Security Risk Estimation and Management in Autonomous Driving Vehicles

Abasi-amefon O. Affia, Raimundas Matulevičius(✉), and Rando Tõnisson

Institute of Computer Science, University of Tartu, Tartu, Estonia
{amefon.affia,rma}@ut.ee

Abstract. Autonomous vehicles (AV) are intelligent information systems that perceive, collect, generate and disseminate information to improve knowledge to act autonomously and provide its required services of mobility, safety, and comfort to humans. This paper combines the security risk management (ISSRM) and operationally critical threat, asset, and vulnerability evaluation (OCTAVE *allegro*) methods to define and assess the AV protected assets, security risks, and countermeasures.

Keywords: Autonomous vehicles · Self-driving cars · Security risk management · ISSRM · OCTAVE · Intelligent information systems

1 Introduction

Information systems (IS) involve in the gathering, processing, storing, and distributing information to perform specific tasks and make decisions [14]. Autonomous vehicles (AV), a.k.a., self-driving cars, infer as the intelligent information systems (IIS) because they perceive, collect, generate and disseminate *sensitive* information to improve knowledge to act autonomously and provide its required services of mobility, safety, and comfort to humans. Therefore, it is necessary to secure data and information against malicious use and its resulting security risks. But the current studies focus on the automotive aspect of the AVs and do not consider the AV as an information system. This shortcoming leads stakeholders to miss vital security-related knowledge of the AV information collection, manipulation and actuation activities to make well-informed decisions regarding security investment in AV systems.

In this study, we analyse *how autonomous vehicles as intelligent information systems can be protected against security risks?* Specifically, we focus on vehicles that have achieved autonomy as defined in the SAE J3016 [28] standard – which is level 4 (semi-AVs that autonomously perform all driving functions under certain conditions, e.g. on a specific type of road or at certain times) and level 5 (full AVs that perform all driving functions under all conditions autonomously). We combine the domain model for IS security risk management (ISSRM) [11] with the operationally critical threat, asset and vulnerability evaluation (OCTAVE

S. Nurcan and A. Korthaus (Eds.): CAiSE Forum 2021, LNBIP 424, pp. 11–19, 2021.
https://doi.org/10.1007/978-3-030-79108-7_2

allegro) [5] to explore the security risks in AVs through the literature study and case analysis.

2 Research Method

There exist different methods aimed at security risk management, relevant for use in AVs: for example, EVITA [13], ETSI Threat, Vulnerability and implementation Risk Analysis (TVRA) standard [12], HEAVENS security model [21], Security-aware hazard and risk analysis method (SAHARA) [17]. These methods provide guidelines when considering AV security; however, they lack a consistent approach for identifying vulnerabilities, threats/attacks, risk assessment, risk estimation, and risk treatment. Consequently, we did not find a standard or method in the literature that apply to AV security, given its complex capabilities and its overlap across multiple information domains.

In this study, we combine the ISSRM domain model [11] and the OCTAVE [5] method. OCTAVE guides the assessment of information security risks; it contains templates to document risk management activities and guide data collection for probability and security risk impact estimation. But OCTAVE does not support an explicit risk analysis process. The ISSRM domain model guides elicitation of assets, attack methods, vulnerabilities, threats, risks, and solutions to mitigate risks. Initially developed for IS risk management, the ISSRM domain model is also applicable in the AV systems as these systems gather, manipulate, interpret and *disseminate* information for the stakeholders. A combination of OCTAVE and ISSRM strengthens the analysis (*i*) with the ISSRM guidance for the security risk definition and (*ii*) with the OCTAVE templates for security risk estimation.

3 Literature Study

Assets. Following [2], the AV system can be decomposed to perception, network and application layers, as illustrated in Fig. 1. The business assets are defined as data and information that are valuable in the system with its security needs to be estimated using the security criteria (i.e., confidentiality, integrity and availability). The *perception layer* includes the system components responsible for collecting *video, location and travel, surrounding environment and other* data [24,27,28]. The collected data is transmitted to the application layers using the *network layer*. The *application layer* uses the collected data to perform tasks, i.e., calculate routes. An actuation module uses these calculations to perform autonomous functions.

