THE UNIVERSITY OF BRITISH COLUMBIA

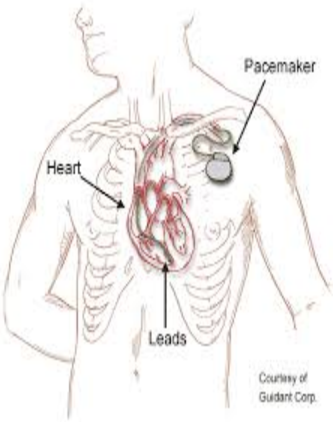
Stopping the Barbarians at the Gate: Protecting End User Devices from Security Attacks

Karthik Pattabiraman

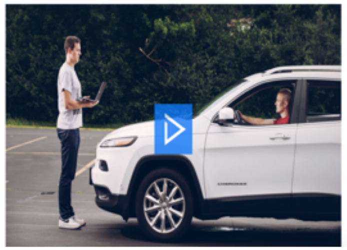
Pritam Dash, Mehdi Karimi, Aarti Kashyap, Zitao Chen, Guanpeng Li, Ekta Aggarwal, Maryam Raiyat, Amita Kamath, Julien Gascon-Samson, Farid Tabrizi, Andre Ivanov

University of British Columbia, Vancouver, Canada

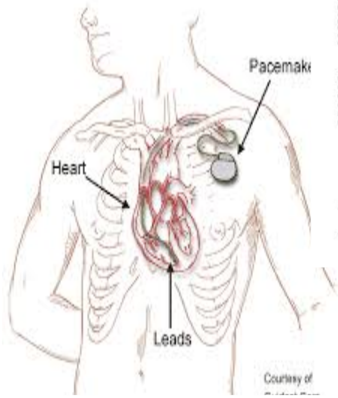
Cyber-Physical Systems (CPS): End User Devices



Cyber-Physical Systems (CPS): End User Devices


HACKERS REMOTELY KILL A JEEP ON THE HIGHWAY—WITH ME IN IT


Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses




Smart meters can be hacked to cut power bills

By Mark Ward
Technology correspondent, BBC News

16 October 2014 | Technology

Share

Smart meters could help people do a better job of managing power use

Smart meters widely used in Spain can be hacked to under-report energy use, security researchers have found.


Nest Thermostat Glitch Leaves Users in the Cold

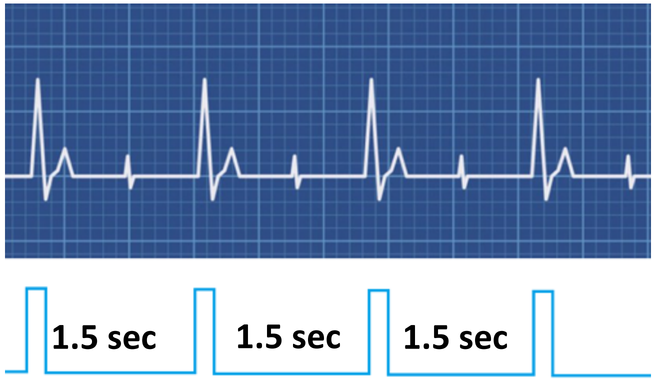
Disruptions
By NICK BILTON JAN. 13, 2016

The Nest Learning Thermostat is dead to me, literally. Last week, my once-beloved "smart" thermostat suffered from a mysterious software bug that drained its battery and sent our home into a chill in the middle of the night.

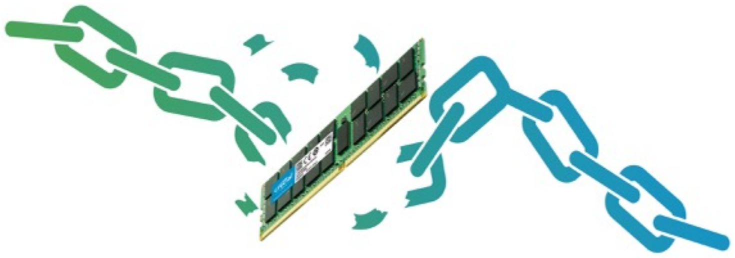
Although I had set the thermostat to 70 degrees overnight, my wife and I were woken by a crying baby at 4 a.m. The thermometer in his room read 64 degrees, and the Nest

CPS Challenges

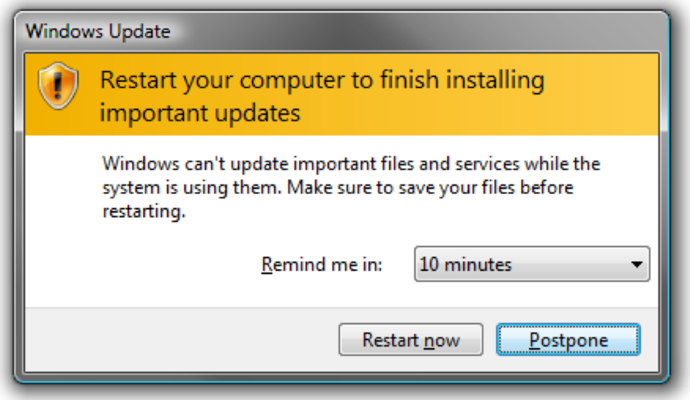
Real-time constraints



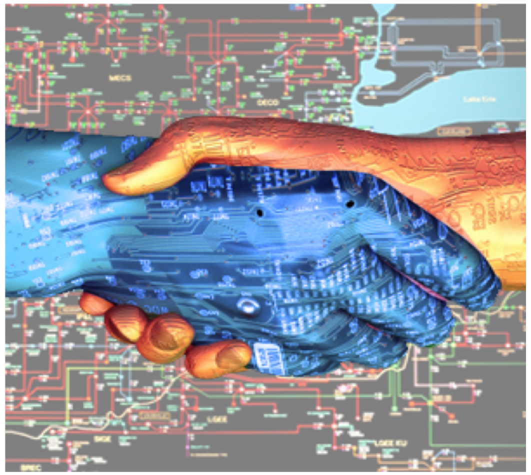
Resource constraints



Hard to Upgrade

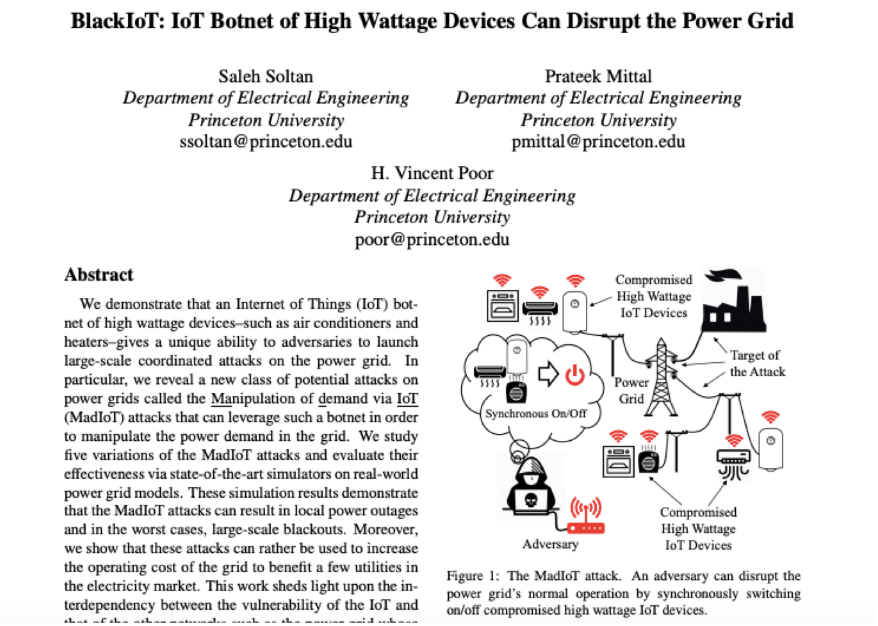


Have human interactions



Why should we care about end device security ?

- Often the first entry point for attackers (weakest link in the trust chain)
- Cause large-scale disruptions by taking over many end-user devices



BlackIoT: IoT Botnet of High Wattage Devices Can Disrupt the Power Grid

Saleh Soltan
Department of Electrical Engineering
Princeton University
ssoltan@princeton.edu

Prateek Mittal
Department of Electrical Engineering
Princeton University
pmittal@princeton.edu

H. Vincent Poor
Department of Electrical Engineering
Princeton University
poor@princeton.edu

Abstract

We demonstrate that an Internet of Things (IoT) botnet of high wattage devices–such as air conditioners and heaters–gives a unique ability to adversaries to launch large-scale coordinated attacks on the power grid. In particular, we reveal a new class of potential attacks on power grids called the Manipulation of demand via IoT (MadIoT) attacks that can leverage such a botnet in order to manipulate the power demand in the grid. We study five variations of the MadIoT attacks and evaluate their effectiveness via state-of-the-art simulators on real-world power grid models. These simulation results demonstrate that the MadIoT attacks can result in local power outages and in the worst cases, large-scale blackouts. Moreover, we show that these attacks can rather be used to increase the operating cost of the grid to benefit a few utilities in the electricity market. This work sheds light upon the interdependency between the vulnerability of the IoT and

Figure 1: The MadIoT attack. An adversary can disrupt the power grid's normal operation by synchronously switching on/off compromised high wattage IoT devices.

