# Autonomous Driving Security: A Comprehensive Threat Model of Attacks and Mitigation Strategies

Mohammad Aminul Hoque
*Dept. of Computer Science*
*University of Alabama at Birmingham*
Birmingham, AL 35294-1241
mahoque@uab.edu

Ragib Hasan
*Dept. of Computer Science*
*University of Alabama at Birmingham*
Birmingham, AL 35294-1241
ragib@uab.edu

***Abstract*—Autonomous vehicles (AVs) are envisioned to enhance safety and efficiency on the road, increase productivity, and positively impact the urban transportation system. Due to recent developments in autonomous driving (AD) technology, AVs have started moving on the road. However, this promising technology has many unique security challenges that have the potential to cause traffic accidents. Though some researchers have exploited and addressed specific security issues in AD, there is a lack of a systematic approach to designing security solutions using a comprehensive threat model. A threat model analyzes and identifies potential threats and vulnerabilities. It also identifies the attacker model and proposes mitigation strategies based on known security solutions. As an emerging cyber-physical system, the AD system requires a well-designed threat model to understand the security threats and design solutions. This paper explores security issues in the AD system and analyzes the threat model using the STRIDE threat modeling process. We posit that our threat model-based analysis will help improve AVs' security and guide researchers toward developing secure AVs.**

***Index Terms*—autonomous vehicles; threat model; security;**

## I. Introduction

Autonomous vehicles (AVs) are rapidly developing and capable of sensing the surrounding environment and operating with little or no human input. Autonomous driving (AD) technology has gained tremendous improvement in recent years. For example, Google's Waymo crosses 20 million miles of autonomous driving [1]. As AVs are considered the future of transportation, ensuring their security is crucial. The AVs require the protection of the sensors from different attacks and robust decision-making algorithms. Hence, a proper understanding of potential threats and vulnerabilities is necessary to ensure AD security, which can be performed by a systematic threat modeling.

AVs are complex cyber-physical systems that impose challenges in ensuring security. The AVs trust the information extracted from sensors without any verification. However, these sensors are vulnerable as the attacker can inject manipulated analog signals [2]. Moreover, the AVs trust the control commands provided by the control algorithms, and the actuators execute them faithfully. Using sensor data without validation is not a sustainable practice as sensor attacks continue to be sophisticated and mature. Poisoning deep learning models by injecting adversarial examples has also become feasible due to the availability of open-source machine learning models for AD and high-quality driving data [3]. Classical security mechanisms such as encryption, authentication, or memory protection are not enough for securing AVs. Hence, the security solution must consider the design and architecture of the system and also meet the security requirements to avoid random usage of security technologies [4]. Considering the unique security challenges and wide range of possible attacks on AVs, a proper threat model-based approach is essential in this regard.

Threat modeling is a systematic approach that analyzes all the security aspects to identify a system's threats and vulnerabilities [5]. First proposed by Microsoft, STRIDE is a popular process for categorizing threats [6]. Threat modeling helps identify the attack vector, profile of the attacker, valuable assets in which the attackers are most interested, and potential mitigation strategies. Thinking from an attacker's perspective and identifying the attacker's motives make the threat modeling problem more challenging. Focusing only on the security issues related to a particular threat might leave a considerable portion of attack space unprotected. The system designers and developers must identify the specific security requirements to ensure the security solution's compatibility with the system architecture. The threat model also helps to validate the assumptions from brainstorming and justify the security solution's countermeasures for solving an issue.

In this paper, we analyze all the components of AD threat modeling. We explain the AD components and explore the attack surfaces to figure out the potential attacks that can cause safety hazards on the road. We also demonstrate the attacker model, identify vulnerabilities, and explore different security enhancement techniques of AD. In this research, we do not leverage the connected vehicle and the CAN bus security; instead, our focus is limited to the components regarding AD. We posit that the threat model can help researchers design better security solutions for AVs in the future.

**Contribution:** The contributions of this paper are as follows:

1) We explore components of a module-based autonomous driving system for AVs to identify security issues.
2) We provide a comprehensive threat model for AD systems and identify assets, entry points, and attacker models.
3) We identify the potential threats, vulnerabilities, and mitigation strategies using STRIDE threat modeling process.

**Organization:** The rest of the paper is organized as follows: Section II provides the background of threat modeling and autonomous driving components. Sections III identifies the valuable assets and IV explores the attacker entry points. Section V explains the attacker model. Section VI presents the threats and vulnerabilities of AVs. Section VII identifies the potential mitigation strategies. We present related works in Section VIII and conclude in Section IX.

