary requirement and the key challenge in autonomous vehicles. Any accidental failures (safety issue) and/or intentional attacks (security issue) may result in severe injury or loss of life. Thus, any missing consideration on either failures or attacks may lead to terrible consequence. Safety and security are inter-related and, therefore, have to be aligned early in the development process. International standards, International Organization for Standardization (ISO 26262), and Society of Automotive Engineers (SAE J3061), have been proposed for vehicle safety and security. However, they do not address all the aspects of autonomous vehicles as they rely on a human driver controlling the vehicle. In high automation vehicles (level 3 or above, as defined by the international standard SAE J3016), the autonomous driving system is fully responsible for driving the vehicle. Thus, different driving automation levels have to be taken into consideration when designing autonomous vehicle safety and security. We propose an approach for aligning safety and security lifecycles, based on SAE J3061, SAE J3016, and ISO 26262 standards at an early development phase. The proposed approach uses the Failure, Attack and Countermeasure (FACT) graph to connect safety failures, security attacks, and the associated countermeasures. The proposed approach is helpful for designing or tailoring the safety and security processes, and selecting appropriate countermeasures for autonomous vehicles taking into consideration the driving automation levels.**

Keywords–Autonomous vehicle; Safety; Security; FACT graph; SAE J3016; SAE J3061; ISO 26262.

I. INTRODUCTION

Autonomous Vehicle (AV) is a vehicle capable of fulfilling the main transportation capabilities of a traditional car. The main difference to a traditional car is a *Driving Automation System (DAS)* designed for AV. DAS provides driving automation to the vehicle platform, thereby offering the possibility of fundamentally changing transportation in order to reduce crashes, energy consumption, pollution, and cost of congestion [1]. Such vehicle attracts lots of attention from academia, industry and government.

AV is a safety critical system. Any failure of AV may result in severe human injuries or even death. Meanwhile, as a cyber physical system, an autonomous vehicle consists of a myriad of heterogeneous components, both cyber and physical, which pose additional security challenges. The complex interactions between these components inside the AV make it difficult to model the system, and to align the safety and security in an autonomous vehicle.

For a cyber physical system, safety aims at protecting the system from accidental failures in order to avoid hazards, while security focuses on protecting the system from intentional attacks [2]. AV's safety and security is shown in Figure 1. Safety of AV includes mechanical system safety and Electrical and Electronic (E/E) system safety. While considering E/E safety, it is composed of DAS safety and vehicle platform safety. Standard ISO 26262 [3] defines the E/E safety for vehicle platform. Similarity, AV security includes physical security and cyber security. For the latter one, DAS security and vehicle security have to be considered. Standard SAE J3061 [4] defines the cyber security for conventional vehicle. Accidental failures may trigger safety losses, such as harm to life, property and environment, and intentional attacks can result in privacy, financial, operational and safety losses. In this paper, we focus on the alignment between E/E system safety and security.

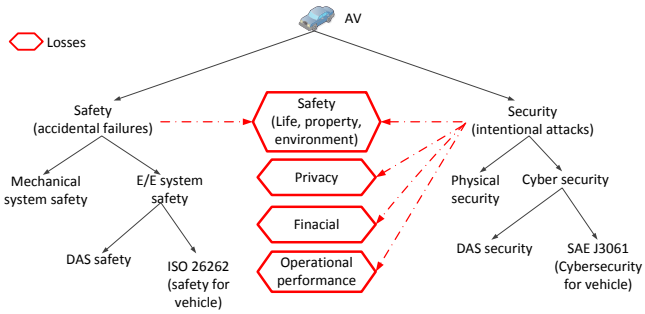

Figure 1. Safety and security in Autonomous Vehicles.

Aligning safety and security is crucial for autonomous vehicles, since any of failures or attacks may lead to safety losses (as seen in Figure 1). The alignment issues for cyber physical system have been discussed in literature [5] [6] [7]. However, such alignment for AVs has not been addressed yet. In SAE J3016 [8], six *levels of driving automation* have been defined. The recent advanced driver assistance systems are only listed around level 1 and 2 (as described in [9]). This will not satisfy the growing demand on driving automation systems. Different level of DAS is corresponding to different driving functions and safety requirements. In addition, different levels will face more potential hazards, threats, and challenges. Thus, it is necessary to consider DAS when we analyse safety and security for AV system, because the selection of safety and security countermeasures for an AV with the same driving function differs depending on its automation level. However, ISO 26262 does not take into consideration driving automation levels and assumes that a human driver is always present [10].

