D2MON: Detecting and Mitigating Real-Time Safety Violations in Autonomous Driving Systems

Bohan Zhang Yafan Huang Rachael Chen Guanpeng Li

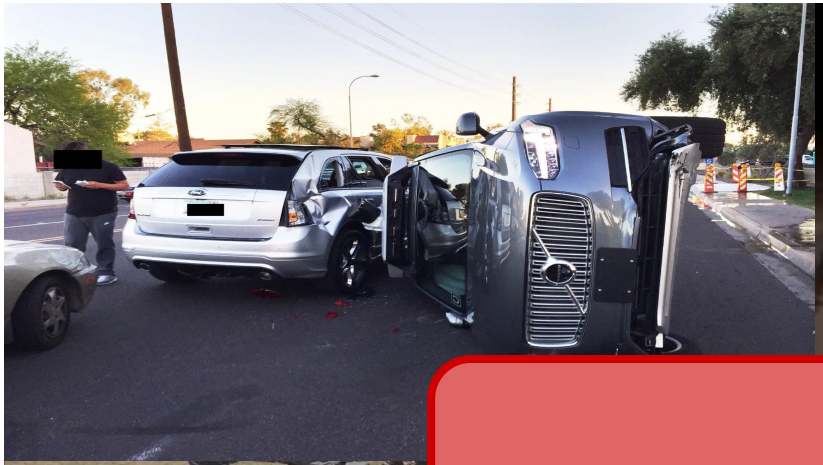
bzhang22@uiowa.edu yafan-huang@uiowa.edu rchen23@students.d125.org guanpeng-li@uiowa.edu

IOWA

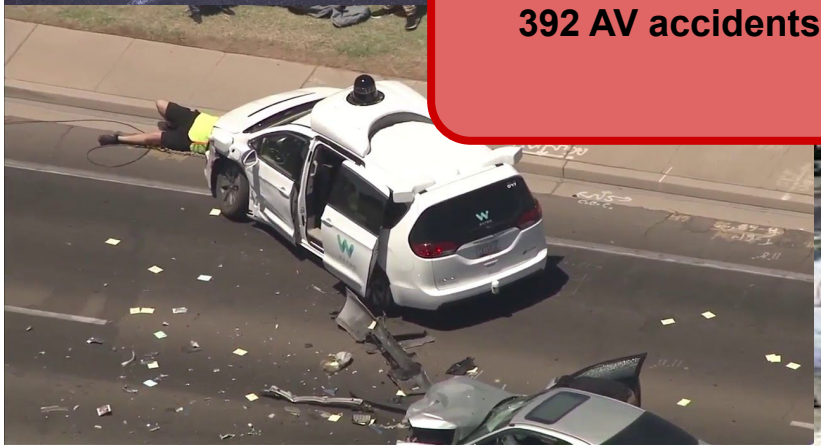
Autonomous Vehicle (AV)

- Every 10 mins, about 26 people will die in automobile accidents globally
  - >94% is drivers' fault that cause the accidents
- AV holds significant potential to improve productivity and quality of life
  - Level-4 autonomous driving system (ADS) requires no human intervention
- By 2019, over 1,400 AVs are in testing by 80+ companies across the US on roads
  - Operate with human-driven cars on public roads
  - Market estimate of AV will reach $60 billion by 2030

AV Accidents



392 AV accidents in past 10 months



Existing Methods:

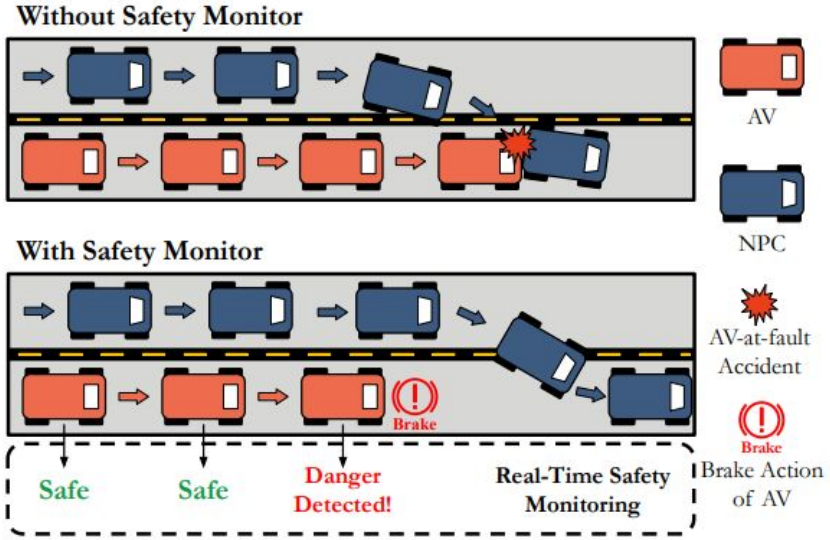
- Road Testing
  - Require million of millages and energy
  - Not efficiency (requrie human to monitoring)

