Investigate and Mitigate the Attacks Caused by Out-of-Band Signals

Dr. Xiali (Sharon) Hei
Associate Professor
The Center for Advanced Computing Studies (CACS)
School of Computing and Informatics
University of Louisiana at Lafayette

Nov. 13, 2023

Bio

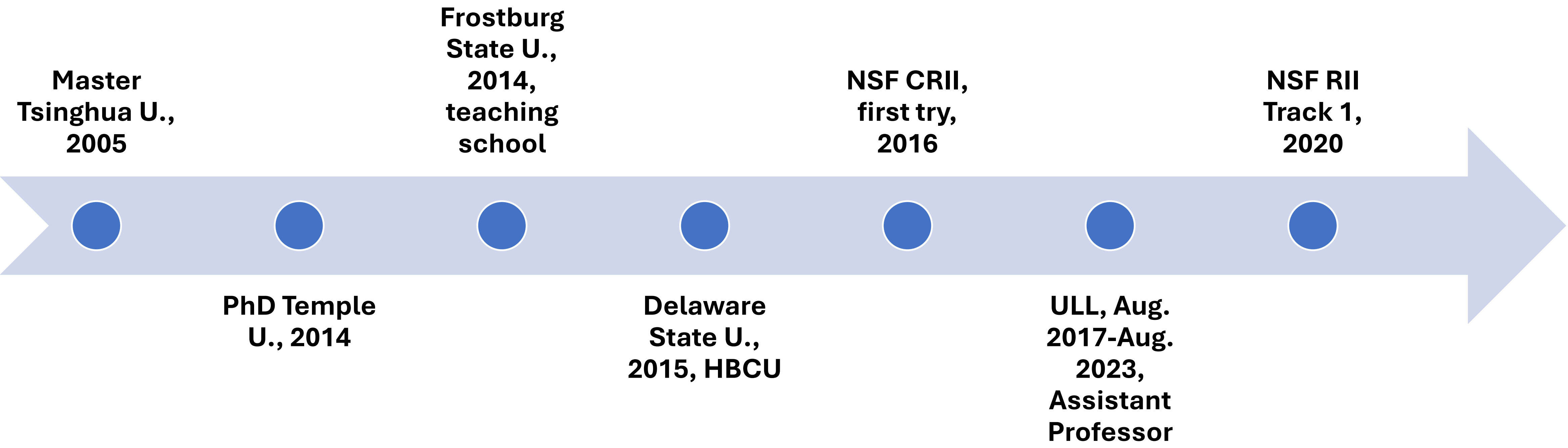
- An endowed associate professor at UL Lafayette, since **August** 2023.
  - Cyber-Physical System (CPS) Security, Side Channel, and Embedded System Security,
  - Privacy, Privacy-preserving learning, AI model security,

- Funds: Secure more than 24 million funds as PI or a co-PI, more than 3 million personal share

- Publications: 60, including peer-reviewed papers, journals, book, and book chapters

Bio (2)

- **Dedicated:** Once I get the opportunity, **I will try my best and succeed.**



Bio (3)

- **Goal-getter: intensive efforts behind each achievement. More than 120 hours on each proposal writing**

Exploring the Physics of Sensing and Embedded Security in Three Pillars

Attacks
Out-of-band Injection

Acoustic

- Real-time Attacks on Inertial Sensors [USENIX Security'18]
- **Physical Adaptive** Process Control on Inertial Sensors [ACM CPSIoTSec'23]

EMI

- Safety-Critical Temperature Systems [ACM CCS'19]
- EEG-Based **ML/DL** Systems [ACM SecTL'23]

Defenses

- Transduction Shield (Detect and Correct EMI) [AsiaCCS'21]
- ADC Bank (Acoustic) [SmartSP'23]
- Actuator Shield (Detect and correct EMI attacks on actuators) [TDSC Under Review]

Side Channel Analysis

- 1st Sensing Side Channel On Unconstrained Keyboard Inputs [Auditory Eyesight] [USENIX'23]
- Physical Feedback Side Channel [Arxiv'22]
- Privacy Risks Posed by Industrial Robots [SmartSP'23]

Machine Learning for Security & Privacy, AI Model security

Understanding and exposing security issues in real-world CAPTCHAs

Attack

A low-cost attack against the hCaptcha system
[WOOT’21]

An object detection-based solver for Google’s image reCAPTCHA v2
[RAID’20]

Design

Audio adversarial CAPTCHAs
[EuroS&P’22]

Live thermal image CAPTCHAs
[AsiaCCS’24]

Addressing security & privacy issues in Machine Learning (ML)

Differentially Private Training of Deep Learning Models on Functional Near-Infrared Spectroscopy Data
[arXiv’23]

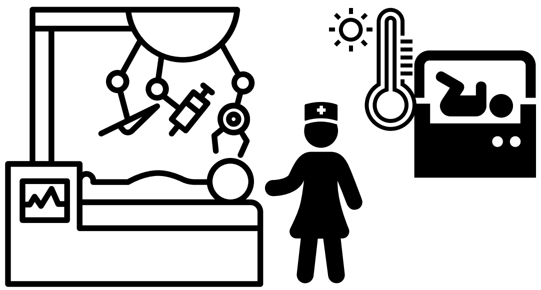
No More Manual DP-SGD Fine-Tuning? Simple heuristic-based strategies provide the best balance between accuracy and privacy
[ICLR’24, Under Review]

XMAM: X-raying models with a matrix to reveal backdoor attacks for federated learning [DCN’23]

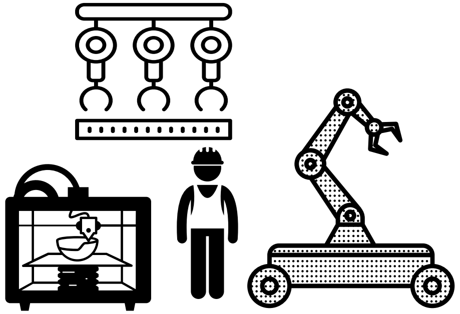
FL-PLAS: Backdoor-Resistant Federated Learning based on Partial Layer Aggregation Strategy [AAAI’24, under review]

SaTC vision

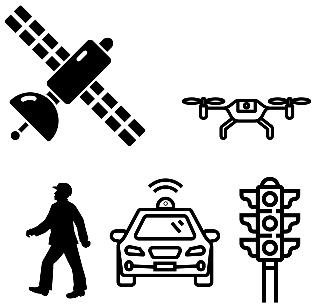
- Motivation:
  - Billions of CPS, IoT systems, and human-computer interfaces rely on **sensors**



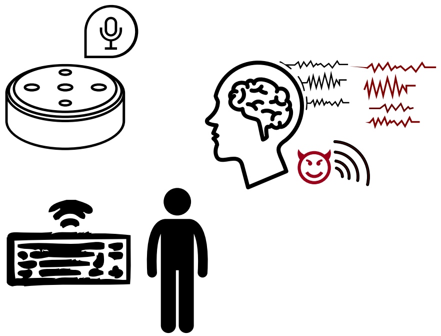
Medical Devices



Manufacturing Systems



Transportation



Human-Computer Interfaces

A World with CPS/IoT Systems: Computations Dealing with the Physical World

- Vision:
  - Decision-making and interaction of **CPS/IoT/HCI** (**Autonomous systems** and **AI**) rely on **sensor-enabled environment perception**
  - Security and safety of CPSs require **threat models** and **computations** resilient to the complex physical world

- Research Questions:
  - How to **anticipate** analog-domain risks in sensing?
  - How to design and conduct **computations** to analyze and mitigate **security**, **privacy,** and **safety** risks?

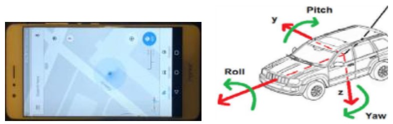
Roadmap

- Research
  - Acoustic attacks on embedded inertial sensor systems
    - Real-time Attacks (**Manual** Attacks) [USENIX’18]
    - **Automatic** Attacks (Using Programs) with Data Feedback [USENIX’18]
    - **Automatic** Attacks (Using Programs) **without** Data Feedback [CPSIoTSec’23]
  - Mitigating safety risks of acoustic attacks on sensors [SmartSP’23]
  - Mitigating safety risks of EMI attacks on sensors [CCS’19, SecTL’23, ASIACCS’21]

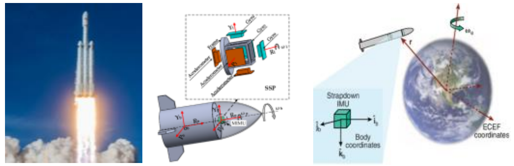
Robots

VR / AR

Navigation



Guidance of Rocket, Aviation



Embedded Inertial Sensor System Security

First Attacks on **Gyroscopes** to **Control Actuation**, Navigation, and AR/VR Systems

1. Injected and Delivered: Fabricating Implicit **Control** over **Actuation Systems** by **Spoofing Inertial Sensors**, In *USENIX Security Symposium, 2018*

How to use acoustic signals to control CPS?

Embedded inertial sensors → **Control actuation in real time**

Yazhou Tu, Zhiqiang Lin, Insup Lee, Xiali Hei. "Injected and Delivered: Fabricating Implicit Control over Actuation Systems by Spoofing Inertial Sensors." ***USENIX Security Symposium***. 2018.

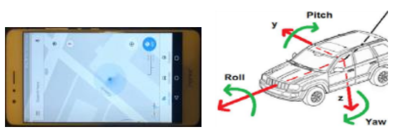
Yazhou Tu, Sara Rampazzi, and Xiali Hei. "Towards Adversarial Process Control on Inertial Sensor Systems with Physical Feedback Side Channels." **CPSIoTSec** 2023.

Jianyi Zhang, et al, "ADC-Bank: Detecting Acoustic Out-of-Band Signal Injection on Inertial Sensors." SmartSP 2023, Best paper award!

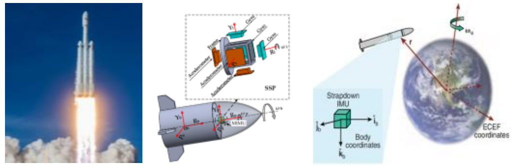
Robots

VR / AR

Navigation



Guidance of Rocket, Aviation



Embedded Inertial Sensor System Security

First Attacks on **Gyroscopes** to **Control Actuation**, Navigation, and AR/VR Systems

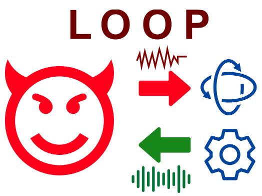
1. Injected and Delivered: Fabricating Implicit **Control** over **Actuation Systems** by **Spoofing Inertial Sensors**, In *USENIX Security Symposium, 2018*

How to use acoustic signals to control CPS?

Embedded inertial sensors → **Control actuation in real time**

First **Automatic** Attacks **without Internal Inertial Sensor Data**

2. Towards Adversarial Process Control on Inertial Sensor Systems with Physical Feedback Side Channels



Yazhou Tu, Zhiqiang Lin, Insup Lee, Xiali Hei. "Injected and Delivered: Fabricating Implicit Control over Actuation Systems by Spoofing Inertial Sensors." ***USENIX Security Symp***. 2018.

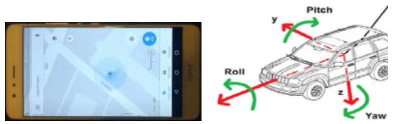
Yazhou Tu, Sara Rampazzi, and Xiali Hei. "Towards Adversarial Process Control on Inertial Sensor Systems with Physical Feedback Side Channels." **CPSIoTSec** 2023.

Jianyi Zhang, et al, “ADC-Bank: Detecting Acoustic Out-of-Band Signal Injection on Inertial Sensors.” SmartSP 2023, Best paper award!

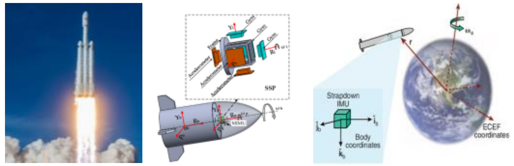
Robots

VR / AR

Navigation



Guidance of Rocket, Aviation

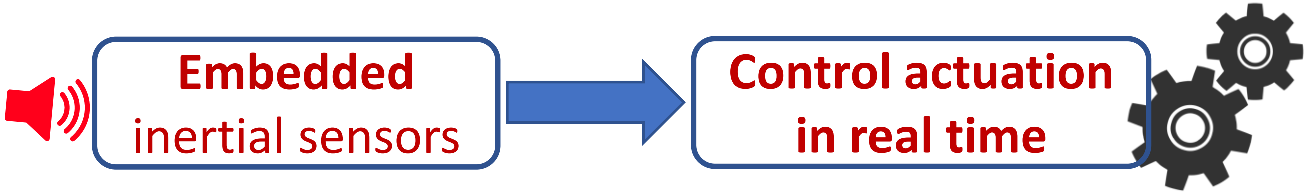


Embedded Inertial Sensor System Security

First Attacks on **Gyroscopes** to **Control Actuation**, Navigation, and AR/VR Systems

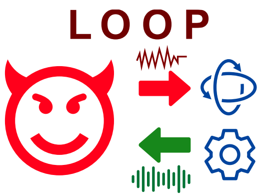
1. Injected and Delivered: Fabricating Implicit **Control** over **Actuation Systems** by **Spoofing Inertial Sensors**, In *USENIX Security Symposium, 2018*

How to use acoustic signals to control CPS?



First **Automatic** Attacks **without Internal Inertial Sensor Data**

2. Towards Adversarial Process Control on Inertial Sensor Systems with Physical Feedback Side Channels
3. ADC-Bank: Detecting Acoustic Out-of-Band Signal Injection on Inertial Sensors



Yazhou Tu, Zhiqiang Lin, Insup Lee, Xiali Hei. "Injected and Delivered: Fabricating Implicit Control over Actuation Systems by Spoofing Inertial Sensors." ***USENIX Security Symp.***. 2018.

Yazhou Tu, Sara Rampazzi, and Xiali Hei. "Towards Adversarial Process Control on Inertial Sensor Systems with Physical Feedback Side Channels." **CPSIoTSec** 2023.

Jianyi Zhang, et al, “ADC-Bank: Detecting Acoustic Out-of-Band Signal Injection on Inertial Sensors.” SmartSP 2023, Best paper award!