## II. Background

In this section, we provide background information regarding threat modeling and AV components.

### A. Threat modeling

A threat model systematically identifies and prioritizes the potential threats and vulnerabilities of a system [7]. There are five steps in threat modeling where each of them is important and complements each other [5], which are:

**Assets:** The attackers always target some assets of a system to gain access to which they are interested in. Before designing a security solution, it is crucial to understand the system's valuable assets, which may attract the attacker.

**Entry points:** The attacker needs to enter into the system to access the targeted assets or launch an attack. They use vulnerable or untrusted points to enter the system, referred to as entry points.

**Attacker model:** The attacker model explains the characteristics of the attackers. It defines who the attackers are, their attack motives, and their capabilities.

**Threats and vulnerabilities:** This step identifies and lists all the potential attacks on a system considering the assets, entry points, and attacker model. Threat modeling processes help to organize the threats and vulnerabilities by categorizing them according to the security properties and requirements.

**Mitigation strategies:** Identifying mitigation strategies is the last step of threat modeling. Mitigation strategies refer to designing security solutions with known security enhancement techniques to mitigate potential attacks and improve the system's security.

### B. Autonomous driving components

AD system has four fundamental pillars for driving autonomously on the road: perception, localization, planning, and control [8]. The AV collects information by leveraging sensors such as cameras, LiDAR, radar, and ultrasonic sensors and extracts knowledge to understand the surrounding environment in the perception module.

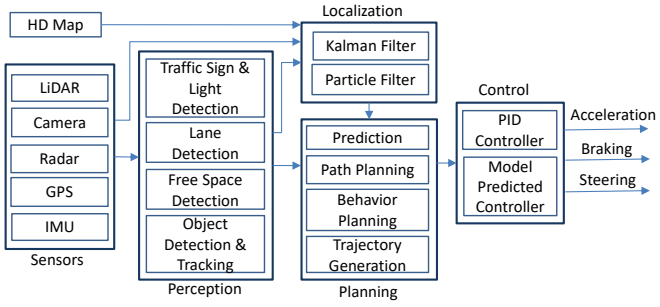


Fig. 1: Autonomous driving components

Understanding the environment and scene representations refer to locating the obstacles, detecting road signs, free space, lane marking, other vehicles, and pedestrians along with their movements. Object detection algorithms perform segmentation, classification, and clustering using sensor data from camera, LiDAR, and radar data. The localization module identifies the vehicle's location on the road with 2-10 centimeter accuracy. This module fuses sensor data from GPS, LiDAR, and IMU using bayesian filters such as Extended Kalman Filter and Particle Filter. Perception and localization modules lead to the next step of planning the path towards the destination. In the planning phase, the AV makes purposeful decisions to reach the destination by avoiding obstacles and maintaining the traffic rules. The planning module consists of prediction, path planning, behavior planning, and trajectory generation components. The initial high-level planned path is followed by generating a trajectory for the next few seconds after predicting other nearby moving agents' behavior. Finally, the control module generates the acceleration, brake, and steering angle using different control algorithms such as the PID controller and model predictive controller and passes the command to mechanical components. Figure 1 shows different AD modules.

## III. Assets of Autonomous Vehicles

Assets are the most important things of AVs that seem attractive to attackers. The most critical assets of AVs are:

**Sensors:** AVs are equipped with multiple sensors such as cameras, radar, LiDAR, and GPS. These sensors are the most critical assets of AVs because they must work correctly for safe autonomous driving. The correctness of sensor data is also essential for AV safety and decision-making algorithms.

**Decision-making algorithms:** The AVs process the sensor data for perception and predict the movement of nearby objects to decide the subsequent actions, which are crucial assets.

**Computation hardware:** Computation hardware, such as NVIDIA Drive PX, NVIDIA AGX, etc., are important assets of AVs which execute the deep learning models on sensor data.

**Log data:** Log data is essential for software debugging and analyzing vehicle behavior in specific circumstances. Log data are also necessary for future forensic investigation.

**Reputation of the manufacturer company:** Multiple technology and automobile companies are currently working on AD. The performance of the AVs may reflect on the reputation of the companies.