- Simulation-Based Software Fuzzing
  - Hard to localize the error (huge code base,

    more than 400000 line of codes)
  - Fix one bug may cause more

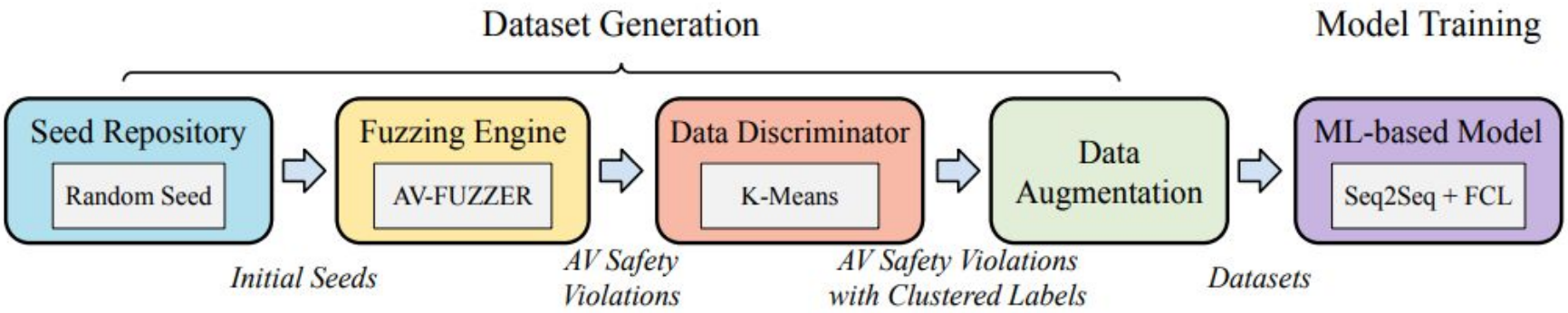
Our Solution

- Key insights
  - NPC (surrounding vehicle)’s trajectoris which will lead the AV safety violations fall into identifiable patterns
- ML-based real-time safety monitor
  - Without debugging the code of ADS
  - Real-time dectection for safey violation
  - Mitigation when danger dectected



Enable Maxium safety

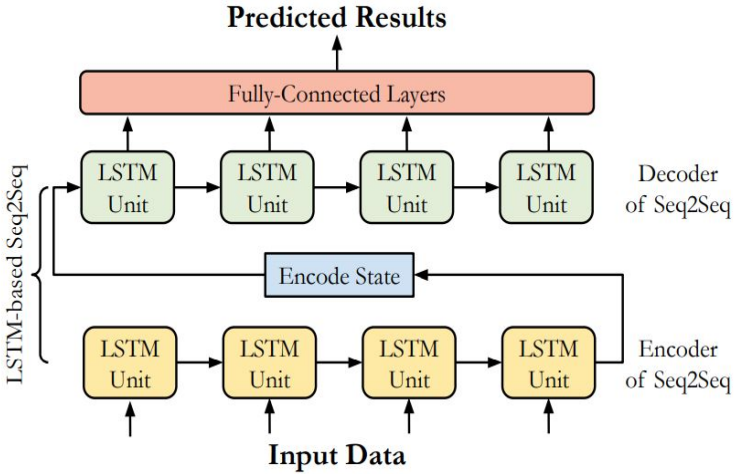
Framework Design: D2MON



ML-Based Model

Input: Past 1.25 second's vehicle running data

- NPC and AV's current speed
- Euclidean distance of NPC and AV's location
- Euclidean distance of NPC and AV's steering angle
- Indicator that judge "is NPC in front of AV?"
- ……