Security Risks. In the literature, we have identified security risks (see Table 1, columns 1–3). Eleven risks (R1–R9, R16) are identified at the perception layer, seven (R10–R11, R19–R23) – at the application layer, and four (R12–R15) at the network layer. R17 can be found at all layers, and R18 is identified at the network and application layers. Details of security risks could also be found in [26]. We do not consider the risks in Table 1 as exhaustive but aim to raise awareness of the AVs' risks.

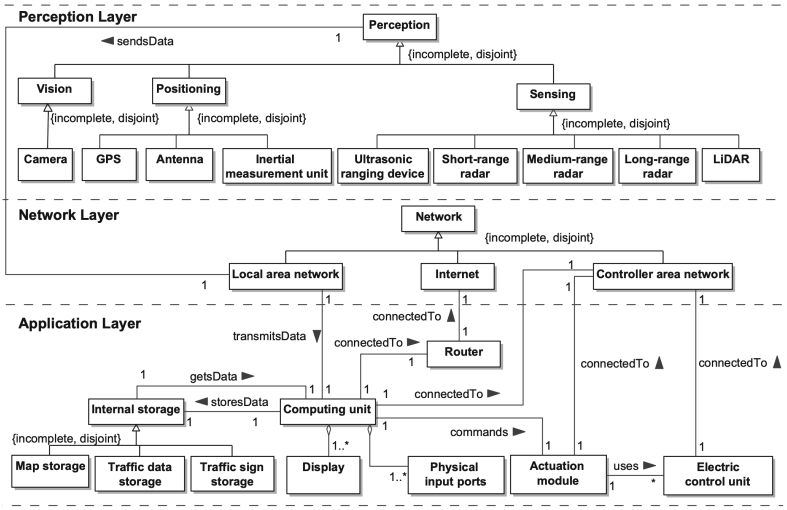

Fig. 1. Literature study: AV system assets (adapted from [26])

Table 1. Literature study: Security risks and countermeasures

Risk ID	Literature	Risk name	Countermeasure	Literature
R1	[19,27,28]	Jamming ultrasonic sensors	Noise detection and rejection; Multiple sensors for redundancy check	[28]
R2	[19,27,28]	Spoofing ultrasound sensors	Noise detection and rejection; Multiple sensors for redundancy check	[28]
R3	[19,23,28]	Acoustic quieting on ultrasound sensors	Multiple sensors for redundancy check	[28]
R4	[19,23,27,28]	Jamming radar	Noise detection and rejection; Multiple sensors for redundancy check	[28]
R5	[19,23,27,28]	Spoofing radar	Noise detection and rejection; Multiple sensors for redundancy check	[28]
R6	[19,24,28]	Blinding cameras	Multiple cameras; Filter to remove harmful light	[24,28]
R7	[24]	Confusing controls using camera inputs	Multiple cameras; Filter to remove harmful light	[24,28]
R8	[19,24]	Relay attack on LiDAR	Multiple LiDAR inputs; Random probing; Shorten pulse period	[24,28]
R9	[19,24,27]	Spoofing LiDAR	Multiple LiDAR inputs; Random probing; Shorten pulse period	[24,28]
R10	[27]	Code modification	Device authentication; Anti-Malware; Isolation	[25,27]
R11	[27]	Code injection	Device authentication; Anti-Malware; Isolation	[25,27]
R12	[25,27]	Packet sniffing	Encryption; Device and user authentication	[25,27]
R13	[19,25,27]	Packet fuzsing	Encryption; Device and user authentication	[25,27]
R14	[23]	Inject CAN messages	Encryption; Device and user authentication	[25,27]
R15	[19,23]	Eavesdropping CAN messages	Encryption; Device and user authentication	[25,27]
R16	[19,22,23]	GPS: Jamming and spoofing	Nullification, Monitoring signals and identification nodes	[20,27]
R17	[23]	EMP attack	Isolation	[25,27]
R18	[23]	Malware injection	Firewall; Anti-Malware; Isolation	[27]
R19	[19,23]	Manipulate map data	User authentication; Device authentication; Isolation	[25,27]
R20	[19]	Extract map data	User authentication; Device authentication; Isolation	[25,27]
R21	[19]	Delete map data	User authentication; Device authentication; Isolation	[25,27]
R22	[19]	Disable actuation module	Isolation; Access control	[25,27]
R23	[19]	Induce bad analysis	Isolation; Access control; Input validation	[25,27]

Countermeasures identified in the literature are presented in Table 1 (see columns 4 and 5). Some key countermeasures at the perception layer are noise detection, rejection, and the use of multiple input sources [24,28]. At the network layer, countermeasures are encryption, special devices and user authentication techniques [25,27]. At the application layer, countermeasures include anti-malware software, firewalls, access control and user authentication.