History Lesson: Barbarians at the Gate (410 AD)

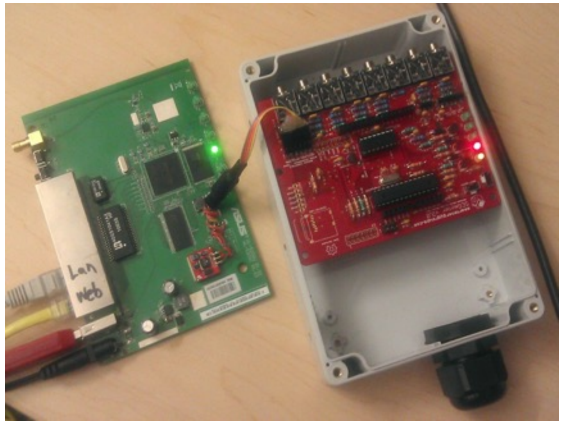
Image source: https://ludwigheinrichdyck.wordpress.com/2018/03/24/barbarians-at-the-gate-the-410-sack-of-rome/

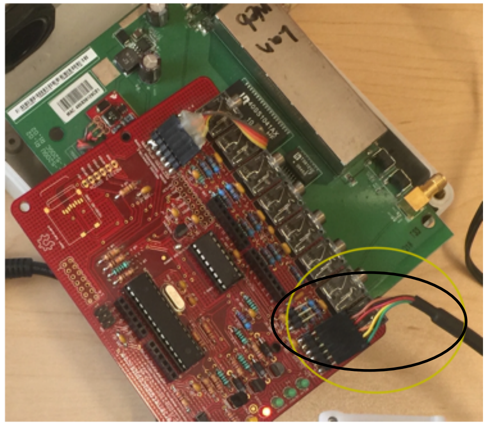
This Talk

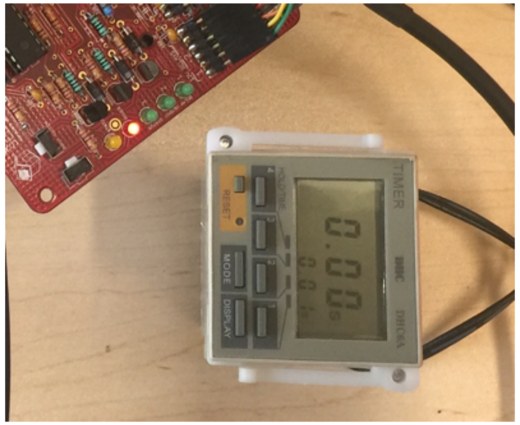
- Motivation
- Attacks on Embedded and IoT devices [ACSAC’19][ACSAC’16]
- Intrusion Detection Systems for Smart Devices [FSE’17][CPS-SPC’18]
- Ongoing work and conclusion

Past Work: Formal Analysis of Smart Meters

- Formally model the states of the CPS [TECS][ACSAC’16]
- Combine with formal attacker models
- Model-check the system for security invariants
  - Identify unsafe states and paths to unsafe states
  - Automatically mount the attacks on the system



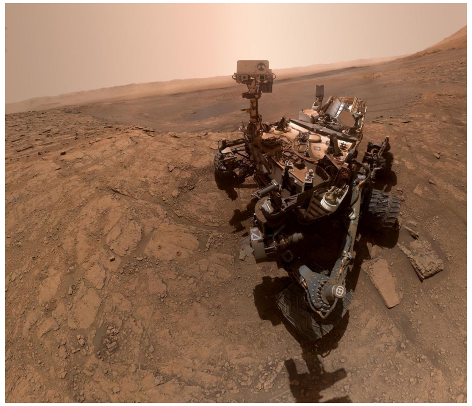




Robotic Vehicles (RV)

- Autonomous UAVs and Rovers.
  - Delivery
  - Warehouse Management
  - Surveillance
  - Cinematography

Autonomous RVs are increasingly becoming popular.
RV missions are time critical.



Motivation

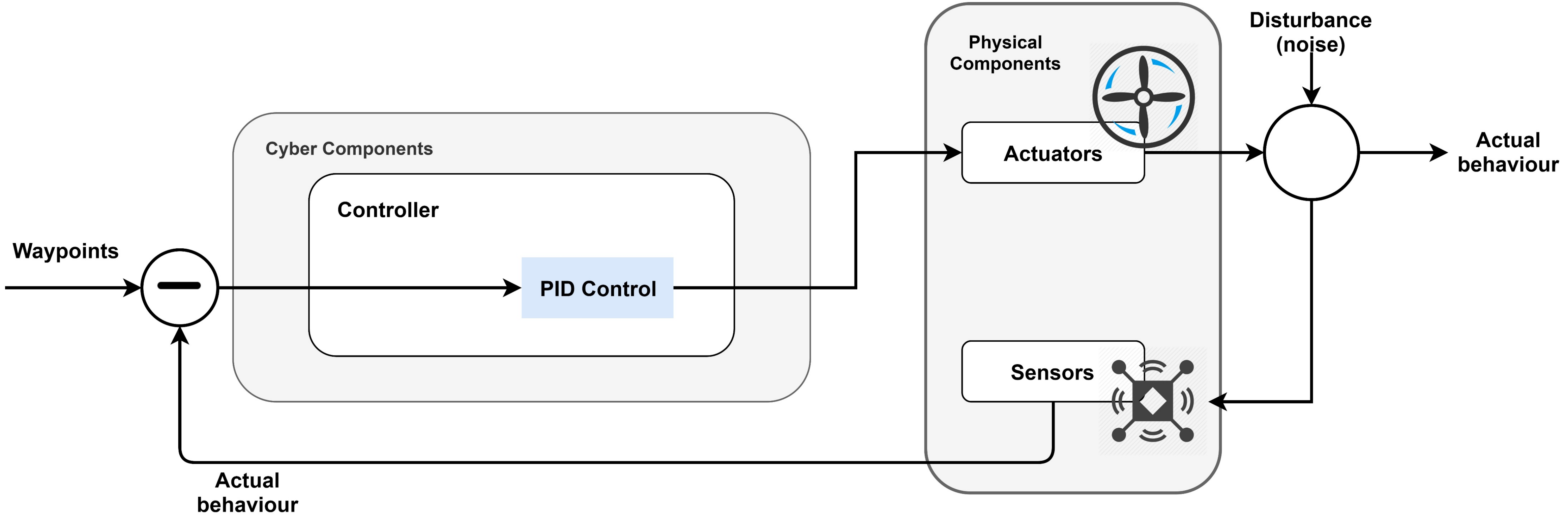
- GPS spoofing [ION GNSS'12], Optical spoofing [CCS'11]
- Acoustic noise injection in MEMS gyroscope [Usenix'15],
- MEMS accelerometer [Euro S&P'17]

However, all these techniques assume there's no protection deployed.

Can an attacker remain stealthy and trigger adversarial actions?

