Cooperative Perception for Safe Control of Autonomous Vehicles under LiDAR Spoofing Attacks

Hongchao Zhang[1], Zhouchi Li[2], Shiyu Cheng[1], and Andrew Clark[1]

[1] Electrical and Systems Engineering Department, McKelvey School of Engineering, Washington University in St. Louis, St. Louis, MO 63130

{hongchao,cheng.shiyu,andrewclark}@wustl.edu

[2] Electrical and Computer Engineering Department, Worcester Polytechnic Institute, 100 Institute Rd, Worcester, MA 01609

zli4@wpi.edu

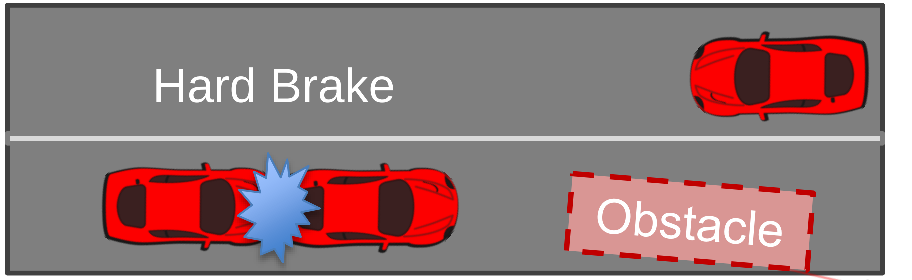
Washington University in St. Louis

Outline

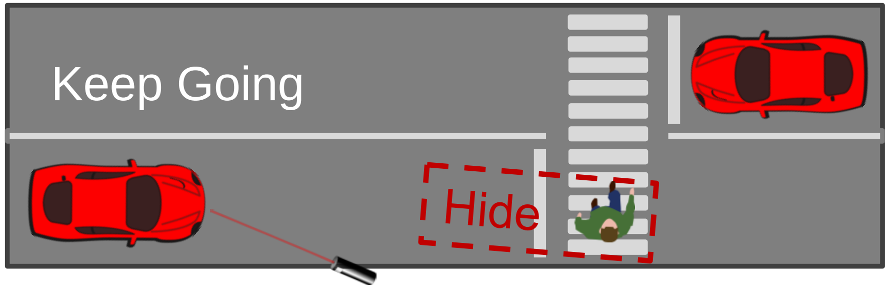
- Introduction
- Related Work
- Threat Model and Analysis
- Proposed Fault Detection, Identification and Isolation
- Case Study
- Conclusion

Introduction

- Autonomous vehicles rely on sensors to observe environment and make decisions.
- LiDAR sensors have been demonstrated to be vulnerable to spoofing attacks, e.g., [1],[2]



