Cyber-Attacks against AI Stack in Autonomous Driving & Intelligent Transportation

Qi Alfred Chen

Assistant Professor, UC Irvine

UCIrvine

AS²Guard

Autonomous & Smart Systems

Guard Research Group

A bit about myself & my group

- Assistant Professor of Computer Science, UC Irvine (2018 -)
  - Ph.D., University of Michigan
- Group: **AS²Guard** (**A**utonomous & **S**mart **S**ystems **Guard**)
- Expertise: **AI/Systems/Network Security**, mainly in **mobile/CPS/IoT**

AS²Guard

Autonomous & **S**mart **S**ystems **Guard** Research Group

Impact: Demo & vulnerability report

cve.mitre.org

NDSS'16

Euro S&P'17

US-CERT
UNITED STATES COMPUTER EMERGENCY READINESS TEAM

IEEE S&P'16

Usenix Sec'14

NDSS'16

Apple

Comcast

Microsoft

WAYMO

apollo

TOYOTA

GM

DAIMLER

BOSCH ...

DEPARTMENT OF TRANSPORTATION
UNITED STATES OF AMERICA

Linux

CCS'15

CCS'17

Usenix Sec'20

NDSS'18

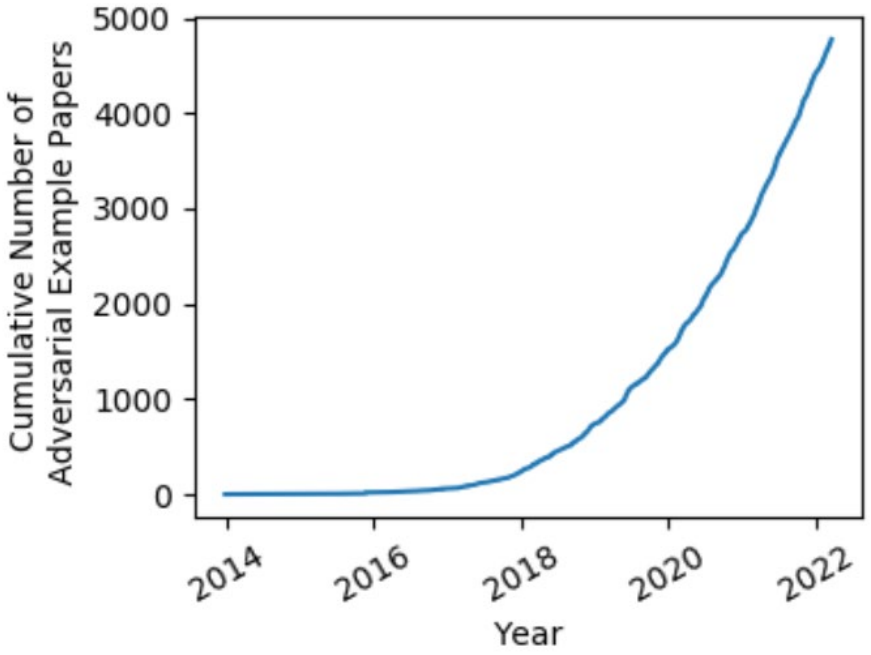
My research so far in mobile/CPS/IoT security

- **CPS AI Security**
  - **Autonomous Driving (AD)** [ACM CCS'19, Usenix Security'20 (a), '20 (b), '21, IEEE S&P'21, NDSS'22, CVPR'22, ICLR'20]
  - **Intelligent transportation** [NDSS'18, TRB'18,'19,'20, ITS'21]
- **Network Security**
  - **Connected Vehicle (CV)** [Usenix Security'21]
  - **Automotive IoT** [Usenix Security'20, NDSS'20]
  - **Network protocol** [ACM CCS'15,'18, IEEE S&P'16]
- **UI (User Interface) Security**
  - **Smartphone** [Usenix Security'14, MobiSys'19]
- **Access Control / Policy Enforcement**
  - **Smartphone** [NDSS'16]
  - **Smart home** [NDSS'17]
- **Side Channel**
  - **Smartphone** [Usenix Security'14]
  - **Network** [ACM CCS'15]

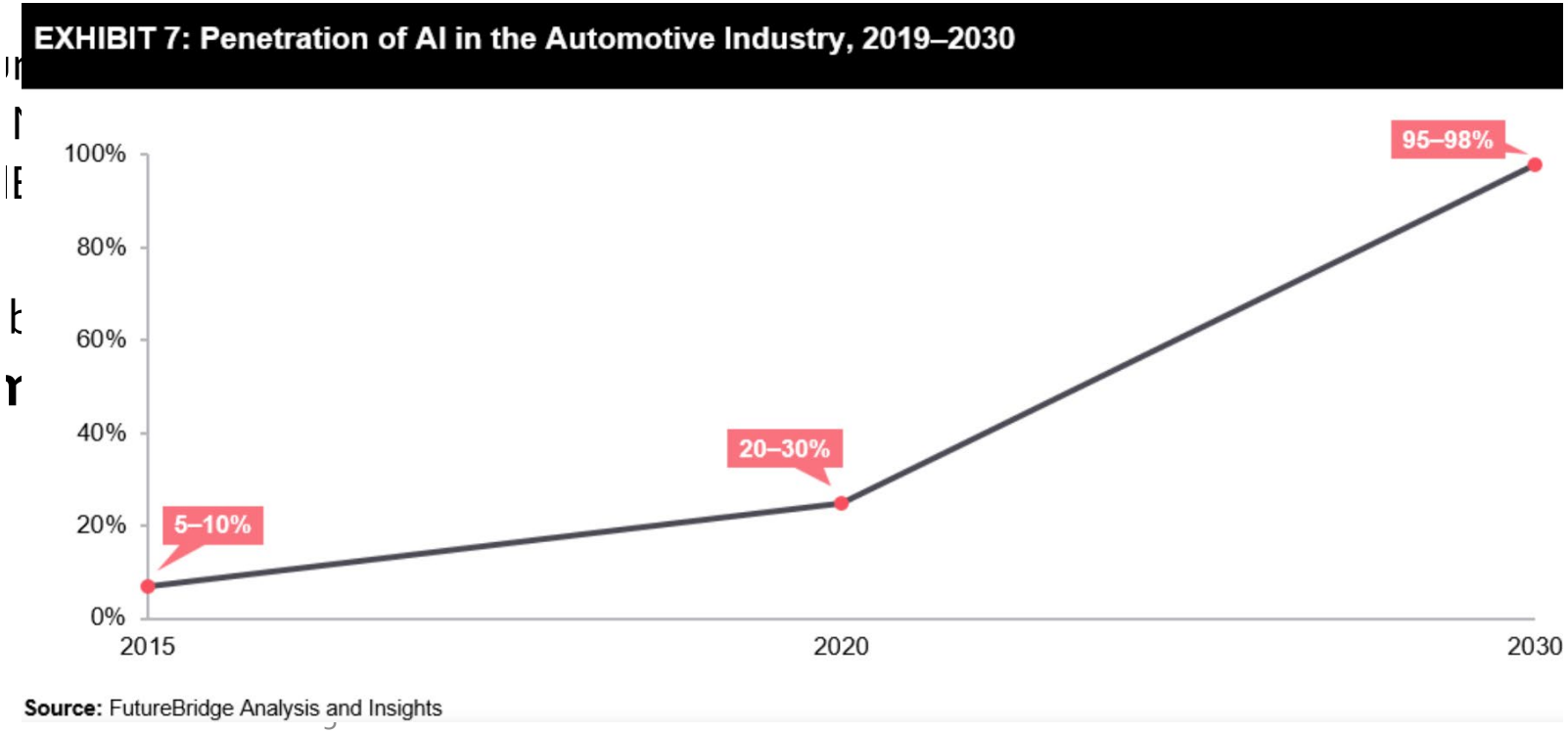
Most recent focus (2018-): CPS AI security

- **CPS AI Security**
  - **Autonomous Driving (AD)** [ACM CCS'19, Usenix Security'20 (a), '20 (b), '21, IEEE S&P'21, NDSS'22, CVPR'22, ICLR'20]
  - **Intelligent transportation** [NDSS'18, TRB'18,'19,'20, ITS'21]

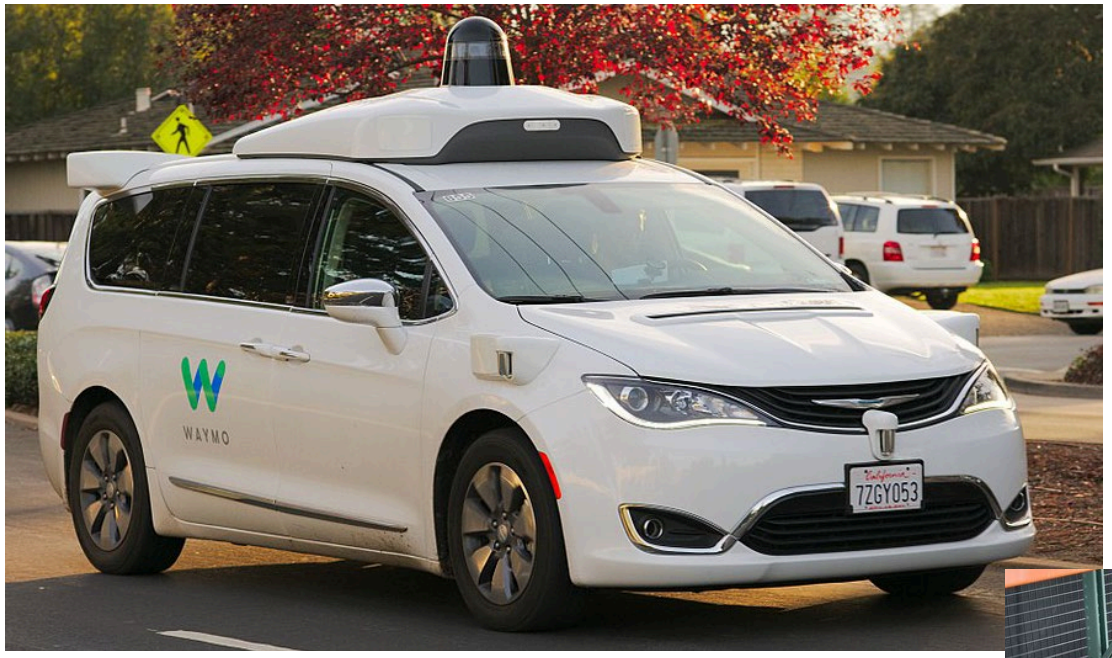
- Relatively new area:
  - AI security: Since 2013 [Szegedy et al., "Intriguing properties of neural networks"]
  - AI penetration in real-world CPS (e.g., since ~2015 in automotive industry)

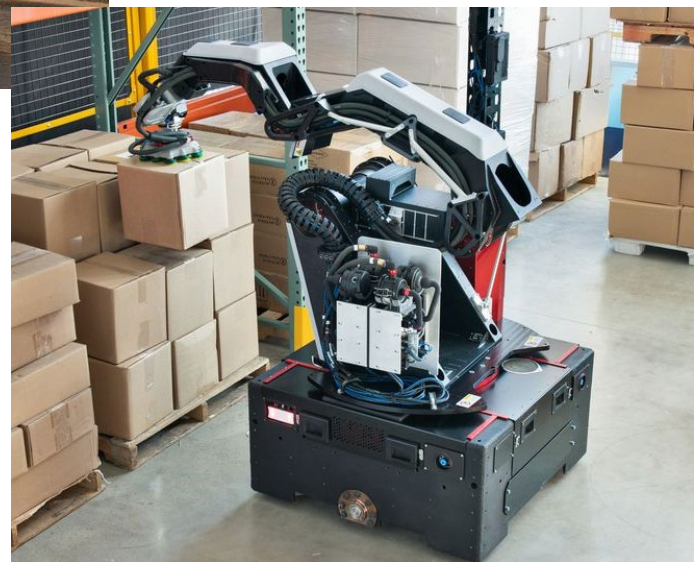


(*Image credit: Nicholas Carlini)



More recently, massive kinds of AI-enabled autonomous systems coming into real life





Current focus (2018-): Automotive & transportation domain

Autonomous Driving (AD)

V2X-based Intelligent Transp.

TESLA

WAYMO

GM

apollo

TOYOTA

ZOOX

Qualcomm

Ford

pony.ai

tusimple

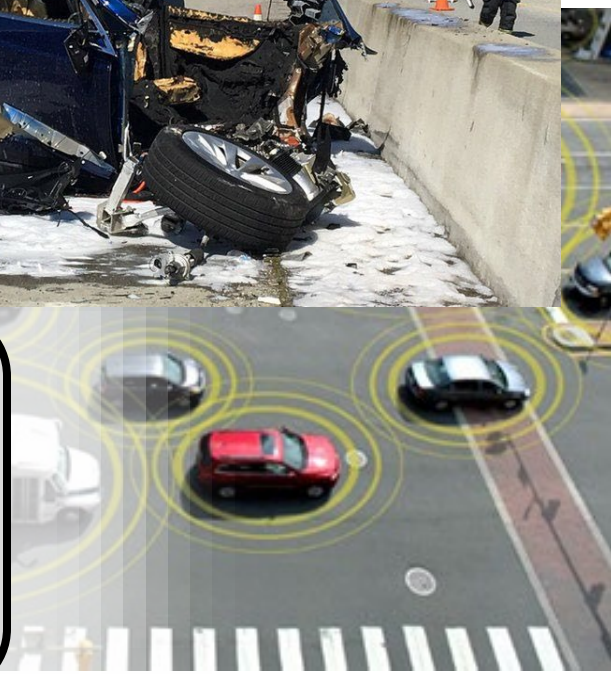
Aurora

ARGO AI

U.S. Department of Transportation

Curre omotiv in

Au



AI stack:

"brain" for autonomous AI decision-making

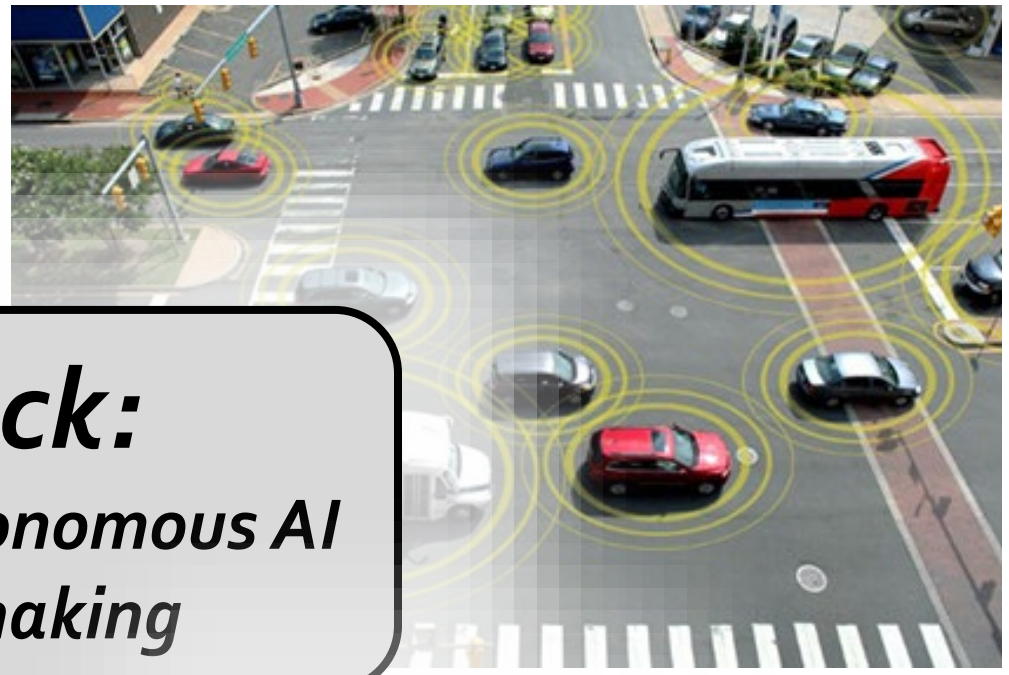
IMPORTANT

U.S. Department of Transportation

Current focus (2018-): Automotive & transportation domain

Autonomous Driving (AD)

V2X-based Intelligent Transp.