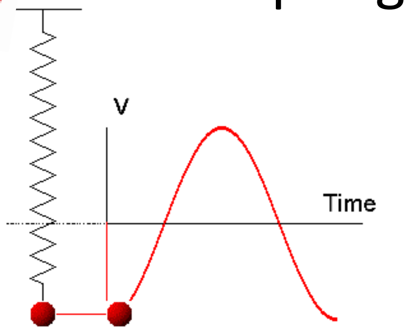
MEMS Inertial Sensors

- Provide motion feedback to control systems
  - Gyroscope: Angular velocity 
  - Accelerometer: Linear acceleration 

- **Miniaturized** mechanical sensing structure
  - Micro-electromechanical systems (MEMS)
  - Transduce inertial stimuli to electrical signals
  - Vulnerable to ***acoustic resonance*** [Dean2007ISIE]

Acoustic Resonance

Sensing structure:
Mass-spring



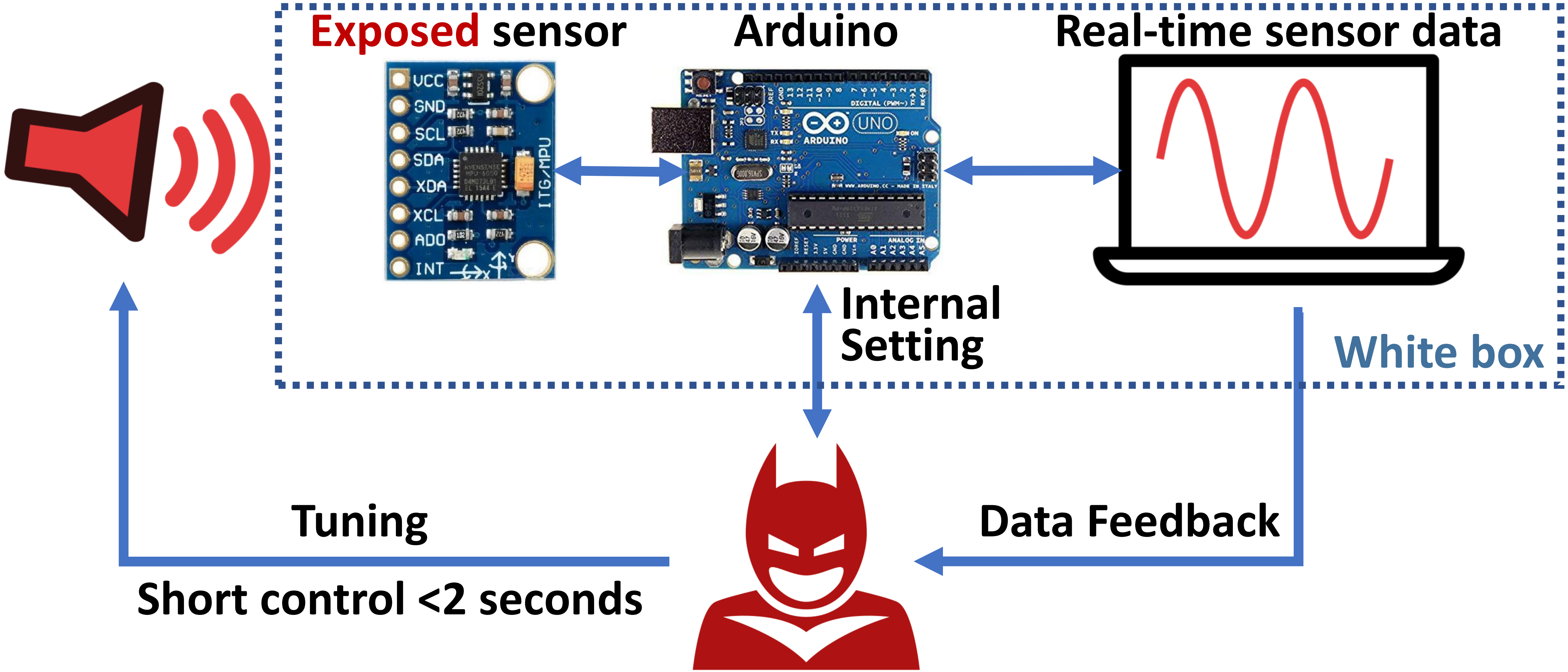
Dean, Robert N., George T. Flowers, A. Scotte Hodel, et al. "On the degradation of MEMS gyroscope performance in the presence of high power acoustic noise." In *2007 IEEE International Symposium on Industrial Electronics*. IEEE, 2007.

Conventional Attacks on MEMS Inertial Sensors

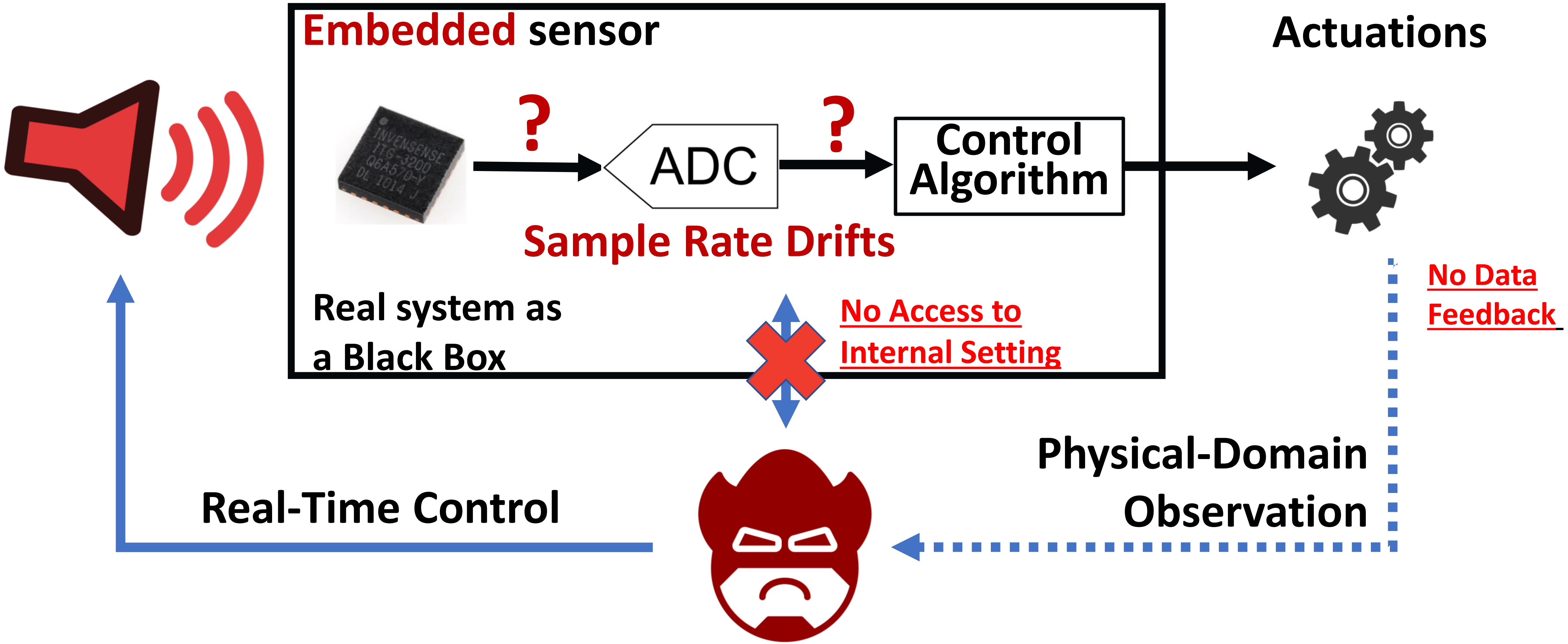
- **DoS** attack on gyroscopes (drones) [Son2015usenix]
- Control the output of exposed accelerometers [Trippel2017EuroS&P]
  - Short control (<2 seconds **after hand tuning**)
  - Require **access** to internal victim hardware (**Arduino**)

Son, Yunmok, et al. "Rocking drones with intentional sound noise on gyroscopic sensors." *24th {USENIX} Security Symposium ({USENIX} Security 15)*. 2015.
Trippel, Timothy, et al. "WALNUT: Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks." *2017 IEEE EuroS&P.*

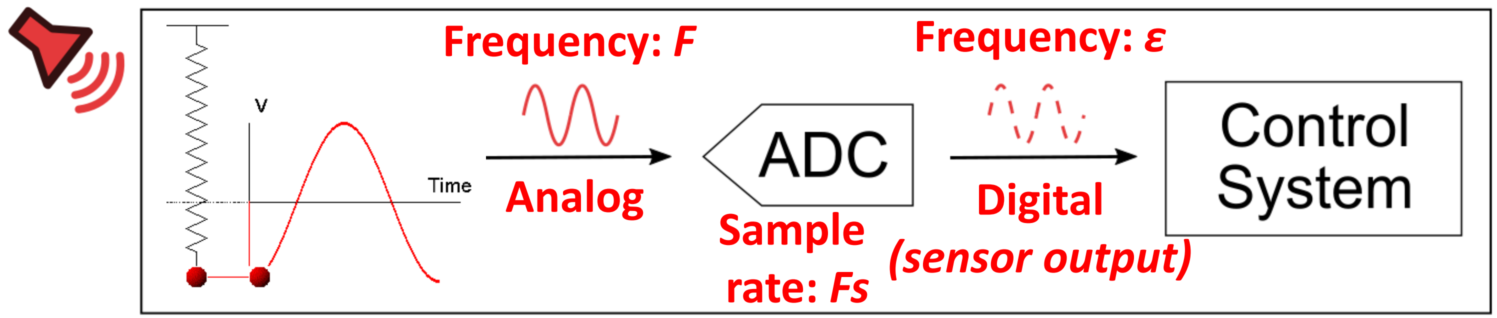
Conventional White-Box Approach



Motivation: A Real System is often a Black Box

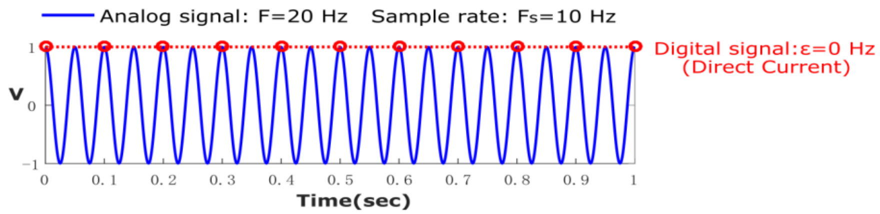


Issues of Acoustic Injection on Inertial Sensors



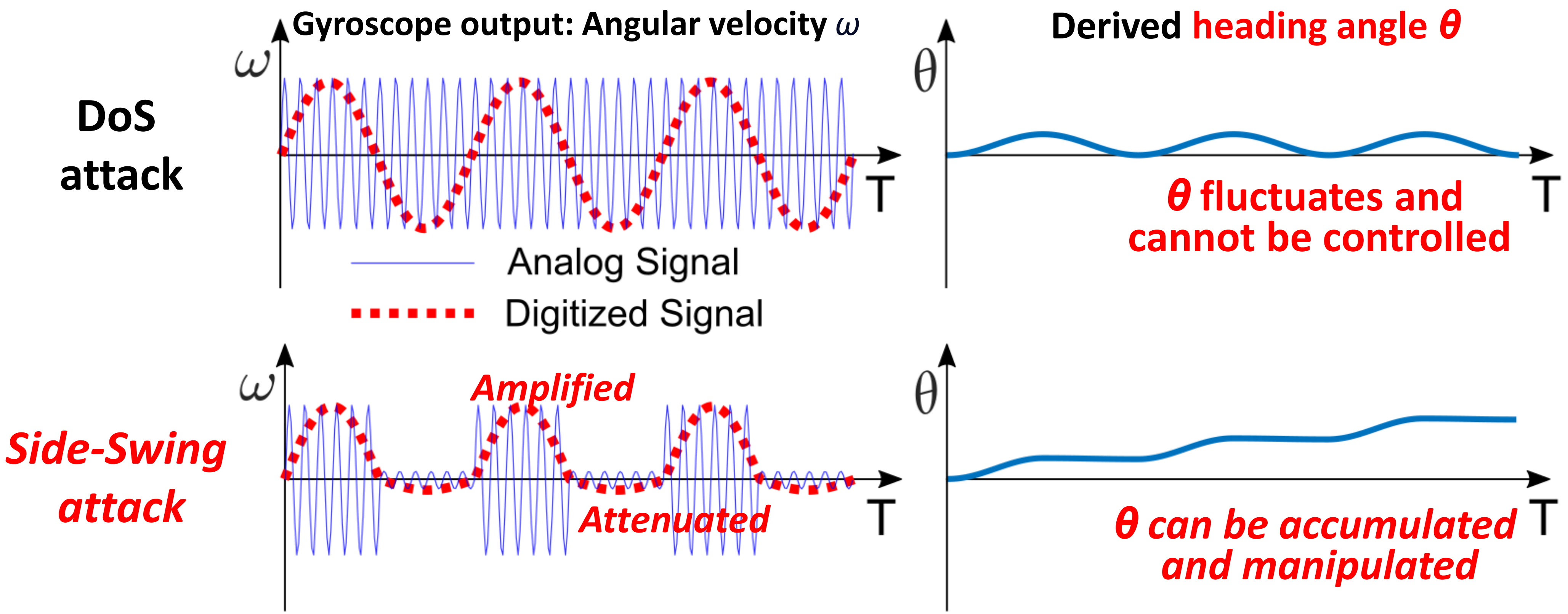
- ***Aliasing*** $$F = n \cdot F_S + \epsilon \quad (-\frac{1}{2}F_S < \epsilon \leq \frac{1}{2}F_S, n \in \mathbb{Z}^+) \quad (1)$$

- When $F = n \cdot F_S$ we have *ε = 0* (**Direct Current**, DC)
  - **Assume a constant sample rate *Fs***



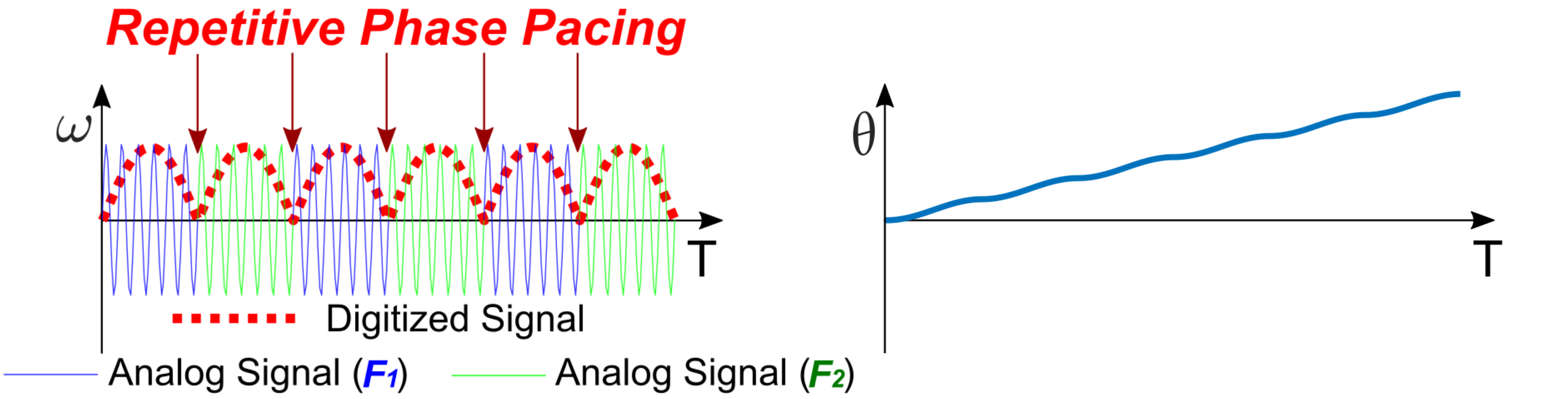
- **In real-world systems: Fs is drifting. The drift is amplified by *n* times (Eq. 1)**

Attack Methods: Side-Swing Attack



- Increase *A[i]* to amplify the induced output in the target direction
- Decrease *A[i]* to attenuate the output in the opposite direction

Switching Attacks



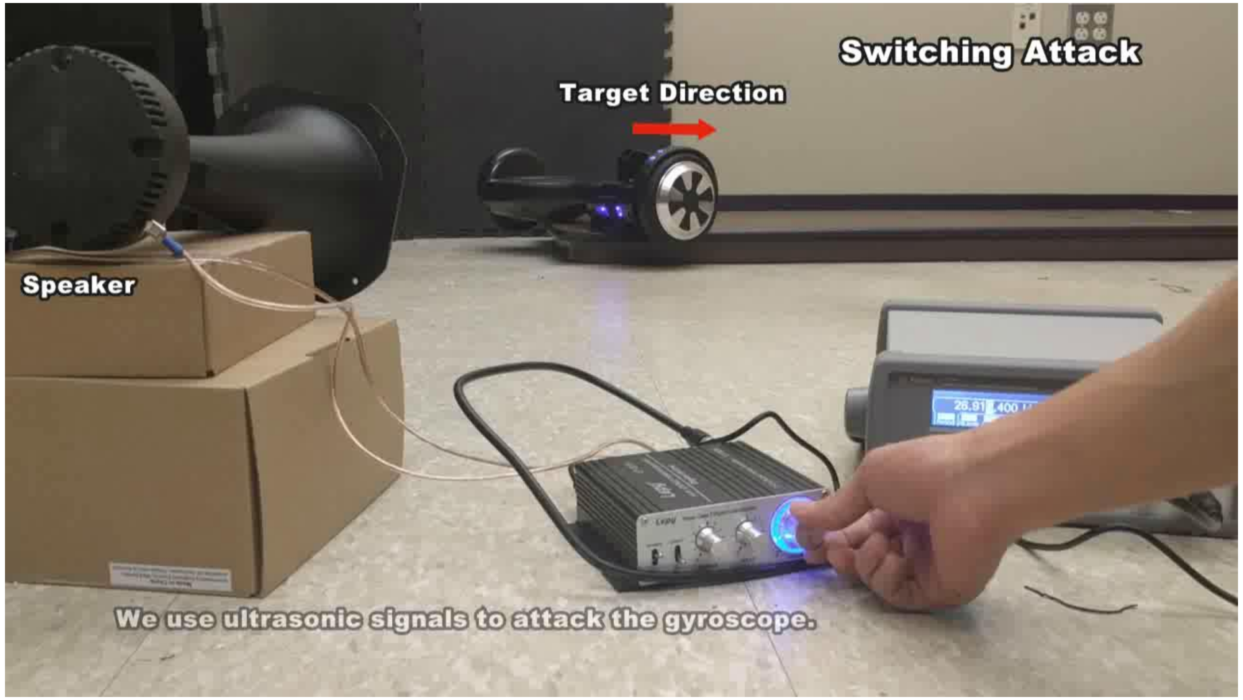
Angular velocity **Derived heading angle**

- **Repetitive *Phase Pacing***
  - Switch *F* between *F1* and *F2* back and forth
  - Basic Idea: Intentionally inducing **positive** and **negative** frequencies in acoustic attacks

Attacks on Closed-Loop and Open-Loop Control Systems

- **Self-balancing Transporter**
  - Side-Swing: https://youtu.be/Y1LLiyhCn9I
  - Switching: https://youtu.be/D-etuH04pms
- **Robot**
  - Side-Swing: https://youtu.be/oy3B1X41u5s
- **Camera Stabilization**
  - Side-Swing: https://youtu.be/FDxaLUtgaCM
  - Switching: https://youtu.be/JcA_WXHrUEs
- **Screwdriver**, **VR headset/controller**, **3D mouses**, **smartphones**, etc.