Falsifying non-existing obstacles



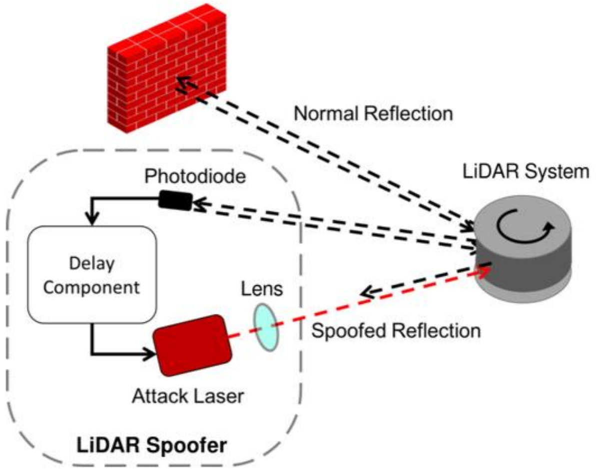
Hiding existing obstacles

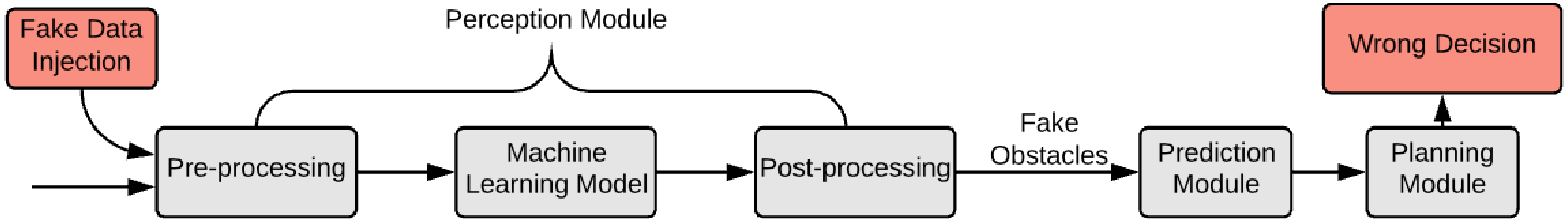
[1] Sun, Jiachen Sun, Yulong Cao Cao, Qi Alfred Chen, and Z. Morley Mao. "Towards robust lidar-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures." In USENIX Security Symposium (Usenix Security'20). 2020.

[2] Cao, Yulong, 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, pp. 2267-2281. 2019.

Types of LiDAR Spoofing

- Goal: causing errors in detection modules.
- Relay attack: spoofer fires laser beams to inject false data [1].
  - Compromise only one sensor and a narrow sector
- Adversarial objects: synthesized 3D printed objects [2]





[1] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao, "Adversarial sensor attack on LiDAR-based perception in autonomous driving," in ACM SIGSAC conference on Computer and Communications Security, 2019, pp. 2267–2281.

[2] Cao, Yulong, et al. "Adversarial objects against lidar-based autonomous driving systems." arXiv preprint arXiv:1907.05418 (2019).

Current State-of-the-art: Detection and Mitigation of LiDAR Spoofing

- Single sensor
  - Random sampling proposed in [1]
  - Randomize the pulses' waveforms [2]
- Redundancy-based approach
  - Fusion and overlapping [3]
- Cooperative perception
  - Connected Automated Vehicles [4]

Increase in cost

Focus on the single-agent case

Leave LiDAR spoofing less studied

[1] Davidson, Drew, et al. "Controlling UAVs with Sensor Input Spoofing Attacks." WOOT. 2016..

[2] Matsumura, Ryuga, Takeshi Sugawara, and Kazuo Sakiyama. "A secure LiDAR with AES-based side-channel fingerprinting." 2018 Sixth International Symposium on Computing and Networking Workshops (CANDARW). IEEE, 2018.

[3] Yeong, De Jong, et al. "Sensor and sensor fusion technology in autonomous vehicles: A review." Sensors 21.6 (2021): 2140.

[4] Bouchouia, Mohammed Lamine, et al. "A Simulator for Cooperative and Automated Driving Security."

Contributions

- Propose a cooperative, multi-vehicle approach to detecting LiDAR spoofing attacks
- We develop a Fault Detection, Identification, and Isolation procedure (FDII) to identify LiDAR attacks and estimate the actual locations of obstacles.
- We propose a controller that guarantees safety based on the updated unsafe region.
- We analyze the correctness of the results from the FDII module.
- We validate our framework in CARLA simulation environment.

Threat Model Analysis

Hide — **Fact 1:** adversary can not remove measured data

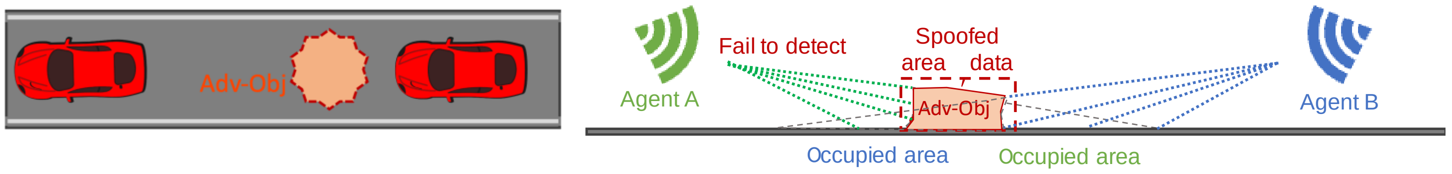
Obstacle — **Fact 2:** the fake obstacle can only be seen by the victim

NEO: Non-Existing Obstacle:

- Agent B cannot see any obstacle
- No overlapping of occupied areas

PRA: Physical Removal Attack:

- Agent B can see obstacle
- Some overlapping of occupied areas



AO: Adversarial Obstacle

- Agent B can see obstacle
- Some overlapping of occupied areas

Proposed Fault Detection, Identification and Isolation

NEO: Non-Existing Obstacle:

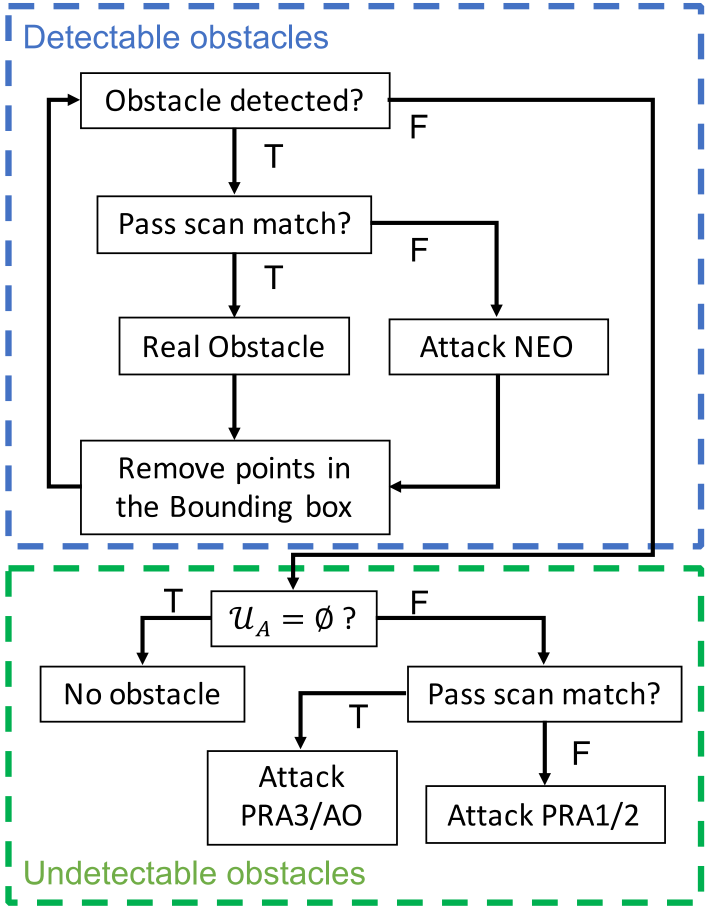
- Agent B cannot see any obstacle
- No overlapping of occupied areas

PRA: Physical Removal Attack:

- Agent B can see obstacle
- Some overlapping of occupied areas
- PRA1/2/3: Full/Partial/No observation on the area affected by the fake obstacle