## IV. Entry points of the attackers

The attacker needs to enter the system to launch an attack or access valuable assets. Entry points are the vulnerable points that the attacker exploits to enter the system. Determining the trust boundary [4] is also essential along with the entry points. Figure 2 shows the potential entry points of attackers in AV.

**Physical access:** AVs are physically accessible on the road, where the attackers can access the outside sensors. They can also perform passive keyless entry [9] to open the door and physically access the hardware and OBD port inside the car.

**Sensors:** Autonomous vehicle sensors are one of the key entry points to launch the attacks. The attacker can use different techniques to jam the sensors or spoof sensor data.

**Deep learning model:** AD perception and path planning depend on the correctness of the deep learning model. Hence, the deep learning models work as another potential entry point for the attackers to perform adversarial attacks.

**Roadside signs:** Attackers can use road signs or traffic lights as an entry point by making them unrecognizable to the AV.

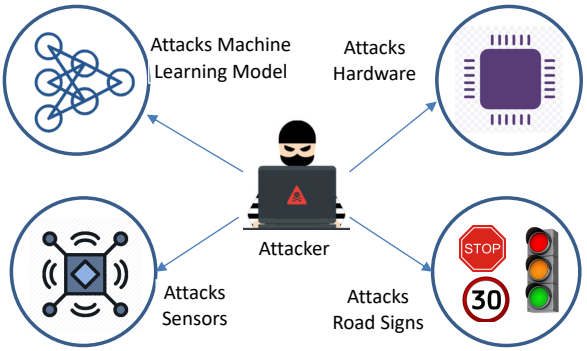

Fig. 2: Attacker entry points in autonomous vehicles

Besides these, an attacker can also access and interfere with sensor data or AD system using CAN bus, vehicular network, or vehicle infotainment system.

## V. Attacker Model

The attacker model defines who the attackers are, their motives, and their capabilities, which are explained below:

### A. Attacker

Attackers are the person or entities who directly launch attacks on the system to achieve any goal. Potential attackers of the AD system are identified and listed as follows:

**Car owner:** Car owners themselves can be the attacker to achieve different gains such as insurance fraud. They have physical access to all the AD components and can tamper with the sensors or deep learning algorithms to damage the vehicle.

**Pedestrian:** Attacker can be a pedestrian who throws lights to blind the camera or laser beam to insert fake 3D point clouds.

**Other cars:** Attackers from other cars can inject out-of-band signals to spoof or jam the sensors such as radar, lidar, and GPS from the same or adjacent lanes. They can also spoof the AV by pretending as emergency vehicles or police cars.

**Infrastructure owner:** The attacker can be an infrastructure owner who executes attacks on AVs in exchange for money. Developing the attack infrastructure can be costly and time-consuming, which may motivate to outsource the resources.

### B. Attack motives

The motives behind an attack can be highly dynamic based on the attacker, system architecture, and the assets that are in the attacker's mind. The potential attack motives can be:

**Traffic collision:** The attacker may try to force the AV to deviate to the lanes of the opposite direction or stop on the shoulder lane. Such kind of incidents can damage the reputation of rival companies and unfairly gain competitive advantages.

**Economic damage:** The purpose of an attack can be causing different economic attacks such as damaging the car, more fuel consumption, routing to toll roads, etc.

**Insurance fraud:** Another motive of attackers behind attacking AVs may be inducing fraudulent activity to gain financial advantage from the insurance company.

**Personal gain:** Attacking the AVs may include different personal gains of attackers, such as forcing them to take alternate routes to make the intended route clear.

**Mass terrorist attack:** Mass terrorist attack on the road may be conducted by forcing collision among the cars, with specific infrastructures, or at a particular place.

### C. Attacker capabilities

Capabilities are the actions an attacker can perform from inside or outside of the system which depends on many factors such as access/privilege of the attacker, resource availability to launch attacks, inside information known to the attacker, etc.

**Targeted and untargeted attack:** The attacker can launch an attack targeting a specific vehicle or broadcasting the attack so that any AV can be the victim.

**Remote attacks:** The attacker is capable of attacking the AD system without physical access, such as jamming the GPS, inserting fake 3D points clouds, creating fake obstacles, etc.

**Manipulating deep learning model:** Attackers know the deep learning model artifacts used in AD, such as the learned weights, activation functions, and model architecture. Moreover, they can also manipulate the model by retraining it with carefully crafted adversarial examples and deploying them into AV [10].

**Manipulating hardware and software states:** Attackers can modify the software and hardware states by manipulating the output variables. They can also design machine learning models to decide when to launch an attack [11], [12].