AI stack:

"brain" for autonomous AI decision-making

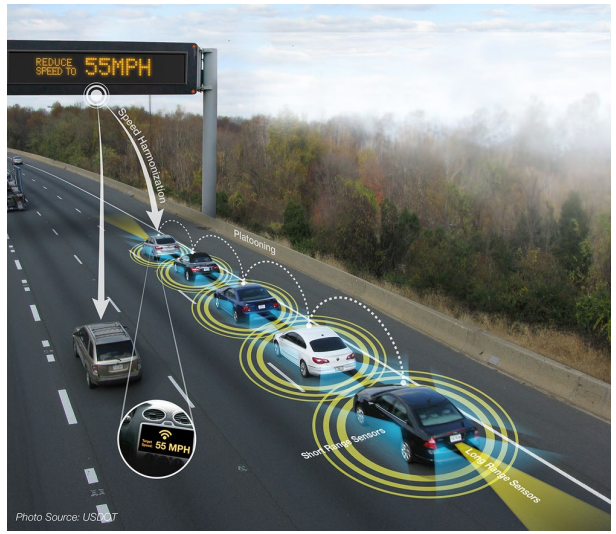
IMPORTANT

U.S. Department of Transportation

Today: Cyber-attack surface to AD & V2X-based transp. AI

GPS

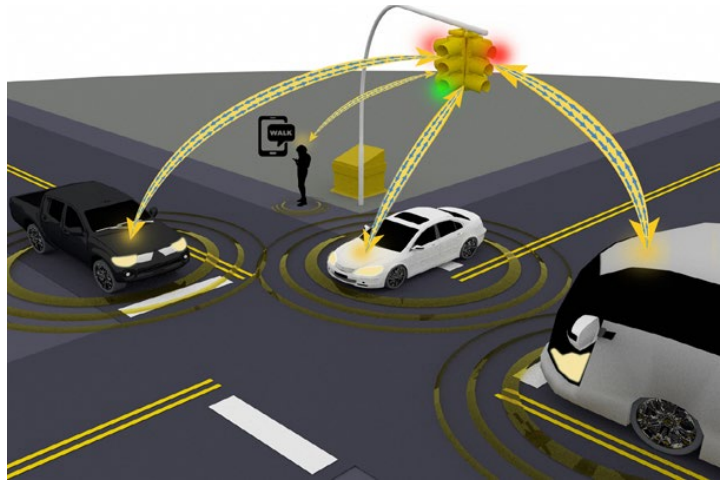
V2V
(vehicle-to-vehicle)



Cooperative Driving Automation (e.g., platoon)

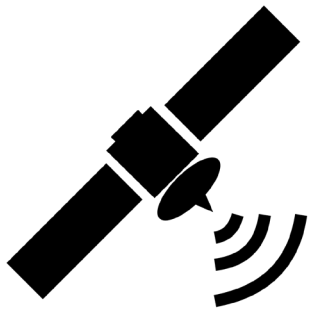
AD vehicle

V2I
(vehicle-to-infrastructure)



Intelligent traffic light

Today: Cyber-attack surface to AD & V2X-based transp. AI



GPS

V2V
(vehicle-to-vehicle)

Cooperative Driving Automation (e.g., platoon)

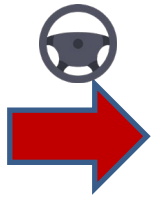
AD vehicle

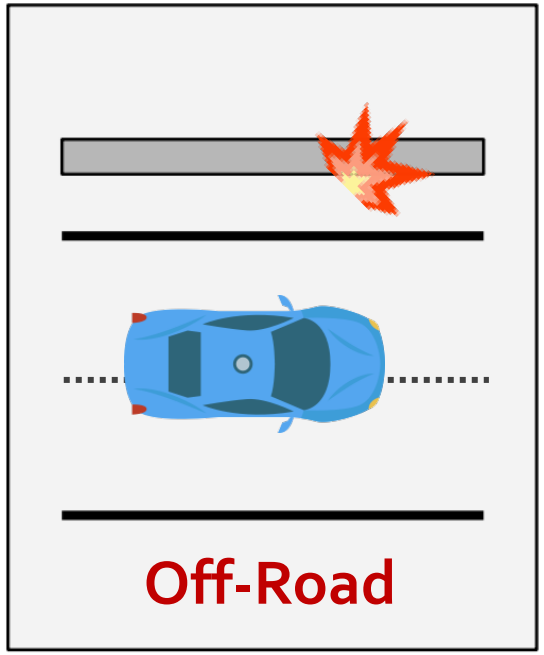
V2I
(vehicle-to-infrastructure)

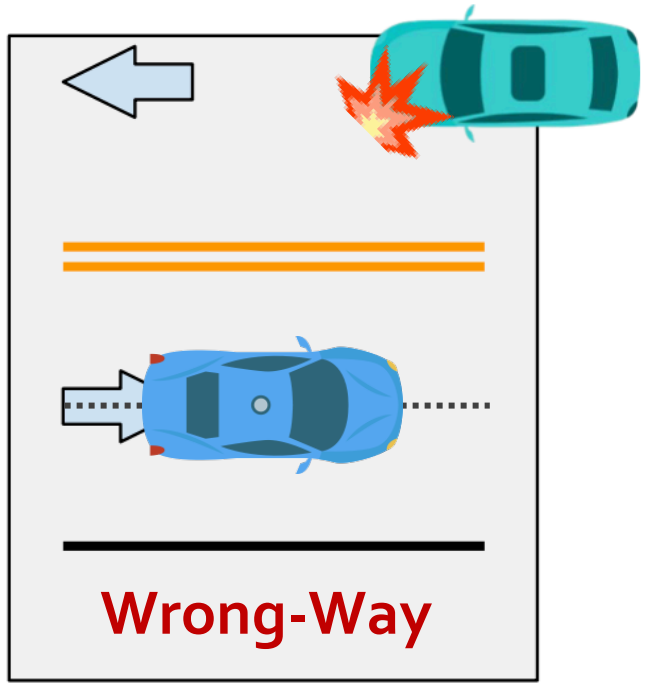
Intelligent traffic light

Localization is safety-critical to AD vehicles

Localization

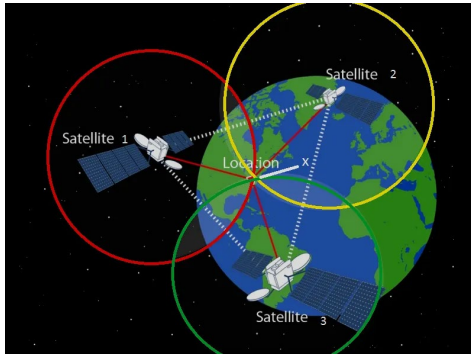


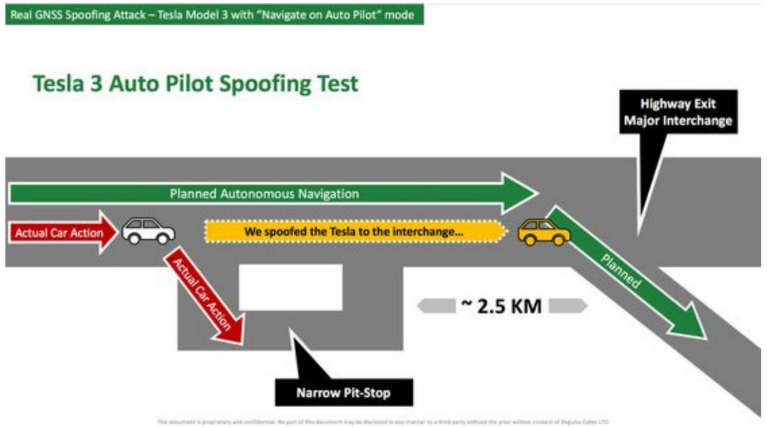




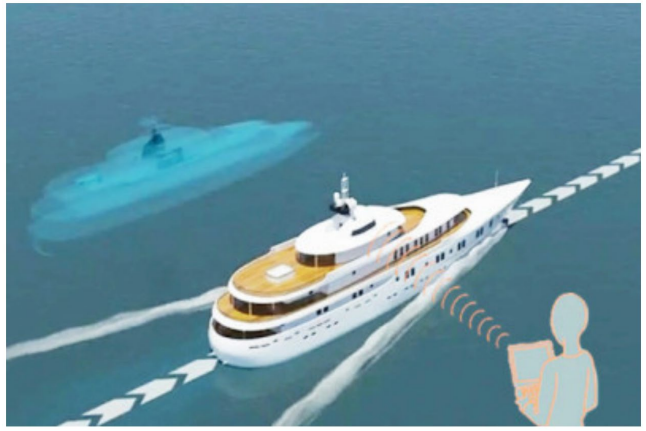
GPS and spoofing attack

- GPS is the *de facto* location input for AD localization
- GPS spoofing attacks
  - Attacker sets **arbitrary position** by sending fake satellite signals
  - Still an **open problem** in civilian GPS
    - Demonstrated on cars, yachts, drones, etc.

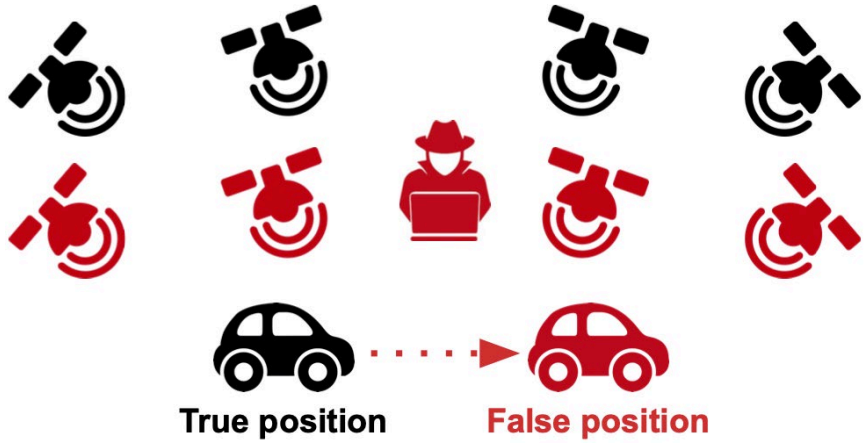




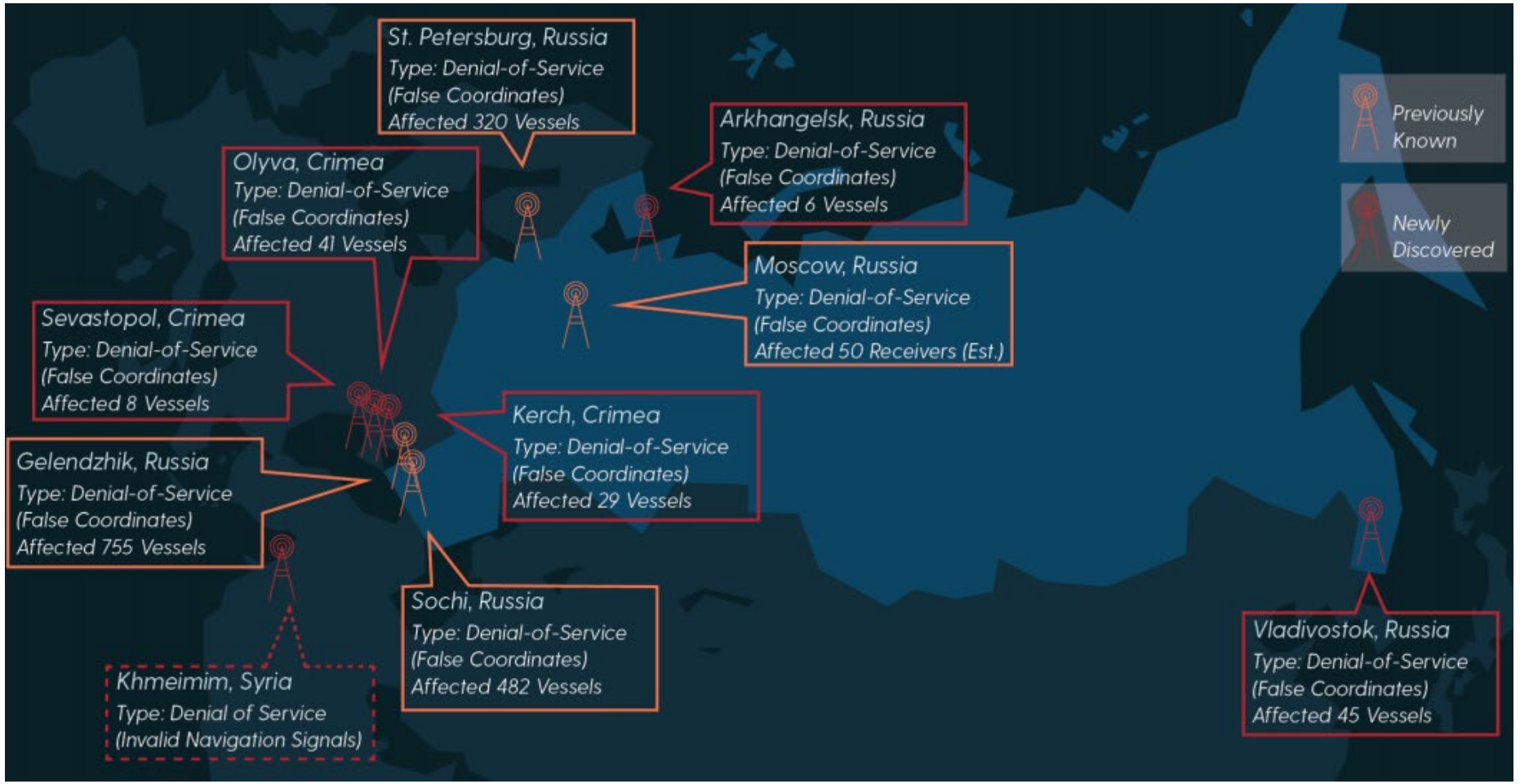
[Regulus Cyber, '19]



[Bhatti et al., NAVIGATION'17]



GPS spoofing is pervasive!