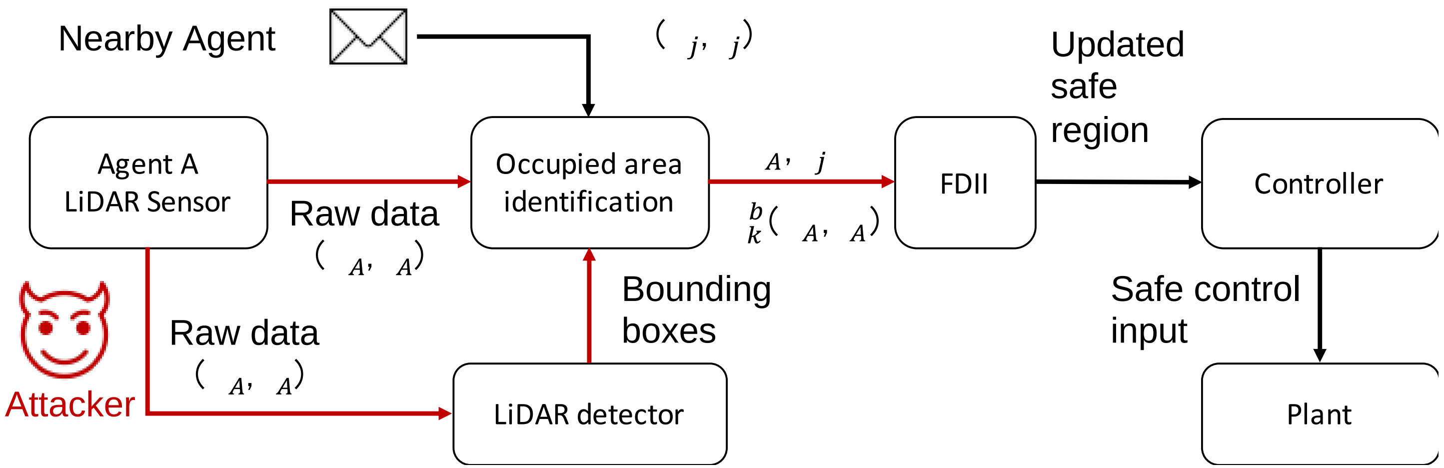
AO: Adversarial Obstacle

- Agent B can see obstacle
- Some overlapping of occupied areas



Proposed Cooperative Framework for Safe Control

- In the paper, we show the correctness of the FDI decision tree.
- Theorem: Suppose we are given the occupied areas U_A and U_B. The obstacle is contained in $U_A \cap U_B$ for any of the attack types NEO, PRA, or AO.

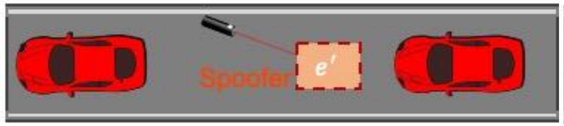


Case Study: Proposed FDII

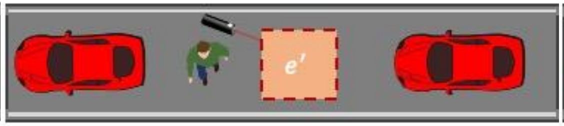
Attack-free | **Attack NEO** | **Attack PRA2** | **Attack PRA3**

Setting

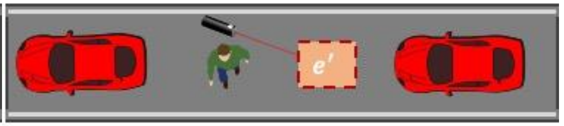
A pedestrian is set between two agents without attack.



An obstacle falsified by spoofer is set in front of Agent A.

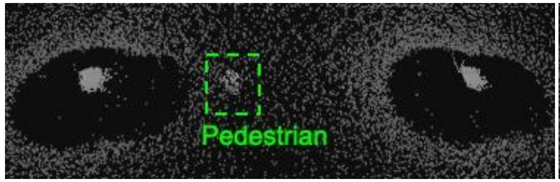


The attack signal is set between Agent A and the pedestrian.



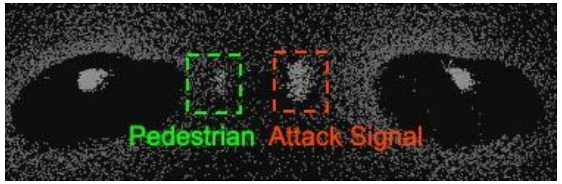
The attack signal affects a relatively smaller area.

LiDAR Point Cloud

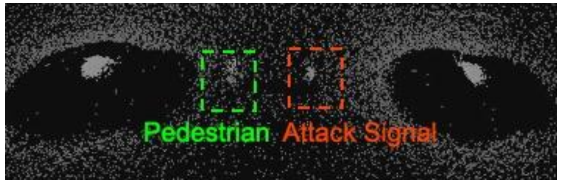


Annotated pedestrian can be detected by both agents with $U_A \neq \emptyset$ and $U_B \neq \emptyset$.

Fake obstacle can be detected only by Agent A with $U_A \neq \emptyset$ and $U_B = \emptyset$.

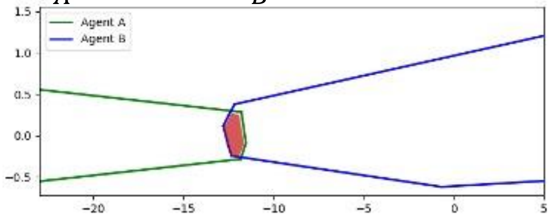


Agent A: Fake obstacle $U_A \neq \emptyset$
Agent B: Pedestrian $U_B \neq \emptyset$

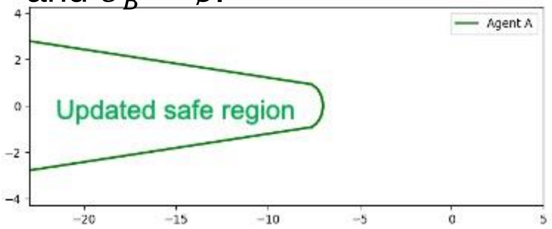


Agent A: Fake obstacle $U_A \neq \emptyset$
Agent B: Pedestrian $U_B \neq \emptyset$

