Poster: Infrastructure-Aided Defense for Autonomous Driving Systems: Opportunities and Challenges

Yunpeng Luo
UC Irvine
yunpel3@uci.edu

Ningfei Wang
UC Irvine
ningfei.wang@uci.edu

Bo Yu
PerceptIn
bo.yu@perceptin.io

Shaoshan Liu
PerceptIn
shaoshan.liu@perceptin.io

Qi Alfred Chen
UC Irvine
alfchen@uci.edu

***Abstract*—Autonomous Driving (AD) is a rapidly developing technology and its security issues have been studied by various recent research works. With the growing interest and investment in leveraging intelligent infrastructure support for practical AD, AD system may have new opportunities to defend against existing AD attacks. In this paper, we are the first to systematically explore such a new AD security design space leveraging emerging infrastructure-side support, which we call Infrastructure-Aided Autonomous Driving Defense (I-A2D2). We first taxonomize existing AD attacks based on infrastructure-side capabilities, and then analyze potential I-A2D2 design opportunities and requirements. We further discuss the potential design challenges for these I-A2D2 design directions to be effective in practice.**

I. Introduction

As Autonomous Driving (AD) technology becomes increasingly deployed and commercialized in the real world, more and more people start to consider the security of AD vehicles. There are a lot of researches trying to create adversarial examples for fooling AI components in AD systems. On the other hand, the AD system design patterns are also evolving recently, with growing interests and investment in leveraging infrastructure-side support. Specifically, a new direction of AD design called Infrastructure-Aided Autonomous Driving (IAAD) is being developed recently, which uses infrastructure side communication and sensing abilities to improve AD reliability while reducing on-board sensing cost [1]. Today, there are many ongoing IAAD testing, and even deployment efforts by companies and institutes. In terms of deployment scenarios, IAAD is found particularly attractive to and thus likely to be first utilized by robo-taxi/ride-hailing services due to the cost considerations [2].

Considering such a new AD design trend, it is important to explore whether and how it may influence the existing AD security design space. Specifically, since the infrastructure-side sensing support can provide extra information about the real-time driving environment/condition to AD vehicles, we are wondering whether there are new opportunities to defend against existing AD attacks. To the best of our knowledge, no prior work has systematically explored how such a new AD design trend can be leveraged for AD defense purposes, and what requirements have to be met for such defense designs to be effective, especially in practical settings. In this paper, we are the first to systematically discuss the opportunities and challenges for such a new AD security design space leveraging emerging infrastructure-side support, which we call Infrastructure-Aided Autonomous Driving Defense (I-A2D2). We survey existing AD attacks and taxonomize them into three categories from the infrastructure-side capability perspective. For each category, we analyze the requirements for the infrastructure to enable defense capabilities, and propose I-A2D2 design directions. After that, we discuss potential challenges for these I-A2D2 design directions to be effective in practice.

II. I-A2D2 Design Opportunities

With infrastructure-side sensing and communication abilities, many existing attacks against AD systems can potentially have new defense opportunities. Due to the different nature of the attacks, different infrastructure-side capabilities can be required for effective and systematic I-A2D2 defense designs. We thus started by performing a comprehensive survey of AD attacks published in recent years, and classified them into 3 categories from such I-A2D2 design requirement perspective:

(A1) Perception of infrastructure-authoritative information. Like human drivers, AD has to follow traffic rules, such as obeying traffic signs and light. Quite some existing works target attacking such information, e.g., hiding or spoofing STOP signs. However, since such information are under authoritative controlled by the government transportation agencies, the infrastructure is able and also authoritative to provide such information. This can at least provide another (if not more trustworthy due to the source authoritativeness) information source to AD vehicles, which thus can enable at least direct attack detection capabilities against all such existing attack vectors in this category.

Requirement: To inform the AD vehicle of authoritative information, the infrastructure must be able to communicate with the AD vehicle in time. Also, since the information to transmit is known in advance, the infrastructure can send it before the AD vehicle needs to react to the information. Thus, if designed properly, the infrastructure should be able to always ensure the information can arrive in time to the AD vehicle side to enable effective defense design opportunities.

(A2) Perception of dynamic road objects. To avoid collision, AD vehicles need to detect dynamic road objects and avoid them proactively. Some attacks aim at the obstacle detection component and try to hide, relocate, or create non-existing objects in front of the victim. The others directly disable the obstacle detection component. To defend against such attacks, the infrastructure-side perception capabilities (e.g., camera and LiDAR) can be leveraged to help perceive maliciously hidden objects or eliminate attacker-introduced fake objects. The AD vehicle side can fuse their own detection results with such infrastructure-side detection results, which can at least detect (if not able to correct) the attacked results.