In this paper, we propose an approach for aligning AV's safety and security at early development phases by synchronizing safety and security lifecycles based on SAE J3061, SAE J3016 and ISO 26262 standards. We use Failure Attack and Countermeasure (FACT) graph [2] to list safety failures, security attacks and the associated countermeasures together, which will avoid the safety losses incurred by either failures

 ISBN: 978-1-61208-605-7

or attacks, thereby guaranteeing the safety of autonomous vehicles. Moreover, this alignment is helpful to design or tailor the safety and security processes for autonomous vehicle considering the driving automation levels, and to support safety and security analysis.

The rest of the paper is organized as follow: we introduce the preliminary information in Section II, and explore AV's safety and security alignment in Section III. Finally, we conclude our work in Section IV.

II. PRELIMINARY

To demonstrate our alignment method, we give some preliminary information in this section.

A. Dynamic Driving Task

The driving task is the function required to operate a vehicle in on-road traffic and includes operational functions (basic vehicle motion control), tactical functions (planning and execution for event/object avoidance and expedited route following) and strategic functions (route and destination timing and selection) [8]. The *Dynamic Driving Task (DDT)* [8] includes the operational and tactical functions, such as (without limitation):

1. Lateral vehicle motion control via steering (operational);
2. Longitudinal vehicle motion control via acceleration and deceleration (operational);
3. Monitoring the driving environment via object and event detection, recognition, classification, and response preparation (operational and tactical);
4. Object and event response execution (operational and tactical);
5. Maneuver planning (tactical);
6. Enhancing conspicuity via lighting, signaling and gesturing, etc. (tactical).

Because the subtasks 3 and 4 are all related to **o**bject and **e**vent **d**etection and **r**esponse, they are collectively referred to as *OEDR*.

When a DDT fails, the response to either re-perform the DDT or reduce the risk of crash is considered as *DDT-fallback*. An example of this is when the adaptive cruise control on a car experiences a system failure that causes the feature to stop performing its intended function. The driver will perform the DDT-fallback by resuming performance of the complete DDT.

B. Levels of driving automation

Driving Automation System, DAS, is the hardware and software that are collectively capable of performing the entire DDT on a sustained basis, which is the key property that can replace a human driver for AV. The levels of driving automation are also classified by the requirements on DAS, which include [8]:

- Level 1, the DAS performs either the longitudinal or the lateral vehicle motion control (subtask 1 **or** 2 of the DDT).
- Level 2, the DAS performs both the longitudinal and the lateral vehicle motion control (subtasks 1 **and** 2 of the DDT simultaneously).
- Level 3, the DAS also performs the OEDR (subtask 3 and 4 of the DDT).
- Level 4, the DAS also performs DDT-fallback.
- Level 5, the DAS is unlimited by **O**perational **D**esign **D**omain (ODD).

Here, the 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, we can design a ODD like this: road way is fixed as express way, the vehicle can hold a speed lower than $35km/h$ driving in the daytime only.

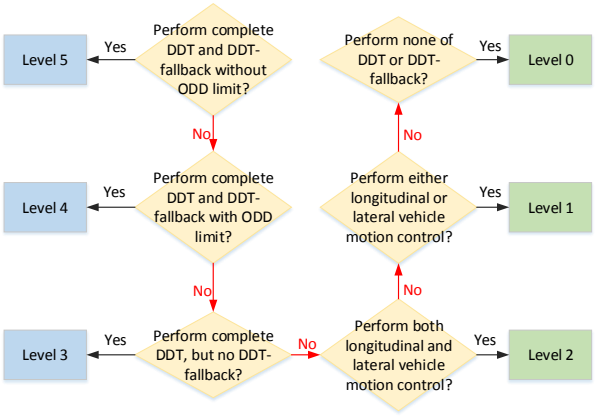

Figure 2. Levels of driving automation.