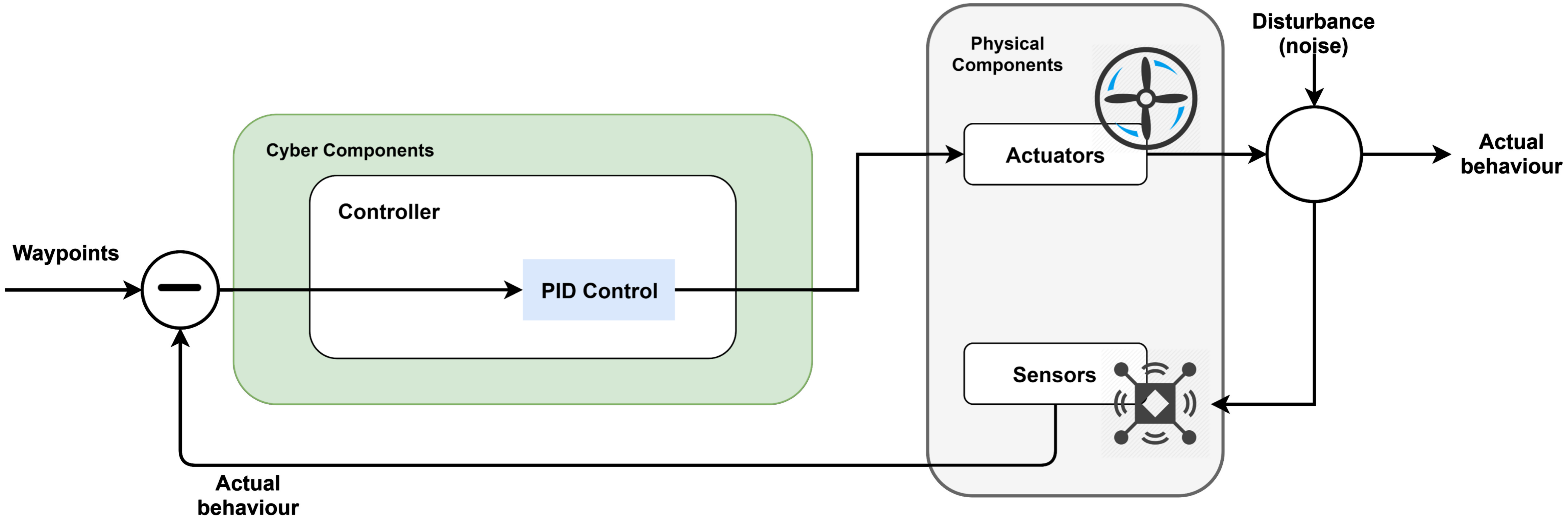
Robotic Vehicle System

- Cyber component
- Physical component



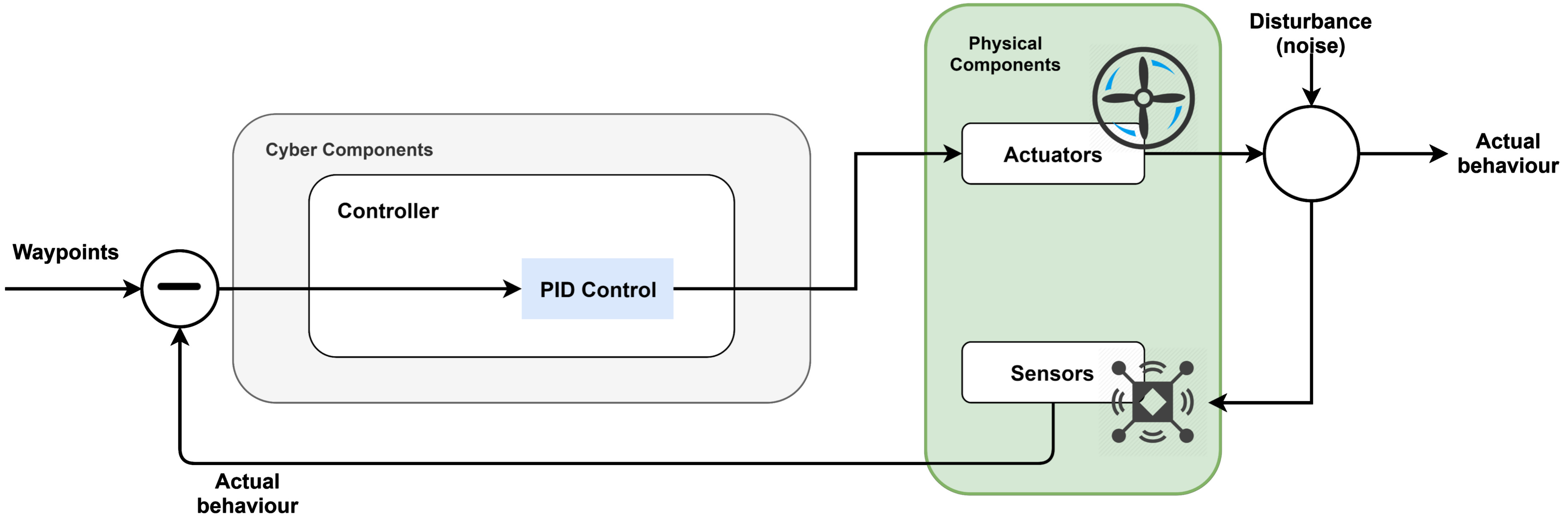
Robotic Vehicle System

- Cyber component
- Physical component



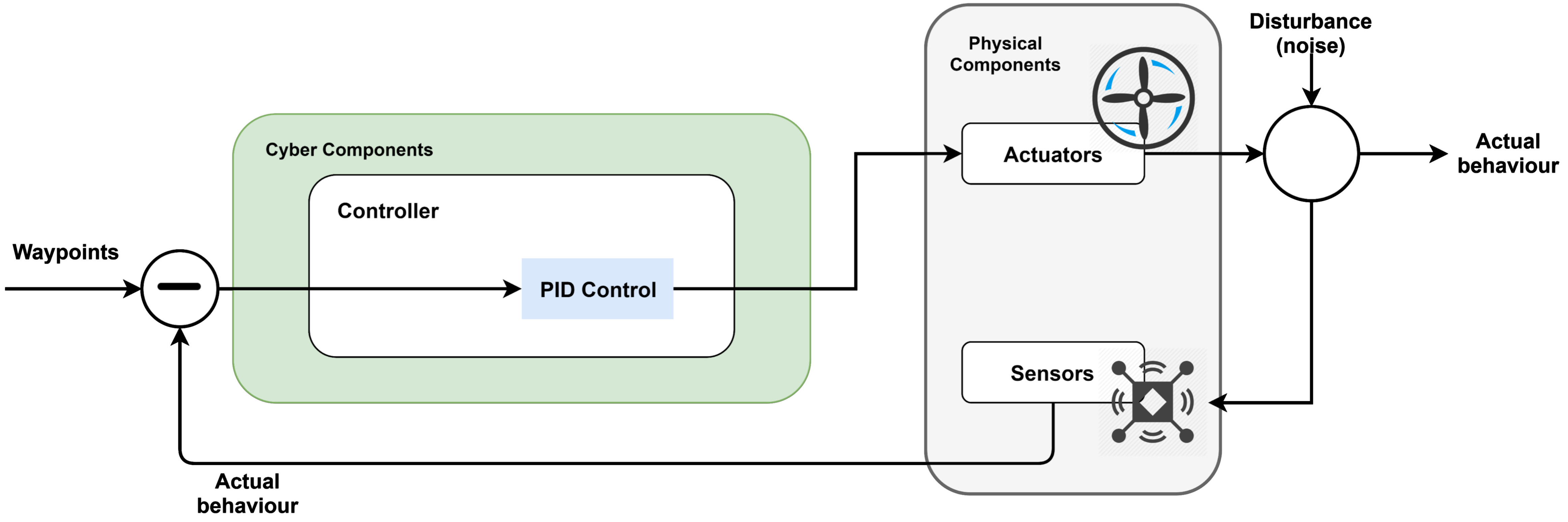
Robotic Vehicle System

- Cyber component
- Physical component



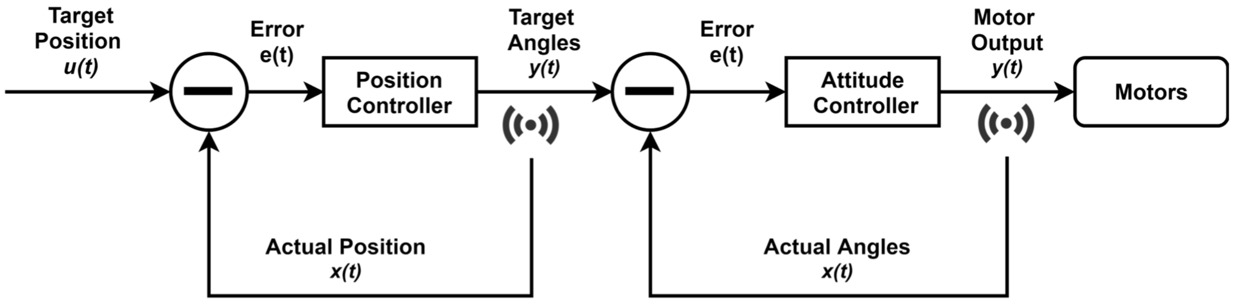
Robotic Vehicle System

- Cyber component
- Physical component

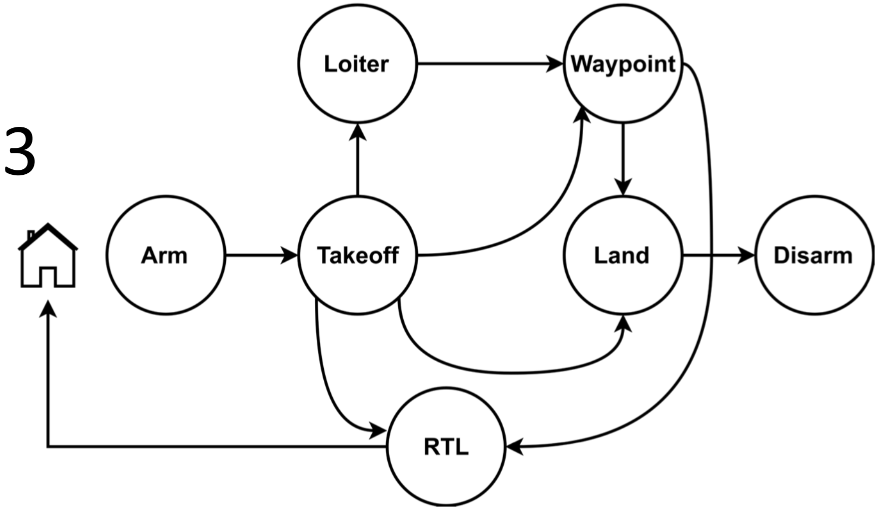


Autonomous Control in RVs

- Control algorithms
  - Position Controller
  - Attitude Controller



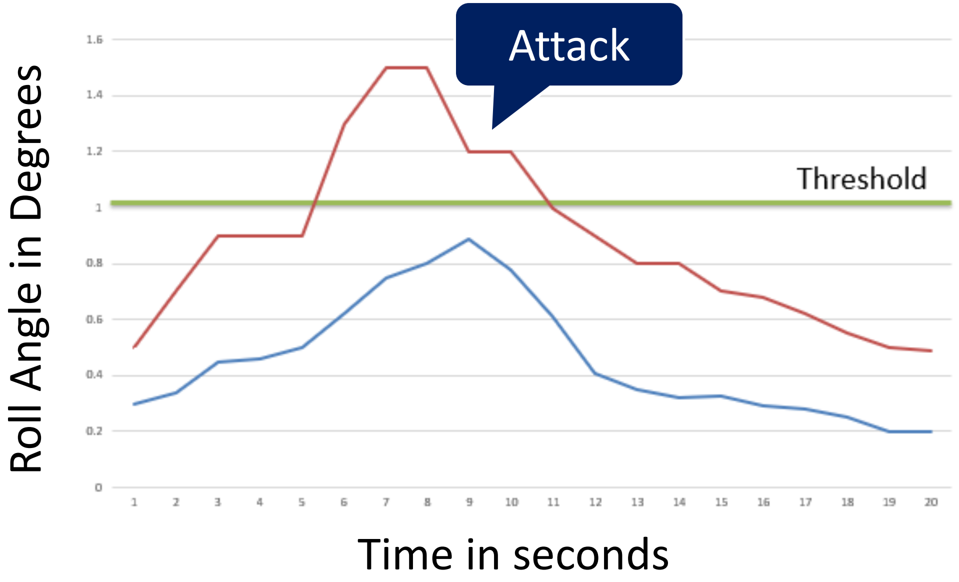
- Modes of Operation
  - A typical drone mission → at least 3 modes.



Control-based Attack Detection Techniques

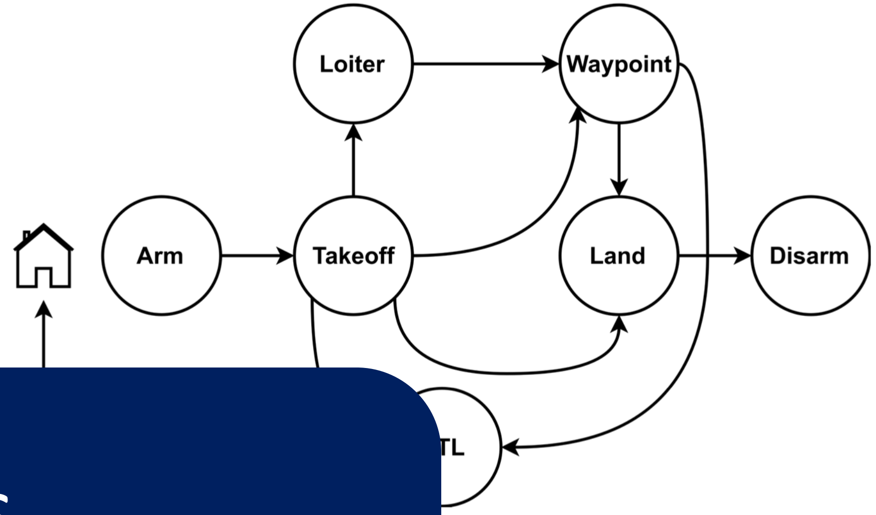
- Control Invariants (CI) [CCS’18]
  - State Space Model to predict target angles.