Requirement: Defending against A2 requires better communication capabilities compared to A1 in both latency and bandwidth. The AD vehicle needs to be informed of the object

in time so that it can have adequate time to make decisions and react (e.g., stop before crash or change lane). Failing to do that, the infrastructure can end up sending obsolete information that can be even misleading in the extreme cases. Besides the communication capability requirements, the infrastructure-side perception also needs to be accurate enough.

(A3) Localization attacks. Localization is a key component for AD to accomplish tasks such as navigation and path planning. In our survey, we find 5 attacks targeting localization, which can cause severe consequences such as driving off the road and even crashing into the incoming vehicle from the opposite direction. To defend against such attacks, the infrastructure side can keep sending information such as locations of all its perceived in-road vehicles to the AD vehicle side. The AD vehicles will find out the location that represents itself for crosschecking. When the attack happens, there will be a mismatch between the AD vehicle's self-localization and such infrastructure-aided localization results; the AD vehicle can thus use the latter as an additional information source to at least perform attack detection. Note that such a design follows the trust-on-first-use (TOFU) assumption, i.e., assuming that the AD vehicle is not under localization attack in the first time it receives the infrastructure-side information.

Requirement: Defending against A3 attacks requires communication capability, as it requires both real-time infrastructure-side perception. In terms of localization accuracy, in the infrastructure-aide localization discussed above, performing localization of AD vehicle is essentially performing object detection like I-A2D2 defenses against A2. We can estimate that with a deviation of $\frac{LaneWidth-VehicleWidth}{2}$, it's possible to cause a vehicle to have lane departure. This is roughly $\frac{2.7-2.12}{2} = 0.29m$ using common lane and car widths. To achieve such infrastructure-aided localization capabilities, it thus requires the infrastructure side to have better sensing ability and more reliable algorithm to perceive objects from the input data in the real time.

III. I-A2D2 Design Challenges

As discussed above, the new infrastructure-aided AD design trend opens quite new and broad defense design spaces for all existing AD attacks. Meanwhile, we also notice several design challenges in this direction, which are discussed below.

Precise self-localization from infrastructure perception. As discussed earlier, to defend against localization attacks (A3), the infrastructure side needs to achieve sufficiently-accurate localization of pass-by vehicles. However, we find that achieving such required accuracy is non-trivial based on our preliminary experiments in a real IAAD-deployed road, which is around 1000 meters long with full IAAD coverage for testing purposes. We experiment with two LiDAR obstacle detection models on the collected data: (1) the built-in segmentation model used in Apollo 5.0 [3], an open source industry-grade AD system, which we denote as "Apollo5"; (2) PIXOR [4] from Uber ATG; (3) PointPillars [5] (only on part of the trace). The distribution of the errors of each frame and also their median are plotted in Fig. 1. As shown, the median error of Apollo5 is 0.68 m, while that of PIXOR is 0.82 m. Thus, both of the two models cannot meet the requirement identified earlier (0.29 m). We only tested PointPillars on part of the trace, where the distance from the infrastructure to the vehicle is also short. The median error of PointPillars is 0.22 m and can meet the requirement in the limited route. We plan to conduct more experiments on PointPillars and also other models.

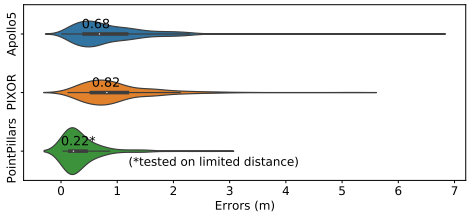


Fig. 1: Error distribution of both the Apollo LiDAR perception model ("Apollo5"), PIXOR and PointPillars.

Adaptive attacks. While I-A2D2 offers various new defense opportunities against existing attacks, once the attacker is aware of such designs, she can also consider I-A2D2-specific adaptive attack designs. We discuss a few such possibilities as follows: (1) *Attack infrastructure-side perception.* Since the sensors and the AI components used for such perception are similar to those used on AD vehicles, the attackers can apply/adapt existing vehicle-side perception attacks to the infrastructure side, or even attack both AD vehicle and infrastructure perception at the same time. (2) *Exploit fixed sensor positions.* Because IAAD sensors are in fixed positions, when facing the same sensor attack, infrastructure can be more vulnerable comparing to AD vehicle (e.g., no need to perform tracking and aiming for laser shooting attacks [6]). In a similar vein, generating adversarial examples can be easier as well, since it no longer requires taking the vehicle motion dynamics into consideration like in [7, 8]. (3) *New cyber-attack surface.* In IAAD/I-A2D2, the communication between infrastructure and AD vehicle can expose AD vehicle's interior system to other devices, which introduces a new cyber-attack surface. This thus calls for careful corresponding protocol designs.