Switching attacks on a self-balancing transporter

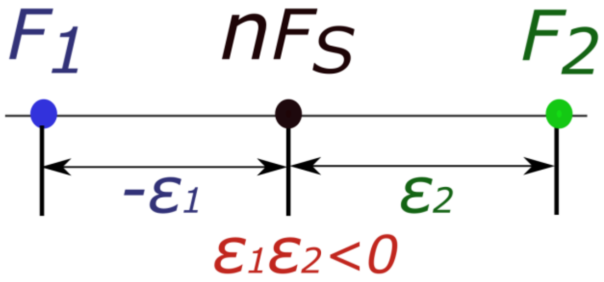


Roadmap

- Research
  - Acoustic attacks on embedded inertial sensor systems
    - Real-time Attacks (**Manual** Attacks) [USENIX'18]
    - **Automatic** Attacks (Using Programs) with Data Feedback [USENIX'18]
    - **Automatic** Attacks (Using Programs) **without** Data Feedback [CPSIoTSec'23]
  - Mitigating safety risks of acoustic attacks on sensors [SmartSP'23]
  - Mitigating safety risks of EMI attacks on sensors [CCS'19, SecTL'23, ASIACCS'21]

Automatic Switching Attack with Feedback

- Motivation:
  - Hand tuning is slow and relies on human reaction
- Program generates and adjusts acoustic signals
  - More effective
  - Active adaptation
  - Controlled with algorithm



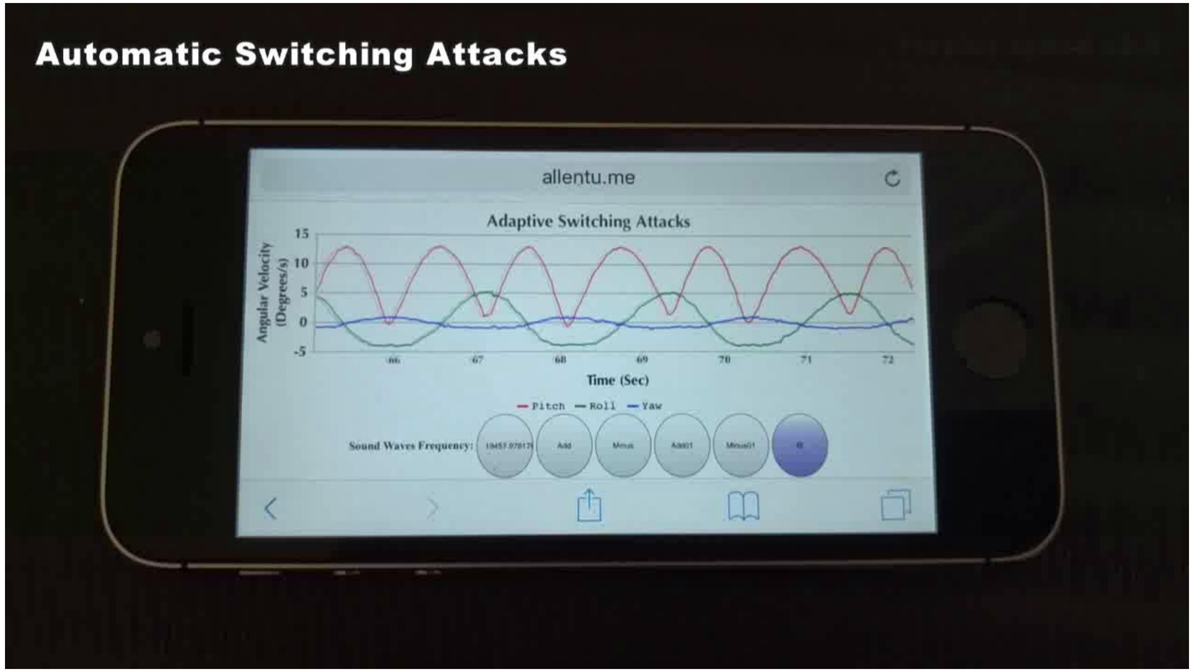
Condition of Phase Pacing

$$
\begin{aligned}
F_1 &= n \cdot F_S + \varepsilon_1 \qquad (-\frac{1}{2}F_S < \varepsilon_1 \leq \frac{1}{2}F_S, n \in \mathbb{Z}^+) \\
F_2 &= n \cdot F_S + \varepsilon_2 \qquad (-\frac{1}{2}F_S < \varepsilon_2 \leq \frac{1}{2}F_S, n \in \mathbb{Z}^+)
\end{aligned} \tag{8}
$$

Example: $nF_S = 20{,}000Hz$,
$F_1 = 19{,}999Hz$, $F_2 = 20{,}001Hz$

Automatic Switching Attacks (with Data Feedback)

- Program **generates** and **adjusts** attack signals in real-time



- We apply our automatic attacks on
  - **Android (Google Maps)**
    - https://youtu.be/dy6gm9ZLKuY
  - **IOS (VR game)**
    - https://youtu.be/kTQFi9CI8R8
  - Web Scripts
    - https://youtu.be/MkpW_j6gd8k

Demo

Rotating the orientation of Google Maps

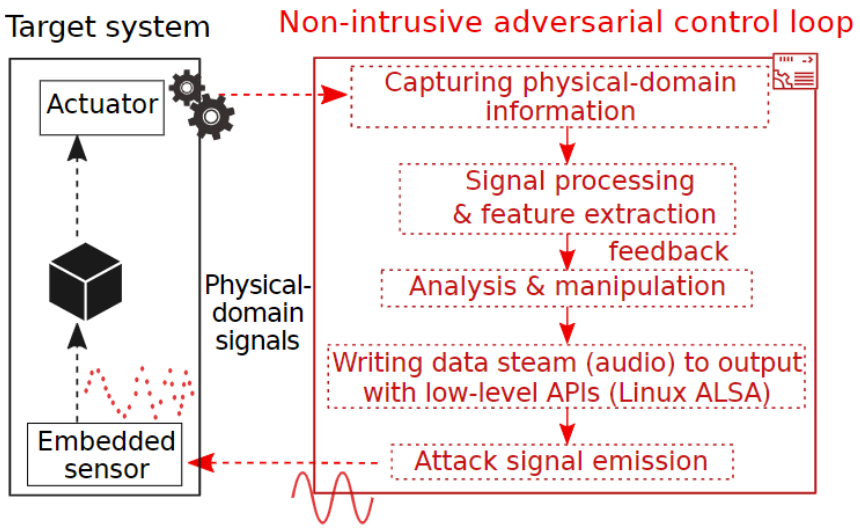
Shooting germs in VR games

Roadmap

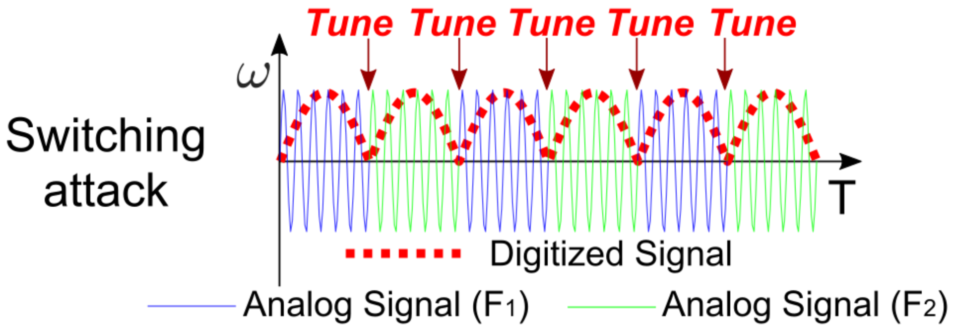
- Research
  - Acoustic attacks on embedded inertial sensor systems
    - Real-time Attacks (**Manual** Attacks) [USENIX’18]
    - **Automatic** Attacks (Using Programs) with Data Feedback [USENIX’18]
    - **Automatic** Attacks (Using Programs) **without** Data Feedback [CPSIoTSec’23]
  - Mitigating safety risks of acoustic attacks on sensors [SmartSP’23]
  - Mitigating safety risks of EMI attacks on sensors [CCS’19, SecTL’23, ASIACCS’21]

Challenges with Automatic Attacks in Black-box Approach

- No digital connections or digital feedback

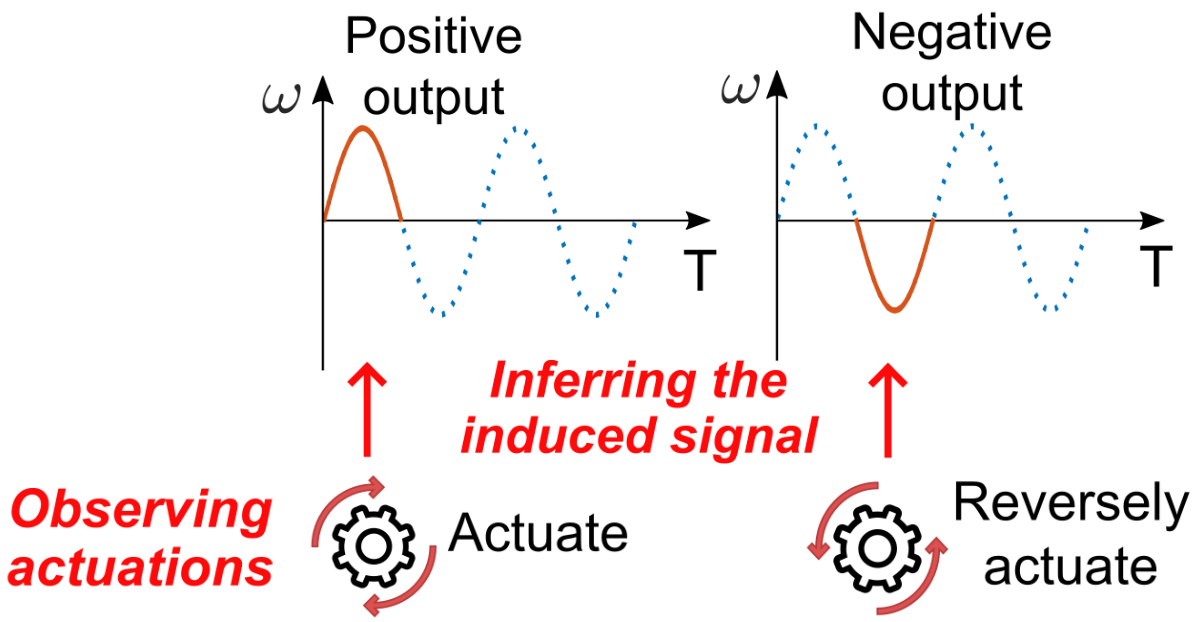


- Problem: Tuning time selection



What if There is No Data Feedback?

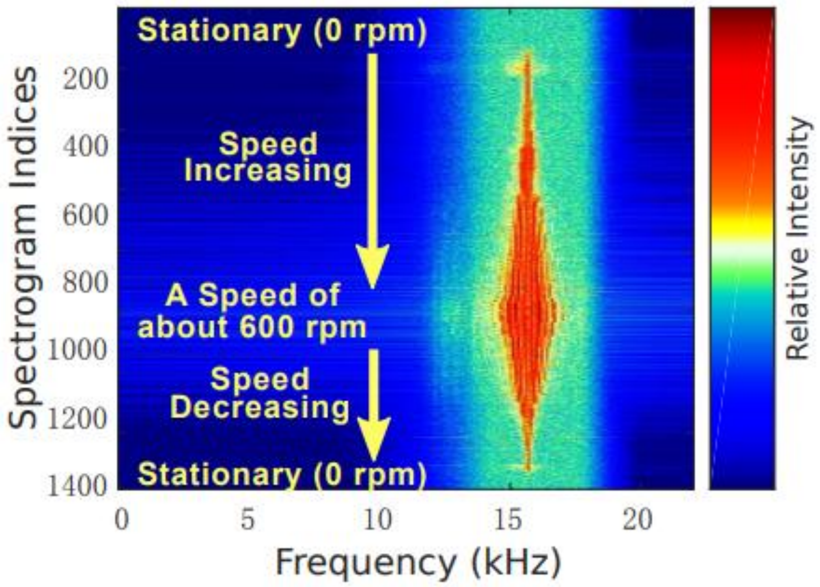
- Reverse Signal Mapping



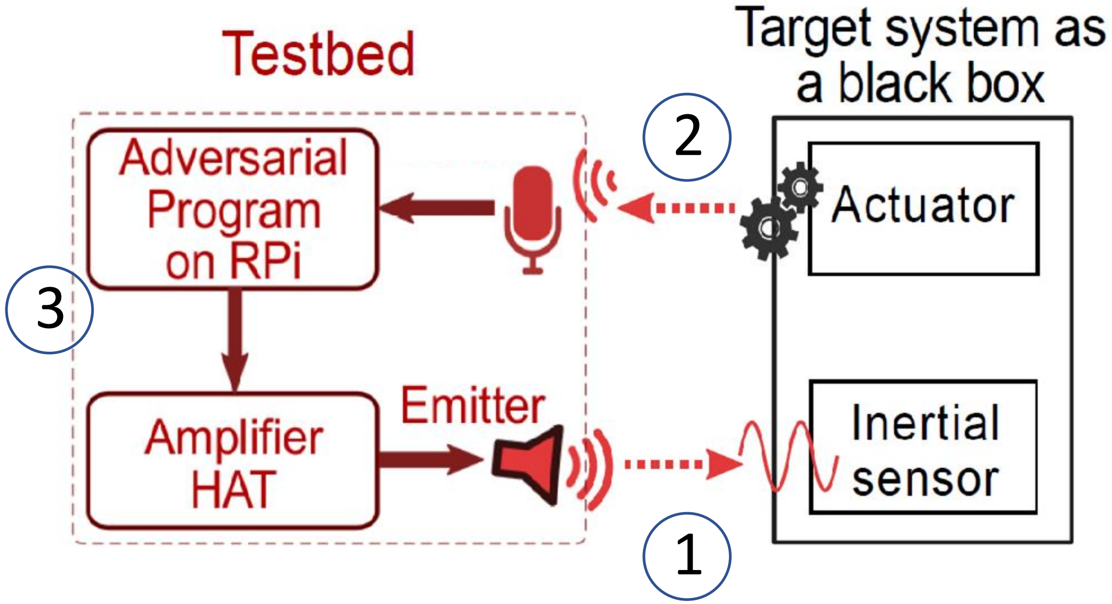
- How to achieve automatic attacks with physical-domain observations?