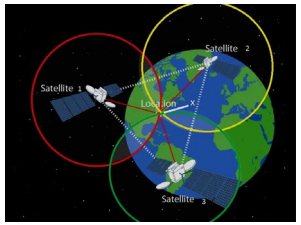


Over 9,883 spoofing events identified; **1,311** civilian vessels affected since Feb. 2016 in Russia.

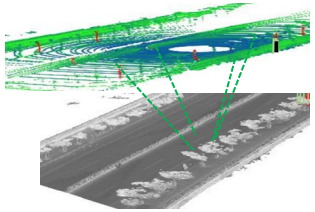
Source: Above Us Only Stars @ C4ADS

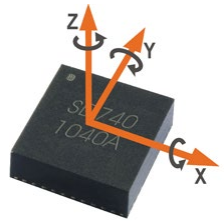
Multi-Sensor Fusion (MSF) based AD localization

- However, production high-level AD systems widely adopt **MSF-based localization** design
  - Baidu Apollo, [ICRA'18] [ITS'16] [IV'16] [Sensors'15] [IROS'13] [IJRR'11], etc.
  - **Leverage strengths & compensate weaknesses** of different sensors to improve ***accuracy*** & ***robustness***
    - Commonly fuse from GPS, LiDAR, and IMU
    - Can achieve ***5.4 cm*** localization accuracy
- In such a design, GPS alone cannot dictate the localization results



GPS


LiDAR locator



IMU

MSF: Generally believed to have potential to defend against GPS spoofing

Sensor Fusion: Resilient estimation algorithms usually assume a variety of multi-modal sensors to achieve their security guarantees. This is also the idea behind sensor fusion, where sensors of different types can help "confirm" the measurement of other sensors [134, 135, 136]. A basic example of sensor fusion in automotive systems is to verify that both the LiDAR readings and the camera measurements report consistent observations.

[Cardenas, CyBOK '19]

Sensor fusion: Combining data from multiple distinct sensors, known as *sensor fusion* [3], significantly raises the difficulty of sensor input spoofing attacks. As an ex-

[Davidson et al., WOOT '16]

We hope the results can help to raise the attention in the community to develop *practically deployable* defense mechanisms (*e.g.*, location verification, signal authentication, sensor fusion) to protect the massive GPS device users and emerging GPS-enabled autonomous systems.

[Zeng et al., USENIX Security '18]

SENSOR FUSION

As should be apparent from earlier discussions, different technologies available for detection and tracking of UAVs have various trade-offs related to cost, accuracy, precision, range, energy efficiency (critical if sensors operate on batteries),
l at other UAVs),
example, while
nly operate very
mputer vision),
NLOS environ-
es). For accurate
JAVs, data fusion
isly use informa-
ors carry critical
for joint use of
coustic sensors,
n optical camer-
as), and this constitutes an open research area.

[Guvenc et al., IEEE Comm '18]

This research presented a statistical approach to the problem of attack detection on the multi-sensor integration of autonomous vehicle navigation systems. Starting with a state-space model of the system under attack, a parametric statistical tool with a multi-sensor integration strategy was developed to identify an attack. Finally, a simulation was designed to verify the proposed detection system and results were presented. A

[Lee et al., SMC '17]

MSF: Generally believed to have potential to defend against GPS spoofing

Sensor Fusion: Resilient estimation algorithms usually assume a variety of multi-modal sensors to achieve their security guarantees. This is also the idea behind sensor fusion, where sensors of different types can help "confirm" the measurement of other sensors [134, 135, 136]. A basic example of sensor fusion in automotive systems is to verify that both the LiDAR readings and the camera measurements report consistent observations.

[Cardenas, CyBOK '19]

Sensor fusion: Combining data from multiple distinct sensors, known as *sensor fusion* [3], significantly raises

SENSOR FUSION

As should be apparent from earlier discussions, different technologies available for detection and tracking of UAVs have various trade-offs related to cost, accuracy, precision, range, energy efficiency (critical if sensors operate on batteries),

l at other UAVs),
example, while
nly operate very
mputer vision),
NLOS environ-
s). For accurate

This research presented a statistical approach to the problem of attack detection on the multi-sensor integration of autonomous vehicle navigation systems. Starting with a state-space model of the system under attack, a parametric statistical

Research Question:

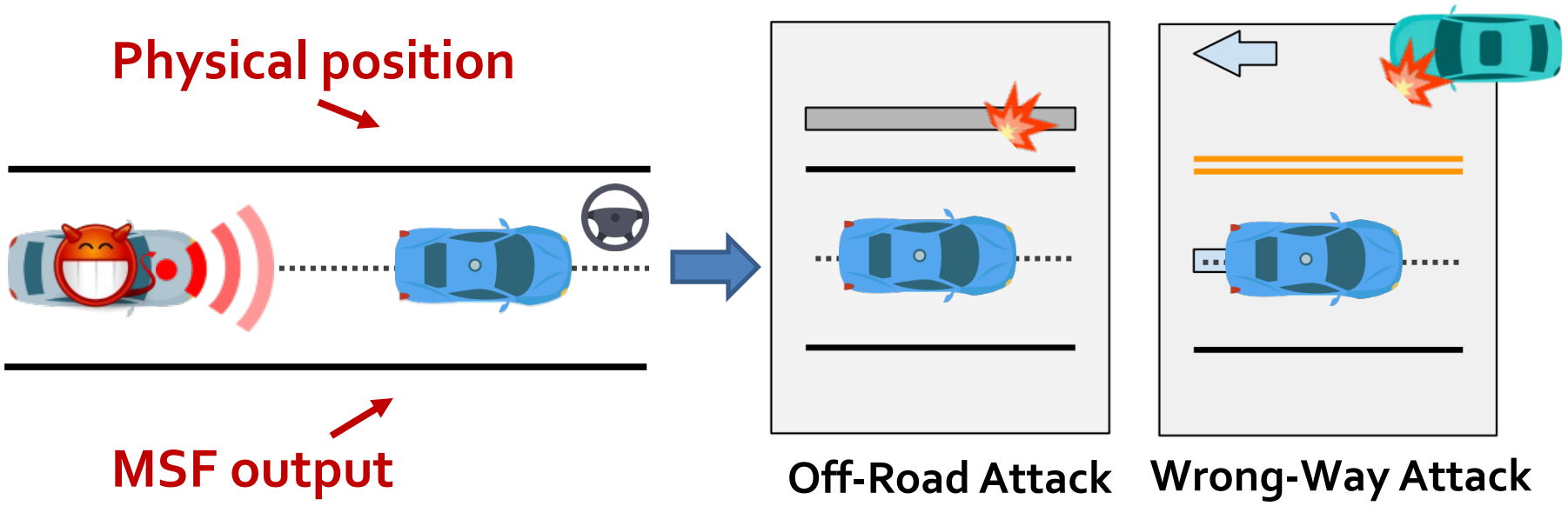
In AD settings, whether state-of-the-art MSF algorithms are indeed sufficiently secure under GPS spoofing?

[Zeng et al., USENIX Security '18]

Our work: FusionRipper

[Usenix Security'20]

- ***First*** to study the security of MSF-based AD localization in practical settings
- Problem formulation
  - Attacker **tailgates** a victim AD vehicle & perform **GPS spoofing**
  - Aim to **maximize lateral deviation** in MSF output w.r.t. no attack
- Attack goals: cause victim to drive **off-road** or onto a **wrong-way**



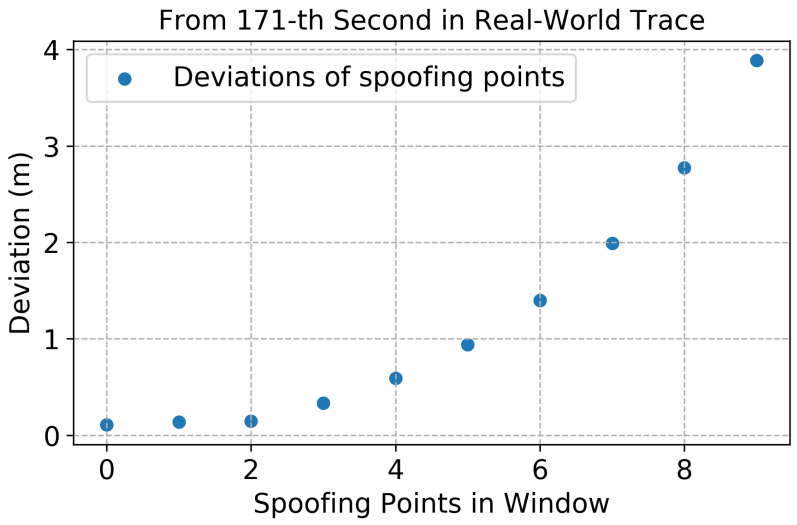
Security analysis

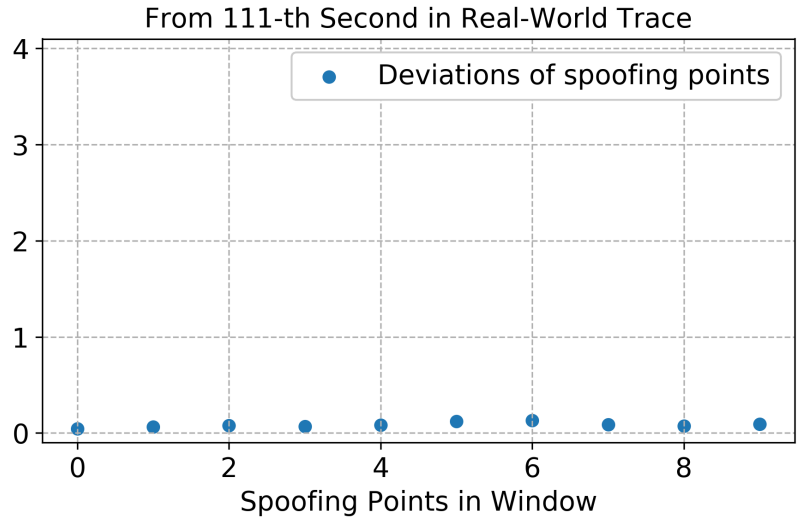
- Aim to find **maximum possible deviation** achievable by spoofing
- Target: Apollo MSF (representative in both design & impl.)
- Dataset: Real-world sensor traces + synthetic trace (w/o noise)
- Methodology: Split trace to attack windows & perform exhaustive search
- Success metric: MSF output deviation
- Results:
  - Synthetic: **100% < 0.076m**
    - Far from reaching any attack goal
    - By design, SOTA MSF is resilient enough to GPS spoofing
  - Real-world: **76% < 0.895m**
    - Majority failed to reach even smallest attack goal
  - Takeaway: ***MSF indeed generally improves security against GPS spoofing***

Attack Goal	Local	Highway
Off-Road	0.895 m	1.945 m
Wrong-Way	2.405 m	2.855 m

Finding: Take-over vulnerability

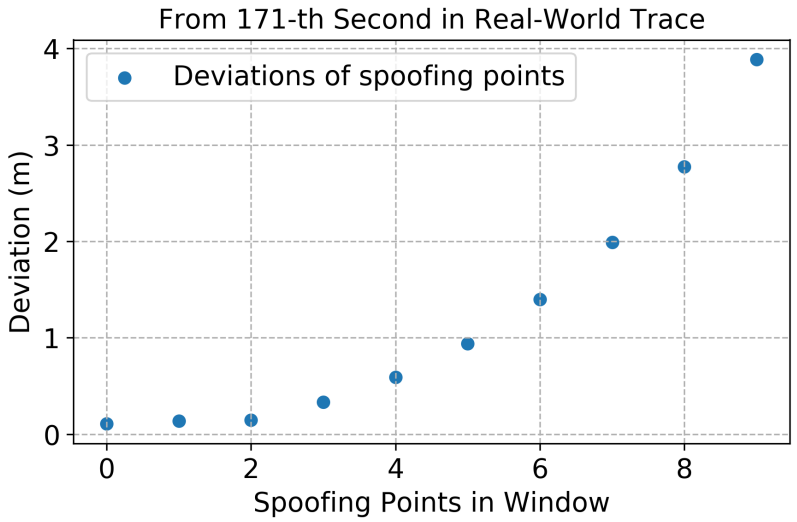
- Still, some windows in real-world trace can achieve large deviations
  - **13% attack windows *satisfy all attack goals (>= 2.855 m)***
- Find that they all exhibit an interesting **take-over effect,** causing an ***exponential growth trend*** of deviations

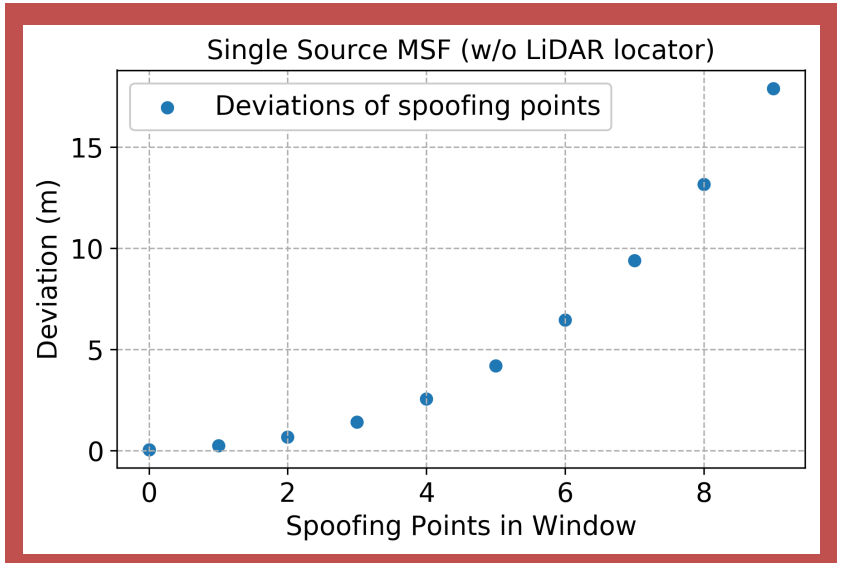




Finding: Take-over vulnerability

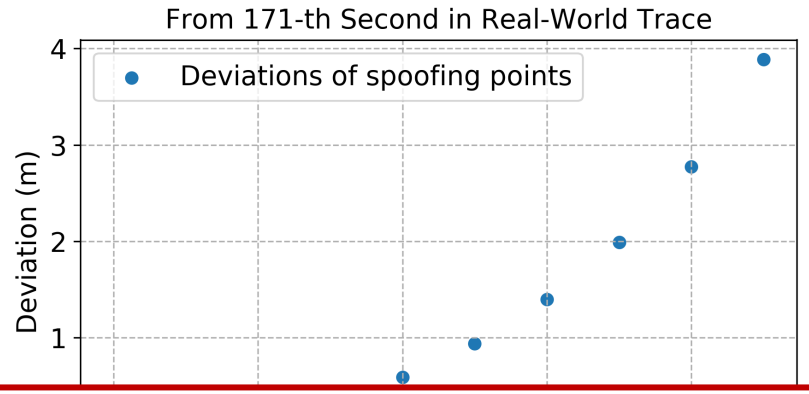
- Still, some windows in real-world trace can achieve large deviations
  - **13% attack windows *satisfy all attack goals (>= 2.855 m)***
- Find that they all exhibit an interesting **take-over effect,** causing an ***exponential growth trend*** of deviations
  - Similar to **when GPS is the only source**
  - Spoofed GPS inputs become **dominating** source to MSF → Later LiDAR becomes outlier!