- Extended Kalman Filter (EKF)
  - Residual analysis → sensor or actuator attacks



Limitations in Control-based Detection

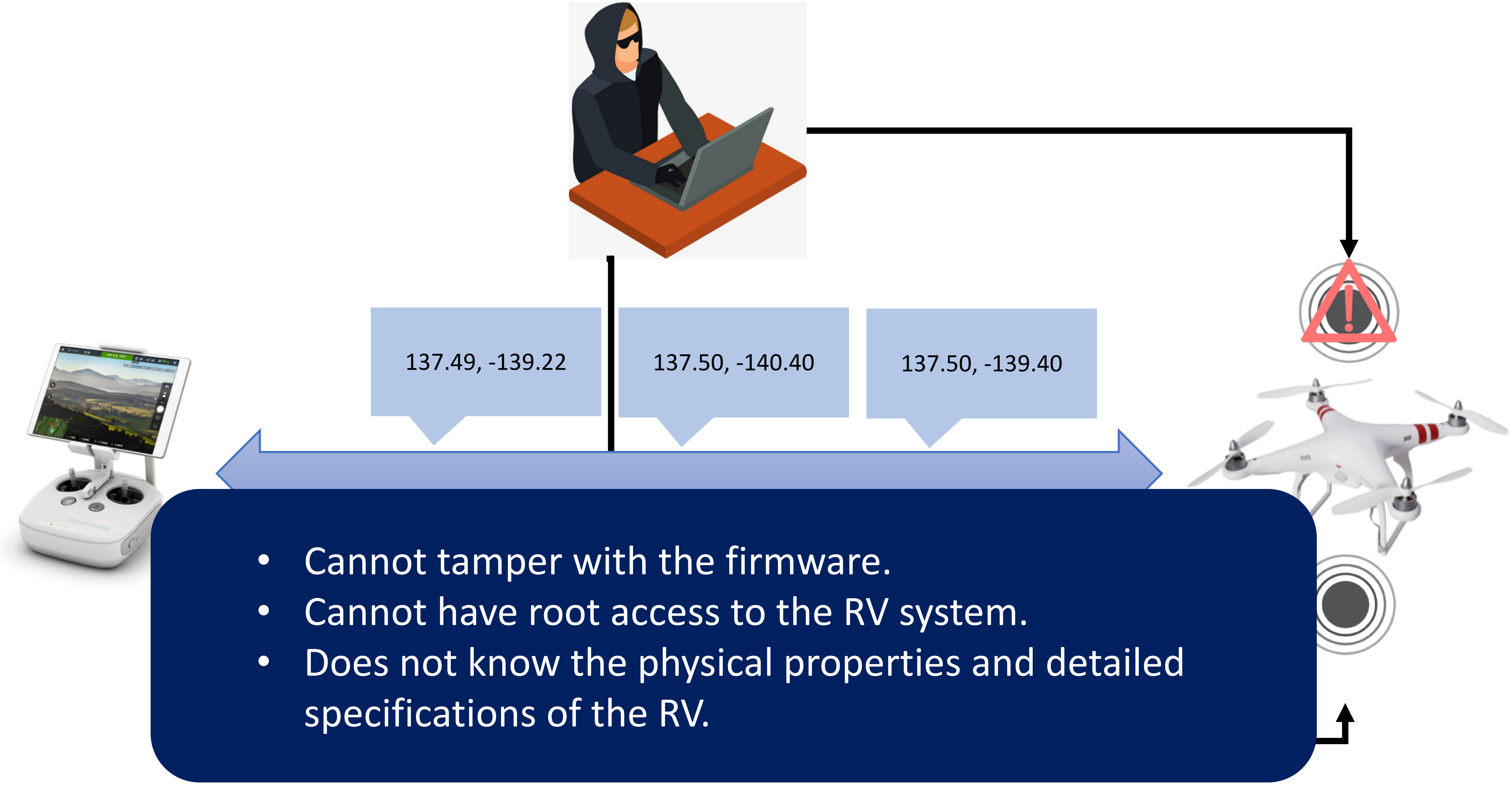
- Fixed threshold
  - Large threshold to reduce False Positives (FP).
    - Environmental factors – friction, wind
    - Sensor faults.
- Fixed Monit
- Often fail to
  - Takeoff
  - Waypoi



Stealthy Attacks

False Data Injection
Artificial Delay
Switch Mode Attack

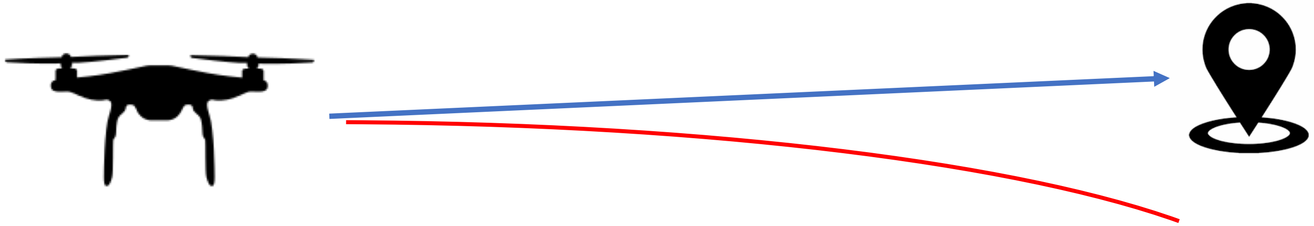
Attack Model



- Cannot tamper with the firmware.
- Cannot have root access to the RV system.
- Does not know the physical properties and detailed specifications of the RV.

Attack 1: False Data Injection Attack

- Tampering sensor measurements
  - Inject false data → sensor
  - Acoustic noise



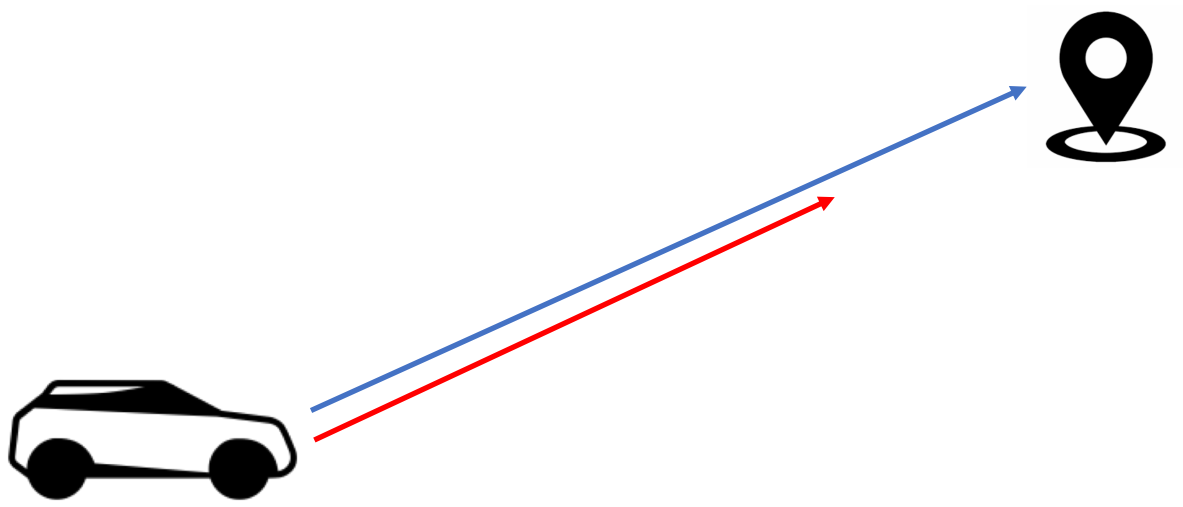
- False Data Injection
  - Delivery at a wrong location
  - Misplacements in warehouse

- [Usenix'15] Son et. al. Rocking Drones with Intentional Sound Noise on Gyroscopic Sensors

Attack 2: Artificial Delay Attack

- Delay system operations
  - Mode changes
  - Motor commands

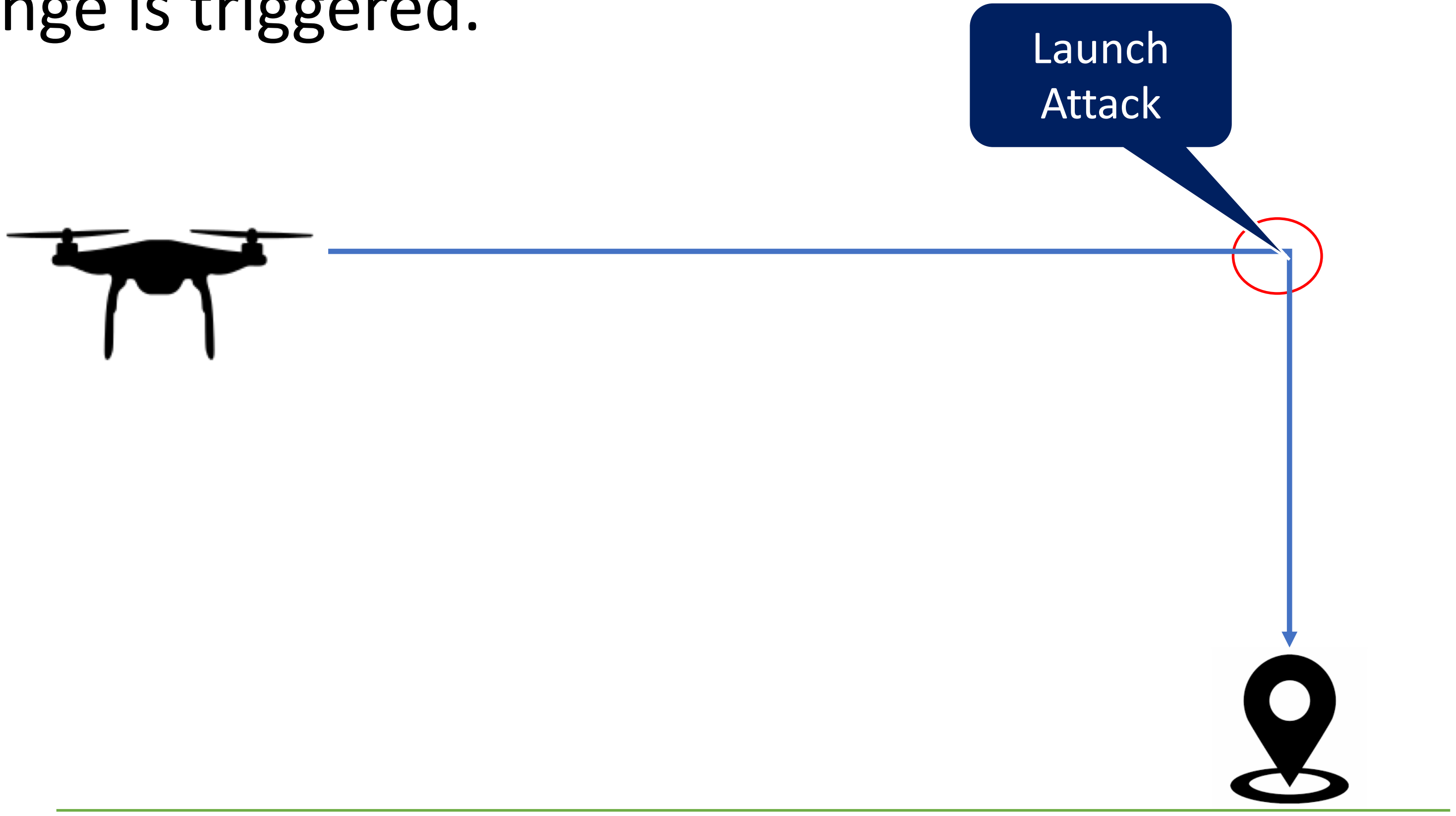
- Artificial delay attack
  - Delay receiving commands
  - Delays RV mission



Attack 3: Switch Mode Attack

- Initiated when a mode change is triggered.
  - Steady-state flight → Land
  - Takeoff → Waypoint

- Switch mode attack
  - Gain elevation instead of landing
  - Potential crash



Results and Evaluation

RQ1 How much effort does the attacker need to expend to derive the state estimation model?