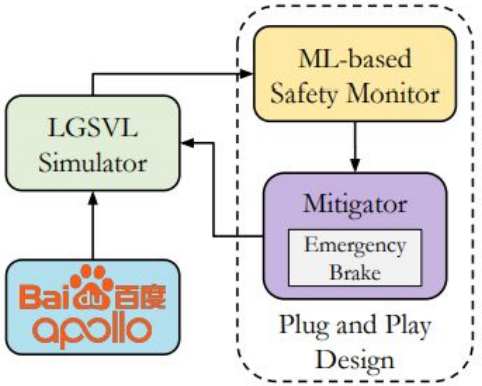


Output: Safety indicators for next 2 seconds

Deployment of Our Technique

- How does the our technique deployed?

- How does safety monitor work?



Experimental Setup

Machine Specification:

- Ubuntu 18.04 machine with 32GB RAM
- AMD 5900X CPU (12-core/24- thread)
- NVIDIA GTX 1080 Ti GPU card

Environment setup

- One npc
- 2-lane straight road

Evaluation: Metrics

Metrics	Formula	Descriptions
FP Val. (False Positive)	(FP Case) / (Total Case)	FP: Predict the safe situation as dangerous.
FN Val. (False Negative)	(FN Case) / (Total Case)	FN: Predict the dangerous situation as safe.
Fault Rate (%)	(Fault Trial) / (Total Trial)	Number of trials that leads to accidents.

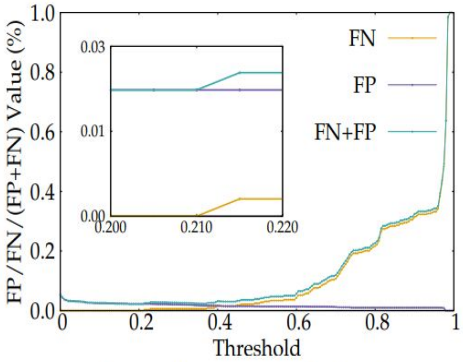
Metrics for evaluate the ML model prediction accuracy

Metrics for evaluate the safety monitor mitigation efficiency

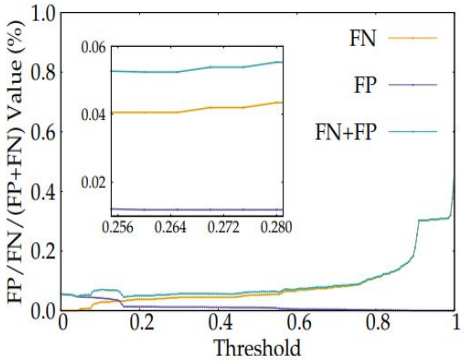
Evaluation: Results

Prediction Accurancy

- FN and FP are 0.0061 and 0.0167 for Cluster 1
- FN and FP are 0.0405 and 0.0118 for Cluster 2



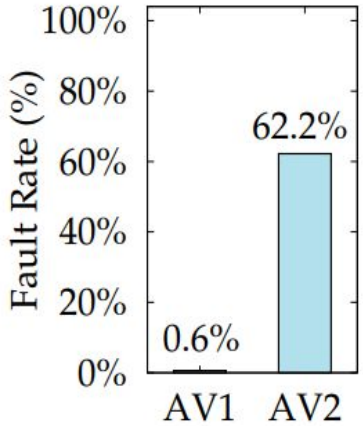
(a) ML prediction accuracy for cluster1.



(b) ML prediction accuracy for cluster2.