Figure 2 shows the levels of driving automation and the corresponding features. For the low driving automation (level 0 to level 2), a driver is needed to perform part or all driving task; while for high automation (level 3 to level 5), DAS can replace the driver to perform the complete DDT. The conventional cars in our daily life are at level 0, *No Driving Automation*. The human driver is necessary to perform the driving task and to respond to all the fallback. Level 1 is *Driver Assistance*, which means a DAS can perform **either** lateral **or** longitudinal control for the car. When the DAS performs **both** lateral **and** longitudinal control, such automation is in level 2, i.e., *Partial Driving Automation*.

For the level 3, i.e., *Conditional Driving Automation*, DAS can perform the whole DDT. But a user of the vehicle who is able to operate the vehicle is expected to be able to resume DDT performance when a DDT system failure occurs or when the DAS is about to leave its ODD. If the DAS also can perform DDT-fallback but with limited ODD, this division of role corresponds to level 4, i.e., *High Driving Automation*. The *Full Driving Automation* (level 5) is the situation when DAS can perform complete DDT and DDT-fallback, and meanwhile, the corresponding ODD is unlimited.

 ISBN: 978-1-61208-605-7

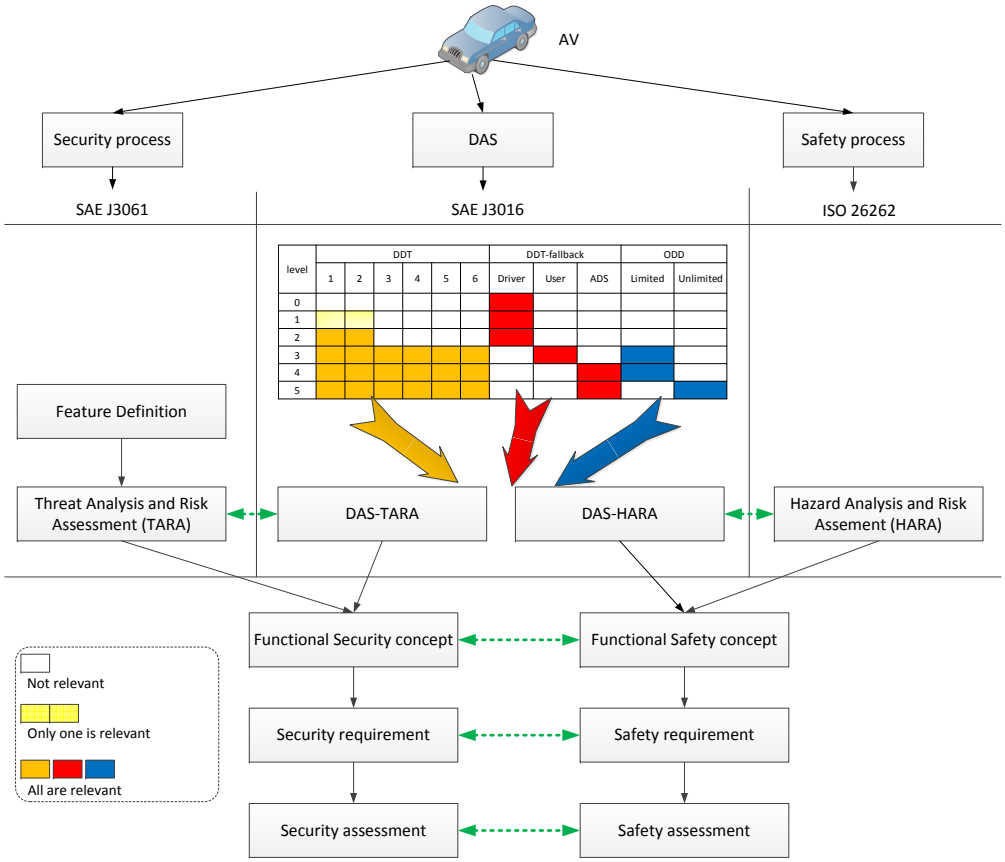

Figure 3. Aligning the safety and security concept phase based on standards SAE J3016, SAE J3061 and ISO 26262.