4 Case Analysis

We have combined the ISSRM and OCTAVE methods to assess the Lexus RX450h AV system in the laboratory environment[1]. The car's architecture is presented in Fig. 2; it is adapted from our literature study (see Fig. 1). The figure also illustrates how risks are associated. Table 2 shows an example of one risk – R6. Here, *an attacker with some expertise and tools sends malicious optical inputs* targeting the AV *cameras*(system assets) because the *cameras are vulnerable to blinding attacks*. If the event happens, it negates the *integrity of video and picture data* leading to unreliable data sensed by the cameras that could

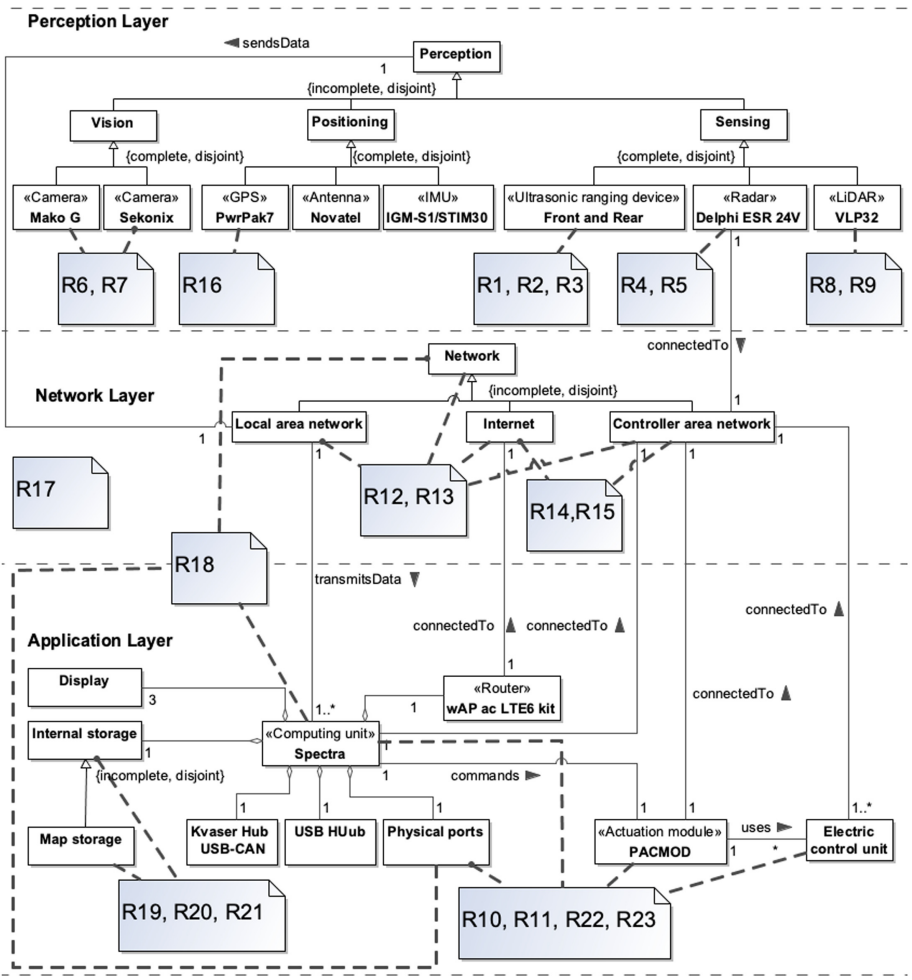

Fig. 2. Case analysis: AV assets and associated security risks (adapted from [26])

[1] https://www.cs.ut.ee/en/autonomous-driving-lab.

provoke wrong decisions when the car is driving/steering. In [26], all considered security risks are explicitly defined using the ISSRM domain model.