**Sensor system knowledge:** The attacker can access the sensor system and understand the underlying properties such as operational frequency, packet format, bandwidth, etc. They can exploit the hardware and manipulate the sensor values or signals such as manipulating LiDAR point cloud [13].

## VI. Threats and Vulnerabilities

This section analyzes the AD system's threats and vulnerabilities according to STRIDE threat modeling process to understand the potentially vulnerable points and severity of the attacks.

### A. STRIDE threat modeling process

The STRIDE threat modeling process was first proposed by Microsoft to identify the security threats of a system [6]. The term STRIDE refers to **S**poofing, **T**ampering, **R**epudiation, **I**nformation disclosure, **D**enial of service, and **E**levation of privilege. Each section of the STRIDE model corresponds to a desirable security property, which are authenticity, integrity, non-repudiability, confidentiality, availability, and authorization.

*1) Spoofing attacks:* Spoofing ruins authenticity by falsifying data to fool the autonomous driving system. In these attacks, the attacker masquerades with a legitimate source of information and inserts fake sensor data to force the vehicle to make wrong decisions. Possible spoofing attacks on AVs can be:

**LiDAR spoofing:** LiDAR spoofing attacks make the obstacles appear closer or farther than the actual distance. Usually, the LiDARs wait to listen for the reflected signals and consider it closer if the reflected signal arrives earlier. The attackers inject counterfeit signals in this wait window. They can delay the original signal before relaying it to take control of the object position [14]. Multiple fake copies of one real object make the attack even worse. LiDAR spoofing attacks can also create fake objects, hide objects, or reduce the performance of object detection algorithms that use LiDAR point clouds. Placing a spoofed object on top of the target causes failure of LiDAR-based object detection algorithms [15].

**Radar spoofing:** Radars are used to determine the obstacle distance by probing a signal in a lower resolution than LiDARs. In a scenario of two moving vehicles, the forward vehicle transmits a fake signal to the following vehicle that provides a wrong idea regarding the distance between them [16]. Hence, the victim car considers the distance as different than the actual.

**Ultrasonic sensor spoofing:** In an ultrasonic sensor spoofing attack, the attacker carefully generates ultrasonic signals that appear legitimate [17]. These signals are identical to the original sensor signals in terms of frequency, amplitude, modulation, etc. The spoofed ultrasonic signal timing is adjusted carefully to create fake obstacles or alter distance as they are considered authentic by autonomous vehicles.

**GPS spoofing:** GPS spoofing misleads GPS receivers by broadcasting realistic and valid but incorrect GPS signals. Initially, the attacker broadcasts a valid GPS signal that is synchronized with the original signals received by the GPS receiver. Then the attacker increases the signal power and gradually modifies the position to spoof the GPS signal. The GPS receiver usually considers the strongest signal available and hence starts using the spoofed GPS signal. Extended Kalman Filter can be exploited using GPS spoofing to perform lane departure attack where spoofed GPS signal overtakes inputs from other sources such as LiDAR [18].

*2) Tampering attacks:* Tampering attacks perform unauthorized updates or alterations to any sensor data. These are attacks on the integrity of the system. Potential tampering attacks are:

**Absorbing laser pulses:** LiDARs use laser pulses to sense the obstacles' distance and depth from the reflection of the pulses. LiDARs can only detect things if the target reflects the light. An attacker can intentionally absorb the laser beam that the LiDARs throw. As no reflected light returns to the LiDAR, it cannot detect the object [14].

**Adversarial attacks:** Adversarial attacks aim to fool the deep learning models where they generate adversarial examples by adding small noise or perturbations to the original data. The adversarial examples can be two types depending on the target. Untargeted attacks force the model to predict any other class than the original one, while the targeted attacks label the adversarial example to a specific class. If the machine learning model is denoted by $M$, an adversarial example is $x^{'}$, and output label is $y$, then $M(x^{'}) \neq y$ denotes untargeted attack and $M(x^{'}) = y^{'}$ expresses the targeted attack. Here, $y^{'}$ denotes the targeted attack label, and $y$ is the true label. LiDAR-based perception can be victim of adversarial attack [10], where adversarial LiDAR point clouds fool the model by minimizing the loss and making the model biased to the spoofed points. Different adversarial point cloud generation techniques [19] have made adversarial attacks more feasible. The adversarial objects may remain undetected by the AD system.