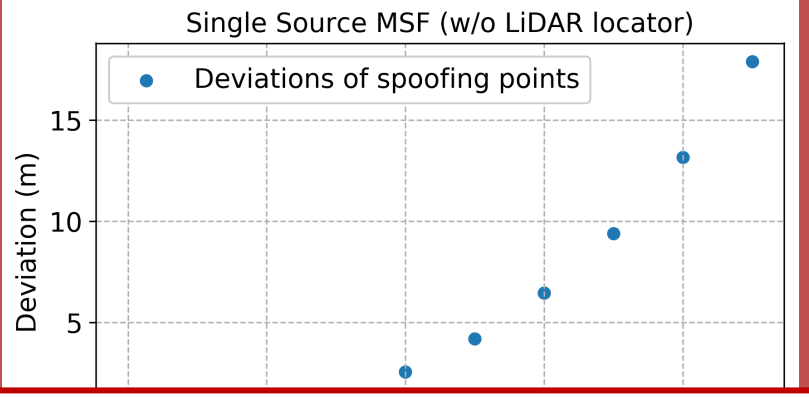




Finding: Take-over vulnerability

- Still, some windows in real-world trace can achieve large deviations
  - **13% attack windows *satisfy all attack goals (>= 2.855 m)***
- Find that they all exhibit an interesting **take-over effect,** causing an ***exponential growth trend*** of deviations
  - Similar to **when GPS is the only source**
  - Spoofed GPS inputs become **dominating** source to MSF → Later LiDAR becomes outlier!





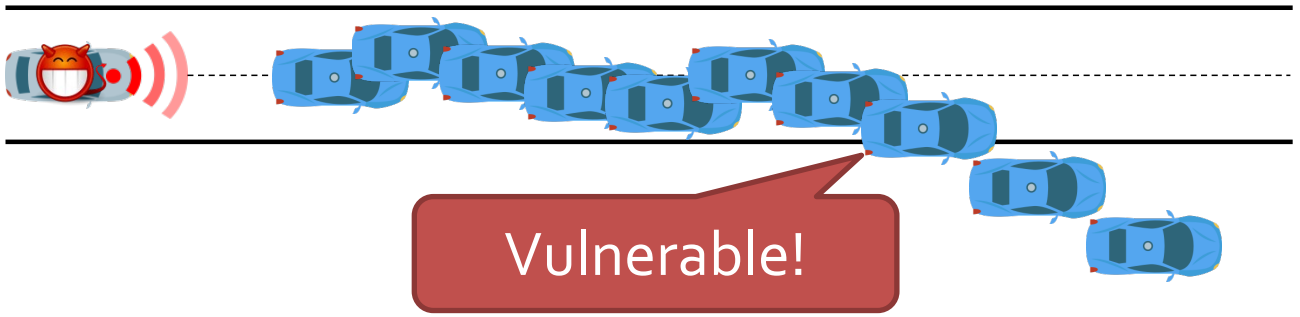
Take-over: fundamentally defeats the **design principle of MSF**!

Cause analysis

- Methodology: Identify possible contributing factors in MSF design, perform correlation analysis to reason causality
- Finding: Mainly appear in time periods when **MSF state & LiDAR localization outputs have *low confidence***
  - In such periods, MSF takes **more update from GPS** → Allows GPS inputs to **dominate** the fusion process
  - Created by **dynamic** & **non-deterministic** factors
    - Sensor noises, algorithm inaccuracies (e.g., LiDAR locator)
    - *That's why it's not observed in noise-free synthetic traces!*
- Even with **high-end AD sensors**, these factors are **large & frequent enough** for GPS spoofing to practically exploit

Exploit take-over vulnerability

- It is **highly attractive** for attacker to exploit take-over vulnerability
  - Attacker can reach **arbitrary** deviation goal
- However, **hard to predict/control** by attacker
- Needs to exploit in an **opportunistic** way
- Design a 2-stage attack: **vulnerability profiling + aggressive spoofing**



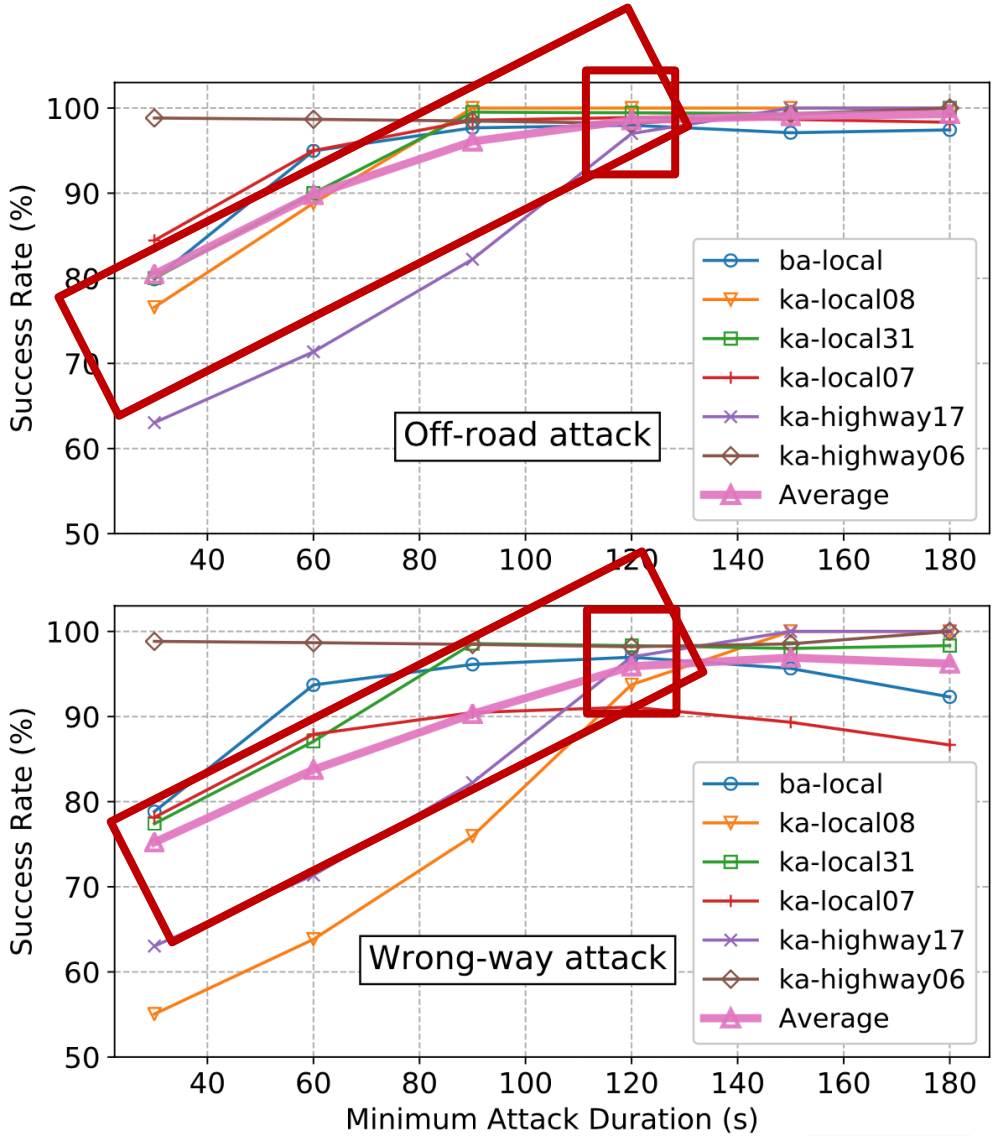
Stage 1: vulnerability profiling **Stage 2: aggressive spoofing**

Evaluation

- Main target: Apollo MSF binary
- Datasets:
  - Apollo trace for MSF localization
  - KAIST Complex Urban
- Success metric:

Attack goal	Local	Highway
Off-Road	0.895 m	1.945 m
Wrong-Way	2.405 m	2.855 m

- Effectiveness results
  - When min. attack duration is 2 min, can achieve **98.6%** & **95.9%** success rates for off-road attack & wrong-way attack
  - Takes only ***~30 sec*** to succeed

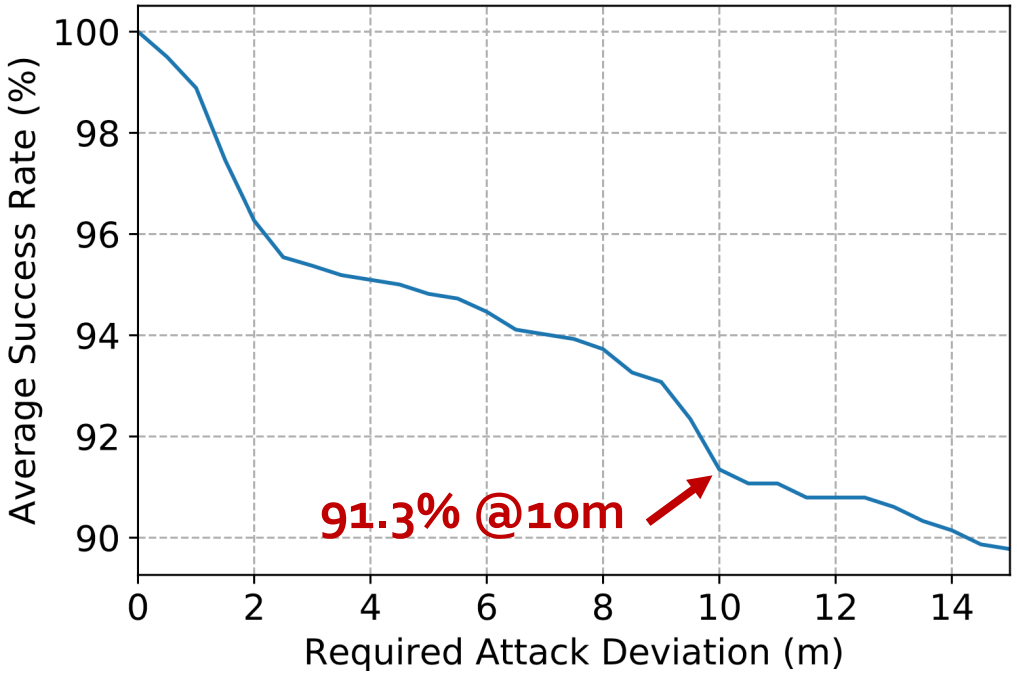


Evaluation

- Main target: Apollo MSF binary
- Datasets:
  - Apollo trace for MSF localization
  - KAIST Complex Urban
- Success metric:

Attack goal	Local	Highway
Off-Road	0.895 m	1.945 m
Wrong-Way	2.405 m	2.855 m

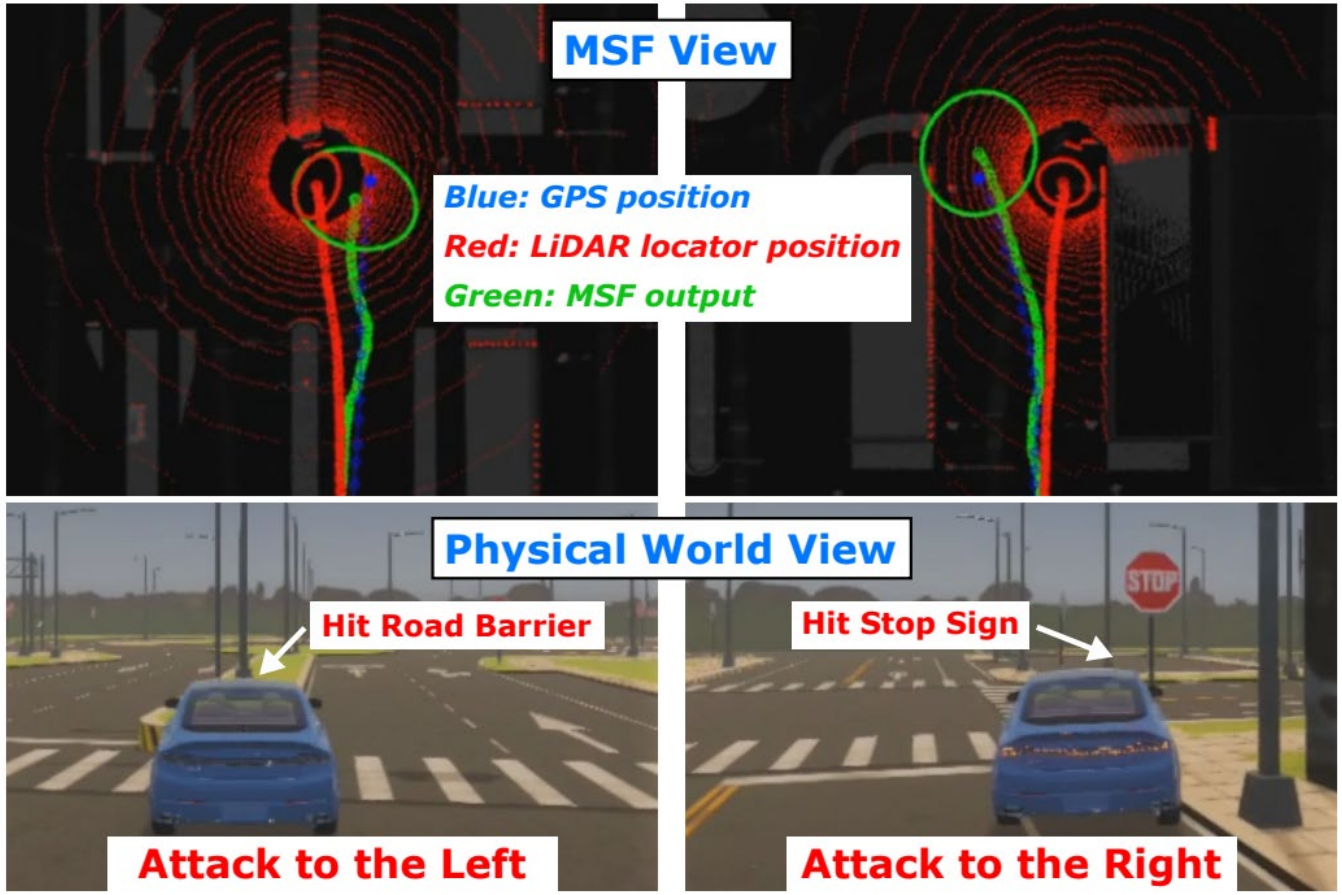
- Effectiveness results
  - When min. attack duration is 2 min, can achieve **98.6%** & **95.9%** success rates for off-road attack & wrong-way attack
  - Takes only ***~30 sec*** to succeed



