Demo: Towards Automated Driving Violation Cause Analysis in Scenario-Based Testing for Autonomous Driving Systems

Ziwen Wan Yuqi Huai Yuntianyi Chen Joshua Garcia Qi Alfred Chen
University of California, Irvine {ziwenw8, yhuai, yuntianc, joshug4, alfchen}@uci.edu
Website: https://sites.google.com/uci.edu/adcauseanalysis

Introduction. The rapid advancement of Autonomous Driving (AD) vehicles, exemplified by companies like Waymo offering 24/7 paid taxi services, highlights the paramount importance of ensuring AD vehicles' compliance with various policies, such as safety regulations, traffic rules, and mission directives. Recent security research [1]–[3] has demonstrated that attackers could cause safety violations in AD systems by exploiting internal bugs via controlling the surrounding environment or other vehicles on the road. Despite significant progress in the development of automated AD security testing tools, there has been a notable absence of research on attributing the causes of driving violations to a certain component execution in such complicated distributed systems to help diagnose the internal bug and thus make the software secure. Counterfactual causality analysis has emerged as a promising approach for identifying the root cause of program failures. While it has demonstrated effectiveness in pinpointing error-inducing inputs, its direct application to the AV context to determine which computation result, generated by which component, serves as the root cause poses a considerable challenge since it is not trivial to straightforwardly eliminate the influence of a specific internal component output to establish the causal relationship between the output of each component and a system-level driving violation. In this work, we propose a novel driving violation cause analysis tool. We design idealized component substitutes to enable counterfactual analysis of ADS components by leveraging the unique opportunity provided by the simulation.

Demonstration Plan. In the demo, we will show an end-to-end visualized process to demonstrate the usage of our novel cause analysis tool, which can analyze the cause of a system-level driving violation from a driving record that includes >1000 different component executions. We will create the demo based on industry-grade AD software and simulator (e.g., Apollo [4], LGSVL [5]).

Symposium on Vehicles Security and Privacy (VehicleSec) 2024
26 February 2024, San Diego, CA, USA
ISBN 979-8-9894372-7-6
https://dx.doi.org/10.14722/vehiclesec.2024.25025
www.ndss-symposium.org

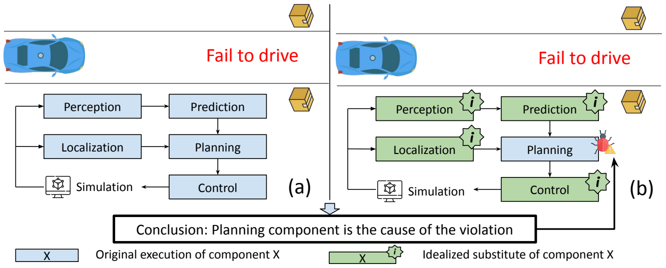

Fig. 1. (a) The original driving trace where AD software fails to execute the driving task. (b) After replacing the components except planning with idealized substitutes (the changes of the execution graph are shown in the figure), the AD software still fails to execute the driving task. This is an indicator that the root cause is the planning component based on the counterfactual causality.

An Illustration of our Tool. We present a potential illustration of our tool based on a vulnerability discovered by *PlanFuzz* [3] in Fig. 1. Two off-road static obstacles controlled by the attackers could trigger the vulnerability and force the AD vehicle to stop permanently and fail to complete the driving tasks (driving task violation). The testing results remain the same after replacing components except for planning with idealized substitutes (our novel design to remove the effects of certain modules or component outputs while still maintaining the test to be valid), indicating that this violation is caused by the outputs generated from the planning component. We further apply this idea to diagnose which specific output is the cause with counterfactual analysis.

Acknowledgments

This research was supported by the NSF CNS-2145493, CNS-1929771, and USDOT UTC Grant 69A3552348327.

References

[1] S. Kim, M. Liu, J. J. Rhee, Y. Jeon, Y. Kwon, and C. H. Kim, "Drivefuzz: Discovering Autonomous Driving Bugs through Driving Quality-Guided Fuzzing," in *CCS 2022*.

[2] R. Song, M. O. Ozmen, H. Kim, R. Muller, Z. B. Celik, and A. Bianchi, "Discovering Adversarial Driving Maneuvers against Autonomous Vehicles," in *USENIX Security 23*.

[3] Z. Wan, J. Shen, J. Chuang, X. Xia, J. Garcia, J. Ma, and Q. A. Chen, "Too Afraid to Drive: Systematic Discovery of Semantic DoS Vulnerability in Autonomous Driving Planning under Physical-World Attacks," in *NDSS'22*.

[4] "Apollo." https://github.com/ApolloAuto/apollo.

[5] LG, "LGSVL Simulator." https://github.com/lgsvl/simulator.