Countermeasures are suggested at all three AV layers – perception, network, and application – to mitigate the security risks. We also elicit the relative countermeasure costs. For example, to reduce risk R6, two controls are suggested: (*i*) multiple sensors for redundancy check estimated with a *low* cost and (*ii*) filter to remove harmful light with a *high* cost. The complete risk management documentation of the Lexus RX450h AV system can be accessed in [26].[2,3]

Limitations of Case Analysis. The case study AV is still in the early development phase (i.e., laboratory settings), so we could not consider the system and business assets throughout its life-cycle management; thus, the risk estimations could change. Additionally, we apply security risks found in the literature in our case study. Although we have discovered added risks, our security risk analysis and estimation is limited by those we identified in the literature.

Table 2. Risk estimation using OCTAVE (R6: Blinding cameras)

Threat	**Business Asset**	Video and picture data									
	Business Asset's Value	Medium – Car can continue driving but can't recognize signs and traffic lights.									
	Area of Concern	An attacker uses their tools to send malicious optical data to the camera causing unwanted blindness, possible hardware damage and loss of integrity of video and picture data.									
	Actor Who would exploit the area of concern or threat?	An attacker with some previous experience and tools to send malicious optical inputs (laser etc.).									
	Means How would the actor do it? What would they do?	An attacker uses their knowledge and malicious optical emitters to send and blind cameras causing unwanted blindness on the cameras and possibly permanently damage the camera sensors.									
	Motive What is the actor's reason for doing it?	Wants to see the car crash and make the company lose reputation.									
	Outcome (choose one) What would be the resulting effect be?	Disclosure:		Destruction:		Modification:		Interruption:	x		
	Security Requirements How would the information asset's security requirements be breached?	Cameras are vulnerable to blinding attacks.									
	Likelihood (choose one)	High:	x	Medium:		Low:					
Consequences What are the consequences to the organisation as a result of the risk?						**Severity** *3 for highest priority, 2 for medium, 1 for lowest					
Blinding attack will cause some blind spots on the image recorded by the cameras. Blind spots can cause not detecting objects and possible accidents because of that. Not having the sensor available without any mitigations will cause the system to not see the outside, possibly other sensors can cover. Using lasers to carry out the attack can permanently damage the camera's lens.						**Impact area**	**Priority***	**Impact**	**Score**		
						Confidentiality	1	Low	1		
						Availability	3	High	9		
						Integrity	2	High	6		
						Relative risk score:			16		
						Total Risk Score (Rel x likelihood):			48		

5 Concluding Remarks

AVs, as IIS, provide a unique perspective on how security risks should be handled while considering the autonomy of data gathering, manipulating, disseminating and actuating functions. In this section, we conclude the paper with the lessons learnt and an overview of the related studies.

[2] Full risk management documentation for R6 https://git.io/JkCBf.

[3] Full risk management documentation for other security risks https://git.io/JkCBY.

5.1 Lessons Learnt

Asset-Related Concepts. We have identified the business assets in the perception, network, and application layers. The study results in the systematic examination of the security risks and helps identify countermeasures. For example, the inclusion of multiple sensors could monitor and analyse data from robust and redundant data sources in AVs, which help develop protection strategies that could identify anomalous inputs and behaviour produced by cyber attacks. However, stakeholders must ensure the presence and trust boundary of the sensors as data sources.

Risk-Related Concepts. The assets, alongside attack methods and threat agents, are dynamic. Thus, our estimated scores in the case analysis would change over periods. The dynamic nature of assets, threat likelihood due to evolving attack tools/ methods, and an evolving environment indicate that security risk management must be an ***iterative*** activity involving dynamic and real-time risk impact estimations as proposed in [16].

We discovered other attacks on the AV not discussed in the literature. One example is carrying out the blinding attack using mirrors to reflect sunlight which could make the likelihood of **R7** (see Footnote 3), *high*. Another is spoofing the *LiDAR* – **R9** (see Footnote 3) with smoke, which increases the likelihood of a spoofing event as it is a low effort attack. Lastly, capability to tamper with the AV code functions used by the *ECU* in the code repository. Efforts must be made to continually improve sensors (i.e. auto-exposure settings for **R7**), algorithms (i.e., improving obstacle detection algorithms for **R9**) and integrity checks for all code used by the *ECU*.

Risk Treatment-Related Concepts. Outlining risk estimations enables treatment prioritisation and return on security on investment analysis in AVs. Using OCTAVE, we can deduce the risk scores based on the relative scores of the impact on the affected assets and threat likelihood while providing a documented risk overview.