Automatic Attacks with Physical Feedback Side Channel

- Approach: **adversarial control loop** using the **physical feedback side channel**
- Non-invasively analyze the target's **response under signal injections**

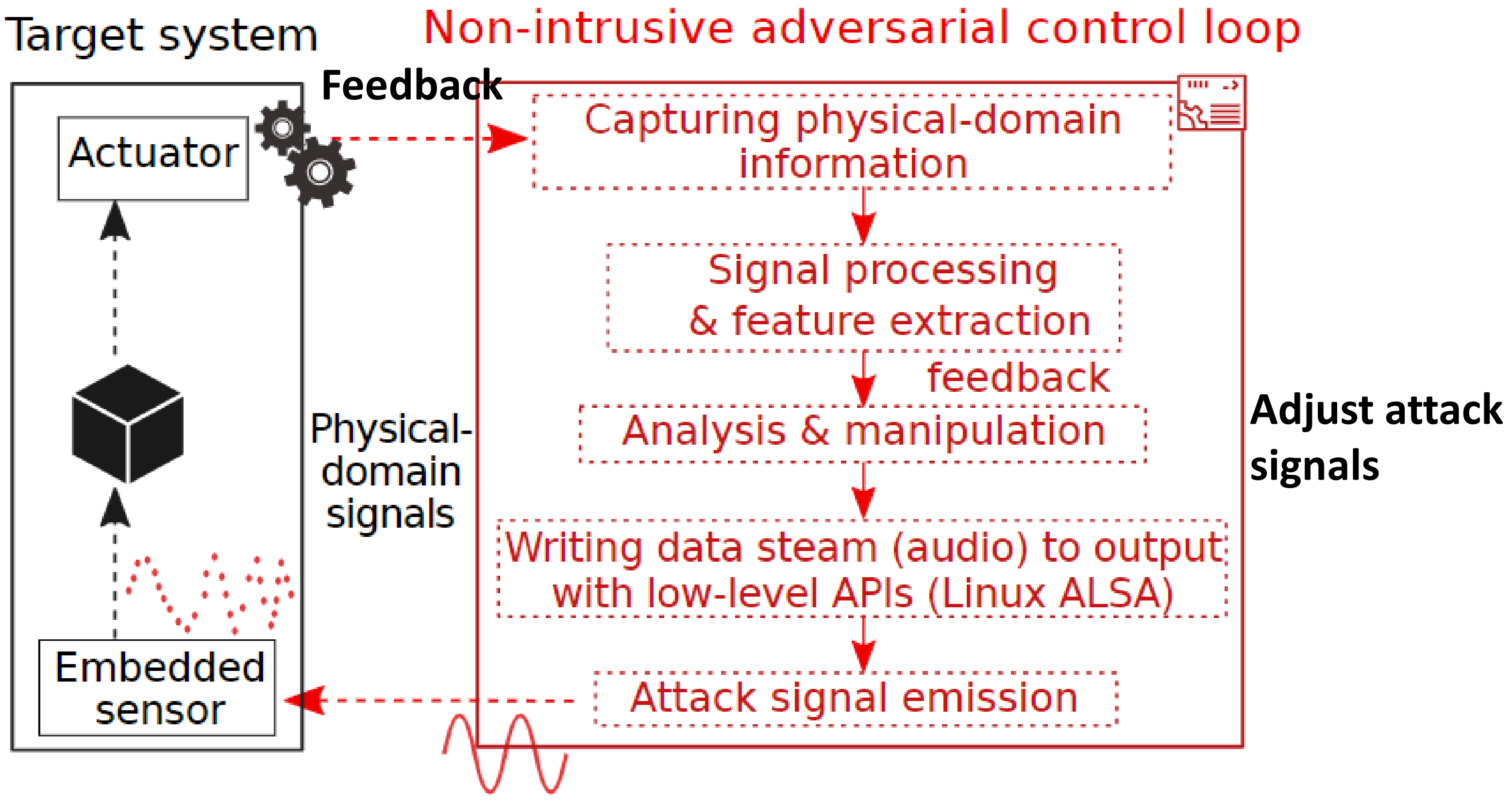


Motor's acoustic emission

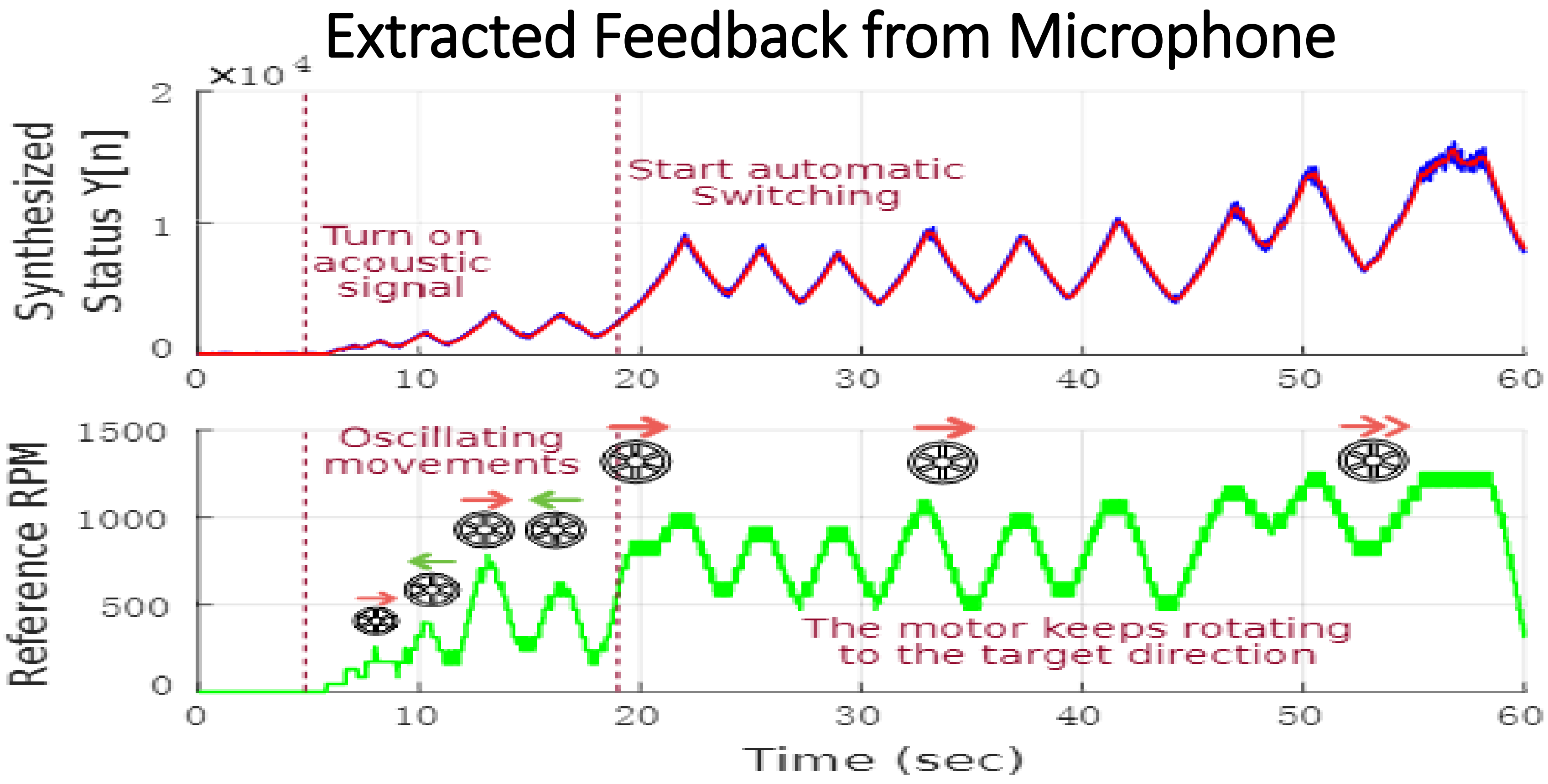


Attack Program Design and Implementation

- The general procedure and basic modules automated with **algorithms**
- **Multi-threaded** program to form **continuous real-time control**



Automatic Switching Attacks on a Self-Balancing Scooter

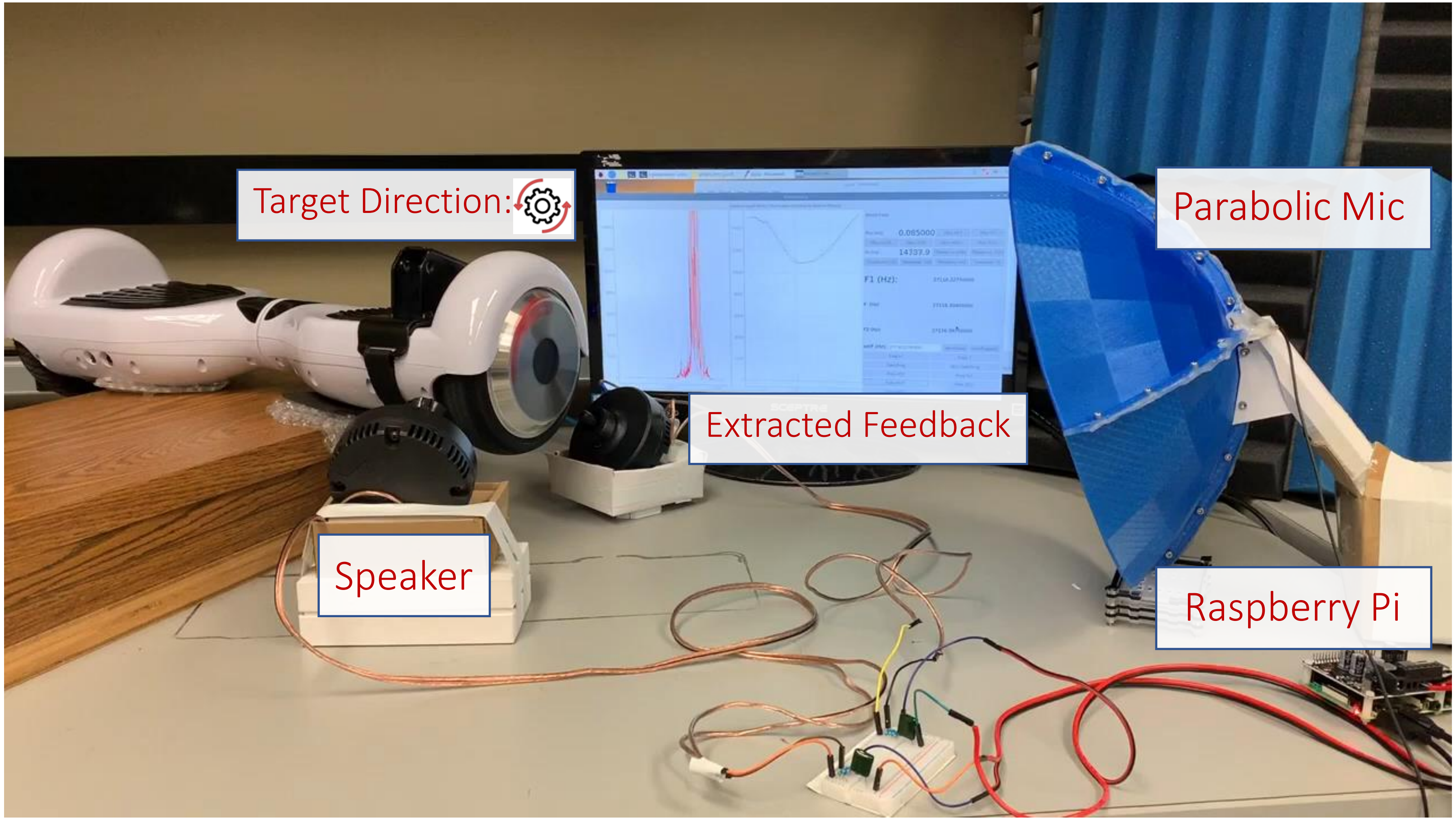


Physical Feedback Side Channel

Attack Effect

Automatic Switching attacks on a self-balancing scooter

Demo: Real-Time Automatic Attack with Physical Feedback Side Channel



Video Code

Contributions

- **Theoretical results:**
  - Sample rate drifts amplification theorem
- ***Non-invasive*** attacks on inertial sensors ***embedded in real systems***
  - *Side-Swing* and *Switching* attacks
  - Evaluated on 25 devices
  - Demonstrate implicit control over different kinds of systems
- ***Automatic*** attacks with/without data feedback
  - The attack system is not connected to any digital interfaces of the system

Roadmap

- Research
  - Acoustic attacks on embedded inertial sensor systems
    - Real-time Attacks (**Manual** Attacks) [USENIX'18]
    - **Automatic** Attacks (Using Programs) with Data Feedback [USENIX'18]
    - **Automatic** Attacks (Using Programs) **without** Data Feedback [CPSIoTSec'23]
  - Mitigating safety risks of acoustic attacks on sensors [SmartSP'23]
  - Mitigating safety risks of EMI attacks on sensors [CCS'19, SecTL'23, ASIACCS'21]

ADC-Bank:
Detecting Acoustic Out-of-Band Signal Injection on Inertial Sensors

Jianyi Zhang, Yuchen Wang, Yazhou Tu, Sara Rampazzi,
Zhiqiang Lin, Insup Lee, Xiali Hei

北京電子科技学院
Beijing Electronics Science & technology Institute
UNIVERSITY of LOUISIANA LAFAYETTE
UF UNIVERSITY of FLORIDA
THE OHIO STATE UNIVERSITY
Penn UNIVERSITY of PENNSYLVANIA

Motivation

- Inertial sensors with **low-pass filters** are still vulnerable to acoustic attacks [1].
- **Shielding** can cause heat dissipation, cost, size, and usability issues.
- **Sampling-based methods like random sampling** or canceling a frequency component by adding two samples based on a delay calculated from a known frequency [2].
- However, it is difficult to cancel injected signals because the sensors are usually vulnerable in one or more frequency ranges instead of a single, previously determined frequency [3; 2; 1].
- The above methods **can not correct** the attack effects.

[1] Tu, Yazhou, et al. "Injected and delivered: Fabricating implicit control over actuation systems by spoofing inertial sensors." USENIX Security 2018.

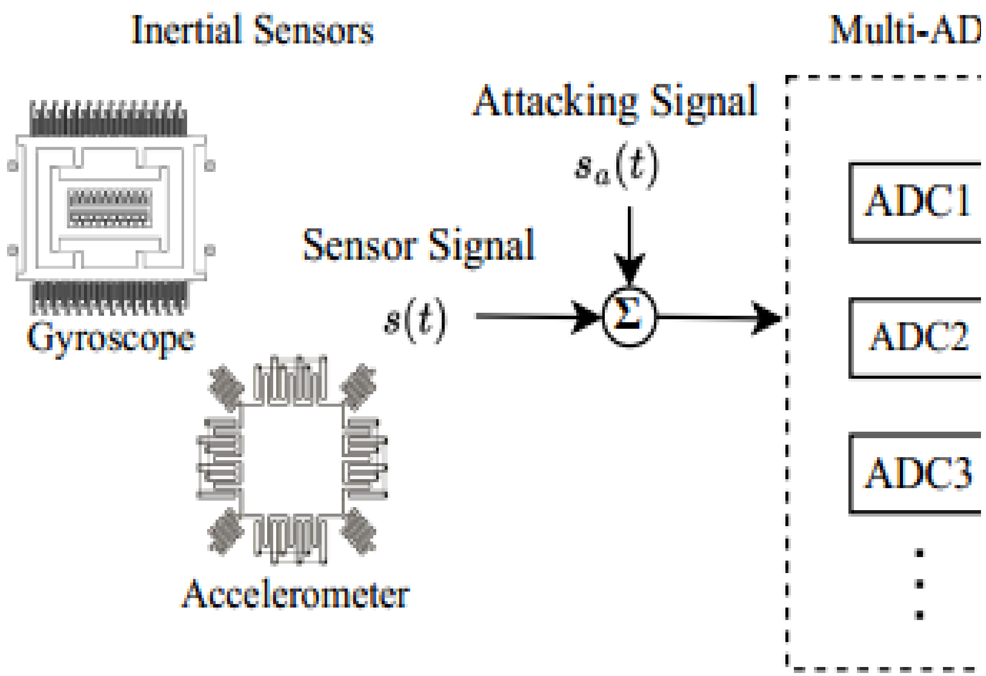
[2] Trippel, Timothy, et al. "WALNUT: Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks." IEEE European symposium on security and privacy (EuroS&P 2017).