Advanced driving assistance systems depend on cameras for driving, which can be vulnerable to subtle adversarial manipulation of images. The attacker maximizes the steering angle and deviation from the road by manipulating the images analyzed by the deep learning models [20].

**Trojan attacks:** In a Trojan attack, the attacker collects a learned deep learning model and inject malicious behavior [21]. The model is retrained with carefully generated trojan trigger examples which force the model to learn a malicious behavior, such as making a u-turn in the middle of the road upon finding a particular sign.

**Fault injection attacks:** The fault injection attacks identify the safety-critical situations and faults that can lead to a potential accident. A machine learning-based fault injector algorithm can identify the scenario to launch an attack through injecting bias to Extended Kalman Filter [12], corrupting the hardware and software states [11], and so on. Fault injection can force the AV to move out, move in, or depart from the lane.

*3) Repudiation attacks:* Repudiation attacks refer to denying after performing an action. The repudiation attack can be performed by the car user or the manufacturing company during a forensic investigation case [22]. Here, the car user may claim that the investigation logs belong to another user. The manufacturing company may also claim that the logs belong to a vehicle from another manufacturer.

*4) Information disclosure attacks:* In information disclosure attacks, the attacker can hide the identity, acquire specific sensitive information regarding the victim, and use them later. Potential information disclosure attacks on AV can be:

**Cache side-channel attack:** An attacker can install malicious software inside the victim AV and perform a cache side-channel attack to predict the destination [23]. The malicious software runs in the same processor as the AD software. It exploits the correlation between physical state and cache access patterns to infer the victim AV's movement.

**Diagnosing deep learning model:** All the information regarding the deep learning model of AD perception is confidential. The attacker diagnoses the model to find and disclose its details, such as its architecture and weights.

**Sensor information disclosure:** Attackers can diagnose the sensors to detect the hardware specifications (i.e., LiDAR beams, camera focal length, etc). Knowledge of the specifications helps the attacker to launch sophisticated attacks.

*5) Denial of service attacks:* The denial of service attacks disrupts the service availability, performance, and efficiency of the system. Attackers can disrupt any of the perception, prediction, or control steps to make the AD system unavailable. Potential denial of service attacks are:

**AV freezing and emergency brake attack:** In an AV freezing attack, a spoofed front near an obstacle can freeze the AV while the vehicle is waiting at the traffic signal. In this attack, the vehicle does not move even after the traffic signal turns green [10]. In the emergency brake attack, the attacker spoofs a front-near obstacle to a moving AV that forces the car to decide to stop within a very short time [10].

**Jamming ultrasonic sensors:** An attacker can jam the ultrasonic sensors by generating fake ultrasonic echo pulses. The generated signals are stronger than the original pulses that cause the sensors to stop working. Usually, the ultrasonic sensors do not expect strong interference, and hence the strong pulses can launch the attack to jam the ultrasonic sensors.

**Camera blinding:** Attackers use a laser beam to beam to partially or fully blind the camera [24] and hide the object consequently. Camera blinding attack depends on environmental light, the light source used in the blinding attack, and the distance between the light source and camera [14].

**LiDAR blinding:** Saturating the LiDAR with sufficient light intensity can blind the LiDAR [25]. Attackers use a light source of the same wavelength as the LiDARs and can focus it onto the target LiDAR.

*6) Elevation of privilege attacks:* Elevation of privilege refers to gaining unauthorized privileges or access in a system that the attacker does not suppose to have. These attacks ruin authorization, which is a critical security property. Possible elevation of privilege attacks on AV are:

**Passive keyless entry and start (PKES):** The attacker can enter and start the car by relaying messages between the smart key and vehicle [9]. Both wired and wireless mechanism is suitable for such kind of attacks.

**Gaining access to camera:** The attacker can access and modify the live feed of the camera used by the perception module of an AV [12]. A man-in-the-middle strategy allows the attacker to gain camera access and modify the data.

**Publish subscribe overprivilege:** Different autonomous driving modules exchange messages among themselves using Robot Operating System (ROS), which works in a publish-subscribe architecture. However, overprivileged publisher and subscriber nodes can be exploited to launch different attacks on AV [26].

### *B. Physical attacks:*

The STRIDE threat modeling process considers the assets are physically protected from the attacker. However, AV assets can be exposed physically to the attacker. Physical attacks on AVs are listed as follows:

**Adding stickers to traffic signs:** One major physical attack on AV system is adding stickers on various traffic signals [27]. In this attack, the attacker does not directly control AV; however, such perturbations may lead to potential accidents.