Mitigation Effiency

- AV with D2MON only cause 3 / 500 accidents compare to the AV without D2MON 303/500 accidents



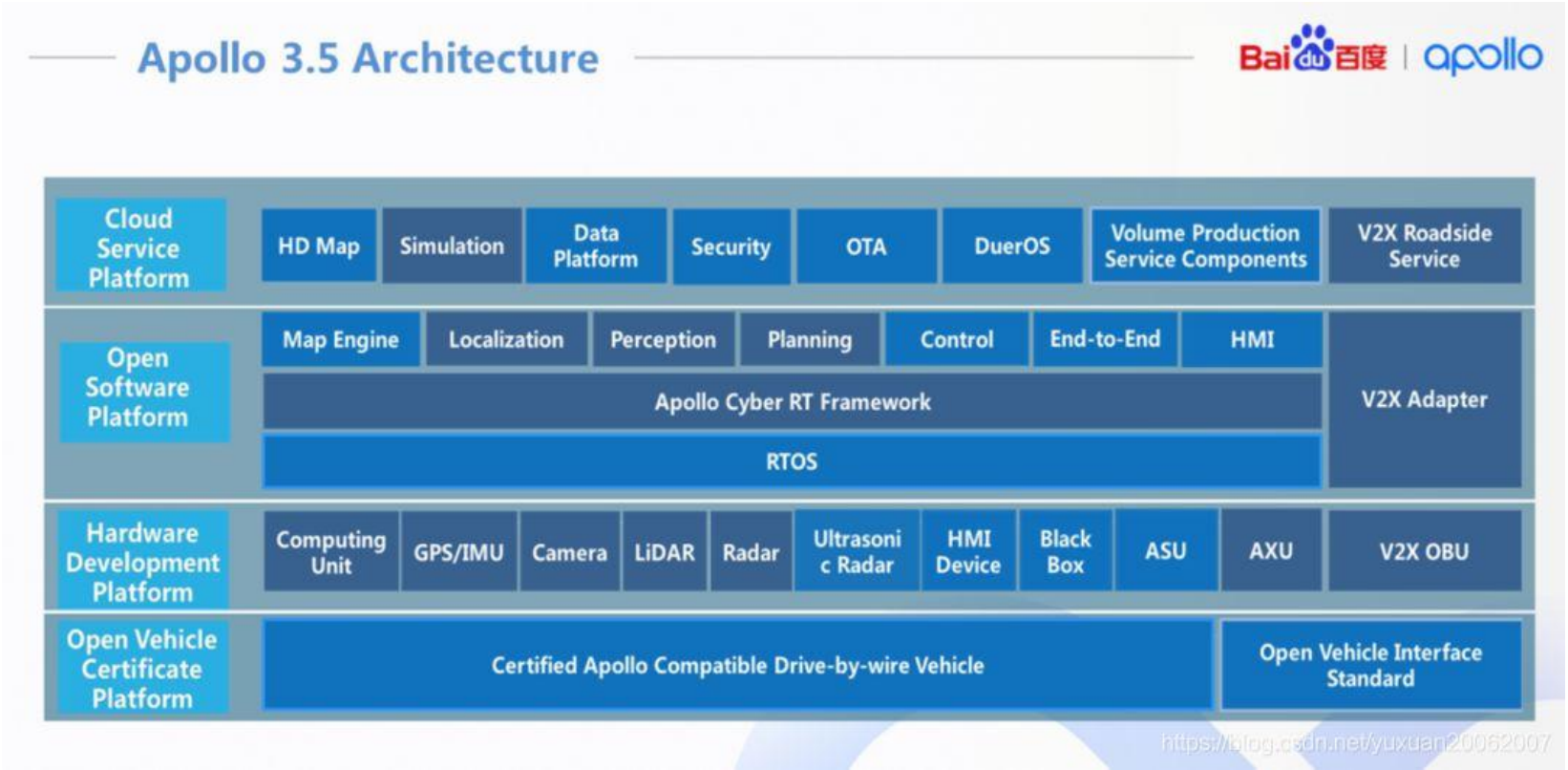
Conclusion & Future Work

- Our goal is to eliminate AV safety violations
  - Proposed a framework for D2MON that able to Creation and Deployment safety monitor
    - NPC’s trajectoris which will lead the AV safety violations fall into identifiable patterns
    - ML-based prediction model
    - Enable the maxium AV-safety

- Future work
  - Use better model for improvement of prediction accuracy
  - Updating the mitigation strategy
  - Test more scenario

Backup Slides

Structure of Autonomnous Driving System



Background: LGSVL

- HD-simulator for AV
- Bridging with Baidu Apollo