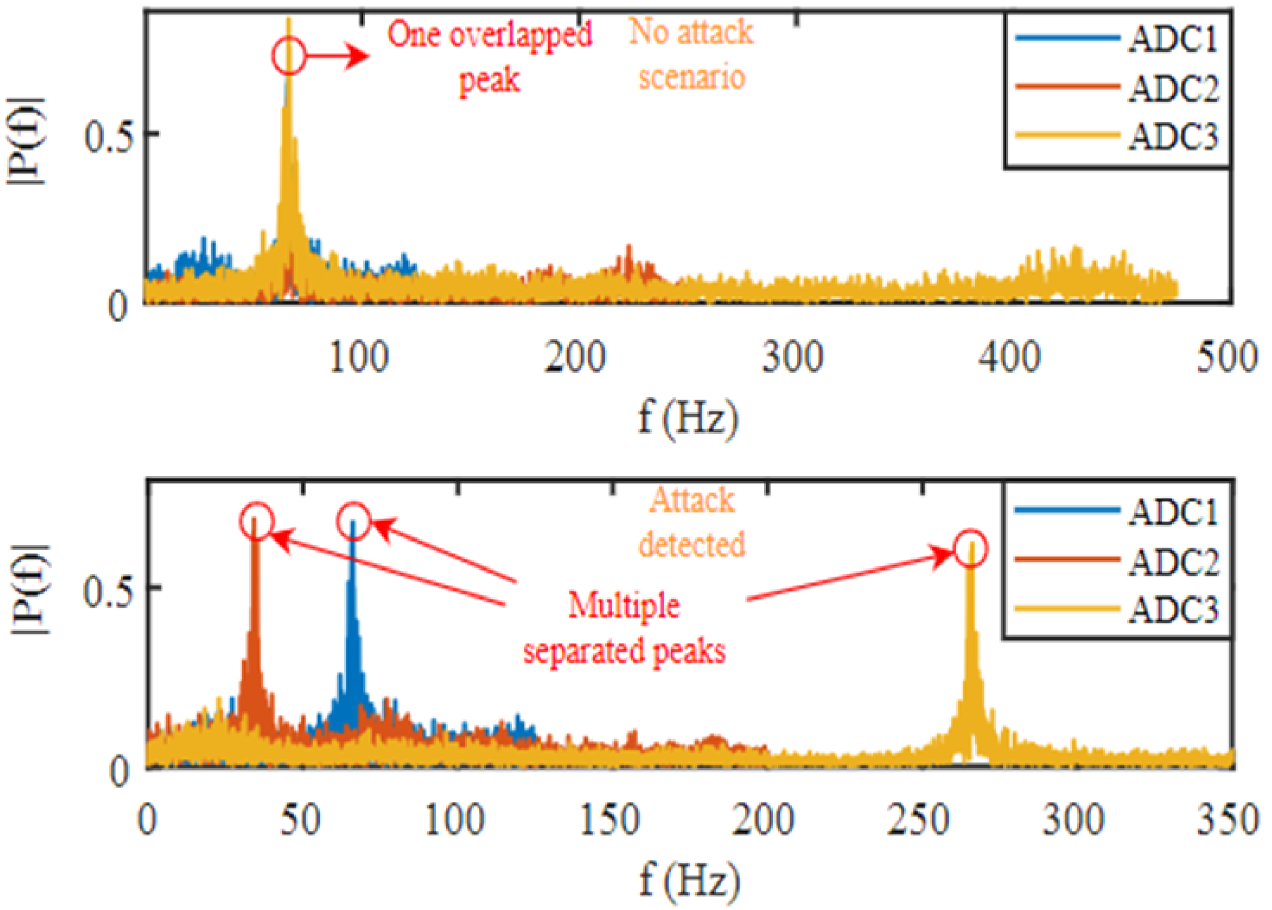
[3] Son, Yunmok, et al. "Rocking drones with intentional sound noise on gyroscopic sensors." USENIX Security 2015.

Defense Approach

- High-level Idea
  - Component redundancy
  - Different configurations and settings
  - Simultaneously measure the physical stimulus
  - Some ADCs have multi-paths, and the sampling rate could be set, easy to integrate, less errors than filters

System Model





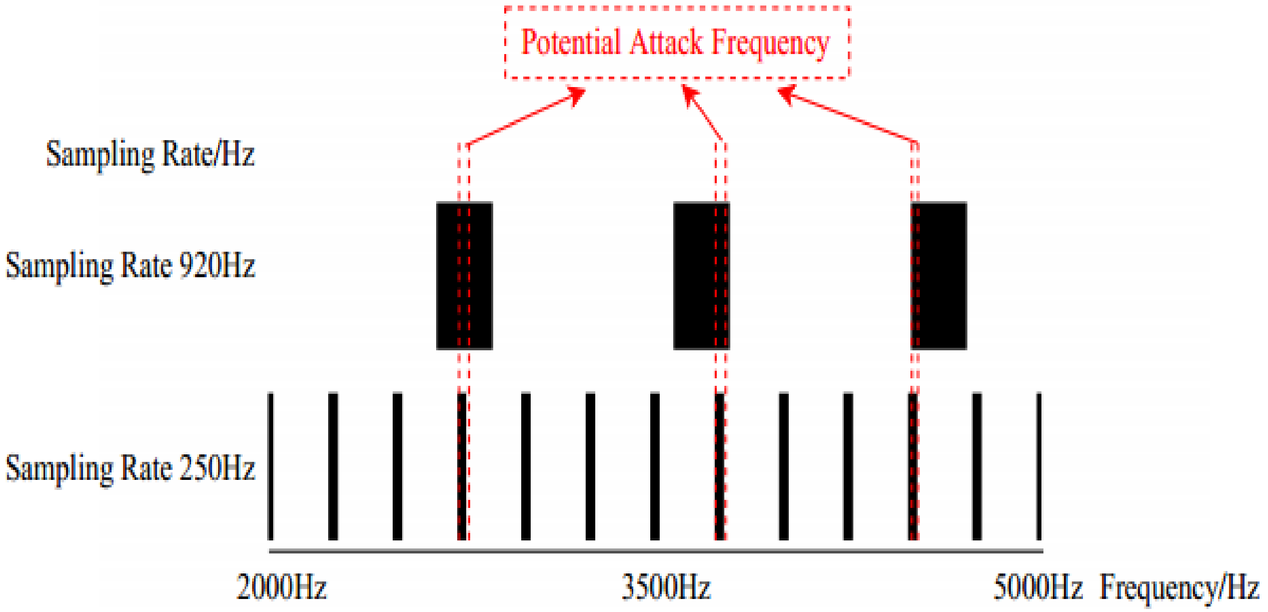
$$F = n \cdot F_S + \epsilon \quad (-\tfrac{1}{2}F_S < \epsilon \leq \tfrac{1}{2}F_S, n \in \mathbb{Z}^+)$$

Attack Frequency Analysis

According to the **peak frequency** obtained from ADC of a certain sampling rate, we can calculate the possible attack frequency ranges of several segments by the following equation:

$$F_a = n_i \cdot F_{Si} + \varepsilon_i \quad (-\frac{1}{2}F_{Si} < \varepsilon_i \le \frac{1}{2}F_{Si}, n_i \in \mathbb{Z}^+)$$

F_{Si} is the sampling rate



Discussion

- Adaptive Attacks and Frequency Drift
  - Acoustic-based Spoofing Attacks
    - Slight frequency drift or sample rate jitter in acoustic-based spoofing attacks can significantly affect sensor output.
  - Challenges for Adaptive Attacks
  - Effect of Increasing ADCs
  - **Limitations** of Frequency Sweeping and Hopping

Conclusion

- We proposed a **component redundancy scheme** to detect acoustic out-of-band signal injection by elaborating and comparing the physical stimulus in different settings.
- We investigated how to **mine the real physical stimulus** from different results of the redundant components.
- We deploy our defense method on off-the-shelf inertial sensors with **commercial ADCs** to evaluate our method.
- We discuss how our strategy can **be used** in future sensors' design and manufacturing.

Roadmap

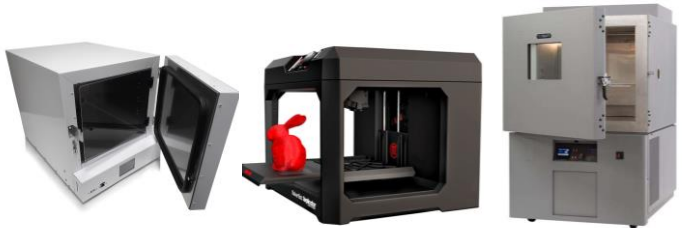
- Research
  - Acoustic attacks on embedded inertial sensor systems
    - Real-time Attacks (**Manual** Attacks) [USENIX’18]
    - **Automatic** Attacks (Using Programs) with Data Feedback [USENIX’18]
    - **Automatic** Attacks (Using Programs) **without** Data Feedback [CPSIoTSec’23]
  - Mitigating safety risks of acoustic attacks on sensors [SmartSP’23]
  - Mitigating safety risks of EMI attacks on sensors [CCS’19, SecTL’23, ASIACCS’21]

Hospitals

Home

Reactors in Labs & Chemical Plants

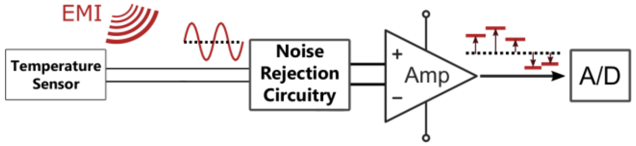
Industrial & Manufacturing Facilities



Mitigating Safety Risks of EMI Attacks on Sensors

Investigating security and **Safety Risks** of EMI Attacks on Temperature Sensors

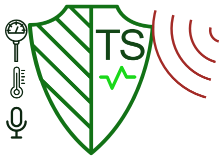
1. *Trick or heat? Manipulating critical temperature-based control systems using rectification attacks*, In *ACM CCS, 2019*



2. *A First Look at the Security of EEG-based Systems and Intelligent Algorithms under Physical Signal Injections". In ACM SecTL, 2023*

Defending against EMI Signal Injections on Sensors

3. *Transduction shield: A low-complexity method to detect and correct the effects of EMI injection attacks on sensors*, In *ACM AsiaCCS*, 2021



Yazhou Tu, Sara Rampazzi, Bin Hao, Angel Rodriguez, Kevin Fu, and **Xiali Hei**. "Trick or heat? Manipulating critical temperature-based control systems using rectification attacks." In *Proceedings of the **ACM CCS***. 2019.

Md Imran Hossen, Yazhou Tu, and **Xiali Hei**. "A First Look at the Security of EEG-based Systems and Intelligent Algorithms under Physical Signal Injections". SecTL 2023.

Yazhou Tu, Vijay Srinivas Tida, Zhongqi Pan, and **Xiali Hei**. "Transduction shield: A low-complexity method to detect and correct the effects of EMI injection attacks on sensors." In *Proceedings of the ACM **AsiaCCS***. 2021.

Trick or Heat? Manipulating Critical Temperature-Based Control Systems Using Rectification Attacks

- A joint work with
- Published at ACM

Session 10A: Cyberphysical Security   CCS '19, November 11–15, 2019, London, United Kingdom

Trick or Heat? Manipulating Critical Temperature-Based Control Systems Using Rectification Attacks

Yazhou Tu*
University of Louisiana at Lafayette
yazhou.tu1@louisiana.edu

Sara Rampazzi*
University of Michigan
srampazz@umich.edu

Bin Hao
University of Louisiana at Lafayette
bin.hao@louisiana.edu

Angel Rodriguez
University of Michigan
angelrod@umich.edu

Kevin Fu
University of Michigan
kevinfu@umich.edu

Xiali Hei
University of Louisiana at Lafayette
xiali.hei@louisiana.edu

ABSTRACT

Temperature sensing and control systems are widely used in the closed-loop control of critical processes such as maintaining the thermal stability of patients, or in alarm systems for detecting temperature-related hazards. However, the security of these systems has yet to be completely explored, leaving potential attack surfaces that can be exploited to take control over critical systems.

In this paper we investigate the reliability of temperature-based control systems from a security and safety perspective. We show how unexpected consequences and safety risks can be induced by

KEYWORDS

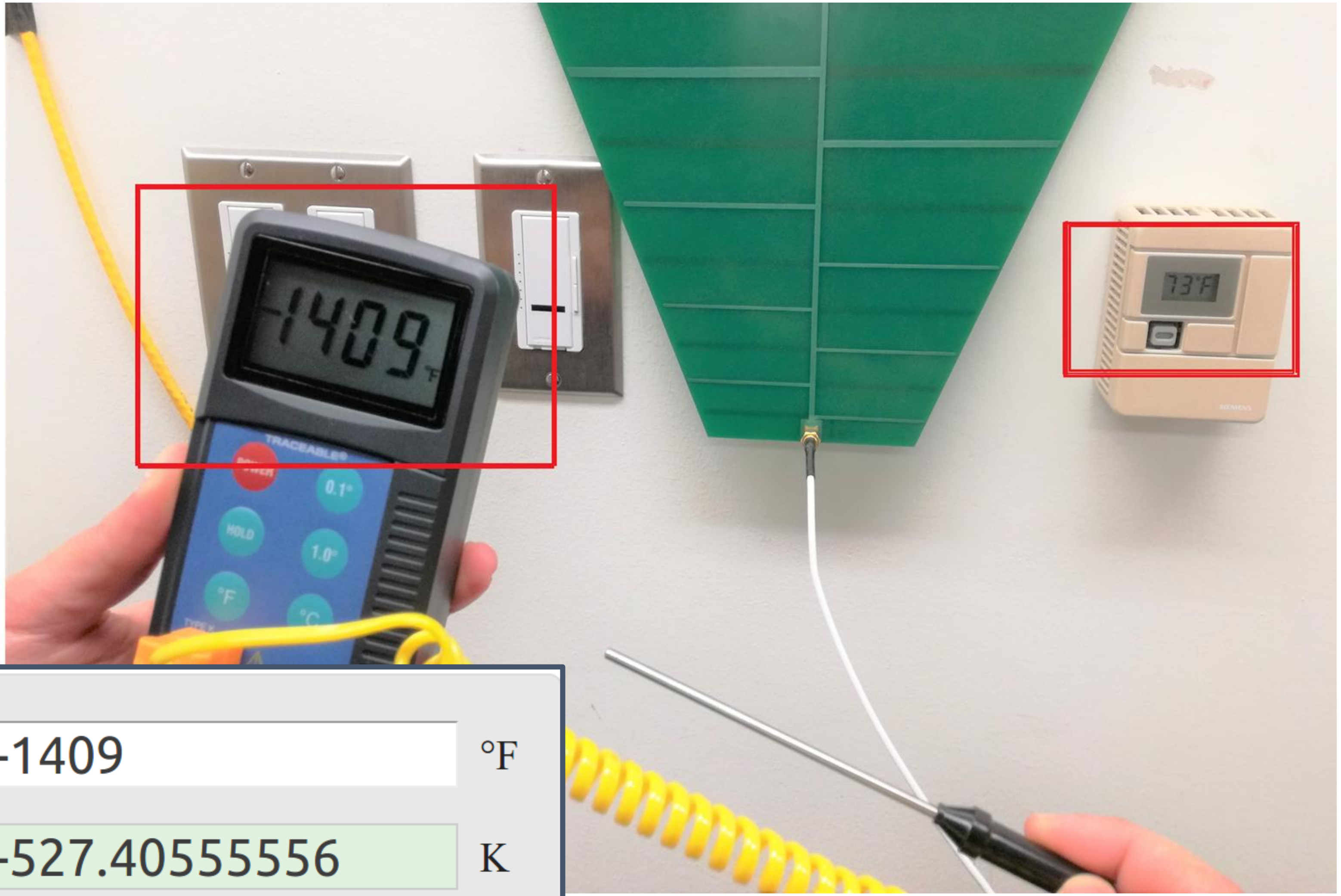
Hardware Security; Safety-Critical Systems; Sensor Signal Injections; Temperature Sensors

ACM Reference Format:

Yazhou Tu, Sara Rampazzi, Bin Hao, Angel Rodriguez, Kevin Fu, and Xiali Hei. 2019. Trick or Heat? Manipulating Critical Temperature-Based Control Systems Using Rectification Attacks. In *2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19), November 11–15, 2019, London, United Kingdom.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3319535.3354195

Can temperature be easily manipulated?