IV. Conclusion

In this paper, we are the first to systematically discuss the opportunities and challenges for the new Infrastructure-Aided Autonomous Driving Defense (I-A2D2) design space. We first taxonomize existing AD attacks based on infrastructure-side capabilities, and then analyze potential I-A2D2 design opportunities and requirements. We further discuss the potential design challenges for these I-A2D2 design directions to be effective in practice. We hope that our discussions and insights can inspire more future research into this promising but currently under-explored defense design space for AD system security.

References

[1] S. Liu, B. Yu *et al.*, "Invited: Towards Fully Intelligent Transportation through Infrastructure-Vehicle Cooperative Autonomous Driving: Challenges and Opportunities," in *DAC*, 2021.

[2] "To Make Self-Driving Cars Safe, We Also Need Better Roads and Infrastructure," https://hbr.org/2018/08/to-make-self-driving-cars-safe-we-also-need-better-roads-and-infrastructure.

[3] "ApolloAuto/apollo: An open autonomous driving platform," https://github.com/ApolloAuto/apollo.

[4] B. Yang, W. Luo *et al.*, "PIXOR: Real-time 3D Object Detection from Point Clouds," in *CVPR*, 2018.

[5] A. H. Lang, S. Vora *et al.*, "PointPillars: Fast Encoders for Object Detection From Point Clouds," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

[6] Y. Cao, C. Xiao *et al.*, "Adversarial Sensor Attack on LiDAR-based Perception in Autonomous Driving," in *CCS*, 2019.

[7] T. Sato, J. Shen *et al.*, "Dirty Road Can Attack: Security of Deep Learning based Automated Lane Centering under Physical-World Attack," in *USENIX Security*, 2021.

[8] Y. Zhao, H. Zhu *et al.*, "Seeing isn't believing: Towards more robust adversarial attack against real world object detectors," in *CCS*, 2019.

Poster: Infrastructure-Aided Defense for Autonomous Driving Systems: Opportunities and Challenges

Yunpeng Luo[1], Ningfei Wang[1], Bo Yu[2], Shaoshan Liu[2], Qi Alfred Chen[1]

[1]University of California, Irvine (UCI) [2]PerceptIn

How can infrastructure help defend against attacks to Autonomous Driving (AD)?

- AD system design patterns are evolving recently: Growing interests & investment in **Infrastructure-Aided Autonomous Driving (IAAD)**
  - Use infrastructure-side communication & sensing abilities to improve AD reliability while reducing on-board sensing cost
  - Already many on-going IAAD testing & deployment efforts (e.g., by Baidu[1], Seoul Robotics[2], BMW[2], …)
  - Highly attractive to and may first be utilized by robo-taxi/ride-hailing services[3]
- Considering the IAAD design trend, important to explore ***whether* & *how* it may influence existing AD security design space**
- **First** to systematically discuss the opportunities & challenges for such a potentially new AD security design space:

I-A2D2

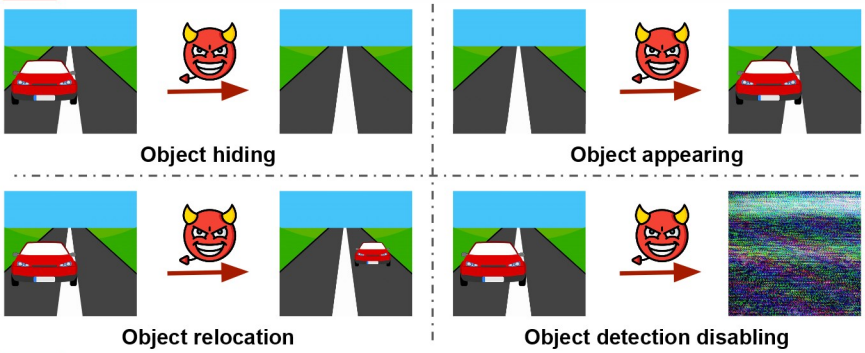
***I*nfrastructure-*A*ided *A*utonomous *D*riving *D*efense**

I-A2D2 Defense Direction: Opportunities

A1 Attack Perception of Infrastructure-Authoritative Information

Sign hiding | Sign appearing | Traffic light changing

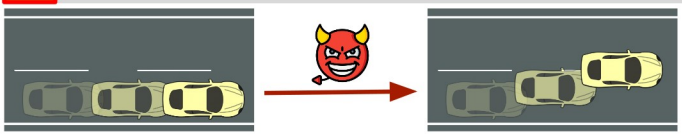
A2 Attack Perception of Dynamic Road Objects