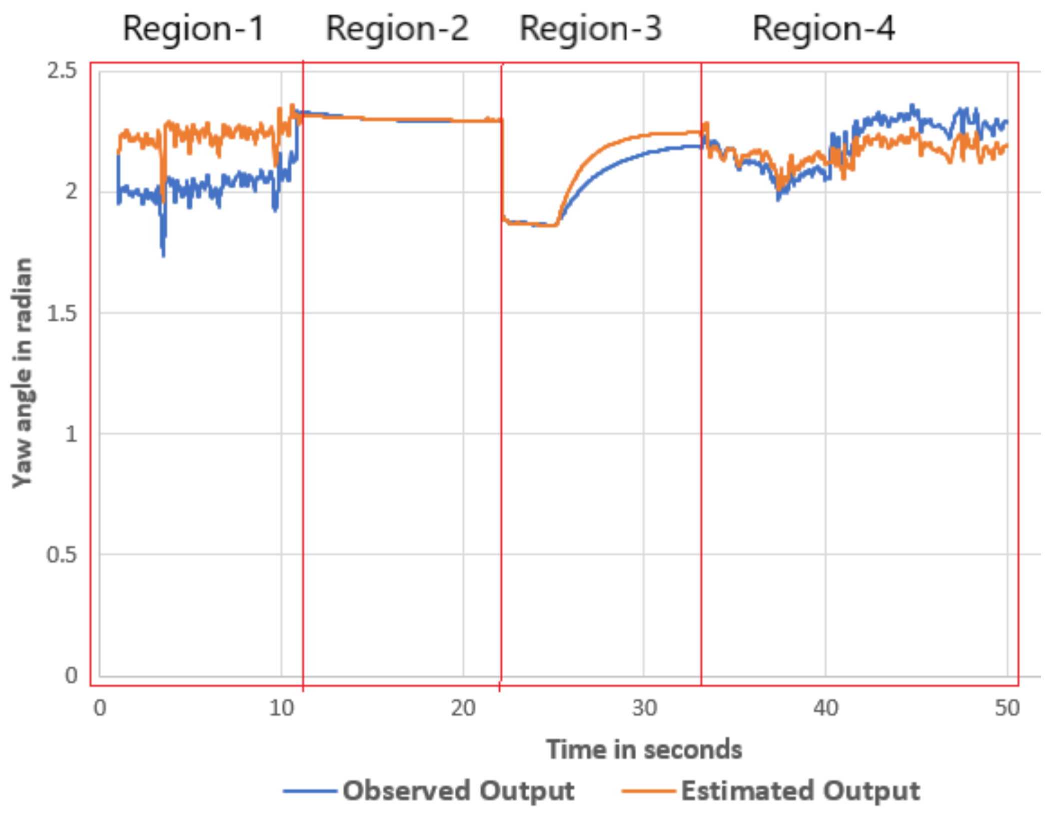
RQ2 What are the impacts of the stealthy attacks on the subject RVs?

RQ3 How effective are the attacks in achieving the attacker’s objectives?

- ArduPilot - http://ardupilot.org/
- Pixhawk - https://pixhawk.org/
- Aion R1 Rover - https://www.aionrobotics.com/r1

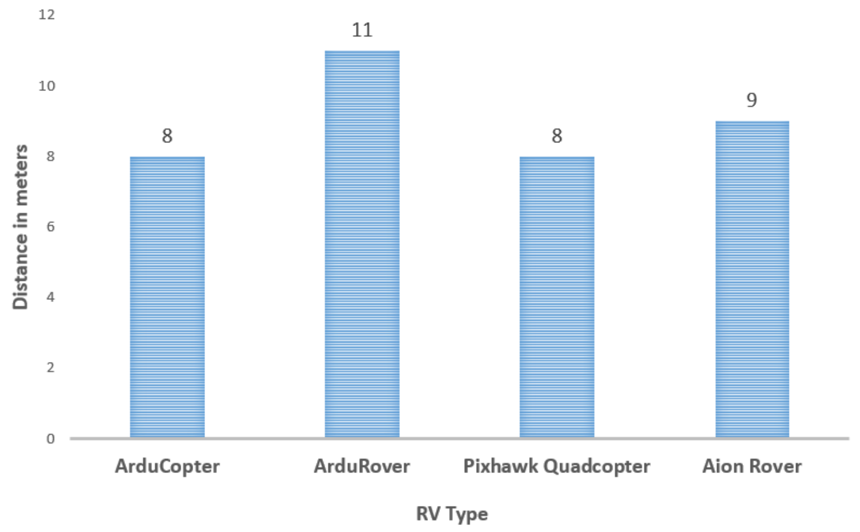
RQ1: Attacker's Effort

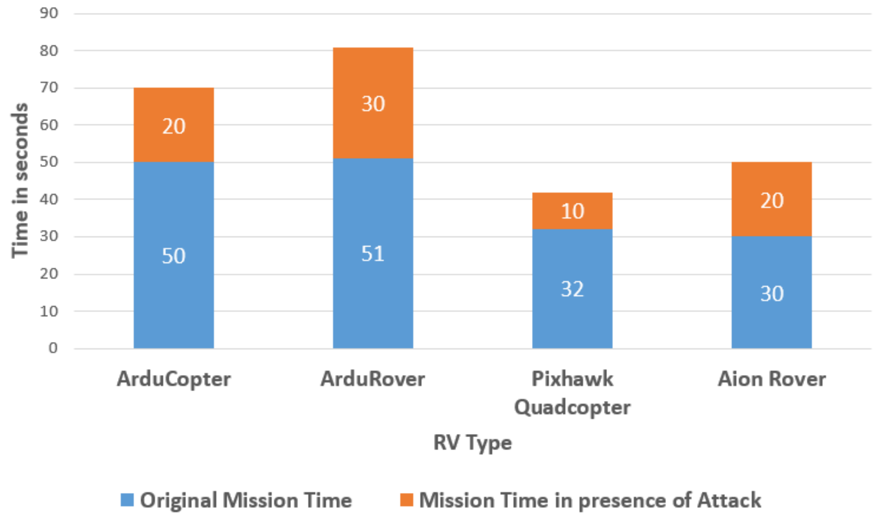
- Attacker's effort in deriving the state estimation model.
- Two Phases
  - Model extraction phase
    - – 15 missions each subject RV.
  - Model testing phase
    - – 5 missions each subject RV.
- Convergence
  - 5-7 missions for all the subject RVs.



R2Q: Impacts of Stealthy Attacks

- False data injection attack
  - Deviates RV from its trajectory.
- Artificial delay attacks
  - Delays mission time
    - Drones → At least 25%
    - Rovers → At least 30%
- Switch mode attack (for drones)
  - Crash landing
  - Land at wrong locations.





Attack Videos

Challenges in Detecting Stealthy Attacks

- Injected manipulations do not cause any immediate observable effects
  - Difficult to differentiate between attacks and drags due to wind or frictions.

- Modelling the dynamic non-linear properties of RV’s controller.
  - e.g., mode changes in during a mission
  - Difficult to consider effects of protracted attacks over a long time

Robotic Vehicles: Summary

- Vulnerabilities in control theory based attack detection techniques
- Demonstrate three types of stealthy attacks on RV systems
  - Attacks deviate a RVs by more than 100 meters, increases duration of RV mission by 25-30%, even result in crashes.
- Demonstrate techniques to automate the attacks on a class of RVs.

Artifacts Evaluated acm Reusable

Artifacts: https://github.com/DependableSystemsLab/stealthy-attacks

This Talk

- Motivation
- Attacks on Embedded and IoT devices [ACSAC’19][ACSAC’16]
- Intrusion Detection Systems for Smart Devices [FSE’17][CPS-SPC’18]
- Ongoing work and conclusion

Motivation

- **Goal: Provide low-cost security for CPS**
  - Satisfying resource and real-time constraints
  - No human intervention needed
  - Is able to detect zero day attacks

Insight: Leverage properties of CPS for intrusion detection

- Simplicity and timing predictability
- Learn invariants based on dynamic execution
- Monitor invariants at runtime for violations

Speed ∝ Distance

$$\text{Speed} \propto \frac{1}{Time}$$

Speed **NOT** ∝ Distance

$$\text{Speed} \propto \frac{1}{Time}$$

Intrusion Detection Systems (IDS)

Signature-based IDSs [CSUR2014]

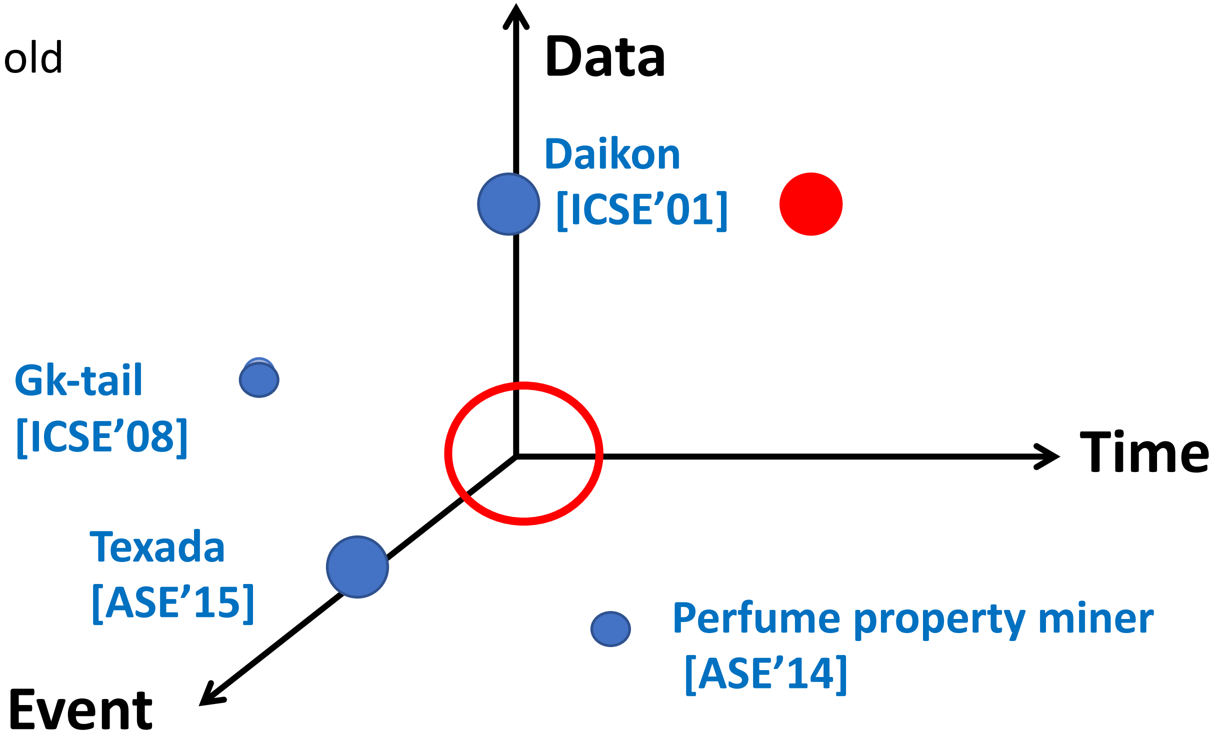
Anomaly-based IDSs [Computers&Security2009]