C. Related safety and security standards

SAE J3061 [4] is a cyber security guidebook for vehicle systems, which defines the lifecycle process framework, and provides guiding principles etc. 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. The concept phase is the first step for the whole lifecycle, which includes the following activities: feature definition, threat analysis and risk assessment, functional security concept, security requirements, and security assessment. The feature definition defines the system being developed to which the cyber security process will be applied, i.e., it defines the boundary of the features. *Threat Analysis and Risk Assessment* (TARA) identifies threats and assesses the risk, and the result of TARA drives all downstream activates. 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. Security assessment is performed to identify the current security posture of the cyber physical vehicle, and it is developed in stages throughout the security lifecycle.

ISO 26262 [3] is an international standard for functional safety of E/E systems in production automobiles defined by the International Organization for Standardization, which provides an automotive safety lifecycle (management, development, production, operation, service, decommissioning) and supports tailoring the necessary activities during these phases. In the development part, similarly to SAE J3061, the safety process is composed of *Hazard Analysis and Risk Assessment* (HARA), functional safety concept, safety requirement and safety assessment.

III. Aligning Safety and Security for AVs

In this section, we introduce an approach to align safety and security for autonomous vehicles.

A. Concept phase of safety and security

Standard SAE J3061 [4] proposes a way to integrate vehicle safety (ISO 26262) and security (SAE J3061) processes by establishing communication paths between safety and cybersecurity concept phase activities, e.g., cybersecurity TARA activity and safety HARA activities, cybersecurity requirement and safety requirement activities. We propose to extend this approach by adding the AV-specific information from SAE J3016 standard, as shown in Figure 3. Additional activities, DAS-TARA and DAS-HARA are added to the integrated safety and security analysis process. Furthermore, communication links are established between DAS-TARA, DAS-HARA, TARA, and HARA activities, as shown in Figure 3.

Figure 3 shows the merged safety and security concept phases, which consists of the phases from different standards. There is no successive order between the activities of safety and security, but for each stage, we need to consider them

 ISBN: 978-1-61208-605-7

simultaneously. We use dotted line with double arrows to depict the simultaneous activities in Figure 3. Because of the automation levels of DAS, TARA and HARA should correspond to each level. A colorful table is used to demonstrate the levels and their properties: yellow denotes the DDT, red represents executor of DDT-fallback, and blue shows ODD constraints. After completion of TARA and DAS-TARA, an activity security concept is performed, which integrates the results of TARA and DAS-TARA, followed by security requirement, and security assessment. In parallel, a functional safety concept activity is performed by DAS-HARA and HARA, followed by safety requirement and safety assessment.

B. Threat analysis and risk assessment for AVs

As mentioned in Section II-C, TARA defines the threats and assesses risks, and derives all the following activities in the security lifecycle. Thus, it is important for the whole security design and development. Most methods for TARA are designed for the automotive domain and are not specific for AVs. In this section, we study automotive TARA cases, and provide a general TARA method, which can also be used for AVs.

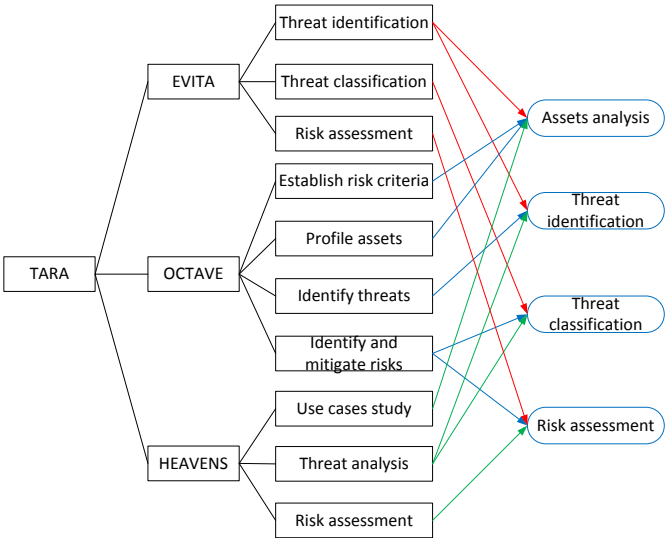

Figure 4. Methods for threat analysis and risk assessment.