Thus, the combination of the ISSRM and OCTAVE methods has filled in the gaps posed by either method's single-use. The ISSRM method refined the security risk management concepts applied in OCTAVE. At the same time, OCTAVE helped provide formal documentation and risk estimation through risk scores based on the relative scores of the affected assets' impact and threat likelihood. The combination provided a useful output to support the AV stakeholders. In [3], Bailey combines the NIST risk assessment process [1] with a probabilistic method and applies the optimisation techniques to recommend the best solutions. Similarly, we have used the countermeasure cost estimation, but we transform the estimates to the qualitative values. This approach allows us to reduce the amount of collected data, but it still supports countermeasure selection decisions.

5.2 Related Work

Related studies analyse security risks in some parts or the whole AV. However, these studies do not consider AVs as information systems. In Table 3, we present

Table 3. Comparing related work

Related work	Risk management				Autonomous focus	Security focus
	Threat	Risk	Solution	Estimation		
Chattopadhyay *et al.* [6]	✓	LD	✓	✗	✓	✓
Parkinson *et al.* [22]	✓	LD	✓	✓	✓	✓
Malik and Sun [18]	✓	LD	✗	✗	✓	✓
Boudguiga *et al.* [4]	✓	✓	✗	✓	LD	SS
Cui *et al.* [7]	✓	LD	✓	✗	✓	SS
Dibaei *et al.* [9]	✓	LD	✓	✗	✓	✓
De La Torre *et al.* [8]	✓	LD	✓	✗	✓	✓
Dominic *et al.* [10]	✓	✓	✓	LD	LD	✓
Kong *et al.* [15]	✓	✓	✗	✗	✓	✓
This study	✓	✓	✓	✓	✓	✓

✓ - Detailed Discussion of concept; LD - Limited Discussion of concept;
✗ - No Discussion of concept; SS - Discusses security and safety.

a comparison illustrating the focus on AV security risk management. It indicates that Dominic *et al.* [10], and Parkinson *et al.* [22] cover the security risk estimation and management in AVs; however, they provide little details on AV system assessment's impact. Other related studies [6–9,15,18] address security risk management, some [4,7] focus on safety and security engineering in AVs. But these studies do not discuss estimations needed for AV stakeholder security investment decisions. In [4], Boudguiga *et al.* TVRA and EVITA methods are combined, but they did not include the cost of countermeasures to assess the severity of the security risks in supporting business decisions. Our study provides a security-focused risk estimation and management analysis on the AV information system, covering the security assets, threats, resulting risks, proposed countermeasures and risk impact estimations based on the mentioned security metrics. We have documented these concepts within the AV case analysis, providing the rationale for making business decisions on securing AVs.

Acknowledgement. This paper is supported in part by EU Horizon 2020 research and innovation programme under grant agreement No 830892, project SPARTA and European Social Fund via Smart Specialisation project with Bolt.

References

1. National Institute of Standards and Technology. SP 800–30 Rev. 1: Guide for Conducting Risk Assessments. https://csrc.nist.gov/publications/detail/sp/800-30/rev-1/final
2. Affia, A.-A.O., Matulevičius, R., Nolte, A.: Security risk management in cooperative intelligent transportation systems: a systematic literature review. In: Panetto, H., Debruyne, C., Hepp, M., Lewis, D., Ardagna, C.A., Meersman, R. (eds.) OTM 2019. LNCS, vol. 11877, pp. 282–300. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33246-4_18