Specification-based IDSs [SmartGridCom2010]

- Static analysis
- Dynamic analysis

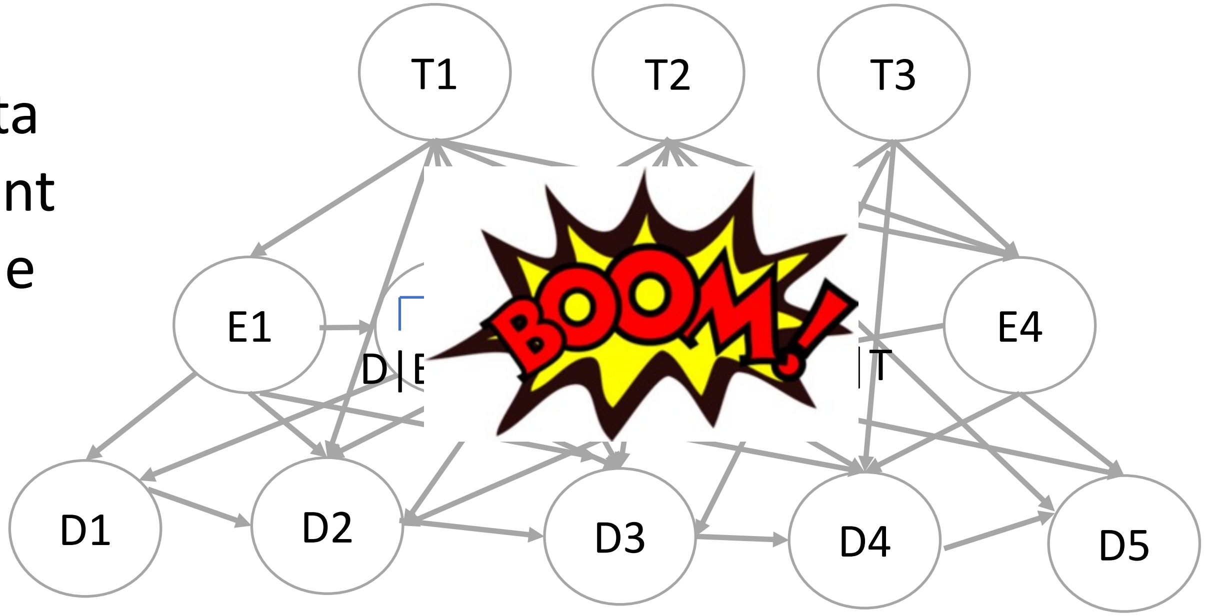
Dynamic Analysis Techniques

- Invariant Examples
  - Energy usage >=0
  - Current – Past <= Threshold



Main Idea

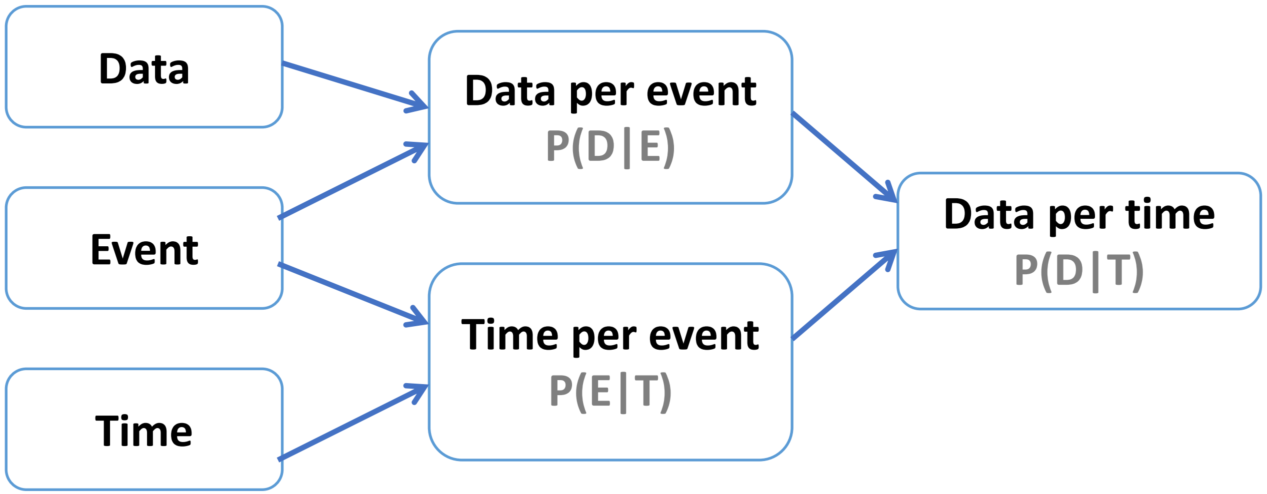
D: Data
E: Event
T: Time



Methodology

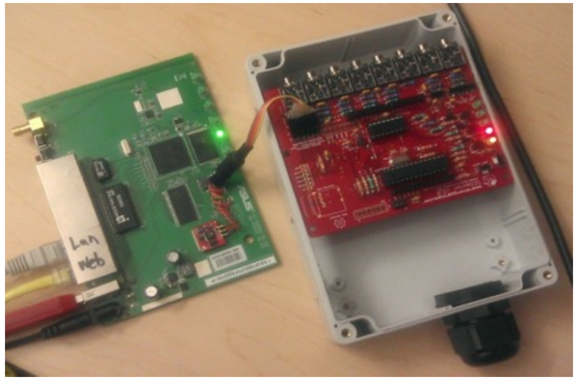
ARTINALI: **A R**eal **T**ime-specific **I**nvariant i**N**ference **AL**gor**I**thm

- 3 dimensions
- 6 classes of invariants

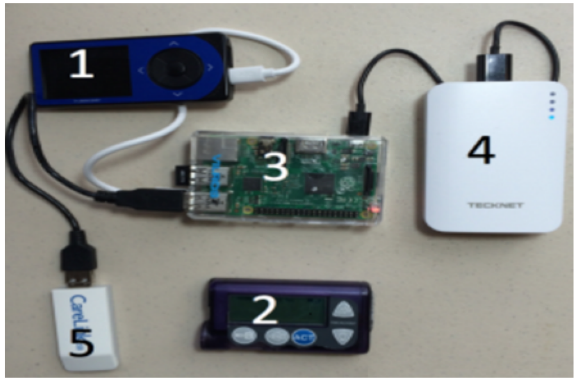


CPS Platforms for Evaluation

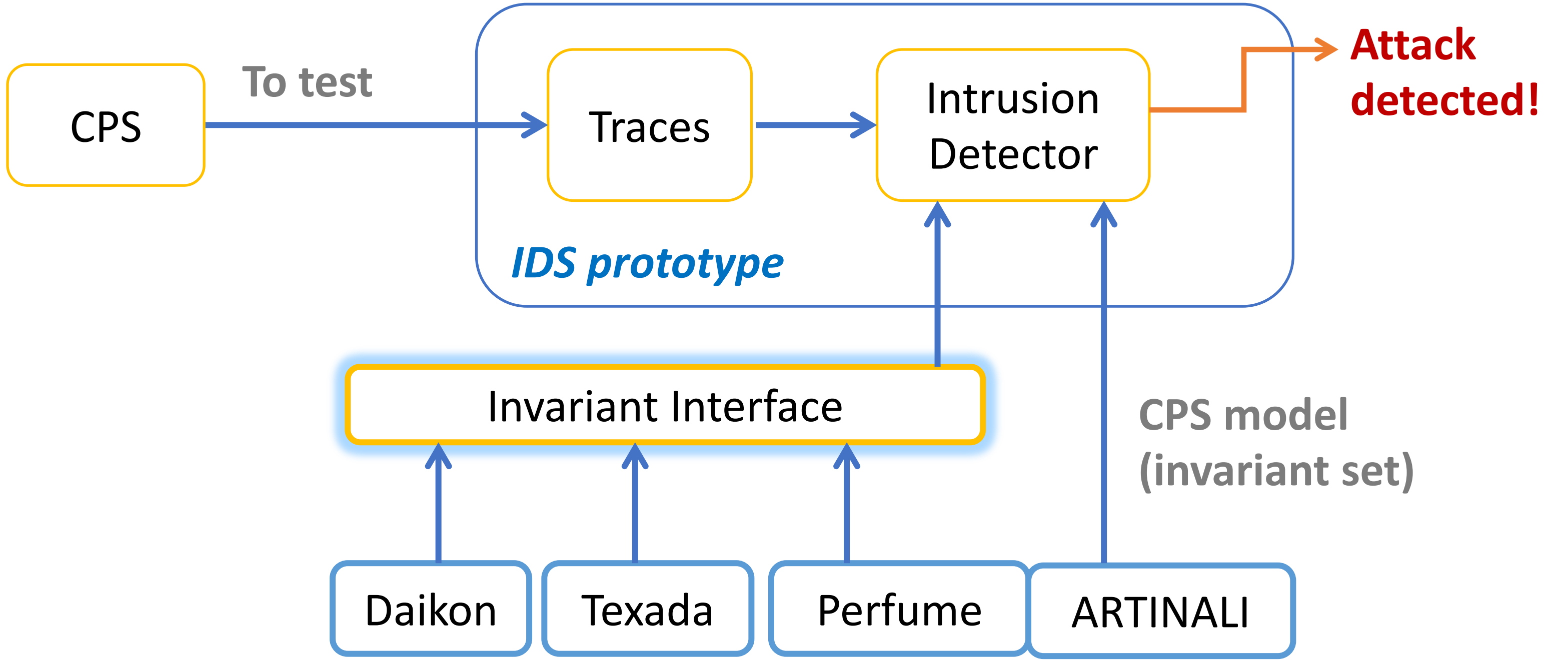
- Advanced metering infrastructure (AMI)
  - SEGMeter
    - http://smartenergygroups.com