EVITA method [12] comes from an European research project EVITA (E-Safety Vehicle Intrusion Protected Applications), which deals with on-board network protection. In EVITA method, TARA phase includes mainly three activities: threat identification, threat classification and risk analysis. Thread identification uses attack trees [13] to identify generic threats; threat classification means classify the threat risk; and risk assessment recommends actions based on the resulting risk classification of the threats.

OCTAVE [14] stands for Operationally Critical Threat, Asset, and Vulnerability Evaluation, which is a process-driven threat/risk assessment methodology. In OCTAVE, TARA phase can be done by such processes: establish risk criteria, profile assets, identify threats, and identify and mitigate risks.

HEAVENS Security Model [15] focuses on methods, processes and tool support for security analysis. In HEAVENS, the main workflow for TARA includes: use case study, threat analysis and risk assessment.

The TARA methods of EVITA, OCTAVE and HEAVENS have different processes (as shown in Figure 4), but these processes have similar functions or similar effects. We classify them into our general method (denoted by arrows in Figure 4). The proposed method has four activities: assets analysis, threat identification, threat classification and risk assessment (rounded rectangles in Figure 4). Assets analysis includes studying use cases, establishing risk criteria, and identifying the assets. Threat identification uses attack trees to identify threats (similar to EVITA). Threat classification classifies the threat risks, and analyzes the mains risks considering use cases. Risk assessment assesses the risks and generates security requirements.

For AVs, the four processes have broader definitions. For assets, besides the visible and information assets on a vehicle, the functional assets (e.g., DDT function) should also be considered. The threats for DAS should be treated as key threats to mitigate, because any functional error of DAS may incur terrible injuries for humans. Thus, the threats which effect DAS should assessed to be of higher risk.

Attack tree [13] is a popular methodology for TARA, which is a graph that describes the steps of the attack process. It uses some basic symbols to demonstrate an attack, e.g., nodes (represent attack events), gates (AND and OR gates) and edges (path of attacks through the system).

C. Hazard Analysis and Risk Assessment for AVs

Following ISO 26262 standard, a HARA is performed to determine the possible hazards, and criticality of the system under consideration. Similar to TARA, the results of HARAs strongly influence the effort to be undertaken in the following activities of ensuring functional safety.

SAHARA [16] is a security-aware HARA method, which expands the inductive analysis of HARA, and encompasses threats from STRIDE model [17], which describes the main security threat categories. SAHARA proposes a security level determination method, and uses it in combination with Automotive Safety Integrity Levels (ASILs) to assess the possible threat.

In [18], the authors propose a HARA method for AV at level 4, i.e., the vehicle is operated on the emergency stopping lane of highway with speed lower than $12km/h$. In this work, ASILs are iteratively refined to achieve specific safety goals for such vehicle.

In summary, conventional HARA is of limited suitability for AVs. But the ASIL is key point that can be used for AVs, because it can be used to assess the threats or hazards impacting DDT or related components of AVs. Fault tree [19] is often used for HARA. Fault trees are similar to attack trees, where the tree nodes represent failure events.

D. Alignment of safety and security

We use FACT graph [2] to combine the safety and security lifecycles. FACT graph is a tree-shaped graph to show system failures, attacks and the associated countermeasures together,