**Damaging sensors:** Most important sensors of autonomous vehicles such as cameras, LiDAR, radar, and ultrasonic sensors are installed outside the AV. The attacker can damage any of these sensors that can force the vehicle not to work properly.

## VII. Mitigation strategies

Defining mitigation strategies is the final step of threat modeling. Mitigation strategies are defined to prevent attacks and mitigate threats and vulnerabilities. Potential mitigation strategies are as follows:

**System level defense:** For system-level defense against AV attacks, the manufacturers can install a separate intrusion detection system in AVs. Using physical invariant can be helpful for such system-level defense [2]. Understanding the AV's physical properties and analyzing the incurred deficiencies in sensor data can help design an online anomaly detection algorithm. Forensics investigation framework for storing and analyzing the events that occur inside the vehicle can help identify the potential reasons behind an incident [22].

Independent localization results from different positioning sources can be cross-checked to reduce the Extended Kalman Filter (EKF) based attacks. For example, a spoofed GPS signal can take over the EKF output due to high noise in LiDAR data and performs a lane departure attack. In that case, it should be detectable by the camera-based lane detection [18].

**Sensor level defense:** Sensor-level defense mechanisms can improve security by fusing different sensor data [28], which undoubtedly increases the attacker's effort. The camera blinding attack can be defended by integrating a removable near-infrared-cut filter into the camera. However, the filter is only effective during the daytime as the AV needs infrared light for night vision. Authentication and integrity verification techniques can be used together to prove the authenticity of the GPS signal source that can protect the GPS device from spoofing attack [29]. Both the radar and LiDAR spoofing attacks can be detected and thwarted by physical challenge-response authentication. Multiple methods can be applied to mitigate the attack, such as the recursive least square method [30] and the Spatio-temporal challenge-response method [16].

**Machine learning-based defense:** The deep learning model can be trained using adversarial examples to make the model more robust against adversarial attacks [31]. For this purpose, adversarial examples can be generated and used with original data to conduct adversarial training. Another possible mitigation strategy is analyzing the distribution of wrongly predicted results. The Trojan attack forces the model to make the wrong decisions for some specific scenarios. Analyzing the wrong predictions' distribution can be highly effective as one of the outputs will be the majority [21].

Cooperative perception in AD is another promising approach for mitigating attacks on the AVs [32]. The edge computing platforms of AVs can share high-level features among them and achieve broader perception [33], [34]. However, cooperative perception requires additional computation complexity and support infrastructure deployments, such as vehicle-to-everything (V2X) communication and roadside edge servers.

## VIII. Related works

Threat model analysis is a process to identify and prioritize the potential threats and vulnerabilities of a particular system. In this paper, we have used the STRIDE threat modeling process. Several other threat modeling approaches have been proposed in the literature, such as the goad-oriented approach [35], attack tree-based approach [36], collaborative attack modeling approach, etc. Threat modeling has been used in different other domains. Engoulou et al. [37] analyzed all the components of threat modeling for the vehicular ad-hoc network (VANET). Hoque *et al.* explored the security, and threat model of fog computing-assisted VANET [32], [38]. Security and privacy issues of cloud and fog computing have also been explored using threat modeling [39].

Security of AVs is an emerging research field as new security vulnerabilities, and attack surfaces are constantly being exposed. Researchers have designed threat models considering specific attacks on AV [10], [13], [18]. However, such attack-specific threat models do not consider all the security aspects of AVs. Besides attack-specific threat modeling, several research works have explored overall AV hardware and software security issues.

Authors of [40] listed the potential cyberattacks on connected and automated vehicles. Ren et al. [41] explored the security issues and attack strategies of AVs. They have considered different attacks on sensors and in-vehicle systems. Liu et al. [42] analyzed different security issues in the context of CAVs. While these research works address various security issues, there is no structured approach towards designing a comprehensive threat model for AVs. This paper addresses the issue by exploring all the components of AV threat modeling.

## IX. Conclusion

Proper threat modeling is the first and most important step to design an effective security solution. Unstructured analysis using only brainstorming, focusing on a particular problem that has just occurred, or depending on previous experiences does not provide a complete threat model for an emerging cyber-physical system such as autonomous vehicles. In this work, we present a comprehensive threat model for the AD system. We analyze the critical assets and the attacker model. We also analyze and categorize the potential attacks on the AD systems based on the STRIDE threat modeling process. Finally, we explore the possible mitigation strategies for the attacks. We hypothesize this threat modeling will help develop secure AVs by enabling the development of better security solutions.

