Name: Xin Liu **Date:** November 6, 2020

Paper: Choi, Hongjun, Wen-Chuan Lee, Yousra Aafer, Fan Fei, Zhan Tu, Xiangyu Zhang, Dongyan Xu, and Xinyan Deng. ”Detecting attacks against robotic vehicles: A control invariant approach.” *In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 801-816, 2018.

Summary: (Problem) Cyber or physical attacks against Robotic vehicles (RVs) may lead to physical malfunction and subsequently disruption or failure of the vehicles' missions. (Solution) The authors present a novel attack detection framework to identify external, physical attacks against RVs on the fly by deriving and monitoring Control Invariants (CI). Furthermore, they propose a method to extract such invariants by jointly modeling a vehicle's physical properties, its control algorithm and the laws of physics. These invariants represented in a state-space form can be implemented and inserted into the vehicle's control program binary for runtime invariant check. (Meaning) Their framework does not require control program source code or per-vehicle control algorithm reverse engineering. Their evaluation of the CI framework with 11 physical or simulated RVs–all running real-world control programs demonstrates high attack detection accuracy and low runtime overhead.

Strengths:

1. The framework in this paper does not aim to check the traditional program-based invariants, but rather control invariants that model both control and physical prop- erties/states of the vehicle.

2. Based on the system identification (SI), the framework works for a wide range of RVs and does not require per-vehicle controller program reverse engineering to derive control invariants.

3. With monitoring window and threshold, it achieves high detection accuracy by filtering out false positive invariant violations.

4. The framework enables software-based detection of physical attacks without hardware modification or addition.

Weaknesses:

1. The behaviors of a vehicle are determined by three factors: physics, control algorithm and parameters, and mission plan and user commands (runtime inputs). But if an attacker knows three of them, it is head to identify and flag the attacks.

2. This framework relies on proper monitoring parameters (monitoring window and accumulated error threshold) to distinguish attacks from transient errors, but false positive detection may happen under unfavorable environmental conditions.

3. Reliable attack response is outside the scope of this paper.