3. Bailey, D.: Quantitative cybersecurity risk management for autonomous vehicle systems. Master's thesis, Technisch Universitat Munchen (2018)
4. Boudguiga, A., Boulanger, A., Chiron, P., Klaudel, W., Labiod, H., Seguy, J.C.: Race: risk analysis for cooperative engines. In: 2015 7th International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–5. IEEE (2015)
5. Caralli, R.A., Stevens, J.F., Young, L.R., Wilson, W.R.: Introducing OCTAVE allegro: improving the information security risk assessment process. Techical report, SEI (2007)
6. Chattopadhyay, A., Lam, K.Y., Tavva, Y.: Autonomous vehicle: security by design. IEEE Trans. Intell. Transp. Syst. (2020)
7. Cui, J., Liew, L.S., Sabaliauskaite, G., Zhou, F.: A review on safety failures, security attacks, and available countermeasures for autonomous vehicles. Ad Hoc Netw. **90**, 101823 (2019)
8. De La Torre, G., Rad, P., Choo, K.K.R.: Driverless vehicle security: challenges and future research opportunities. Futur. Gener. Comput. Syst. **108**, 1092–1111 (2020)
9. Dibaei, M., et al.: Attacks and defences on intelligent connected vehicles: a survey. Digit. Commun. Netw. (2020)
10. Dominic, D., Chhawri, S., Eustice, R.M., Ma, D., Weimerskirch, A.: Risk assessment for cooperative automated driving. In: Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy, pp. 47–58 (2016)
11. Dubois, É., Heymans, P., Mayer, N., Matulevičius, R.: A systematic approach to define the domain of information system security risk management. In: Nurcan, S., Salinesi, C., Souveyet, C., Ralyté, J. (eds.) Intentional Perspectives on Information Systems Engineering, pp. 289–306. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12544-7_16
12. ETSI TS: 182 027: Telecommunications and internet converged services and protocols for advanced networking (TISPAN). IPTV Architecture (2011)
13. EVITA: E-safety Vehicle Intrusion Protected Applications (2011). http://www.evita-project.org/
14. Gupta, E.: Information system. Bajaj, Ankit 197 Bakry, Mohamed Abd El Latif 28 Bala, Shashi 414 Baporikar, Neeta 118, p. 97 (2000)
15. Kong, H.K., Hong, M.K., Kim, T.S.: Security risk assessment framework for smart car using the attack tree analysis. J. Ambient. Intell. Humaniz. Comput. **9**(3), 531–551 (2018)
16. Leite, F.L., Schneider, D., Adler, R.: Dynamic risk management for cooperative autonomous medical cyber-physical systems. In: Gallina, B., Skavhaug, A., Schoitsch, E., Bitsch, F. (eds.) SAFECOMP 2018. LNCS, vol. 11094, pp. 126–138. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99229-7_12
17. Macher, G., Sporer, H., Berlach, R., Armengaud, E., Kreiner, C.: Sahara: a security-aware hazard and risk analysis method. In: 2015 Design, Automation and Test in Europe Conference & Exhibition (DATE), pp. 621–624. IEEE (2015)
18. Malik, S., Sun, W.: Analysis and simulation of cyber attacks against connected and autonomous vehicles. In: 2020 International Conference on Connected and Autonomous Driving (MetroCAD), pp. 62–70. IEEE (2020)
19. Maple, C., Bradbury, M., Le, A.T., Ghirardello, K.: A connected and autonomous vehicle reference architecture for attack surface analysis. Appl. Sci. **9**(23) (2019)
20. O'Hanlon, B.W., Psiaki, M.L., Bhatti, J.A., Shepard, D.P., Humphreys, T.E.: Real-time GPS spoofing detection via correlation of encrypted signals. Navigation **60**(4), 267–278 (2013)

21. Olovsson, T.: HEAling Vulnerabilities to ENhance Software Security and Safety (HEAVENS) 2015, besucht am 20, pp. 33–89 (2018). https://research.chalmers.se/en/project/5809
22. Parkinson, S., Ward, P., Wilson, K., Miller, J.: Cyber threats facing autonomous and connected vehicles: future challenges. IEEE Trans. Intell. Transp. Syst. **18**(11), 2898–2915 (2017)
23. Petit, J., Shladover, S.E.: Potential cyberattacks on automated vehicles. IEEE Trans. Intell. Transp. Syst. **16**(2), 546–556 (2014)
24. Petit, J., Stottelaar, B., Feiri, M., Kargl, F.: Remote attacks on automated vehicles sensors: experiments on camera and lidar. In: Black Hat Europe (2015)
25. Scalas, M., Giacinto, G.: Automotive cybersecurity: foundations for next-generation vehicles. In: 2nd International Conference on new Trends in Computing Sciences (ICTCS), pp. 1–6 (2019)
26. Tõnisson, R.: Security Risk Management in Autonomous Driving Vehicles: Architecture Perspective. University of Tartu (2020)
27. Thing, V.L., Wu, J.: Autonomous vehicle security: a taxonomy of attacks and defences. In: IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), pp. 164–170 (2016)
28. Yan, C., Xu, W., Liu, J.: Can you trust autonomous vehicles: contactless attacks against sensors of self-driving vehicle. In: DEFCON (2016)