## Acknowledgement

This research was supported by the National Science Foundation through awards ACI-1642078 and CNS-1351038.

## References

[1] Fortune, "Waymo reaches 20 million miles of autonomous driving," 2020. [Online]. Available: https://fortune.com/2020/01/07/googles-waymo-reaches-20-million-miles-of-autonomous-driving/

[2] R. Quinonez, J. Giraldo, L. Salazar, and E. Bauman, "Savior: Securing autonomous vehicles with robust physical invariants," *29th Usenix Security Symposium*, 2020.

[3] S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu, "A survey of deep learning techniques for autonomous driving," *Journal of Field Robotics*, vol. 37, no. 3, pp. 362–386, 2020.

[4] S. Myagmar, A. J. Lee, and W. Yurcik, "Threat modeling as a basis for security requirements," in *Symposium on requirements engineering for information security (SREIS)*, vol. 2005. Citeseer, 2005, pp. 1–8.

[5] R. Hasan, S. Myagmar, A. J. Lee, and W. Yurcik, "Toward a threat model for storage systems," in *Proc. of the 2005 ACM workshop on Storage security and survivability, VA, USA*. ACM, 2005, pp. 94–102.

[6] Microsoft, "The stride threat model," 2009. [Online]. Available: https://docs.microsoft.com/en-us/previous-versions/commerce-server/ee823878(v=cs.20)

[7] F. Swiderski and W. Snyder, *Threat Modeling*. Microsoft Press Redmond, WA, USA, 2004.

[8] S. D. Pendleton, H. Andersen, X. Du, X. Shen, M. Meghjani, Y. H. Eng, D. Rus, and M. H. Ang, "Perception, planning, control, and coordination for autonomous vehicles," *Machines*, vol. 5, no. 1, p. 6, 2017.

[9] A. Francillon, B. Danev, and S. Capkun, "Relay attacks on passive keyless entry and start systems in modern cars," in *Proc. of NDSS*, 2011.

[10] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, and Z. M. Mao, "Adversarial sensor attack on lidar-based perception in autonomous driving," in *Proceedings of the 2019 ACM CCS*, 2019, pp. 2267–2281.

[11] S. Jha, S. Banerjee, T. Tsai, S. K. Hari, S. W. Keckler, and R. K. Iyer, "Ml-based fault injection for autonomous vehicles: a case for bayesian fault injection," in *2019 49th Annual IEEE/IFIP DSN*. IEEE, 2019.

[12] S. Jha, S. Cui, T. Tsai, Z. Kalbarczyk, and R. Iyer, "Ml-driven malware that targets av safety," *arXiv preprint:2004.13004*, 2020.

[13] J. Sun, Y. Cao, Q. A. Chen, and Z. M. Mao, "Towards robust lidar-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures," in *29th USENIX Security Symposium*, 2020.

[14] J. Petit, B. Stottelaar, M. Feiri, and F. Kargl, "Remote attacks on automated vehicles sensors: Experiments on camera and lidar," *Black Hat Europe*, vol. 11, p. 2015, 2015.

[15] J. Tu, M. Ren, S. Manivasagam, M. Liang, F. Cheng, and R. Urtasun, "Physically realizable adversarial examples for lidar object detection," in *Proceedings of the IEEE/CVF CVPR*, 2020, pp. 13 716–13 725.

[16] P. Kapoor, A. Vora, and K.-D. Kang, "Detecting and mitigating spoofing attack against an automotive radar," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2018, pp. 1–6.

[17] W. Xu, C. Yan, X. Ji, and J. Liu, "Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles," *IEEE IoT Journal*, vol. 5, no. 6, pp. 5015–5029, 2018.

[18] J. Shen, J. Y. Won, Z. Chen, and Q. A. Chen, "Drift with devil: Security of multi-sensor fusion based localization in high-level autonomous driving under {GPS} spoofing," in *29th {USENIX} Security Symposium*, 2020.

[19] C. Xiang, C. R. Qi, and B. Li, "Generating 3d adversarial point clouds," in *Proc. of IEEE CVPR*, 2019, pp. 9136–9144.