Achievable deviation is not limited to wrong-way driving on highway

Attack demo

- Setup: Apollo 5.0 + LGSVL
- Demo location: Our website (https://sites.google.com/view/cav-sec/fusionripper)

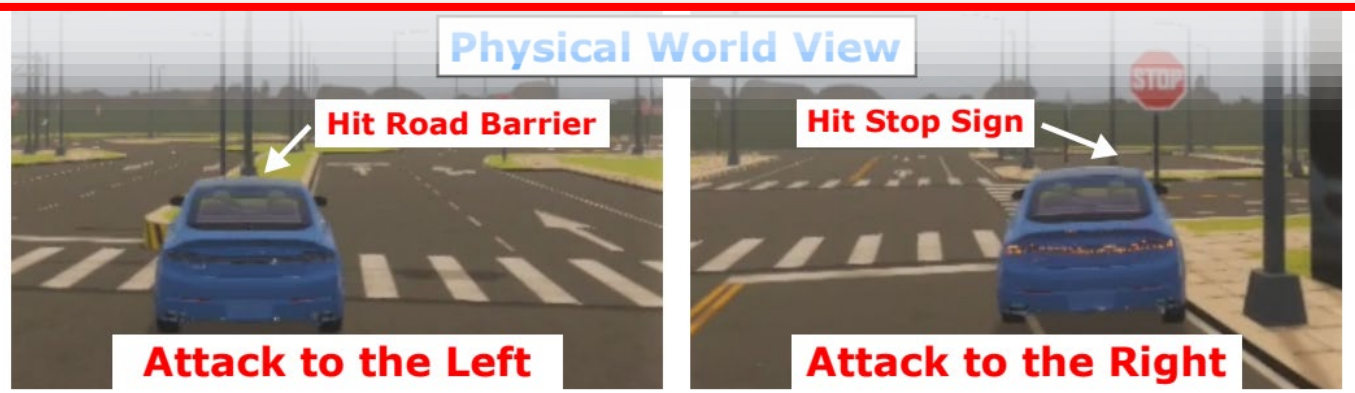


Attack demo

- Setu
- Den

per)

- All materials: https://sites.google.com/view/cav-sec/fusionripper
- Also have evaluations for ablation study, robustness, generality (w/ 2 other MSFs), comparison w/ naive attack, black-box attack design (profiling cost <= half a day), etc.
  - See our paper for more details (can be found on the website as well)
- Defense?
  - Fundamental solutions are not immediately deployable
    - E.g., prevent GPS spoofing, improve sensing and AD localization tech
  - GPS spoofing detection: Can make attack harder, but not a solved problem yet
    - All existing techniques have known evasion methods [Psiaki & Humphrey, Proc. IEEE'16]
  - Thus, the AI stack should always be prepared for GPS spoofing
    - ***Call for defense designs at the AI stack!***



Today: Cyber-attack surface to AD & V2X-based transp. AI

GPS

V2V
(vehicle-to-vehicle)

Cooperative Driving Automation (e.g., platoon)

AD vehicle

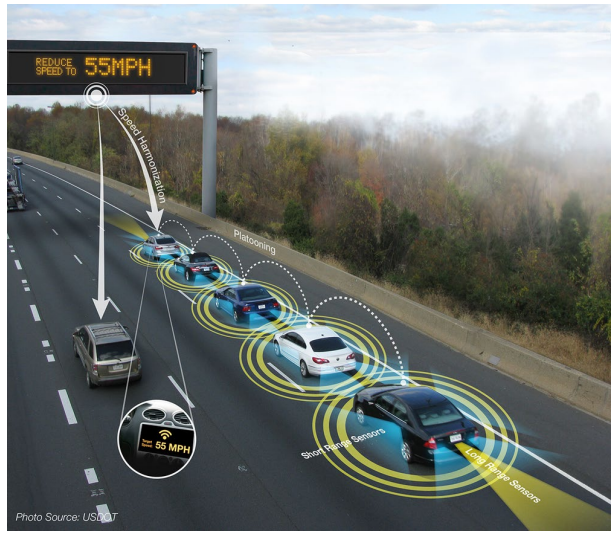
V2I
(vehicle-to-infrastructure)

Intelligent traffic light

Today: Cyber-attack surface to AD & V2X-based transp. AI

GPS

V2V
(vehicle-to-vehicle)



Cooperative Driving Automation (e.g., platoon)

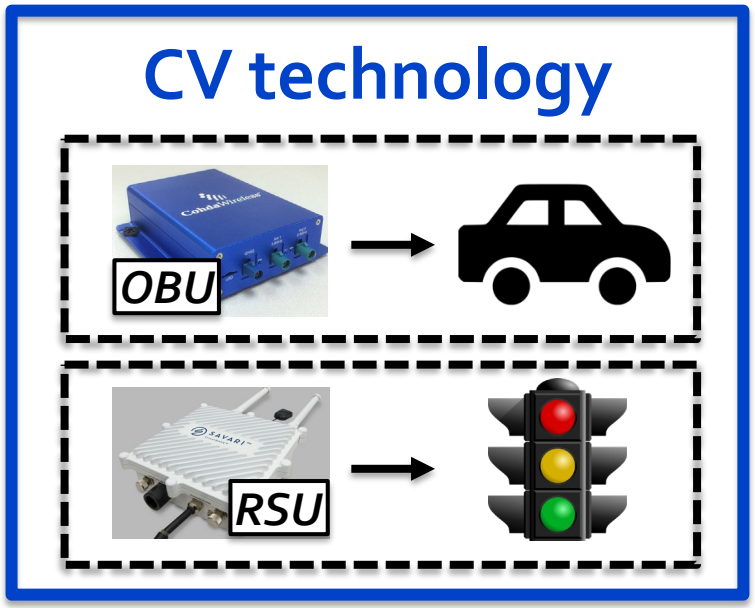
AD vehicle

V2I
(vehicle-to-infrastructure)

Intelligent traffic light

Background: CV (Connected Vehicle)/V2X (Vehicle-to-Everything) technology

- Wirelessly connect vehicles & infrastructure to dramatically improve **mobility**, **safety**, & **convenience**
- Expect to **soon** transform transportation systems today
  - 2016.9, USDOT launched ***CV Pilot Program***



Under deployment & testing

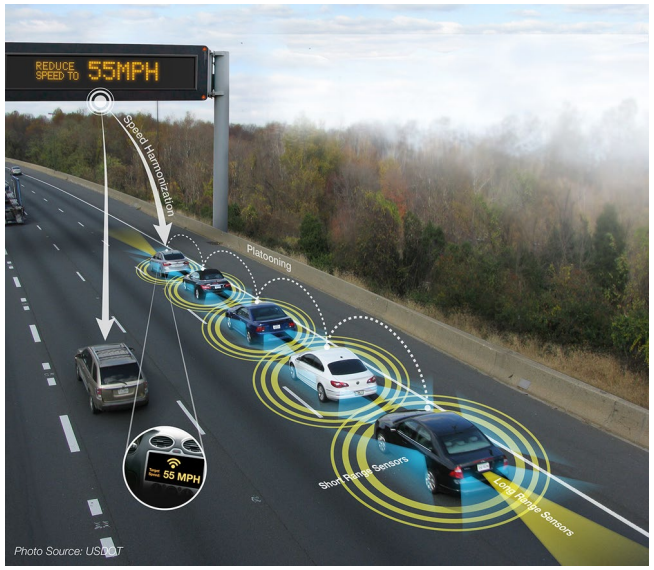
NEW YORK CITY

City of Tampa

WYOMING

CV/V2X-enabled transportation AI

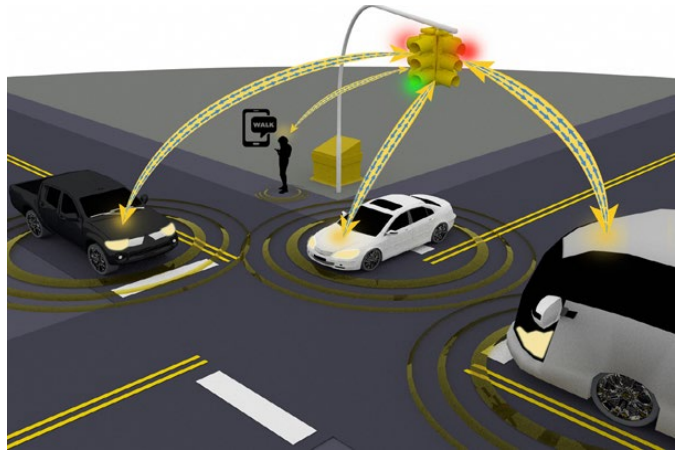
Safety warnings (e.g., forward collision warning)



Cooperative Driving Automation (e.g., platoon)

GPS

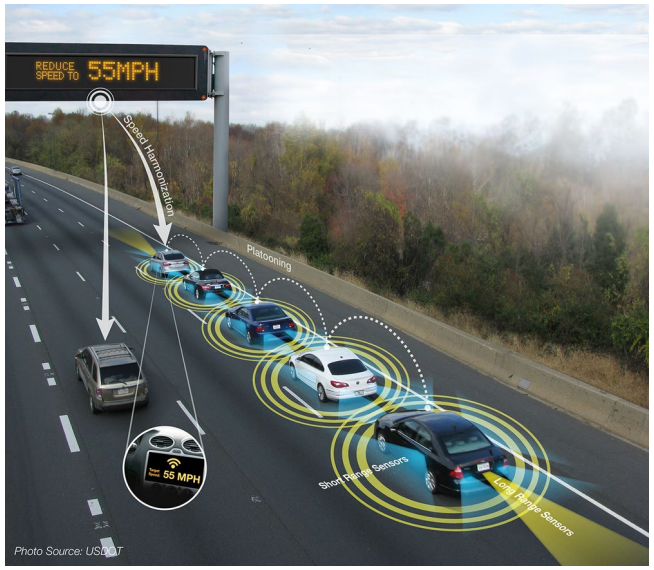
CV technology



Intelligent traffic light

CV/V2X-enabled transportation AI security

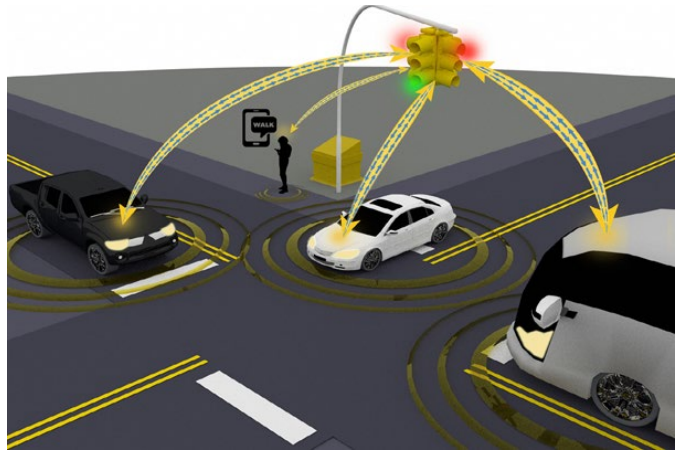
Safety warnings (e.g., forward collision warning)



Cooperative Driving Automation (e.g., platoon)

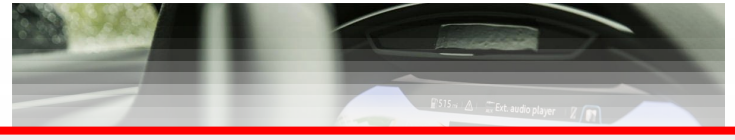
CV technology

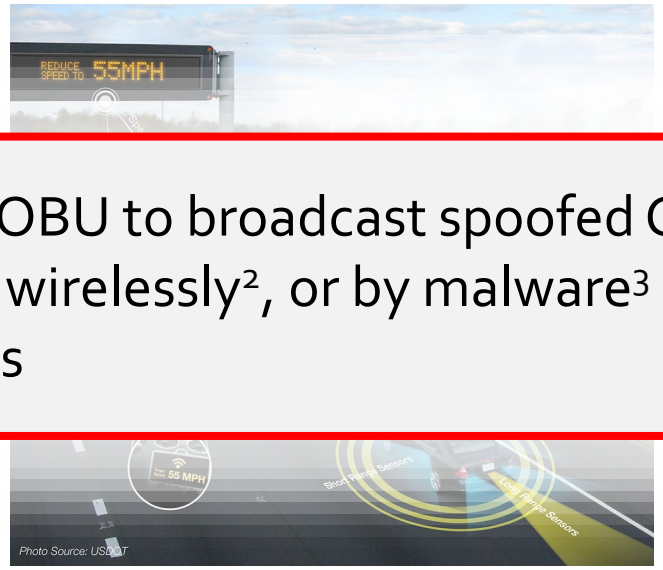
GPS



Intelligent traffic light

CV/V2X-enabled transportation AI security





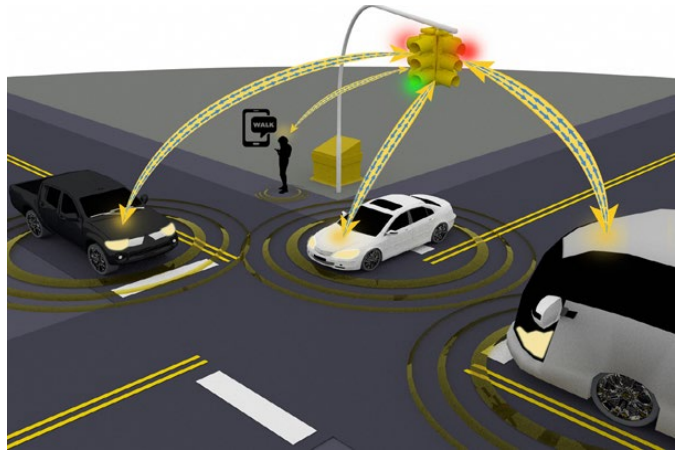
Safety wa
(e.g., fo
collision w

Malicious vehicle owners deliberately control OBU to broadcast spoofed CV data

- OBU itself can be compromised physically[1], wirelessly[2], or by malware[3]
- Compromise OBU input using sensor attacks