- Smart Artificial Pancreas (SAP)
  - OpenAPS
    - https://openaps.org/



Experimental Setup



Targeted Attacks

CPS Platform	Targeted attack	Attack entry point
AMI (SEGMeter)	Meter spoofing [ACSAC2010]	Deception on A
	Sync. Tampering [ACSAC2010]	Deception on D
	Message dropping [CCNC2011]	DoS on A
SAP (OpenAPS)	CGM spoofing [Healthcom2011]	Deception on A
	Stop basal injection [BHC2011]	Deception and DoS on C
	Resume basal injection [BHC2011]	Deception and DoS on C

Take away :

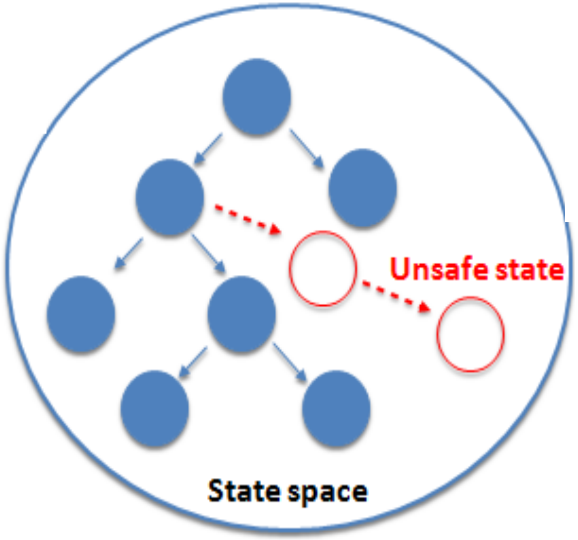
ARTINALI detected all the targeted attacks

Arbitrary Attacks

Data mutations

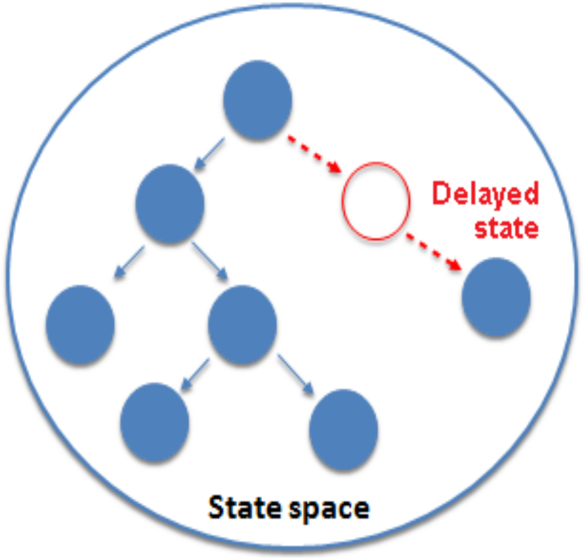
Smart facial recognition system (CVE-2016-1516)

Branch flipping



CGM spoofing in SAP, [BHC2011]

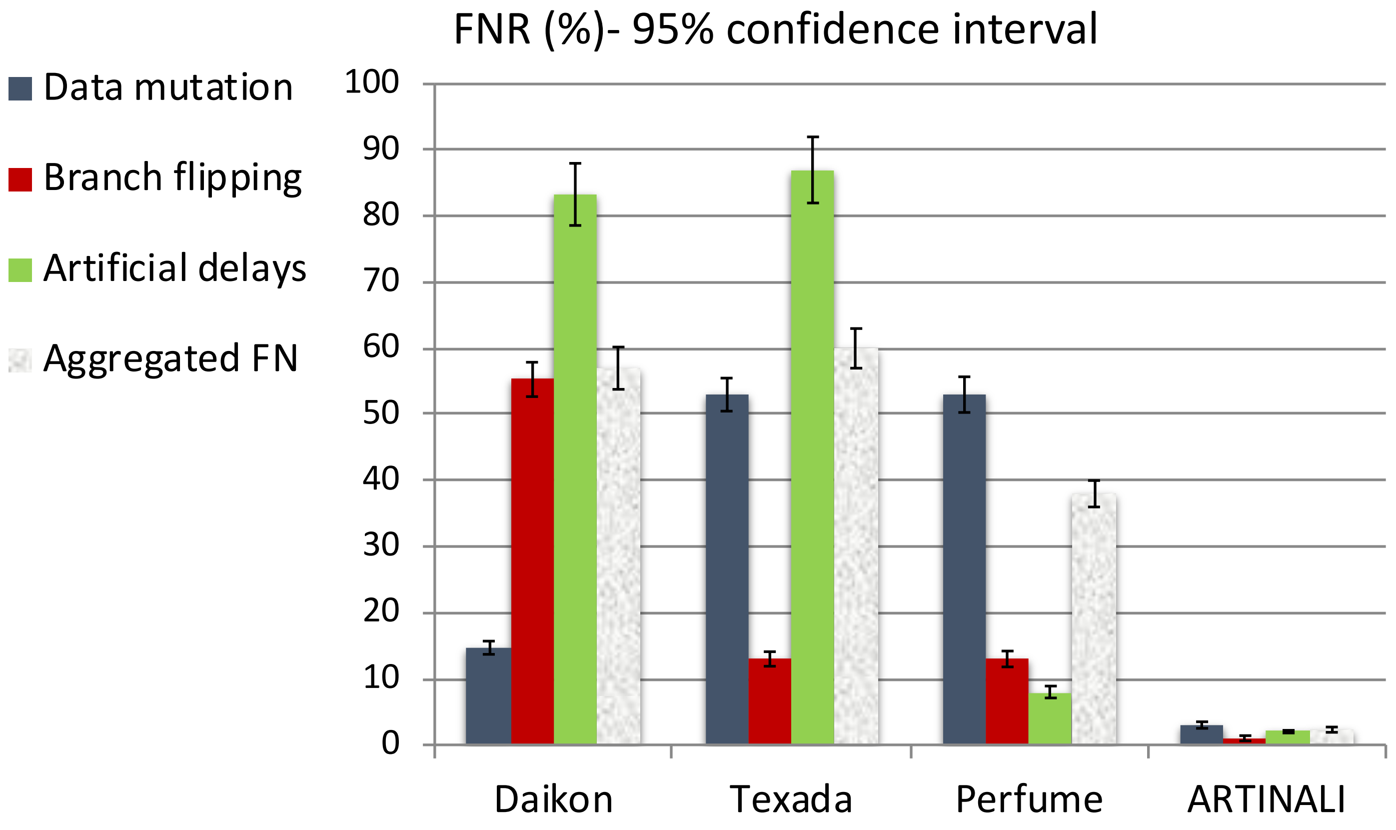
Artificial delay insertion



Synchronization tampering in smart meter, [ACSAC2010]

False Negative (FN) Rate

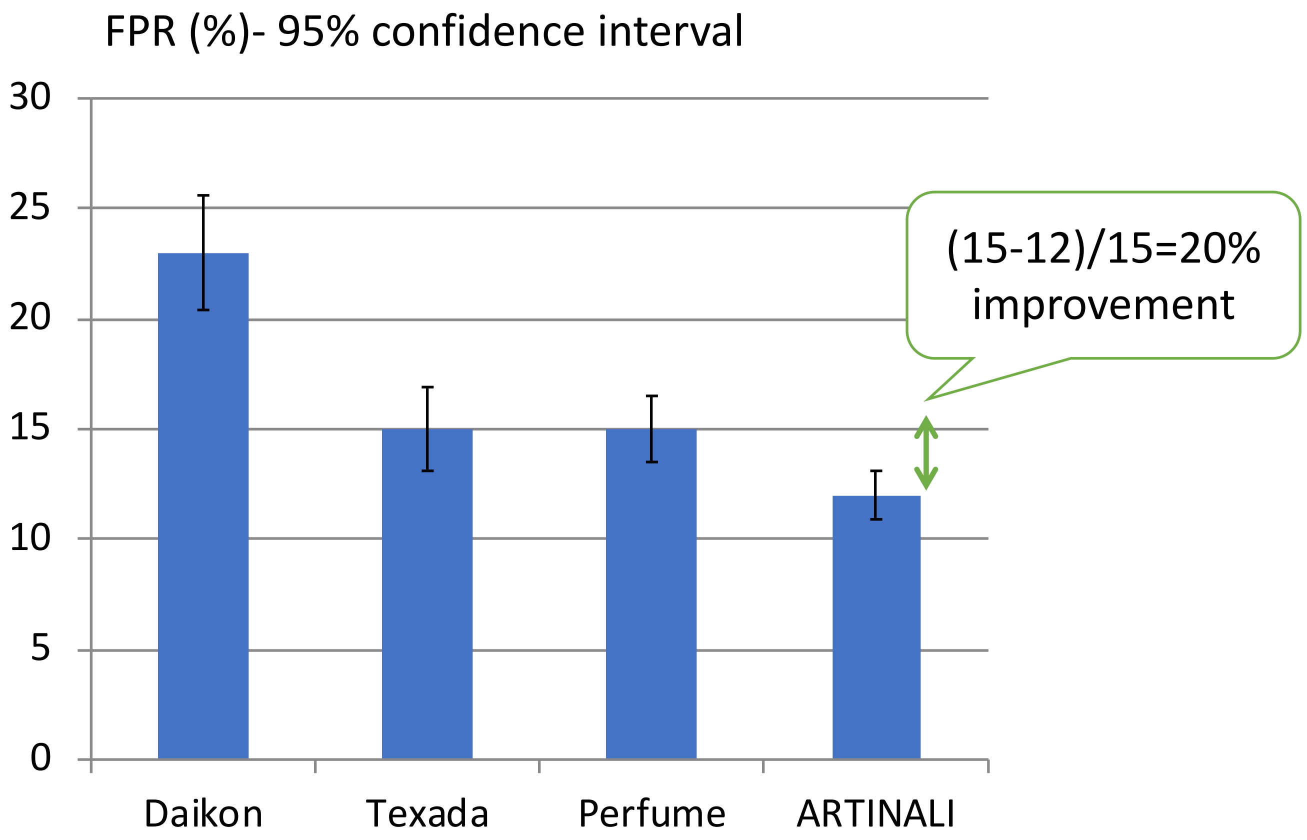
- ARTINALI-based IDS reduces the ratio of FN by 89 to 95% compared with the other tools across both platforms.
  - SEGMeter



False Positive (FP) Rate

- ARTINALI-based IDS reduces the ratio of FP by 20 to 48% compared with the other tools across both platforms.

