Addressing Vulnerability of Sensors for Autonomous Driving

Yi Zhu[1], Foad Hajiaghajani[1], Changzhi Li[2], Lu Su[3], Wenyao Xu[1], Zhi Sun[4], and Chunming Qiao[1]
[1]State University of New York at Buffalo, Buffalo, NY USA [2]Texas Tech University, Lubbock, TX USA
[3]Purdue University, West Lafayette, IN USA [4]Tsinghua University, Beijing, China

In recent years, connected and autonomous vehicles has gain increasing attentions. The industries have been started to test and operate their autonomous vehicles on real roads. The automation levels of the autonomous vehicles can be divided into 6 ranges: from fully manual (Level 0) to fully autonomous (Level 5). Among them, the autonomous vehicles in Level 3 and 4 are especially attractive. In autonomous vehicles with Level 3 and 4, a fundamental part is the perception systems. The perception systems collect data from various sensors such as GPS, camera, radar, and LiDAR, to understand the surrounding environments. The autonomous vehicles then make driving decisions based on the perception results. Thus, it is essential to study the vulnerability of the perception systems.

Recent studies have found that the sensors of the perception systems can be spoofed by injecting signals and the adopted deep neural networks (DNNs) can be fooled by adversarial attacks [6]. For example, the GPS sensor can be spoofed to generate wrong locations by injecting special GPS signals [9, 11]. The visual perception systems such as traffic sign perception systems, can be fooled by placing some stickers on the traffic signs [4]. In addition to GPS and camera, radar and LiDAR are playing increasingly important role in AV's perception systems, and their vulnerabilities are also studied by some recent works. In [3, 2], radar is spoofed by injecting false signals to change the detection results. The LiDAR perception systems can be attacked by both laser-based method and object-based method. In laser-based methods [1, 7], the authors combine the spoofing attack and the adversarial attack to fool the LiDAR perception systems. The attacker can inject some fake points in the LiDAR point cloud, and make these fake points wrongly recognized as a vehicle by the perception system. In object-based method [10], the attacker can generate a specially-shaped object, and place it on the rooftop of a vehicle. This object can fool the DNNs and hide the vehicle from the LiDAR perception system.

For radar perception systems, although existing attacks can spoof the radar sensors, the adversarial impact to the driving decisions of the autonomous vehicle is not well studied. In our recent work [8], we performed end-to-end analysis on the vulnerability of radar detection system at driving decision level in real-world driving scenarios. We proposed a spoofing attack method, where the attacker can inject some special signals into the radar sensor using a radar jammer. By spoofing the radar sensors, a fake obstacle can be created at any locations, and the location of an existing obstacle can be changed in the radar perception. The decision-level analysis shows that the autonomous vehicles make dangerous driving decisions after the attacks, which demonstrates the potential security threats of the proposed attacks. To mitigate the attacks, we further proposed a challenge-response authentication scheme and a RF fingerprinting scheme to detect the spoofing signals.

We also propose to improve the security of millimeter-wave radar based perception systems in the presence of advanced attacks against fast-chirp radars. Specifically, we proposed to use a RF mixer to introduce a frequency shift to the RF signal, and create a false target [5]. A hybrid-chirp FMCW approach was proposed as a countermeasure. We also demonstrated that advanced attacks can be launched using not only adaptive finite state machine based approaches that combine inter- and intra-chirp-sequence synchronization, but also machine learning techniques to efficiently adjust the attacker's waveform parameters. To counter such attacks, we explore the use of unique RF fingerprint based physical-layer authentication to identify attacks.

For LiDAR perception systems, existing attacks face many challenges when performed in practice. In laser-based methods, shooting laser signals to a moving LiDAR precisely is not easy. And in existing object-based methods, the adversarial objects have to be in some specific shapes and sizes, which limits the attack flexibility. To address these challenges, in [13], we proposed a flexible attack method that can be easily performed in real driving scenarios and can hide a vehicle from the LiDAR object detection system. In our proposed attack, the attacker can use arbitrary objects such as drones as the adversarial objects, instead of specially-shaped objects. As shown in Figure 1, after controlling these drones to hover at some adversarial locations around a target vehicle, the DNN of LiDAR detection model is fooled and the target vehicle is hidden. Our experiments in physical world show that such attacks is easy to perform and effect in fooling the detection results. To mitigate the attacks, we discussed potential defense strategies such as sensor fusion. And based on the characteristics of adversarial locations, we also proposed a novel defense method to mitigate the attacks.