YES.



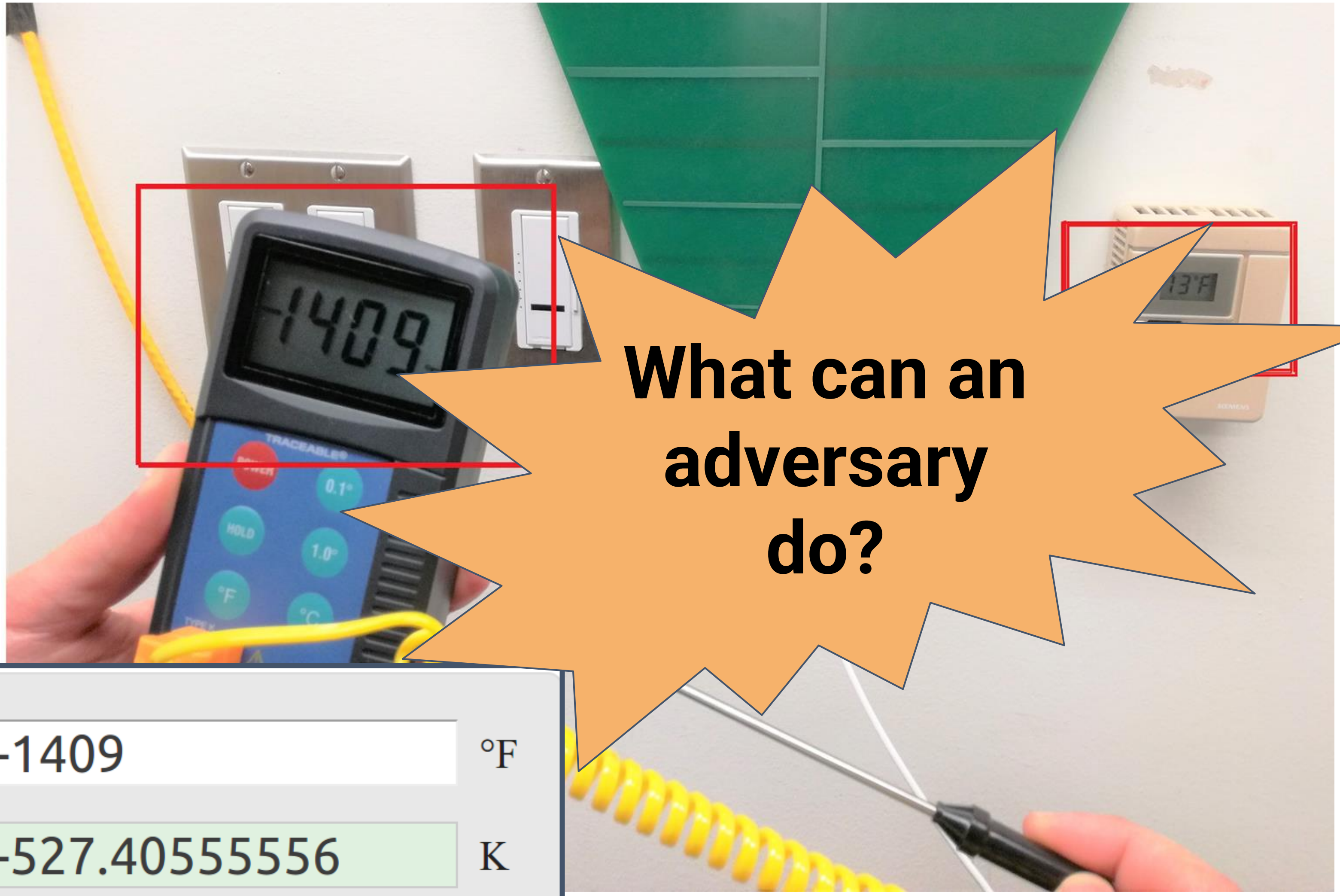
Fahrenheit: -1409 °F

Kelvin: -527.40555556 K

Absolute zero
- 459.67 °F

Can temperature be easily manipulated?

YES. Also using a simple walkie-talkie



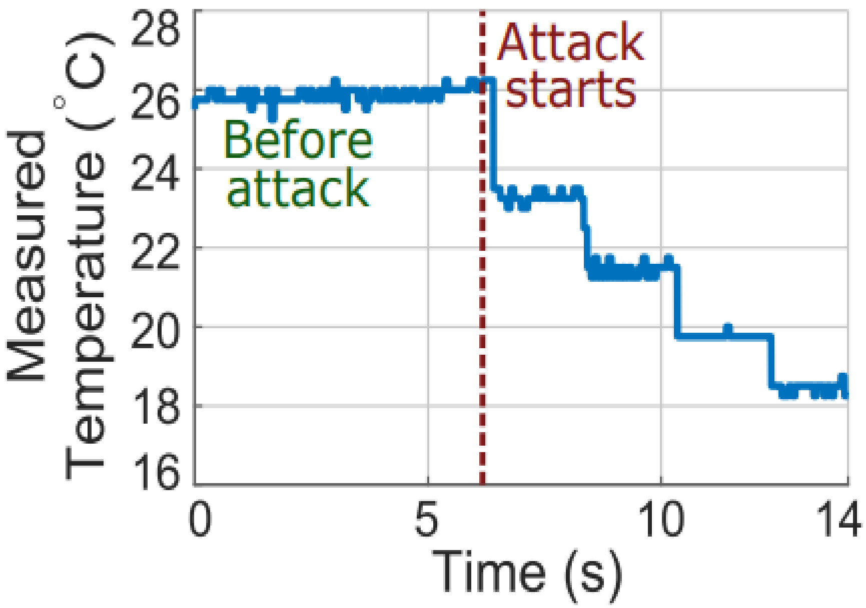
What can an adversary do?

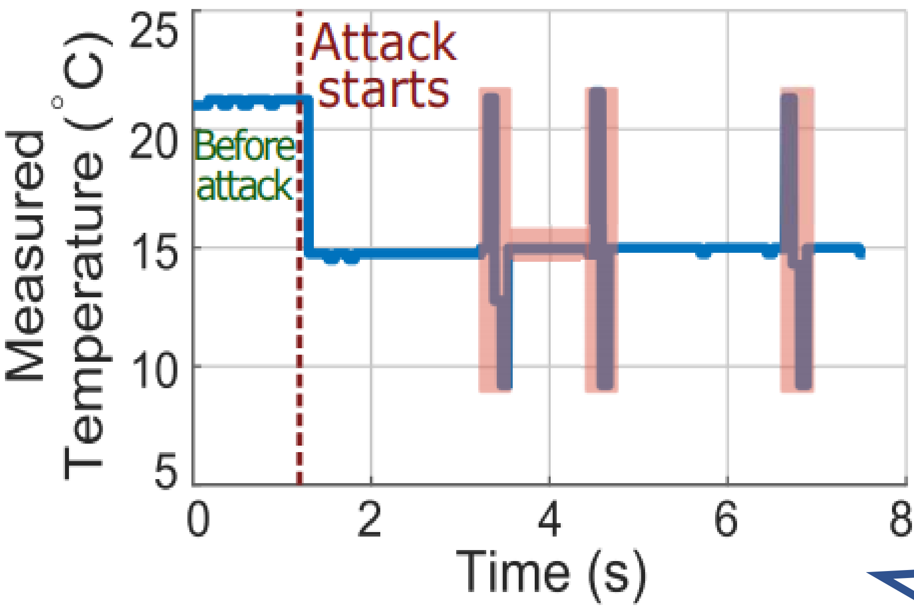
Absolute zero
- 459.67 °F

Fahrenheit:	-1409	°F
Kelvin:	-527.40555556	K

Controlling the Temperature Sensor Output

K-Type Shielded Thermocouple
1 m distance
3.08 W power

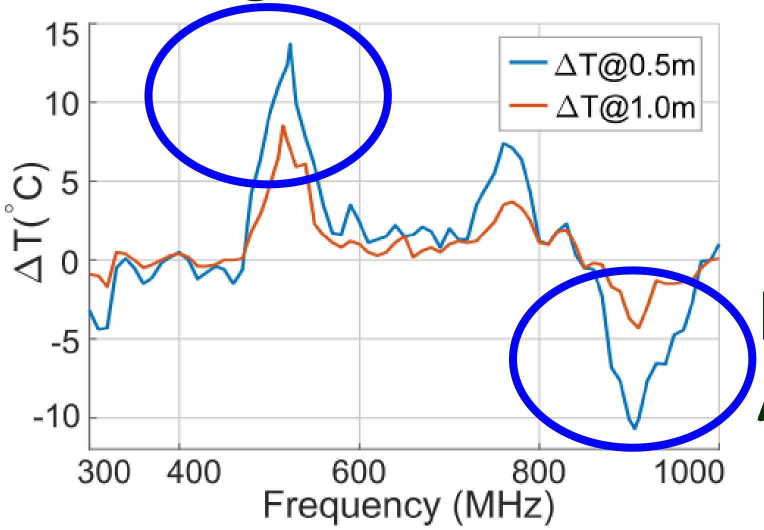




Real-World Implications

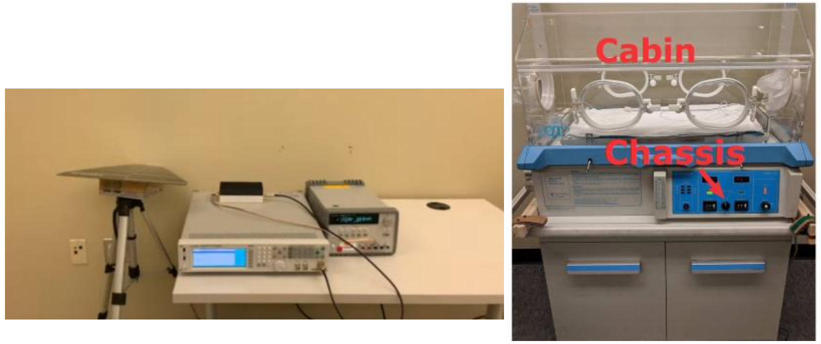
- Medical Devices, Cold Chain of Vaccine[1] and Blood, and Biomedical Applications

Cooling Attack



Heating Attack

EMI Injections on Temperature Sensor of Infant Incubator

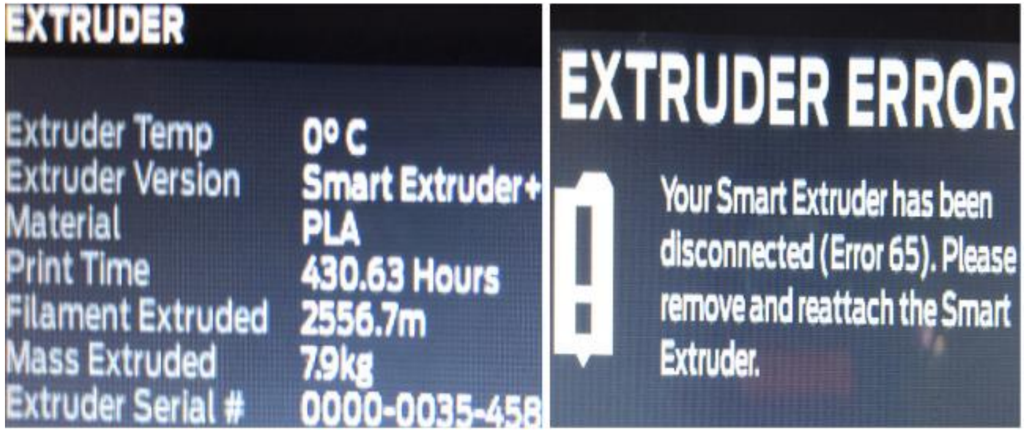


Attack can be launched from adjacent room (15 cm wall)

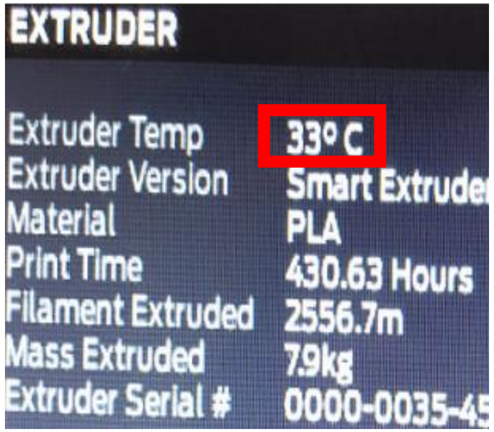
1. Long et al. "Protecting COVID-19 Vaccine Transportation and Storage from Analog Cybersecurity Threats." Biomedical Instrumentation & Technology 55.3 (2021): 112-117.

Real-World Implications

- EMI-Based Adversarial Temperature Control in Additive Manufacturing



EMI: 400 MHz



EMI: 1 GHz

Frequency: 1 GHz

Power: 3.08 W

The actual temperature: 23◦C

Hospitals

Home

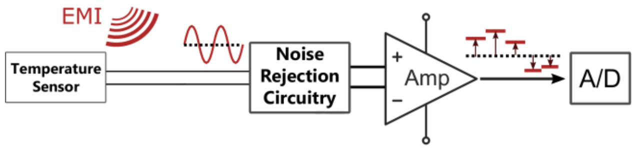
Reactors in Labs & Chemical Plants

Industrial & Manufacturing Facilities

Mitigating Safety Risks of EMI Attacks on Sensors

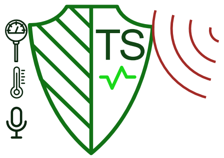
Investigating security and **Safety Risks** of EMI Attacks on Temperature Sensors

1. *Trick or heat? Manipulating critical temperature-based control systems using rectification attacks*, In *ACM CCS, 2019*
2. *A First Look at the Security of EEG-based Systems and Intelligent Algorithms under Physical Signal Injections". In ACM SecTL, 2023*



Defending against EMI Signal Injections on Sensors

3. *Transduction shield: A low-complexity method to detect and correct the effects of EMI injection attacks on sensors*, In *ACM AsiaCCS*, 2021



Yazhou Tu, Sara Rampazzi, Bin Hao, Angel Rodriguez, Kevin Fu, and **Xiali Hei**. "Trick or heat? Manipulating critical temperature-based control systems using rectification attacks." In *Proceedings of the **ACM CCS***. 2019.

Md Imran Hossen, Yazhou Tu, and **Xiali Hei**. "A First Look at the Security of EEG-based Systems and Intelligent Algorithms under Physical Signal Injections". SecTL 2023.

Yazhou Tu, Vijay Srinivas Tida, Zhongqi Pan, and **Xiali Hei**. "Transduction shield: A low-complexity method to detect and correct the effects of EMI injection attacks on sensors." In *Proceedings of the ACM **AsiaCCS***. 2021.