  - SEGMeter



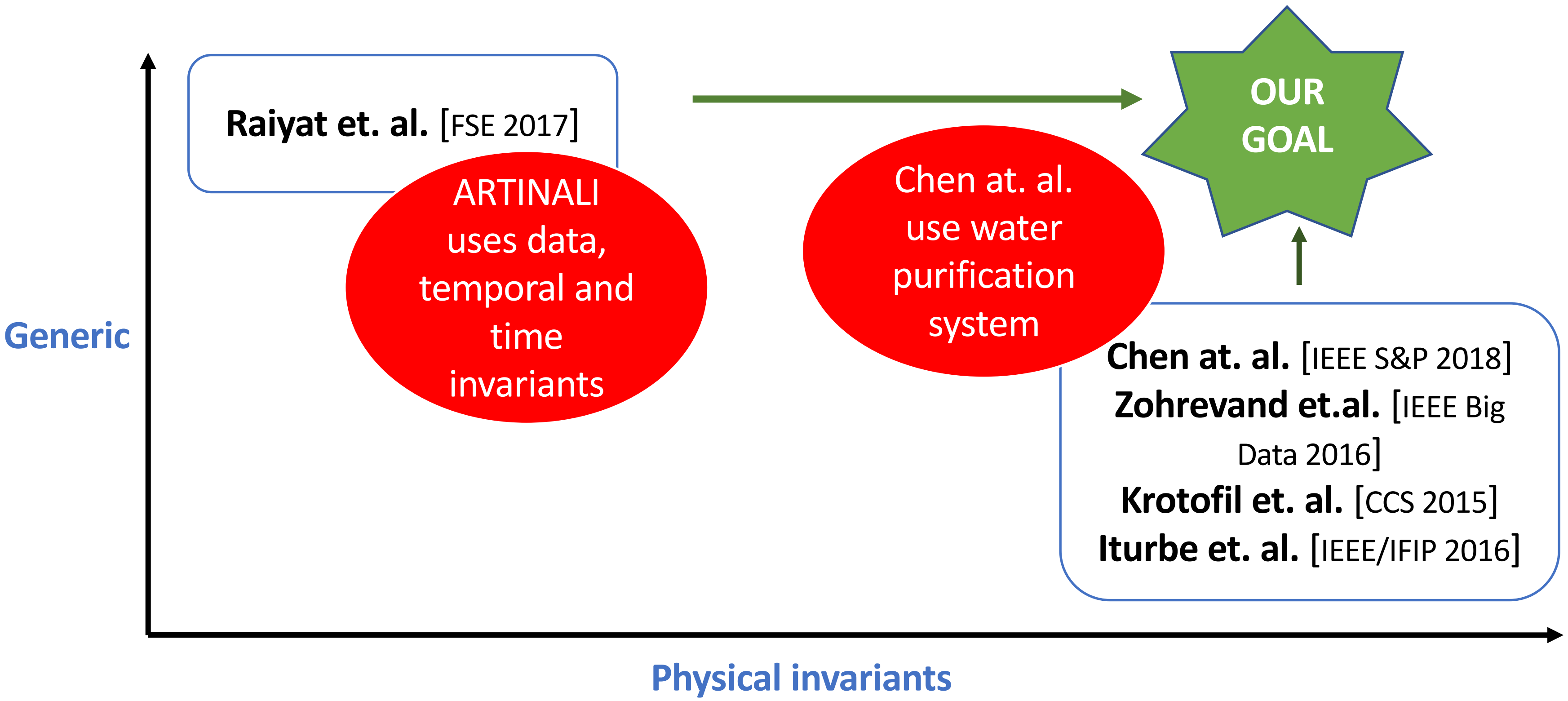
Summary of ARTINALI

- ARTINALI: A Multi-Dimensional model for CPS
  - Captures *data-event-time* interplay

- *Compared to other techniques*
  - Increases the *coverage* of IDS
  - Decreases the rate of *false positives*

- However, ARTINALI still has high false-positives (FPS)
  - Can we reduce FPs further ?

CORGIDS: Correlation-Based Detection



Hidden Markov Model (HMM)

Finite model used to **describe probability** distribution over possible sequences of a given system.

Example: Reinforcement learning and pattern recognition such as speech, handwriting and gesture recognition

HMM

- **Finding correlations** in multidimensional, **non-linear time series** systems like **CPS**.
- **Likelihood of data** belonging from a dataset.

Experimental setup

- **Unmanned Aerial Vehicle (UAV)**

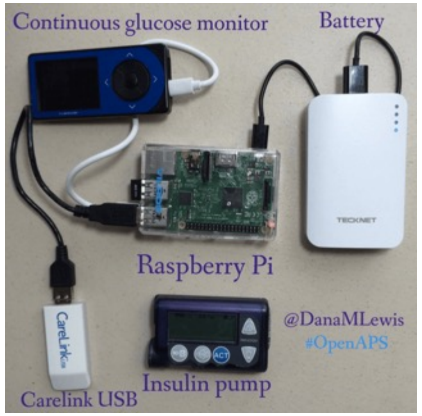
  ArudPilot's Software in the Loop (SITL)

  (http://ardupilot.org/dev/docs/sitl-simulator-software-in-the-loop.html)

- **Smart Artificial Pancreas (SAP)**

  Open Artificial Pancreas System (OpenAPS)

  (https://openaps.org/)



Evaluation

TESTBED	TARGETED ATTACKS	FP (%)	FN (%)
UAV	Battery Tampering	0.0	12.20
	Flooding	0.0	11.30
	Distance Spoofing	0.0	12.80
SAP	Insulin Tampering	5.60	4.20
	Glucose Spoofing	2.80	8.40

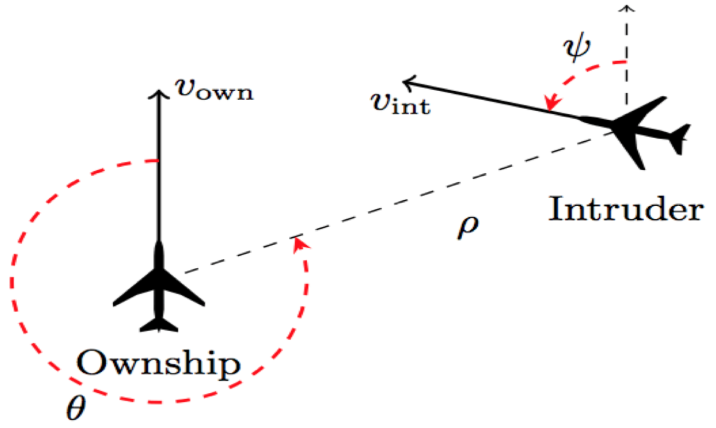
Summary of CORGIDS

- **Physical properties** of CPS are **indicative** of its behavior.
- **HMM** are **good** at finding **correlations among properties**.
- **CORGIDS** had **higher Precision** and **Recall** than ARTINALI

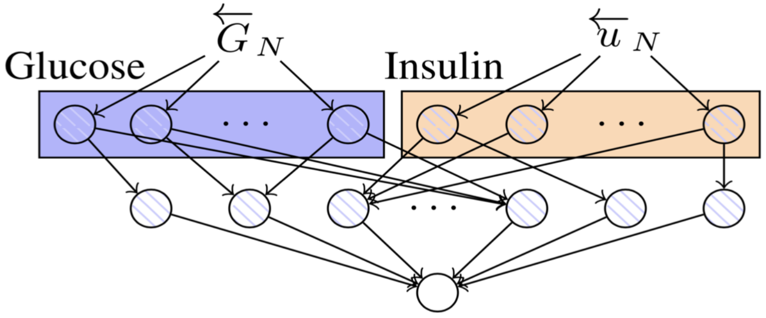
This Talk

- Motivation
- Attacks on Embedded and IoT devices [ACSAC’19][ACSAC’16]
- Intrusion Detection Systems for Smart Embedded Devices using Dynamic Invariants [FSE’17][CPS-SPC’18]
- Ongoing work and conclusion

DNN based CPS are Replacing PID controllers



ACAS Xu (Airborne collision avoidance system X manned) - DNN



Artificial Pancreas System - DNN

-> Small changes to the original inputs can result in crashes.

-> The boundary values on which the DNN is trained can result in

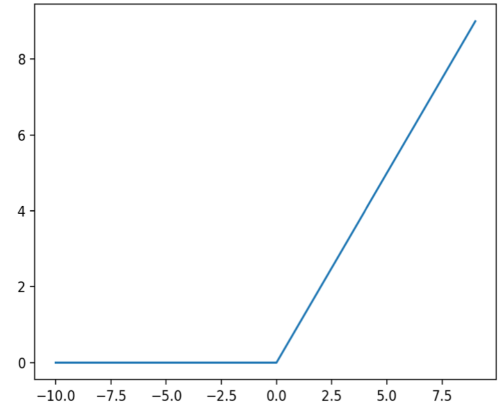
1) Crashes of Unmanned vehicles
2) wrong amounts of insulin delivery to patients

ReLUSyn: Synthesizing Data Ranges for Attacks

- Encoding the DNN as a 0-1 MILP Problem

- Allows to build a query mechanism to find the FDI attacks
  - Providing speed up over brute force

- ReLU activation function--**non-linear function**
  - Cannot be encoded as an ILP Problem
  - ReLU is however piecewise linear
  - 0-1 MILP allows to represent ReLU as piecewise linear



Preliminary Results: Brute force vs 0-1 MILP

Artificial Pancreas System- (1 layer + 50 neurons/layer)

Brute force	0-1 MILP
Time = 10 sec + Search time to find right inputs	< 1 sec/ attack

ACAS Xu- (5 layer + 50 neurons/layer)

Brute force	0-1 MILP
Time = Timeout	~ 10 secs/ attack

Brute Force	0-1 MILP
Running through all the iterations	Simple query mechanism
Time out in ACAS Xu	Optimal FDI
No automations	Fully automated

Conclusions

- **End Devices in CPS are important to be protected from attacks**
  - Provide a conduit for attackers to get a foot-hold into the system
  - Can cause large-scale disruptions of critical infrastructures

- **Attackers can remain stealthy by leveraging properties of the CPS**
  - Knowledge and physical access to the CPS
  - Need host-based intrusion detection systems for security

- **Host-based IDS for end-user devices**
  - Leverage invariants and machine learning to learn CPS behaviors
  - Detect attacks proactively with low false-positives

Questions? karthikp@ece.ubc.ca