 ISBN: 978-1-61208-605-7

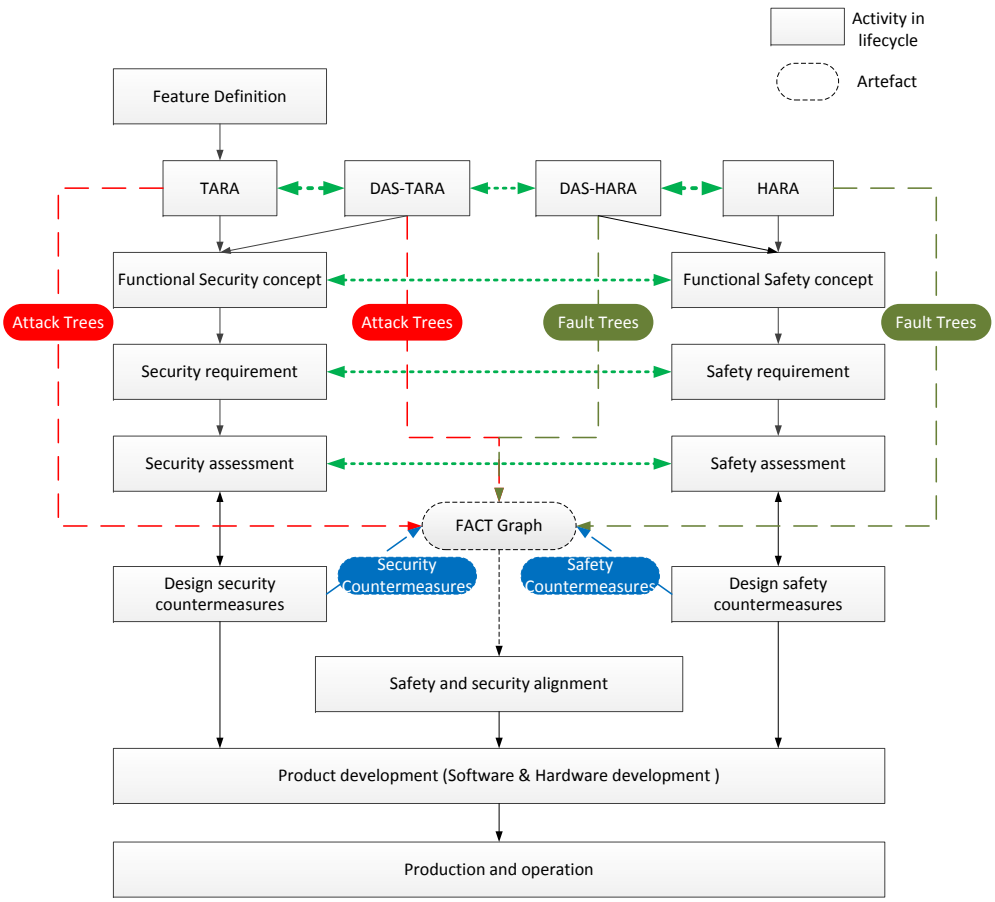

Figure 5. Safety and security alignment for autonomous vehicles.

which is formed throughout several activities of the merged safety and security lifecycle.

The alignment approach is shown in Figure 5, where we use rectangles to denote the activity in lifecycle, and rounded rectangles to present the artefact (i.e., the methodology used for activity). The concept phase comes from Figure 3. We can see that DAS-TARA and TARA constitute the threat analysis and risk assessment part for autonomous vehicles. This is followed by the security concept, the security requirement and the security assessment. Simultaneously, DAS-HARA and HARA should be achieved from a safety view, followed by functional safety concept, safety requirement and safety assessment. After the concept phase part, design of safety and security countermeasures is added to provide mitigation approaches. This activity is not only served for alignment purpose (proposing countermeasure to FACT graph), but also served for the next phases, such as production development, and production and operation as defined in standard SAE J3061.

Figure 6 depicts a simple example of AV FACT graph, which includes Global Positioning System (GPS) failures. GPS data is very important for autonomous vehicles, which is used for localizing the car. If this data is wrong, the consequences could be disastrous. For example, wrong GPS data may lead to traffic disturbance or crash hazard [20]. Here, we consider GPS data on AVs to be the target of an attacker. The associated FACT graph is formed using the following steps (as shown in Figure 6):

- 1. Add safety failures as a subtree of the attack goal (e.g., GPS error). In this situation, a functional fail is considered as a type of safety failure.
- 2. Add security attack as a subtree of GPS error. We consider two types of intentional attacks: spoofing and jamming. Spoofing attacks will modify GPS data, while jamming attacks will prevent AV from receiving GPS data.
- 3. Add safety countermeasures (if any) to associated safety failure. For functional failures, we can consider periodic inspection as one of mitigation technique.
- 4. Add security countermeasures (if any) to corresponding security attack. To avoid spoofing GPS data, we can consider to set the authentication before reading the GPS data. To mitigate jamming GPS data, we can use anti-jam GPS techniques [20]. They are marked as SEC_1 and SEC_2 in Figure 6 respectively.