[20] A. Boloor, K. Garimella, X. He, C. Gill, Y. Vorobeychik, and X. Zhang, "Attacking vision-based perception in end-to-end autonomous driving models," *Journal of Systems Architecture*, p. 101766, 2020.

[21] Y. Liu, S. Ma, Y. Aafer, W.-C. Lee, J. Zhai, W. Wang, and X. Zhang, "Trojaning attack on neural networks," 2017.

[22] M. A. Hoque and R. Hasan, "Avguard: A forensic investigation framework for autonomous vehicles," in *IEEE ICC*. IEEE, 2021.

[23] M. Luo and A. C. Myers, "Stealthy tracking of autonomous vehicles with cache side channels," in {*USENIX*} *Security 20*, 2020, pp. 859–876.

[24] C. Yan and J. Liu, "Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle," *DEF CON*, vol. 24, 2016.

[25] H. Shin, D. Kim, Y. Kwon, and Y. Kim, "Illusion and dazzle: Adversarial optical channel exploits against lidars for automotive applications," in *International Conference on CHES*. Springer, 2017, pp. 445–467.

[26] D. K. Hong, J. Kloosterman, and Z. M. Mao, "Avguardian: Detecting and mitigating publish-subscribe overprivilege for autonomous vehicle systems," in *2020 EuroS&P*. IEEE, 2020, pp. 445–459.

[27] K. Eykholt, I. Evtimov, E. Fernandes, T. Kohno, and D. Song, "Robust physical-world attacks on deep learning visual classification," in *Proceedings of the IEEE CVPR*, 2018, pp. 1625–1634.

[28] K. Yang, R. Wang, Y. Jiang, H. Song, C. Luo, Y. Guan, X. Li, and Z. Shi, "Sensor attack detection using history based pairwise inconsistency," *Future Generation Computer Systems*, vol. 86, pp. 392–402, 2018.

[29] V. L. Thing and J. Wu, "Autonomous vehicle security: A taxonomy of attacks and defences," in *IEEE iThings*. IEEE, 2016, pp. 164–170.

[30] R. G. Dutta, X. Guo, and Y. Jin, "Estimation of safe sensor measurements of autonomous system under attack," in *Proc. of 54th Annual Design Automation Conference*, 2017, pp. 1–6.

[31] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," *arXiv preprint arXiv:1412.6572*, 2014.

[32] M. A. Hoque and R. Hasan, "Towards an analysis of the architecture, security, and privacy issues in vehicular fog computing," in *2019 SoutheastCon*. IEEE, 2019.

[33] Q. Chen, X. Ma, Q. Yang, and S. Fu, "F-cooper: feature based cooperative perception for autonomous vehicle edge computing system using 3d point clouds," in *Proceedings of the 4th ACM/IEEE SEC*, 2019, pp. 88–100.

[34] M. A. Hoque and R. Hasan, "R-cav: On-demand edge computing platform for connected autonomous vehicles," in *IEEE 7th WF-IoT*. IEEE, 2021.

[35] E. A. Oladimeji, S. Supakkul, and L. Chung, "Security threat modeling and analysis: A goal-oriented approach," in *Proc. SEA '06*. ACTA Press, 2006, pp. 178 – 185.

[36] V. Saini, Q. Duan, and V. Paruchuri, "Threat modeling using attack trees," *Journal of Computing Sciences in Colleges*, vol. 23, pp. 124–131, 2008.

[37] R. G. Engoulou, M. Bellaïche, and A. Quintero, "Vanet security surveys," *Journal of Network and Computer Applications*, vol. 44, pp. 1–13, 2014.

[38] M. A. Hoque and R. Hasan, "Towards a threat model for vehicular fog computing," in *IEEE UEMCON*. IEEE, 2019, pp. 1051–1057.

[39] Y. Karim and R. Hasan, "Towards a threat model for fog computing," in *IEEE UEMCON*. IEEE, 2019, pp. 1110–1116.

[40] J. Petit and S. E. Shladover, "Potential cyberattacks on automated vehicles," *IEEE Transactions on ITS*, vol. 16, no. 2, pp. 546–556, 2014.

[41] K. Ren, Q. Wang, C. Wang, Z. Qin, and X. Lin, "The security of autonomous driving: Threats, defenses, and future directions," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 357–372, 2019.

[42] S. Liu, L. Liu, J. Tang, B. Yu, Y. Wang, and W. Shi, "Edge computing for autonomous driving: Opportunities and challenges," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1697–1716, 2019.