Object hiding | Object appearing

Object relocation | Object detection disabling

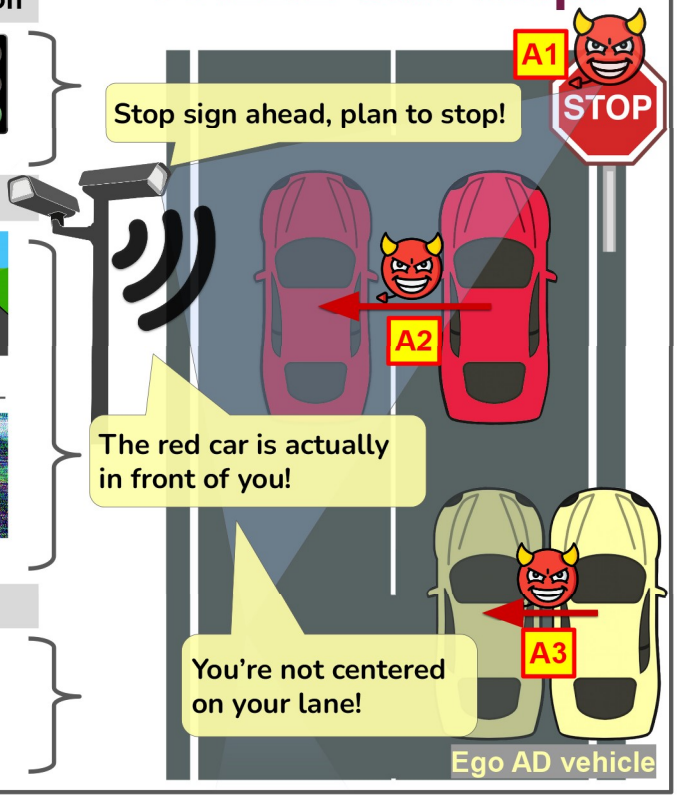
A3 Attack Localization



Trajectory shifting

The attacker can shift the victim's localization result and lead to wrong veering.

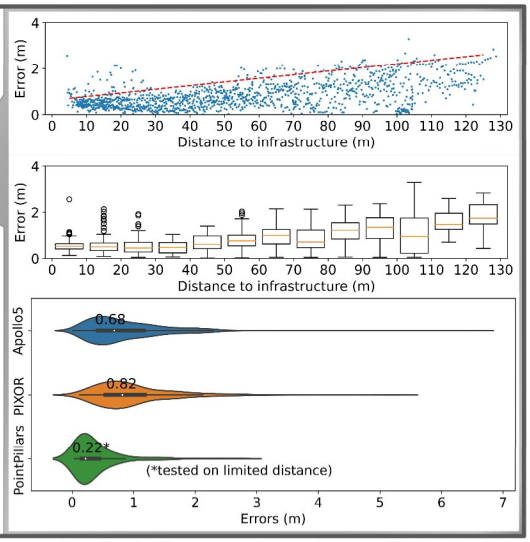
I-A2D2 can help!



I-A2D2 Defense Direction: Challenges

Precise Self-Localization from Infrastructure Perception

- Current object detection may not be precise enough for self-localization from infrastructure.
- Tested in real-world road deployed with IAAD (1,000 meters long, with LiDAR & camera)
- Initial results
  - Not enough for defending against A3.
  - Less precise for objects far away. Laser points become more sparse.
  - Less precise for nearby objects. The LiDAR is mounted in a higher position than the vehicles, requiring laser channels with large pitching angle.



Adaptive Attacks

- **Attack infrastructure-side perception**
  - Sensors and AI components used on infrastructure are the same as for AD vehicles → possible to apply the same attacks to infrastructure.
- **Exploit fixed sensor positions**
  - IAAD sensors are in fixed positions. Can be more vulnerable comparing to AD vehicle.
- **New cyber-attack surface**
  - Communication between infrastructure & AD vehicle can expose AD vehicle's interior system.

[1] "Baidu and Tsinghua U Introduce Apollo Air to Empower Autonomous Driving with Roadside Sensing," https://medium.com/apollo-auto/baidu-and-tsinghua-u-introduce-apollo-air-to-empower-autonomous-driving-with-roadside-sensing-99b1/b0bc/1.
[2] "Seoul Robotics' autonomous 'Control Tower' remotely manages self-driving vehicle fleets," https://www.engadget.com/the-level-5-control-tower-is-a-puppet-master-for-autonomous-vehicle-fleets-140041909.html.
[3] "To Make Self-Driving Cars Safe, We Also Need Better Roads and Infrastructure," https://hbr.org/2018/08/to- make- self- driving- cars- safe-we-also-need-better-roads-and-infrastructure.

Contact: Yunpeng Luo <yunpel3@uci.edu>