With the use of FACT graph, any misalignment between safety and security countermeasures can be identified, as well as countermeasure duplicates and missing means of protection. Furthermore, safety and security countermeasures are associated to the relevant faults and attacks, thus, it is easy to analyze

 ISBN: 978-1-61208-605-7

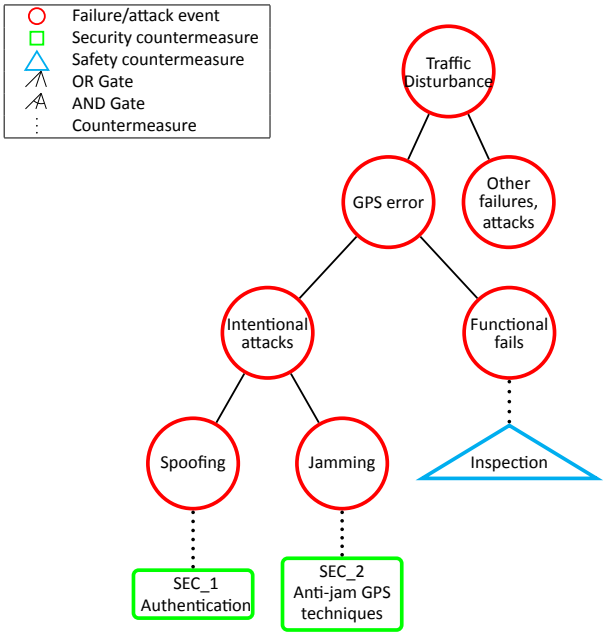

Figure 6. Forming an FACT graph considering GPS error.

the potential failure and attack, and then analyze safety and security requirements.

IV. Conclusion

Safety is the primary target when designing autonomous vehicles. Any accidental failures (safety issues) and/or intentional attacks (security issues) for such vehicle may result in severe safety losses, e.g., human injuries or even death. Thus, the effective alignment of safety and security for AVs is of great importance.

The main difference between AV and conventional vehicle is that there are different levels of driving automation in AV that define which operational and tactical functions are performed by a human driver and by the driving automation system. Thus, the selection of safety and security countermeasures for an AV with the same functions differs depending on its automation level. In this paper, we have proposed an approach for aligning autonomous vehicle safety and security at early development phases considering the levels of driving automation. The proposed approach suggests a way to integrate safety and security lifecycle process phases, defined by SAE J3016, SAE J3061 and ISO 26262 standards. Using this approach, practitioners may align AV's safety and security activities, by following the merged safety and security lifecycle process.

Our proposal can be used for analyzing safety and security of existing AVs, as well on designing new AVs. In the future, we will extend our alignment framework to enable more comprehensive AV safety-security analysis.

References

[1] J. M. Anderson, K. Nidhi, K. D. Stanley, P. Sorensen, C. Samaras, and O. A. Oluwatola, Autonomous vehicle technology: A guide for policymakers. Rand Corporation, 2014, ISBN:978-08-33-08-39-82.

[2] G. Sabaliauskaite and A. P. Mathur, "Aligning cyber-physical system safety and security," Complex Systems Design & Management Asia, 2015, pp. 41–53, ISBN:978-33-19-12-54-42.

[3] International Organization for Standardization (ISO), ISO-26262: Road Vehicles - Functional safety, Dec 2016.

[4] Society of Automotive Engineers (SAE), SAE-J3061: Cybersecurity Guidebook for Cyber-Physical Vehicle Systems, Jan 2016.

[5] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee, and S. K. S. Gupta, "Ensuring safety, security, and sustainability of mission-critical cyber–physical systems," Proceedings of the IEEE, vol. 100, 2012, pp. 283–299, ISSN:0018-9219.