GPS

CV technology

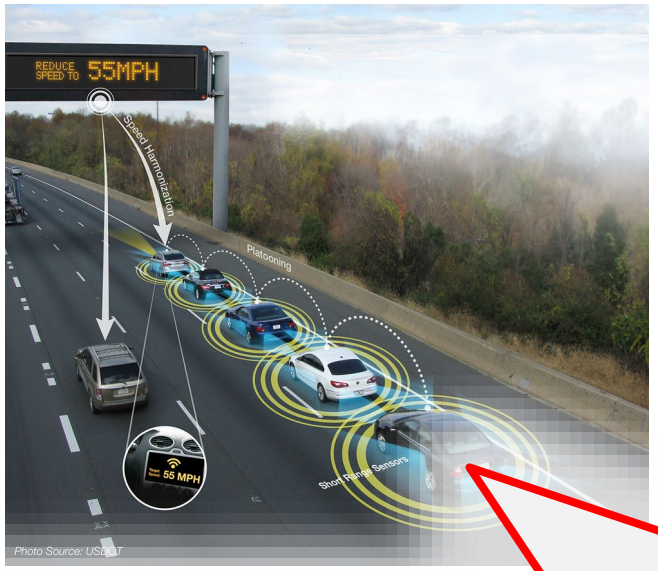


Intelligent traffic light

[1] Koscher et al. @IEEE S&P'10 [2] Checkoway et al. @Usenix Security'11 [3] Mazloom et al. @Usenix WOOT'16

CV/V2X-enabled transportation AI security

Safety warnings (e.g., forward collision warning)



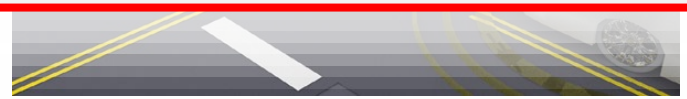
Cooperative Driving Automation (e.g., platoon)

GPS

CV technology

Prior works: Discovered that spoofing attacks can cause **collision** or **significant traffic flow instability** [IEEE Comm. Mag.'15, …, RAID'19]

- However, ***all relying on manual analysis*** --- *time-consuming*, *incomplete*, & *error-prone*



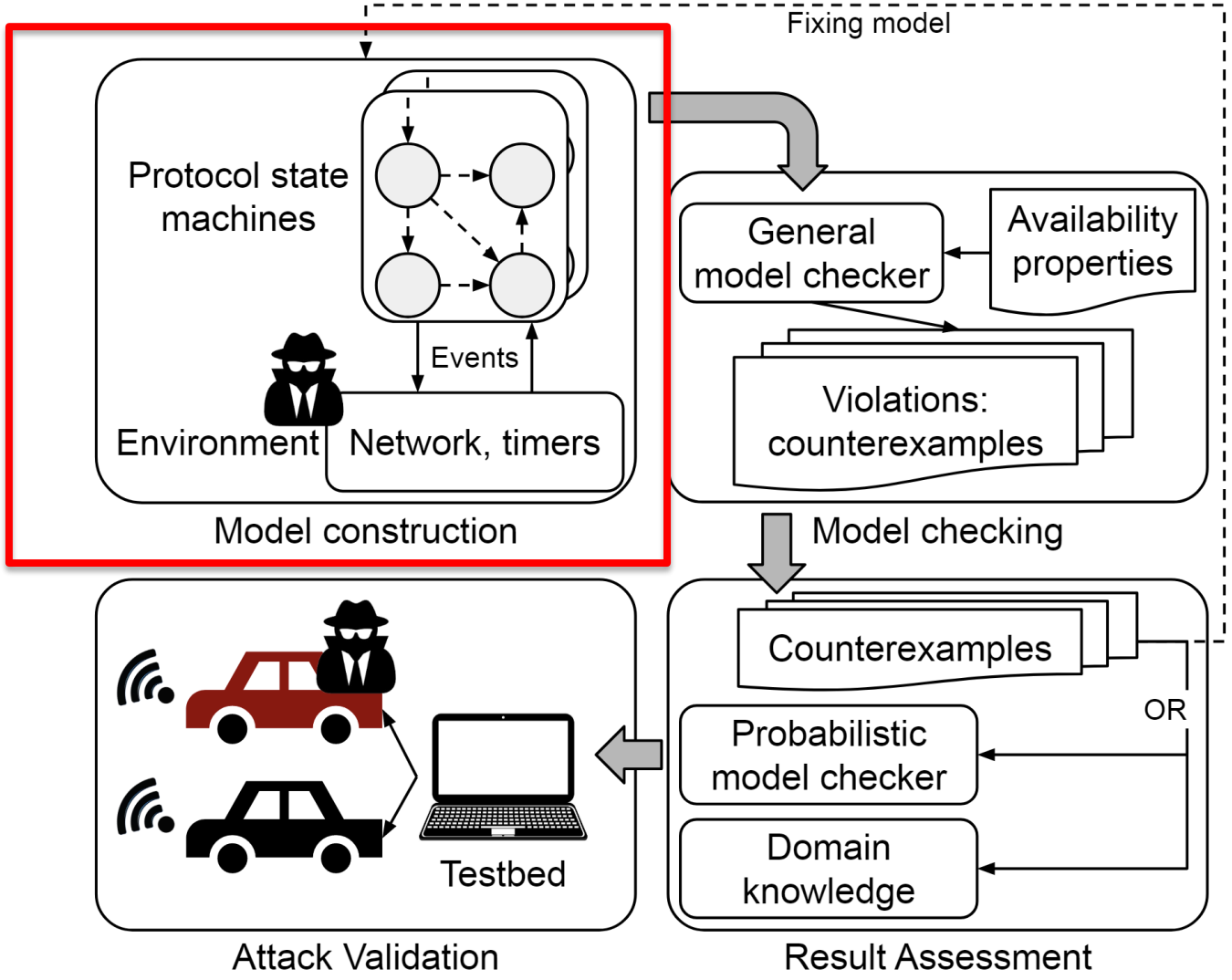
Intelligent traffic light

Our work: CVAnalyzer

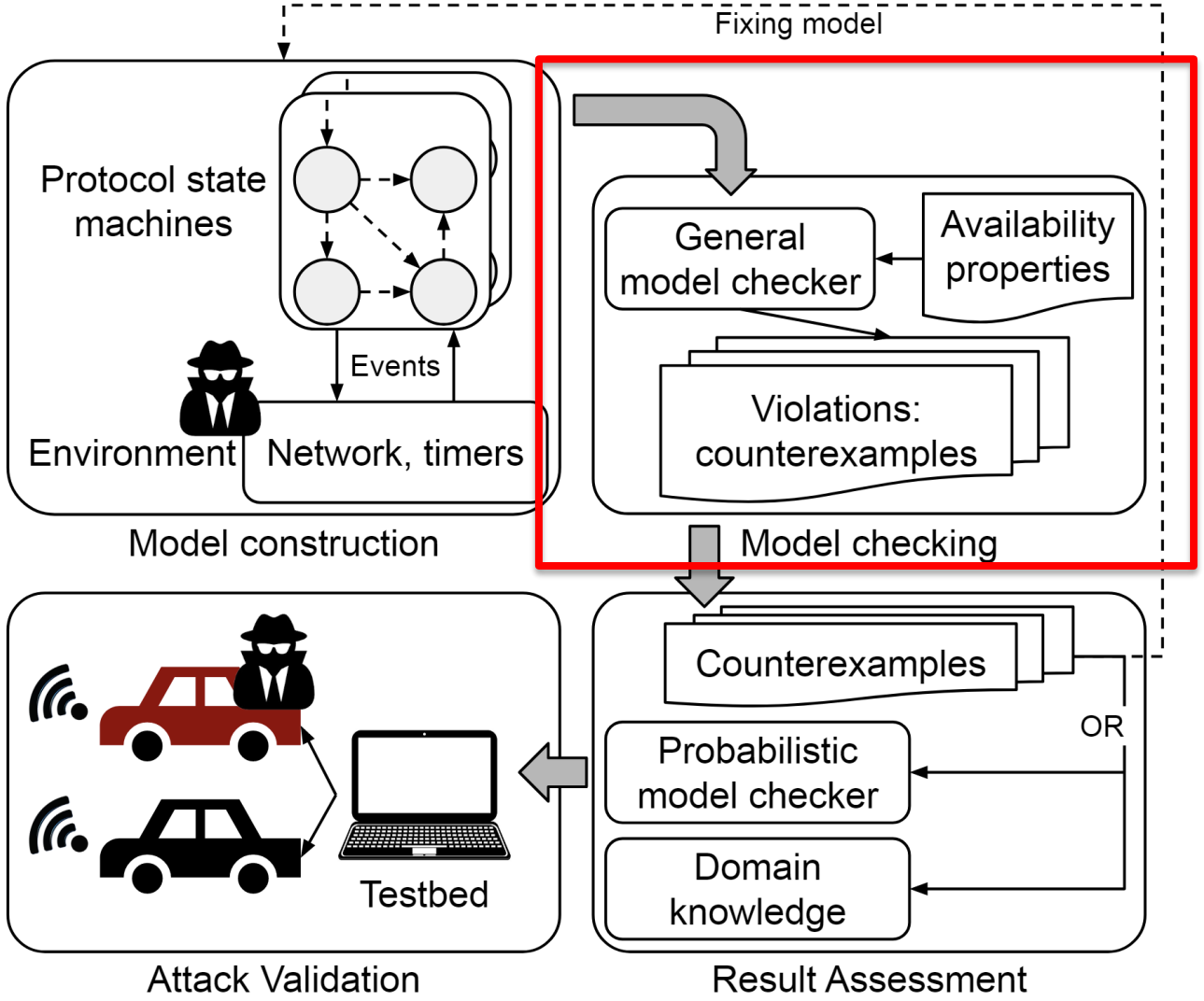
[Usenix Security'21]

- ***First* automatic vulnerability discovery** method in CV protocols using ***model checking***
  - Applicable to both ***network-layer CV protocols*** (e.g., IEEE 1609 protocol family) & ***application-layer protocols*** (e.g., cooperative driving AI protocol such as platoon management)
  - Focused on ***availability*** property
    - I.e., application layer should be always able to consume valid incoming packets
      - E.g, all CV devices should eventually learn unknown certificates; all platoon members should eventually switch to idle state
    - ***Important*** since its violations can prevent legitimate protocol participants from accessing critical services
      - E.g., can delay/prevent the receiving of safety-critical CV messages (e.g., forward collision warning) → collisions