Secure and Trustworthy Deep Learning Systems (SecTL) 2023

A First Look at the Security of EEG-based Systems and Intelligent Algorithms under Physical Signal Injections

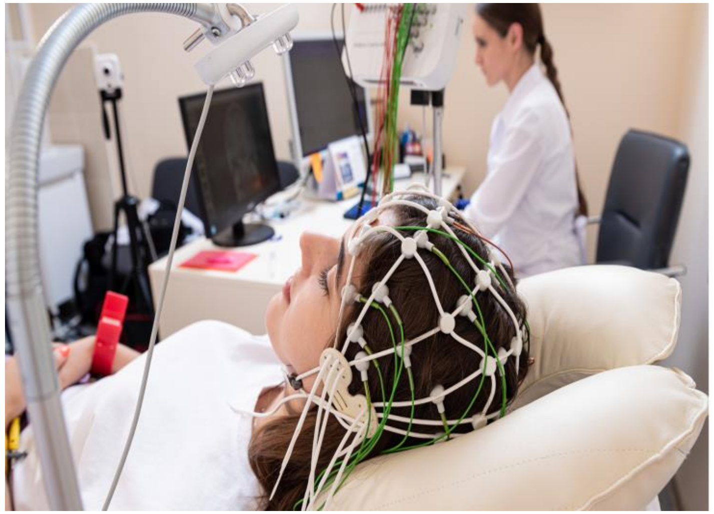
Md Imran Hossen, Yazhou Tu and Xiali Hei
The Center for Advanced Computer Studies
University of Louisiana at Lafayette
July 10, 2023

UNIVERSITY of LOUISIANA LAFAYETTE

UNIVERSITY OF LOUISIANA AT LAFAYETTE
SCHOOL OF COMPUTING & INFORMATICS
Ray P. Authement College of Sciences

EEG Data in Everyday Life

Healthcare



Diagnosing and monitoring **neurological disorders**, such as epilepsy and sleep disorders

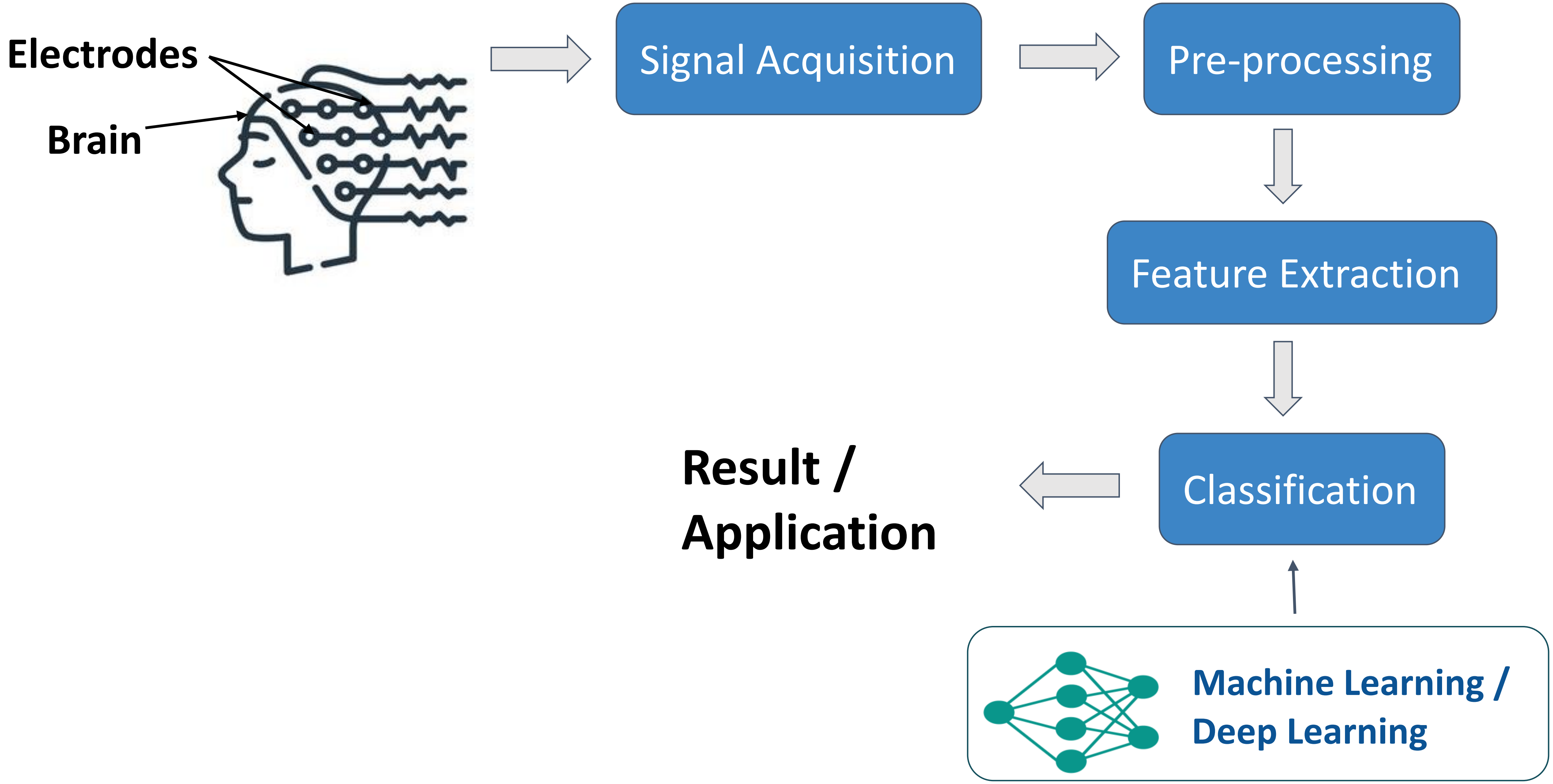
Research

Cognitive neuroscience research to study **brain functions** and **mental states**

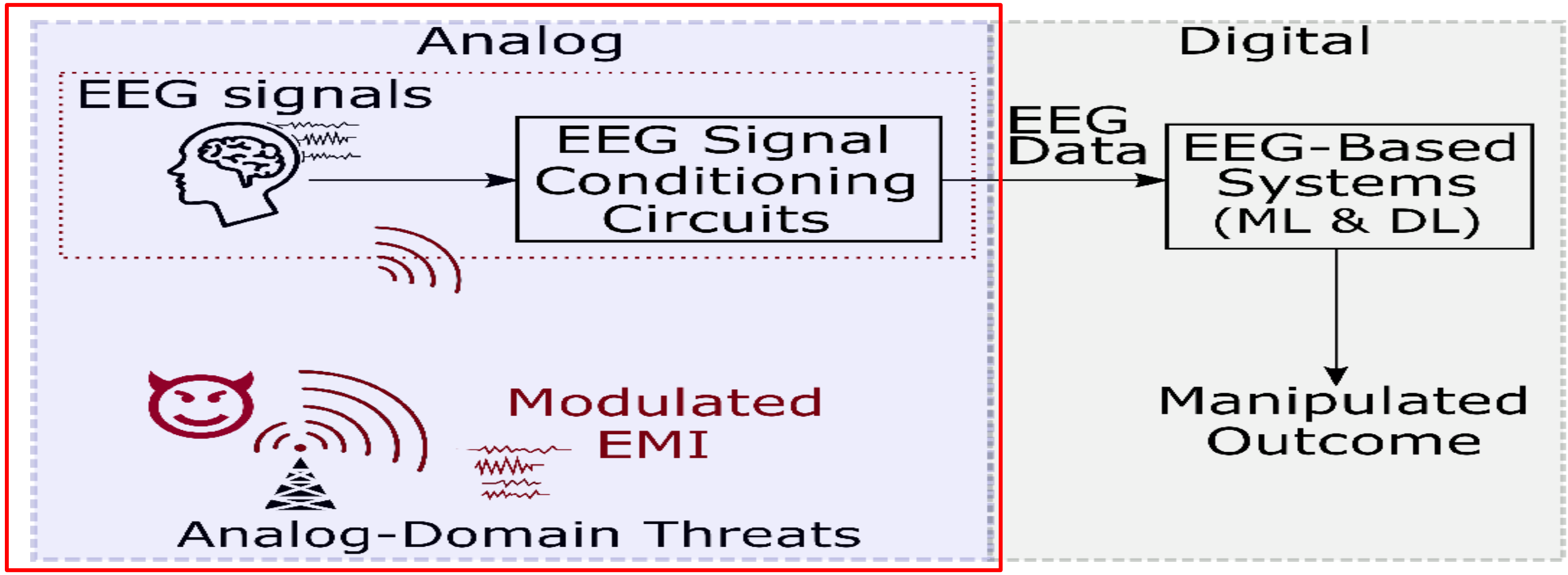
Entertainment

Brain-computer interfaces (BCIs) for **gaming** and **virtual reality** applications

EEG-based Intelligent Systems



Robustness of EEG-based Systems under **Analog-Domain Threats**



EMI → **Electromagnetic Interference**

ML → **Machine Learning**

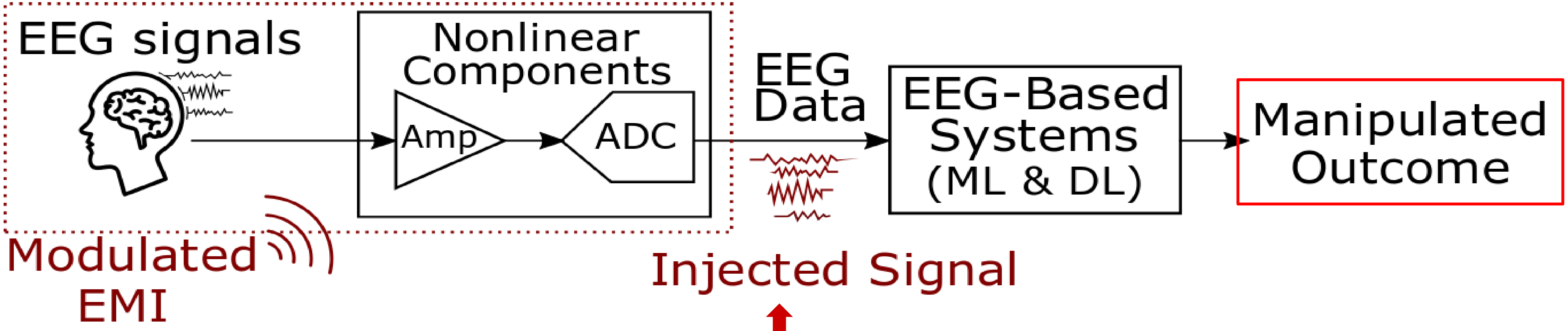
DL → **Deep Learning**

→ **Consequences of the attacks can be severe:**

- Attacks can lead to **misclassification** of EEG signals, resulting in **incorrect diagnoses or interpretations**
- Manipulating the models can cause the system to provide **incorrect recommendations or actions** based on compromised EEG data

Physical Signal Injection Attacks

- **Novelty: First demonstration** of **physical signal injection** attacks on ML and DL models utilizing EEG data
- **Attack Methodology: Non-invasively** injecting signals into EEG recordings



Simple Patterns:
Potentially **easier** to filter

Sine Wave
Square Wave

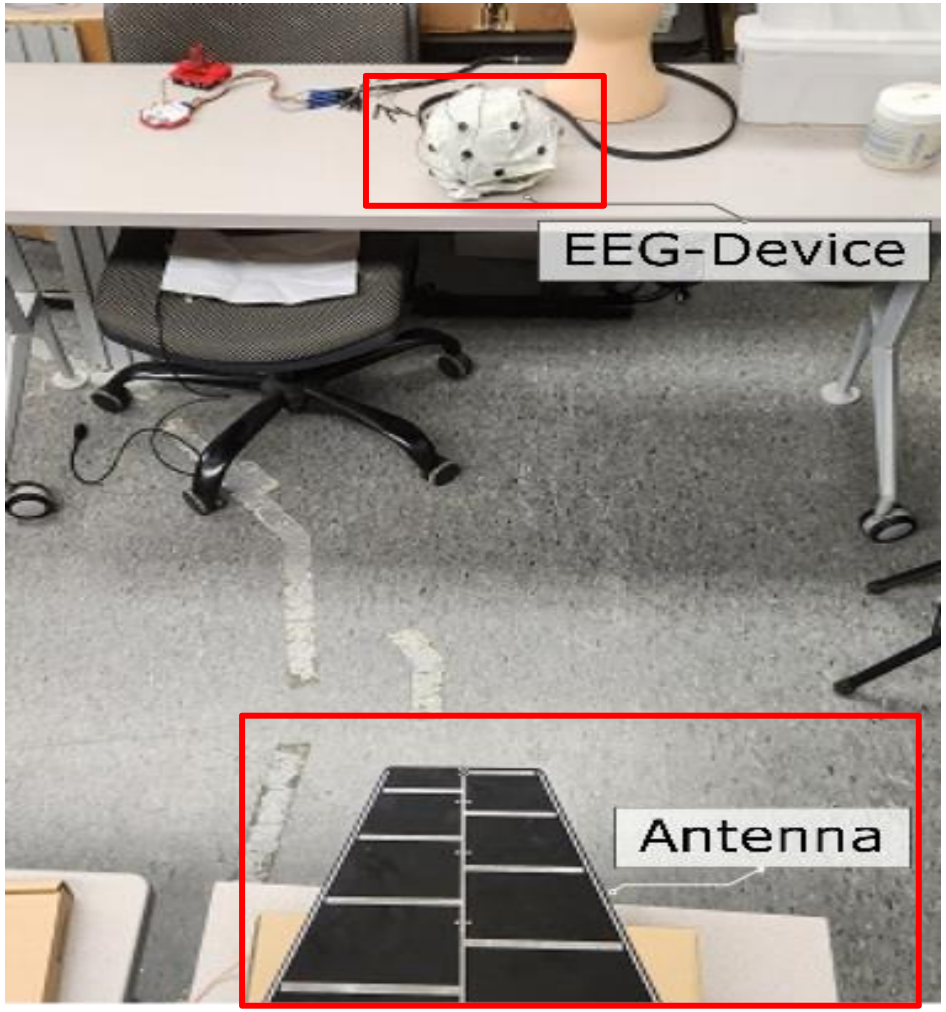
Brain Waves (Theta, Alpha, Mu, etc.)