[6] L. Pièrre-Cambacédès and M. Bouissou, "Cross-fertilization between safety and security engineering," Reliability Engineering & System Safety, vol. 110, 2013, pp. 110–126, ISSN: 0951-8320.

[7] T. Novak and A. Treytl, "Functional safety and system security in automation systems-a life cycle model," in IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) Sept. 15-18, 2008, Hamburg, Germany, Sep. 2008, pp. 311–318, ISSN:1946-0740, URL: http://ieeexplore.ieee.org/document/4638412/ [accessed: 2017-09-20].

[8] Society of Automotive Engineers (SAE), SAE-J3016: Taxonomy and Definitions for terms Related to Driving Automation Systems for On-ROad Motor Vehicles, Sep 2016.

[9] H. Martin, K. Tschabuschnig, O. Bridal, and D. Watzenig, Functional Safety of Automated Driving Systems: Does ISO 26262 Meet the Challenges? Springer International Publishing, Sep 2017, chapter 16, pp. 387–416, in Automated Driving, ISBN:978-33-19-31-89-50.

[10] F. Warg, M. Gassilewski, J. Tryggvesson, V. Izosimov, A. Werneman, and R. Johansson, "Defining autonomous functions using iterative hazard analysis and requirements refinement," in International Conference on Computer Safety, Reliability, and Security September 20-23, 2016, Trondheim, Norway, September 2016, pp. 286–297, ISBN:978-33-19-45-48-01, URL: https://doi.org/10.1007/978-3-319-45480-1_23 [accessed: 2017-09-20].

[11] NHTSA, "Federal automated vehicles policy," 2016, URL: https://www.transportation.gov/AV [accessed: 2017-09-20].

[12] R. A. et al., "Deliverable d2.3: Security requirements for automotive on-board networks based on dark-side scenarios," Tech. Rep., 2008, URL: https://rieke.link/EVITAD2.3v1.1.pdf [accessed: 2017-09-20].

[13] B. Schneier, Attack trees. Wiley Publishing, Inc., Oct 2015, chapter 21, pp. 318–333, in Secrets and Lies, ISBN:978-11-19-18-36-31.

[14] C. J. Alberts, S. G. Behrens, R. D. Pethia, and W. R. Wilson, "Operationally critical threat, asset, and vulnerability evaluation (octave) framework, version 1.0," DTIC Document, Tech. Rep., 1999, URL:https://resources.sei.cmu.edu/asset_files/TechnicalReport/1999_005_001_16769.pdf [accessed: 2017-09-20].

[15] M. I. et al., "Deliverable d2 security models," Tech. Rep., 2014, URL:https://research.chalmers.se/en/project/5809 [accessed: 2017-09-20].

[16] G. Macher, H. Sporer, R. Berlach, E. Armengaud, and C. Kreiner, "Sahara: a security-aware hazard and risk analysis method," in Design, Automation Test in Europe Conference Exhibition (DATE) March 09-13, 2015, Grenoble, France, Mar 2015, pp. 621–624, ISSN:1530-1591, URL: http://ieeexplore.ieee.org/document/7092463/ [accessed: 2017-09-20].

[17] M. Corporation, "The stride threat model," 2005, URL: https://msdn.microsoft.com/en-us/library/ee823878(v=cs.20).aspx [accessed: 2017-09-20].

[18] T. Stolte, G. Bagschik, A. Reschka et al., "Hazard analysis and risk assessment for an automated unmanned protective vehicle," in IEEE Intelligent Vehicles Symposium (IV) June 11-14, 2017, Los Angeles, CA, USA, June 2017, pp. 1848–1855, ISBN:978-15-09-04-80-45, URL: http://ieeexplore.ieee.org/document/7995974/ [accessed: 2017-09-20].

[19] E. Ruijters and M. Stoelinga, "Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools," Computer science review, vol. 15, 2015, pp. 29–62, ISSN:1574-0137.

[20] J. Petit and S. E. Shladover, "Potential cyberattacks on automated vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 2, 2015, pp. 546–556, ISSN:1524-9050.

 ISBN: 978-1-61208-605-7