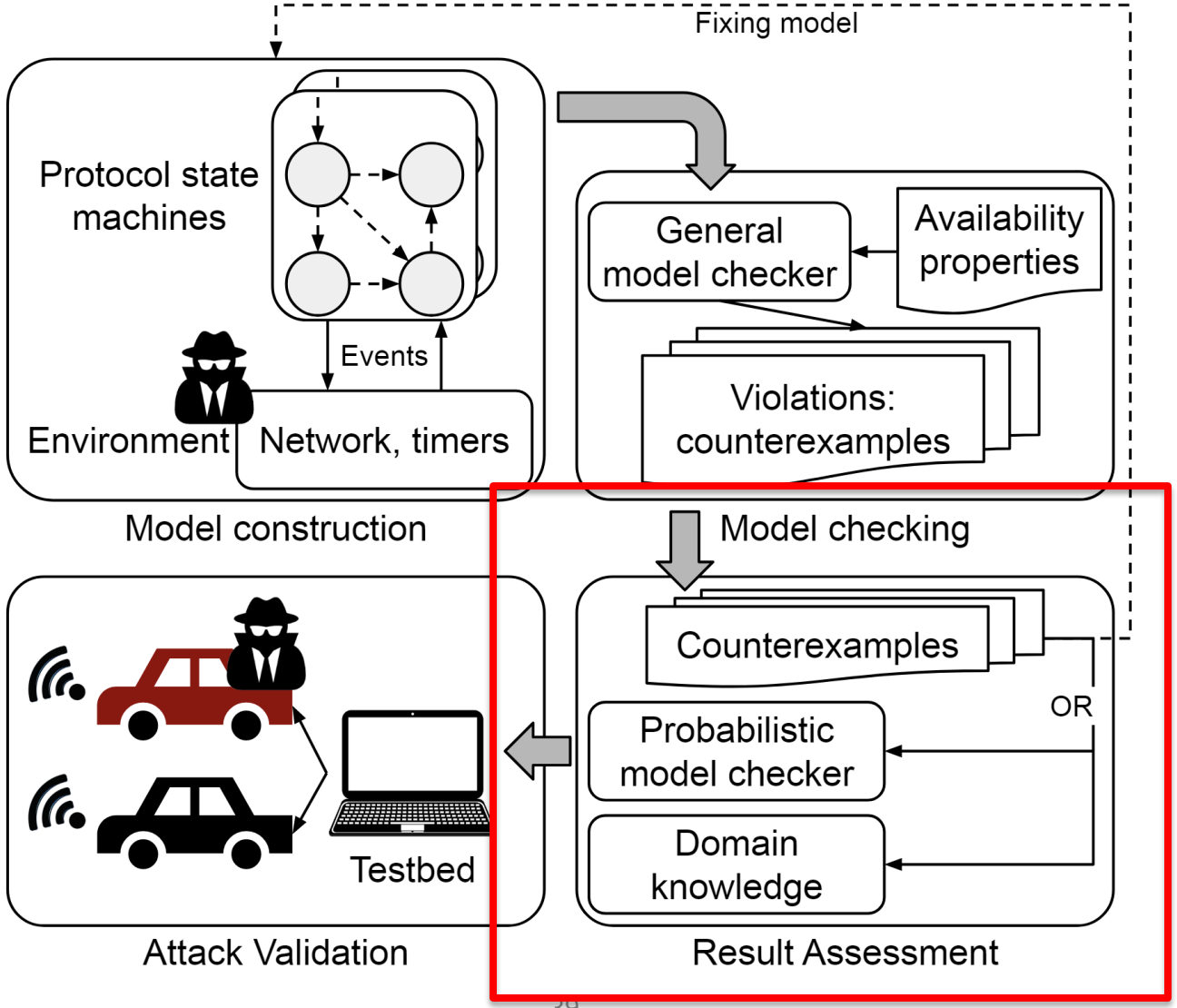
CVAnalyzer design



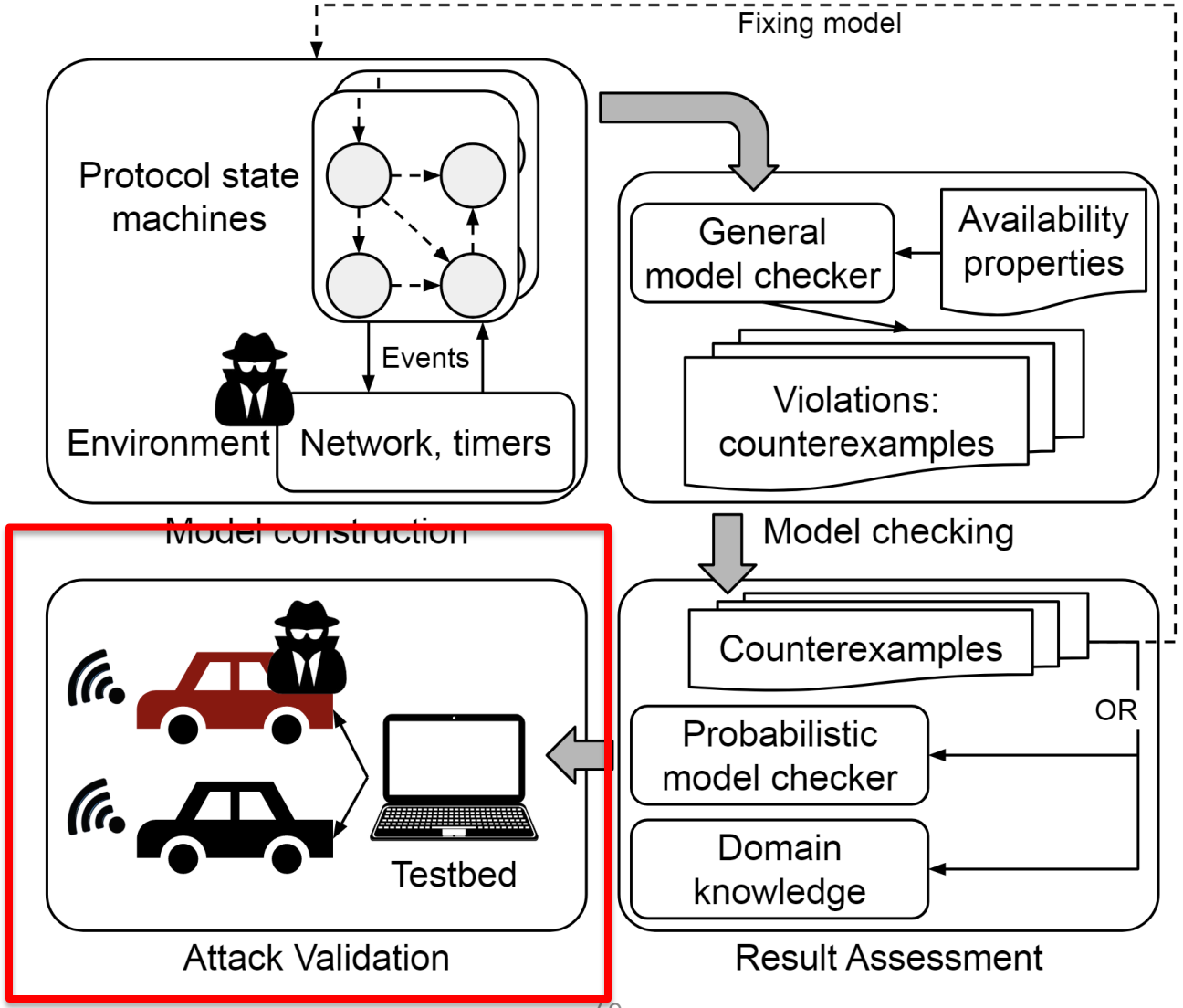
CVAnalyzer design



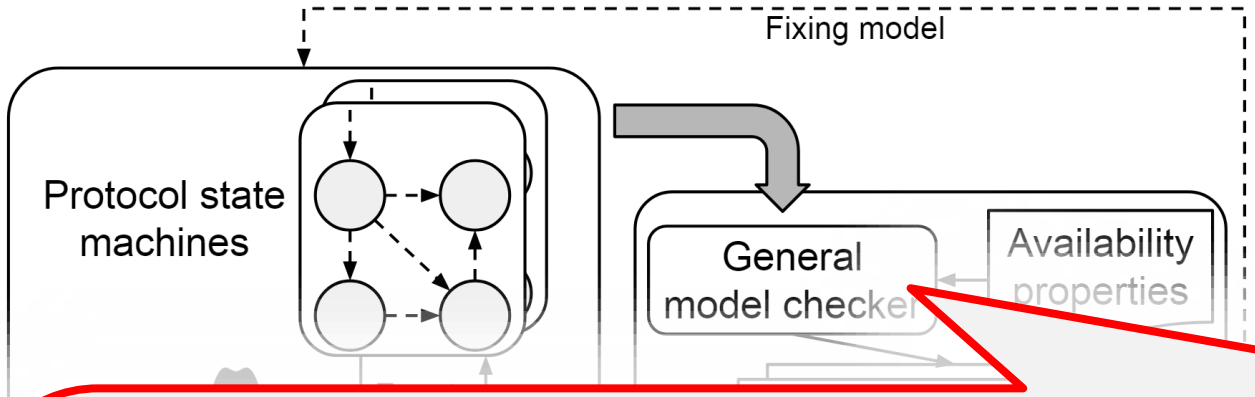
CVAnalyzer design



CVAnalyzer design



CVAnalyzer design



- Main challenge: State explosion
- Solution: Identify **problem-specific state reduction strategies** to eliminate unnecessary states while still preserving soundness
  - Strategy #1: Find **equivalent classes in inputs** that will by design trigger the same state transitions → Reduce the state input space
    - E.g., the use of 3-byte hashes to match certificates in CV
  - Strategy #2: Leverage differences between **attacker's action space & those of benign vehicles** → Reduce state space
    - E.g., attackers can send arbitrary fake certificate ids, but benign vehicle will only send its own
  - Effectiveness improvement: ***unfinished after >24 hrs → finish < 2hrs***

Results

- **19 discovered vuln (*18 new* *compared to manual discovery in prior works!*)**
  - ***4 (all new)*** from P2PCD (Peer-to-Peer Certificate Distribution) protocol in IEEE 1609
  - ***15 (14 new)*** from 2 popular platoon protocols (VENTOS, PLEXE)!

ID	Name	Implications
N1	Response Mute	Stop the CV device from sending learning responses
N2, N3	Request Mute	Stop the CV device from sending learning requests
N4	Numb	Stop the CV device from recording unknown certificates
A1, A2	(Prerequisites)	Cause traffic collision [1], lead to A3-15
A3, A4	Split Trigger	Interfere the traffic flow stability, decrease efficiency and safety
A5-14	PMP Block	Prevent platoon members from performing any maneuvers
A15	Inconsistency	Lead to failures of the split maneuver and the leader/follower leave maneuver

N*: CV network protocol, P2PCD A*: CV application, PMP

[1] Abdo et al. Application level attacks on connected vehicle protocols. RAID 2019

Results

- **19 discovered vuln (*18 new* *compared to manual discovery in prior works!*)**
  - ***4 (all new)*** from P2PCD (Peer-to-Peer Certificate Distribution) protocol in IEEE 1609
  - ***15 (14 new)*** from 2 popular platoon protocols (VENTOS, PLEXE)!

ID	Name	Implications
N1	Response Mute	Stop the CV device from sending learning responses
N2, N3	Request Mute	Stop the CV device from sending learning requests
N4	Numb	Stop the CV device from recording unknown certificates
A1, A2	(Prerequisites)	Cause traffic collision [1], lead to A3-15
A3, A4	Split Trigger	Interfere the traffic flow stability, decrease efficiency and safety
A5-14	PMP Block	Prevent platoon members from performing any maneuvers
A15	Inconsistency	Lead to failures of the split maneuver and the leader/follower leave maneuver

N*: CV network protocol, P2PCD A*: CV application, PMP

[1] Abdo et al. Application level attacks on connected vehicle protocols. RAID 2019

Results

- **19 discovered vuln (*18 new* *compared to manual discovery in prior works!*)**
  - ***4 (all new)*** from P2PCD (Peer-to-Peer Certificate Distribution) protocol in IEEE 1609
  - ***15 (14 new)*** from 2 popular platoon protocols (VENTOS, PLEXE)!

ID	Name	Implications
N1	Response Mute	Stop the CV device from sending learning responses
N2, N3	Request Mute	Stop the CV device from sending learning requests
N4	Numb	Stop the CV device from recording unknown certificates
A1, A2	(Prerequisites)	Cause traffic collision [1], lead to A3-15
A3, A4	Split Trigger	Interfere the traffic flow stability, decrease efficiency and safety
A5-14	PMP Block	Prevent platoon members from performing any maneuvers
A15	Inconsistency	Lead to failures of the split maneuver and the leader/follower leave maneuver

N*: CV network protocol, P2PCD A*: CV application, PMP

[1] Abdo et al. Application level attacks on connected vehicle protocols. RAID 2019

Results

- **19 discovered vuln (*18 new* compared to manual discovery in prior works!)**
  - ***4 (all new)*** from P2PCD (P
  - ***15 (14 new)*** from 2 popu

ID	Name	Im
N1	Response Mute	St
N2, N3	Request Mute	St
N4	Numb	St
A1, A2	(Prerequisites)	Ca
A3, A4	Split Trigger	Interfere the traffic flow stability, decrease efficiency and safety
A5-14	PMP Block	Prevent platoon members from performing any maneuvers
A15	Inconsistency	Lead to failures of the split maneuver and the leader/follower leave maneuver

Representative causes:

- Use **short hash** size for certificate matching
  - E.g., ***3 bytes*** in P2PCD for performance purposes → only ***10k*** **offline** certificate generation to find a collision due to the birthday paradox!
- Allow **unicast** message when the design **assumes broadcast** messages (e.g., message volume throttling)
- Lack of handling for **non-responding receiver**
- Lack of consistency-checking for **global states** (e.g., whether a platoon member lies about its position)

N*: CV network protocol, P2PCD A*: CV application, PMP

[1] Abdo et al. Application level attacks on connected vehicle protocols. RAID 2019

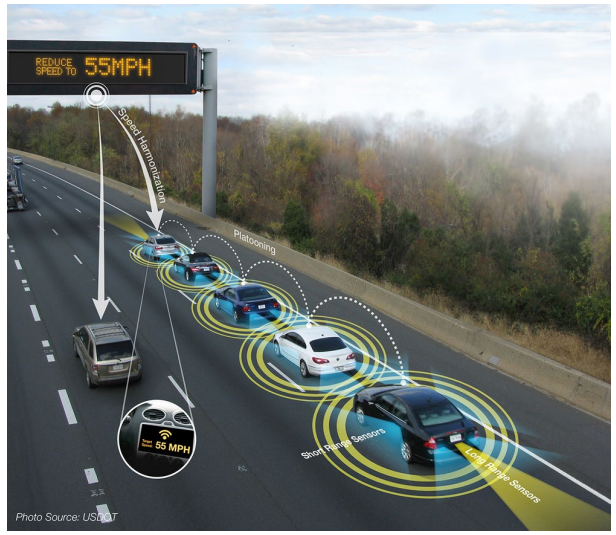
Result validation & reporting

- Successfully validate all attacks in a **real-world testbed**
  - OBU: Ubuntu 16.04 + closed source IEEE 1609.x
- Interestingly, some protocol implementation details **makes the attack easier**:
  - N1 and N2: indefinitely block communication
  - N1: only require 3 malicious packets rather than 4
  - N2: only require 3-byte hash collision instead of 8-byte collision
- Vulnerability report
  - **Reported** to & received **vuln acknowledgements** for *all 4 newly-discovered P2PCD vulns* from **IEEE 1609 Working Group**
    - Now discussion mitigation solutions, **planned to be integrated into the next version of IEEE 1609.2**

Today: Cyber-attack surface to AD & V2X-based transp. AI

GPS

V2V
(vehicle-to-vehicle)



Cooperative Driving Automation (e.g., platoon)

AD vehicle

V2I
(vehicle-to-infrastructure)

Intelligent traffic light

Today: Cyber-attack surface to AD & V2X-based transp. AI

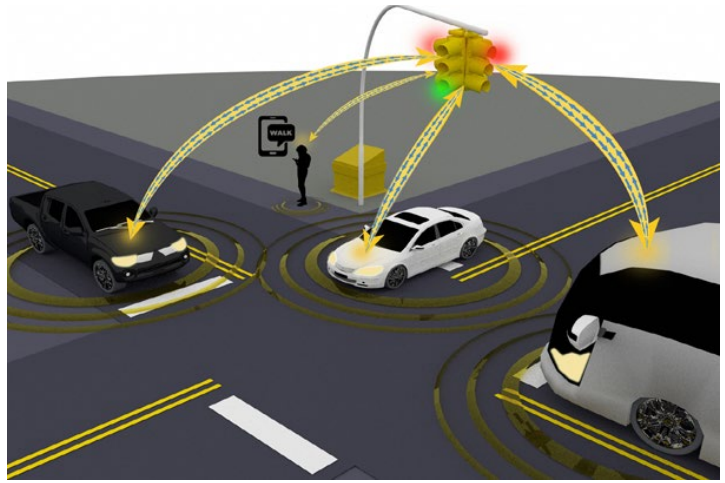
GPS

V2V
(vehicle-to-vehicle)

Cooperative Driving Automation (e.g., platoon)

AD vehicle

V2I
(vehicle-to-infrastructure)



Intelligent traffic light

Today: Cy… …o. AI

First to study security of infrastructure-side CV systems [NDSS'18]

- Target: USDOT Intelligent Traffic Signal (I-SIG) system
- Attack vector: CV data spoofing
- Impact: ***One single attack vehicle can create massive traffic jams!***
  - Root cause: New security vuln at ***traffic control algorithm*** level
  - Demo: https://sites.google.com/view/cav-sec/congestion-attack