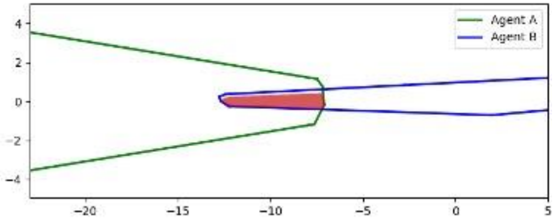
Unsafe Region



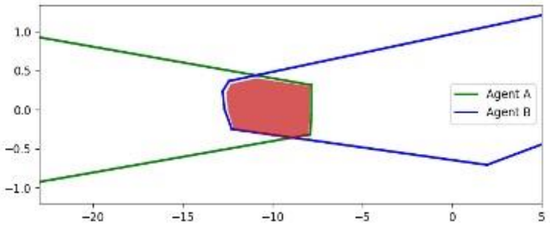
Detect No attack;
unsafe region $U_A \cap U_B$



Detects NEO attack
No unsafe region $U_A \cap U_B = \emptyset$



Detects PRA2 attack;
unsafe region $U_A \cap U_B$.



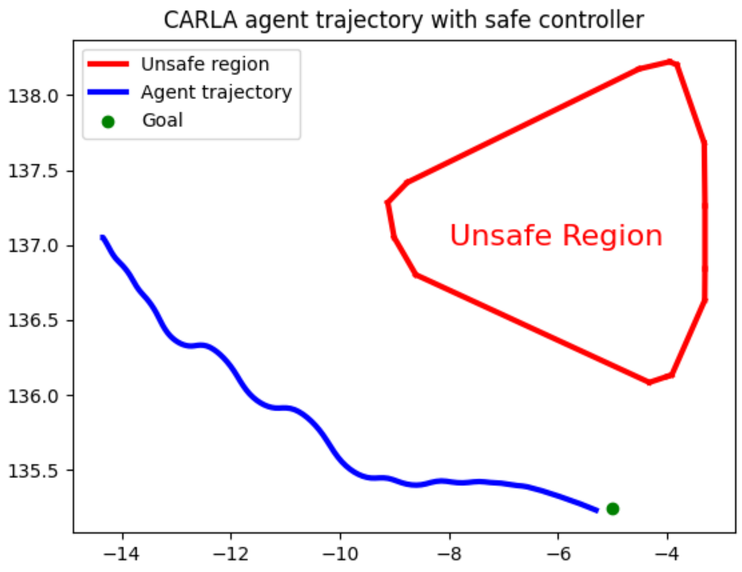
Detects PRA3 attack;
unsafe region $U_A \cap U_B$.

Case Study: Safe Control

- Safe Control
  - Unsafe Region updated by proposed FDII
  - Translate the unsafe region to a set of half-plane safe constraints.
  - Controller compute control input to satisfy constraints.
- Simulation in CARLA
  - We define an MPC controller for a linearized vehicle dynamics:

$$\begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & 0 & 0.03 & 0 \\ 0 & 1 & 0 & 0.03 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k} + \begin{bmatrix} 0.0045 & 0 \\ 0 & 0.0045 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta v_x \\ \Delta v_y \end{bmatrix}$$

  - We realize our controller with do-mpc [1], which calls CasADi [2] and IPOPT [3] for nonlinear programming.



[1] Lucia, Sergio, et al. "Rapid development of modular and sustainable nonlinear model predictive control solutions." Control Engineering Practice 60 (2017): 51-62.

[2] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: a software framework for nonlinear optimization and optimal control," Mathematical Programming Computation, vol. 11, no. 1, pp. 1–36, 2019.

[3] Wächter, Andreas, and Lorenz T. Biegler. "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming." Mathematical programming 106 (2006): 25-57.

Conclusion

- We developed a Fault Detection, Identification, and Isolation procedure that identifies non-existing obstacle, physical removal, and adversarial object attacks, while also estimating the actual locations of obstacles.
- We proposed a control algorithm that guarantees that these estimated object locations are avoided.
- We validated our framework using a CARLA simulation, in which we verify that our FDII algorithm correctly detects each attack pattern.

Thank You

Thank you for your attention

Thanks to our sponsor

Contacts

{hongchao, cheng.shiyu, andrewclark}@wustl.edu

zli4@wpi.edu