In LiDAR perception systems, another important task is LiDAR semantic segmentation, which classifies each point to a certain class in the LiDAR point cloud. It provides point-level information about the surrounding environments and has many applications such as obstacle detection and road boundary extraction. Thus, the security of LiDAR

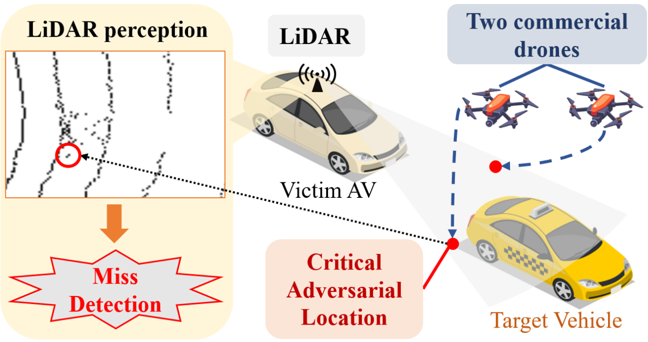
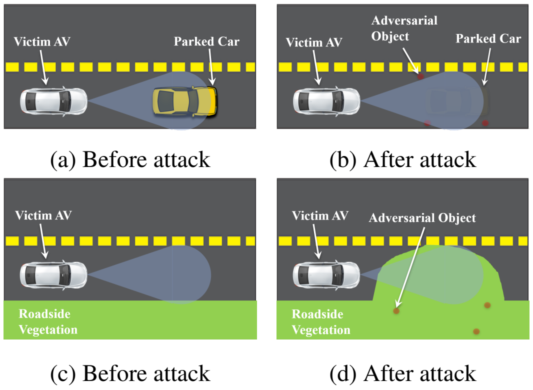


Figure 1: An example of the proposed attack in [13].



(a) Before attack

(b) After attack

(c) Before attack

(d) After attack

Figure 2: Vehicle hiding attack (first row) and road surface changing attack (second row) in [12].

semantic segmentation has broad impacts. In [12], we proposed the first study on the vulnerability of LiDAR semantic segmentation. In the proposed attacks, the attacker maliciously places some traffic signs and cardboard at some specific locations. Such attacks can hide a vehicle or change the road to vegetables from the LiDAR semantic segmentation systems, as shown in Figure 2. To mitigate the attacks, we proposed a model aggregation method, which trains multiple models under different parameter initialization and aggregates the outputs of these models as the final outputs.

To conclude, we investigated the vulnerability of both radar and LiDAR perception systems. We proposed spoofing attacks against radar perception system that can create fake obstacles and change the obstacles' locations. We proposed to use arbitrary objects to hide an existing vehicle from the LiDAR object detection systems. And we proposed the attacks against LiDAR semantic segmentation systems by using traffic signs and cardboard. We investigated potential defense strategies to mitigate these attacks.

References

[1] Yulong Cao, Chaowei Xiao, Benjamin Cyr, Yimeng Zhou, Won Park, Sara Rampazzi, Qi Alfred Chen, Kevin Fu, and Z Morley Mao. Adversarial sensor attack on lidar-based perception in autonomous driving. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 2267–2281, 2019.

[2] Ruchir Chauhan. *A platform for false data injection in frequency modulated continuous wave radar*. Utah State University, 2014.

[3] Ruchir Chauhan, Ryan M Gerdes, and Kevin Heaslip. Demonstration of a false-data injection attack against an fmcw radar. *Embedded Security in Cars (ESCAR)*, 2014.

[4] Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Robust physical-world attacks on deep learning visual classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1625–1634, 2018.

[5] Prateek Nallabolu and Changzhi Li. A frequency-domain spoofing attack on fmcw radars and its mitigation technique based on a hybrid-chirp waveform. *IEEE Transactions on Microwave Theory and Techniques*, 69(11):5086–5098, 2021.

[6] Fabio Pierazzi, Feargus Pendlebury, Jacopo Cortellazzi, and Lorenzo Cavallaro. Intriguing properties of adversarial ml attacks in the problem space. In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 1332–1349. IEEE, 2020.

[7] Jiachen Sun, Yulong Cao, Qi Alfred Chen, and Z Morley Mao. Towards robust lidar-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures. In *29th {USENIX} Security Symposium ({USENIX} Security 20)*, pages 877–894, 2020.

[8] Zhi Sun, Sarankumar Balakrishnan, Lu Su, Arupjyoti Bhuyan, Pu Wang, and Chunming Qiao. Who is in control? practical physical layer attack and defense for mmwave-based sensing in autonomous vehicles. *IEEE Transactions on Information Forensics and Security*, 16:3199–3214, 2021.

[9] Nils Ole Tippenhauer, Christina Pöpper, Kasper Bonne Rasmussen, and Srdjan Capkun. On the requirements for successful gps spoofing attacks. In *Proceedings of the 18th ACM conference on Computer and communications security*, pages 75–86, 2011.

[10] James Tu, Mengye Ren, Sivabalan Manivasagam, Ming Liang, Bin Yang, Richard Du, Frank Cheng, and Raquel Urtasun. Physically realizable adversarial examples for lidar object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13716–13725, 2020.

[11] Kexiong Curtis Zeng, Shinan Liu, Yuanchao Shu, Dong Wang, Haoyu Li, Yanzhi Dou, Gang Wang, and Yaling Yang. All your {GPS} are belong to us: Towards stealthy manipulation of road navigation systems. In *27th {USENIX} security symposium ({USENIX} security 18)*, pages 1527–1544, 2018.

[12] Yi Zhu, Chenglin Miao, Foad Hajiaghajani, Mengdi Huai, Lu Su, and Chunming Qiao. Adversarial attacks against lidar semantic segmentation in autonomous driving. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, 2021.

[13] Yi Zhu, Chenglin Miao, Tianhang Zheng, Foad Hajiaghajani, Lu Su, and Chunming Qiao. Can we use arbitrary objects to attack lidar perception in autonomous driving? In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS '21)*, 2021.