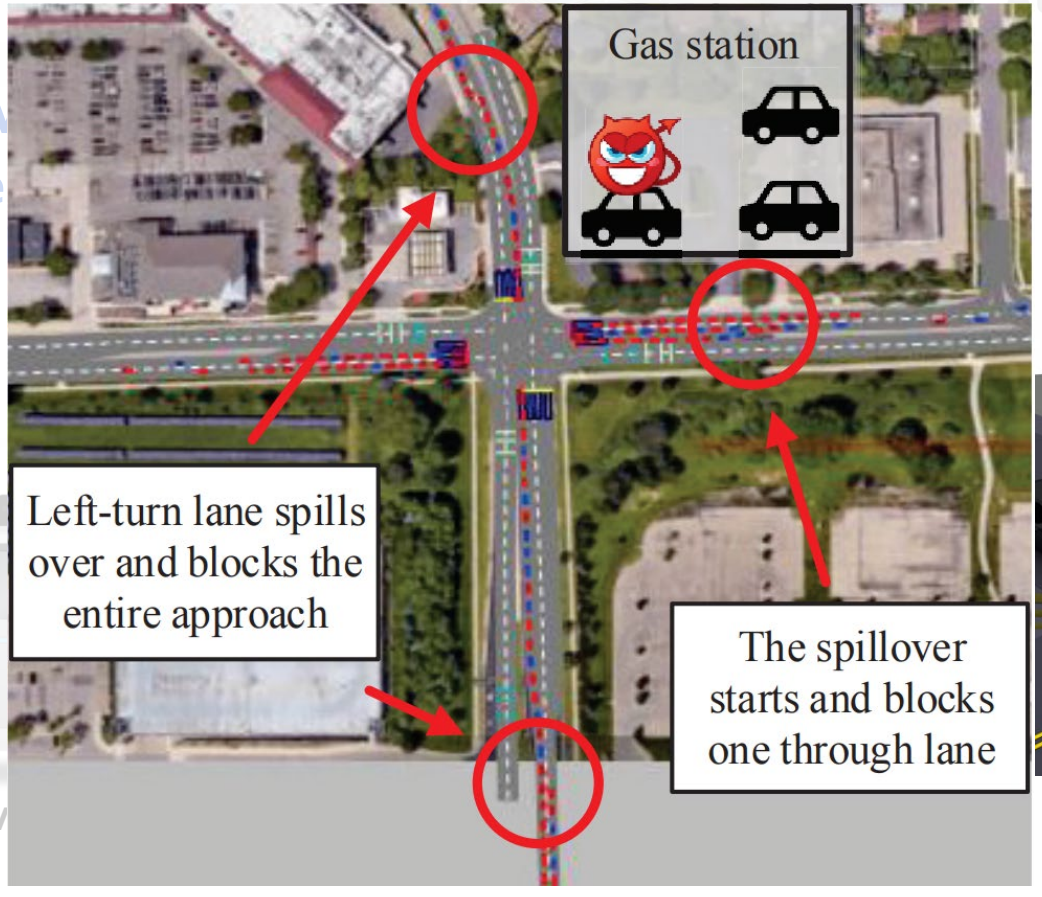
GPS

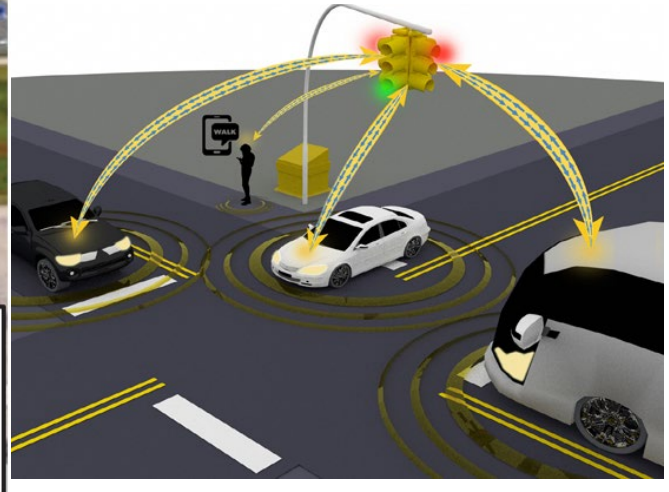
(vehicle…

…riving

…platoon)

AD v…





Intelligent traffic light

Today: Cy… …o. AI

First to study security of infrastructure-side CV systems [NDSS'18]

- Target: USDOT Intelligent Traffic Signal (I-SIG) system
- Attack vector: CV data spoofing
- Impact: ***One single attack vehicle can create massive traffic jams!***
  - Root cause: New security vuln at ***traffic control algorithm*** level
  - Demo: https://sites.google.com/view/cav-sec/congestion-attack

Defenses:

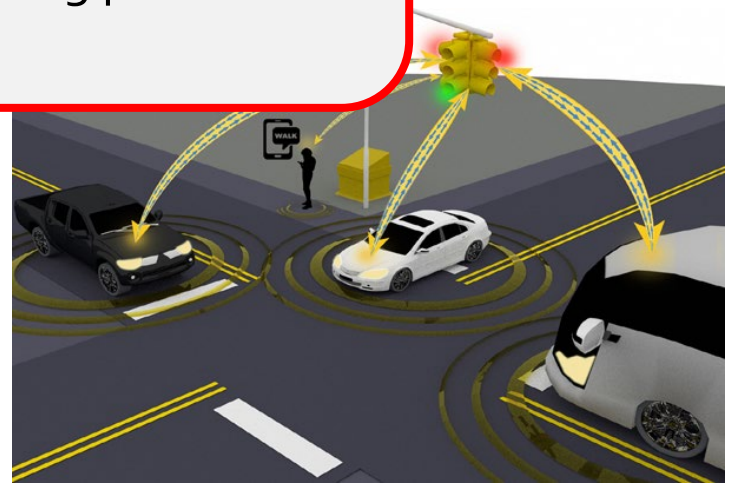
- [TRB'19] Trajectory-based attack detection at ***transportation infrastructure*** side
- [AutoSec'20 ***Best Paper Award***] Hardware-based spoofing prevention at ***vehicle*** side

GPS

AD vehicle

V2I

(vehicle-to-infrastructure)

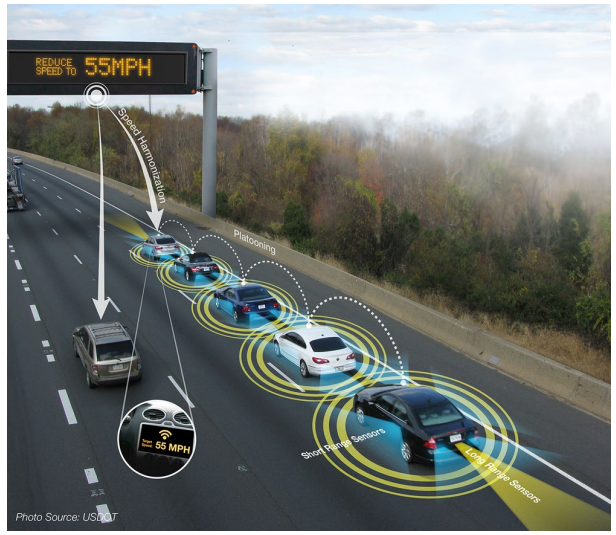


Intelligent traffic light

So far, cyber-attack surface to AD & V2X-based transp. AI

GPS

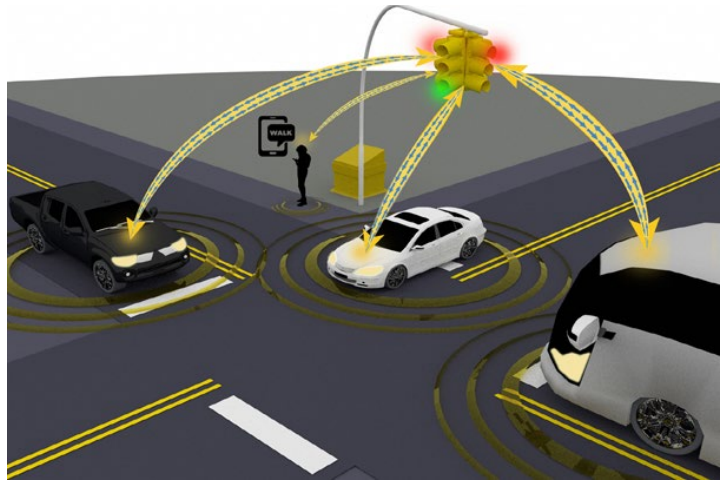
V2V
(vehicle-to-vehicle)



Cooperative Driving Automation (e.g., platoon)

AD vehicle

V2I
(vehicle-to-infrastructure)



Intelligent traffic light

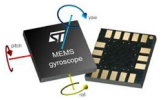
Physical-layer attack surface to AD & V2X-based transp. AI

Vehicle-side sensors

Camera

LiDAR

RADAR



IMU

Cooperative Driving Automation (e.g., platoon)

V2V

*Also large attack surface at **physical layer**!*

- ***Sensor attacks** (e.g., spoofing/jamming) & **physical-world attacks** (e.g., add stickers to physical objects)*

Infrastructure-side sensors

Camera

LiDAR

RADAR

Intelligent traffic light

V2I

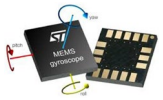
(vehicle-to-infrastructure)

Vehicle-side sensors

Camera

LiDAR

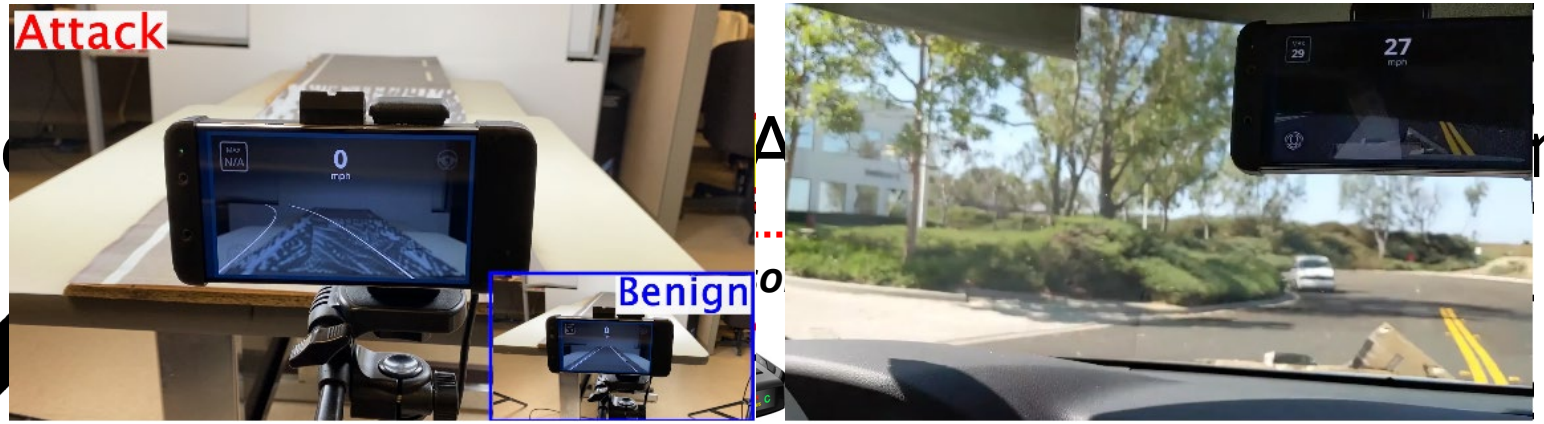
RADAR



IMU

AD vehicle

Physi … A … m my group

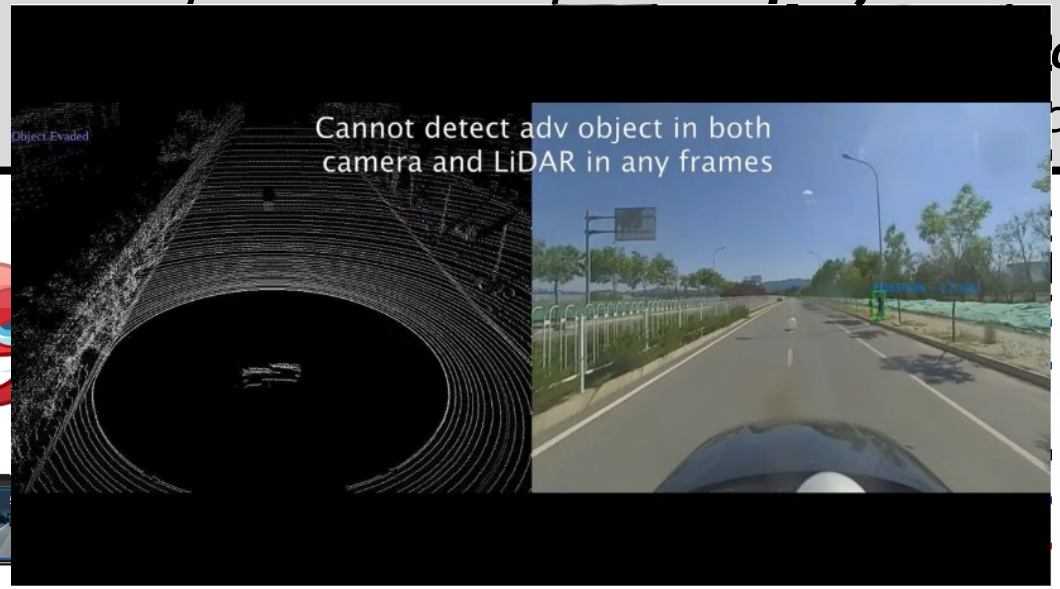


[Sato et al., Usenix Security'21 (NDSS'20 Best Poster)]

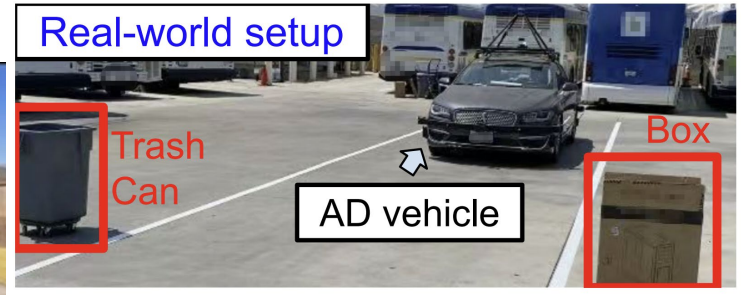
All demos are at https://sites.google.com/view/cav-sec

*Also large attack surface at **physical layer***

- … cal-
… cts)



[Cao et al., IEEE S&P'21]



[Wan et al., NDSS'22]

Conclusion

- **My group:** Actively researching **AI stack security** in AD & intelligent transportation, under both *cyber-* & *physical-layer attack vectors*
  - Collection of our efforts: https://sites.google.com/view/cav-sec
- ***Only the beginning*** of this research problem space
  - For example, now mostly on attack side, need more on ***defense*** side
  - To facilitate community building & broader impacts:
    - Co-found ***ACM/ISOC AutoSec (Automotive & Autonomous Vehicle Security) Workshop (2019 -)***, co-located **w/ *NDSS'21*** & ***'22***
    - Co-created ***DEF CON's first AutoDriving-themed hacking competition*** in 2021 (one of world's most famous hacker convention)
    - Served on **NIST focused group & panel on *AD AI test standards & metrics***

National Institute of Standards and Technology

Sponsors:

Qualcomm

Conclusion

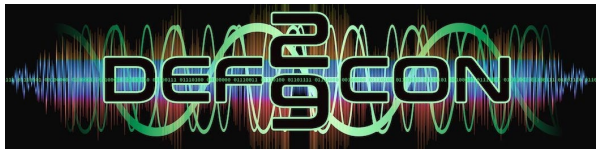
- **My group:** Actively researching **AI stack security** in AD & intelligent transportation, under both *cyber-* & *physical-layer attack vectors*
  - Collection of our efforts: **https://sites.google.com/view/cav-sec**
- ***Only the beginning*** of this research problem space
  - For example, now mostly on attack side, need more on ***defense*** side
  - To facilitate community building & broader impacts:
    - Co-found ***ACM/ISOC AutoSec (Automotive & Autonomous Vehicle Security) Workshop (2019 -)***, co-located **w/ *NDSS'21*** & ***'22***
    - Co-created ***DEF CON's first AutoDriving-themed hacking competition*** in 2021 (one of world's most famous hacker convention)
    - Served on **NIST focused group & panel on *AD AI test standards & metrics***
  - **Happy to chat more & form collaborations!**

Contact

Alfred Chen (alfchen@uci.edu)

Homepage:https://www.ics.uci.edu/~alfchen/

AS²Guard Autonomous & Smart Systems Guard Research Group



Sponsors:

NSF

DEPARTMENT OF TRANSPORTATION UNITED STATES OF AMERICA

gm

TOYOTA

Qualcomm