Brain wave components:
Challenging to detect and filter

Threat Model

- **Adversary Capabilities:**
  - Can **non-invasively** inject signals using **antennas** within a range of one to several meters
  - Can **amplify the transmitting power** and employ directional antennas for signal injection from **longer distances**
  - Could also use **portable** EMI-emitting devices (e.g., off-the-shelf software-defined radio (SDR))

- **Black-Box Attack Setting:**
  - The adversary **does not have the knowledge** of the internal structure or parameters of the **target models**
  - The adversary **does not directly modify** the data to launch the attacks



Experimental Setting

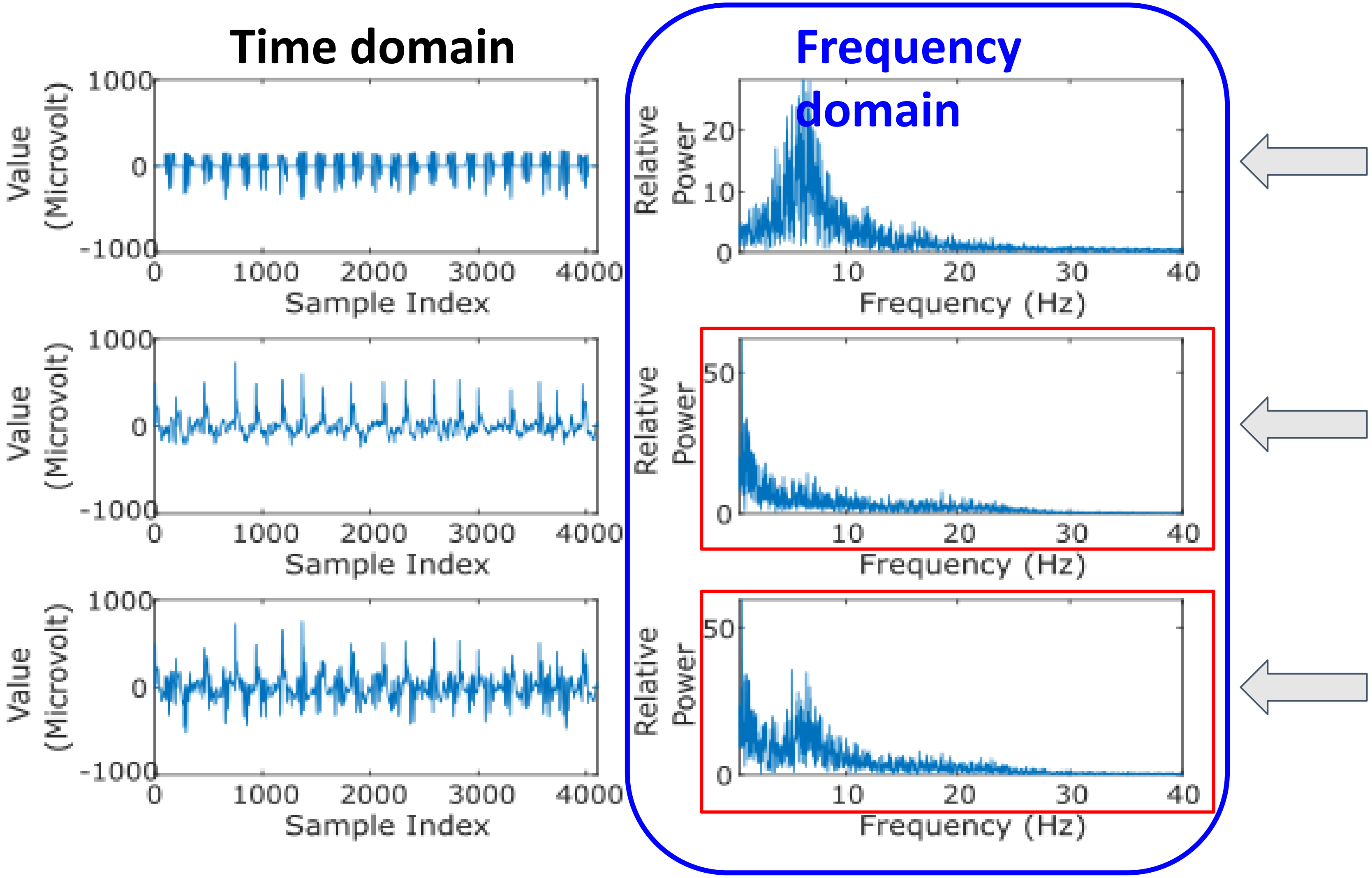
- **ML Task:**
  - **Seizure detection** (classifying EEG recordings into seizure or non-seizure)
- **Dataset:**
  - The **Bonn** University EEG dataset
- **Models:**
  - Logistic Regression **(LR)**
  - Support Vector Machine **(SVM)**
  - Decision Trees **(DT)**
  - Random Forest **(RF)**
  - K-nearest Neighbours **(KNN)**
  - CNN-based DL model **(ConvNet1D)**

Performance of Models **Without** Attack

Model	Accuracy	Precision	Recall	F1
LR	0.9920	0.9923	0.9920	**0.9921**
SVM	0.9920	0.9923	0.9920	**0.9921**
DT	0.9760	0.9765	0.9760	**0.9762**
RF	0.9920	0.9923	0.9920	**0.9921**
KNN	0.9680	0.9678	0.9680	**0.9675**
ConvNet1D	0.9920	0.9923	0.9920	**0.9921**

All models achieve **an F1-score over 96%** on the test dataset for the **seizure** detection

EEG Signal Characteristics **Under** **Attack**



EMI Injected data (**Theta wave** band 4-7 Hz)

Original signal

Attack signal
[original signal + injected data]

Performance of Models **Under** Attack

Injection	Model	Accuracy	Precision	Recall	F1
Sine wave (10 Hz)	LR	0.4720	0.8549	0.4720	**0.4922**
	SVM	0.2880	0.8439	0.2880	**0.2305**
	DT	0.8800	0.9030	0.8800	0.8863
	RF	0.8320	0.9087	0.8320	0.8470
	KNN	0.7920	0.8874	0.7920	0.8115
	ConvNet1D	0.3680	0.8481	0.3680	**0.3552**
Square wave (10 Hz)	LR	0.5040	0.8575	0.5040	**0.5299**
	SVM	0.2400	0.8417	0.2400	**0.1452**
	DT	0.8160	0.8758	0.8160	0.8308
	RF	0.7600	0.8909	0.7600	0.7838
	KNN	0.6800	0.8769	0.6800	0.7111
	ConvNet1D	0.2560	0.8424	0.2560	**0.1746**
Brain-wave-band noise (Theta wave band 4-7 Hz)	LR	0.6720	0.8758	0.6720	0.7036
	SVM	0.2720	0.8431	0.2720	**0.2030**
	DT	0.6160	0.8685	0.6160	**0.6494**
	RF	0.2000	0.0400	0.2000	**0.0667**
	KNN	0.2720	0.8431	0.2720	**0.2030**
	ConvNet1D	0.2000	0.0400	0.2000	**0.0667**

Significant degradation

Performance of Models **Under Attack** (Cont'd)

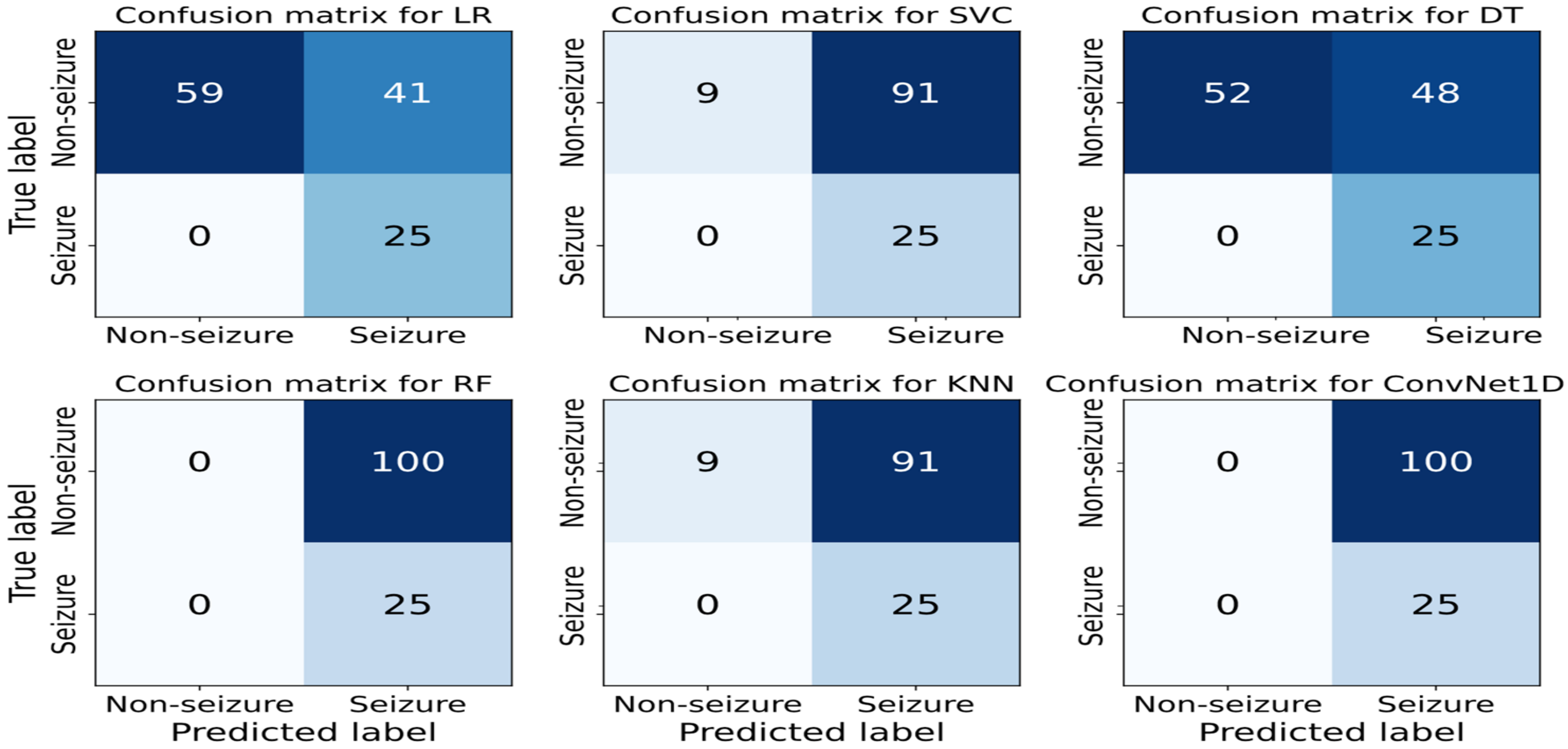


Figure: Confusion matrices for different models with EEG data **corrupted** using **EMI-injected brain-wave-band noise (Theta-wave band 4-7 Hz)**

Contributions

- Conducted the **first study** on **non-invasive physical injection** attacks on EEG-based systems
- Showed how attackers can **degrade the performance** of ML and DL models by injecting external signals into EEG recordings **without requiring access** to the original data, compromising the reliability of EEG-based systems
- Highlighted the need for **trustworthy** bioelectric-signal-based **measuring, processing, and decision-making** to enhance safety and reliability

Transduction shield: A low-complexity method to detect and correct the effects of EMI injection attacks on sensors

Yazhou Tu, Vijay Srinivas Tida, Zhongqi Pan, Xiali Hei

UNIVERSITY of LOUISIANA LAFAYETTE

Transduction Shield: A Low-Complexity Method to Detect and Correct the Effects of EMI Injection Attacks on Sensors

Yazhou Tu
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
yazhou.tu1@louisiana.edu

Vijay Srinivas Tida
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
vijay-srinivas.tida1@louisiana.edu

Zhongqi Pan
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
zhongqi.pan@louisiana.edu

Xiali Hei
University of Louisiana at Lafayette
Lafayette, Louisiana, USA
xiali.hei@louisiana.edu

ABSTRACT

The reliability of control systems often relies on the trustworthiness of sensors. As process automation and robotics keep evolving, sensing methods such as pressure sensing are extensively used in both conventional systems and rapidly emerging applications. The goal of this paper is to investigate the threats and design a low-complexity defense method against EMI injection attacks on sensors.

To ensure the security and usability of sensors and automated processes, we propose to leverage a matched dummy sensor circuit that shares the sensor's vulnerabilities to EMI but is insensitive to legitimate signals that the sensor is intended to measure. Our method can detect and correct corrupted sensor measurements without introducing components or modules that are highly complex compared to an original low-end sensor circuit. We analyze and evaluate our method on sensors with EMI injection experiments using different attack parameters. We investigate several attack scenarios, including manipulating the DC voltage of the sensor

ACM Reference Format:

Yazhou Tu, Vijay Srinivas Tida, Zhongqi Pan, and Xiali Hei. 2021. Transduction Shield: A Low-Complexity Method to Detect and Correct the Effects of EMI Injection Attacks on Sensors. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security (ASIA CCS '21), June 7–11, 2021, Virtual Event, Hong Kong*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3433210.3453097

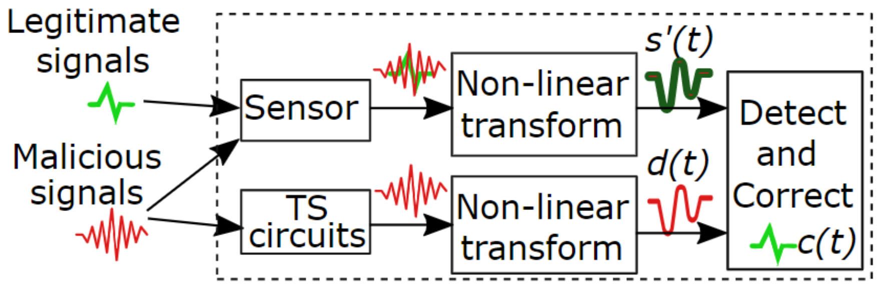
1 INTRODUCTION

Sensors measure physical properties in the real world and provide feedback to guide automated processes in industrial and robotic applications. For instance, pressure sensors are widely utilized in monitoring and control processes in industrial applications, aircraft systems, critical facilities such as the nuclear power plant, and emerging applications such as the sensorization of humanoids to measure physical properties, including strain, weight, flow rate, air and fluid pressure [12, 15, 21, 43, 47, 49].

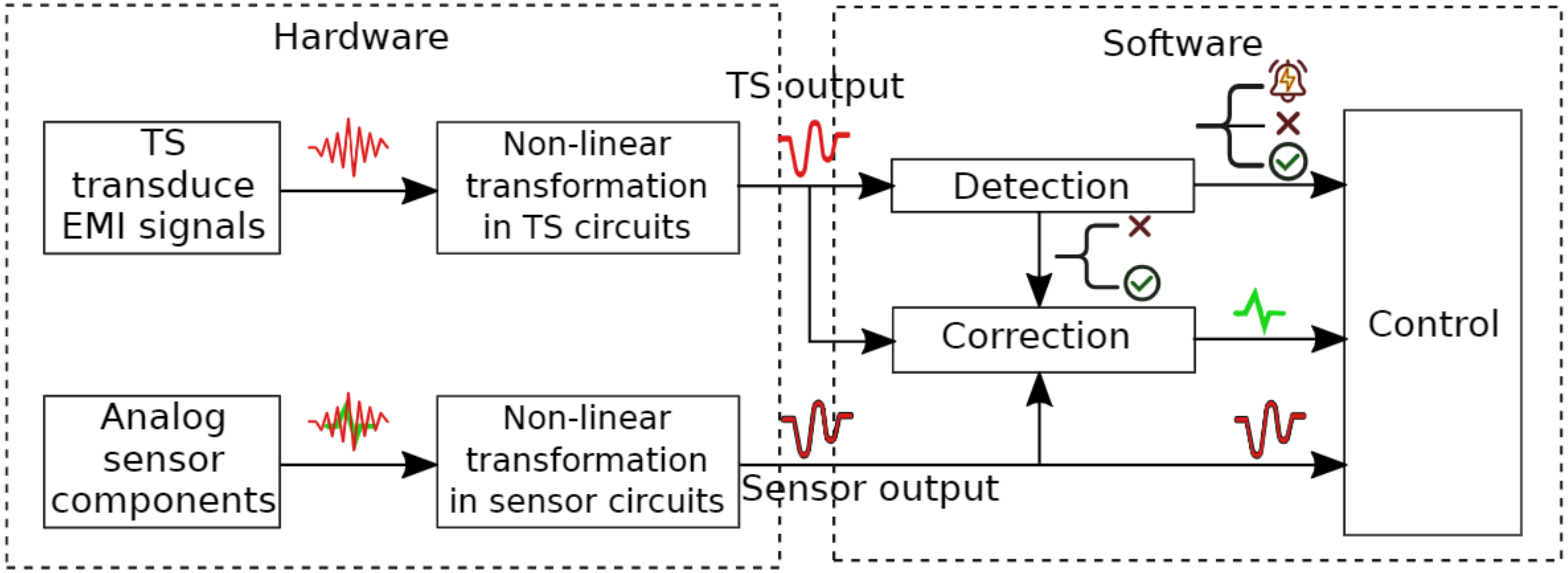
Ideally, sensors should only be sensitive to specific physical stim-

Transduction Shield: Detect and Correct the Effects of EMI Attacks on Sensors

- Leverage a matched dummy sensor circuit
  - shares the sensor's vulnerabilities to EMI
  - insensitive to legitimate signals
  - low-cost

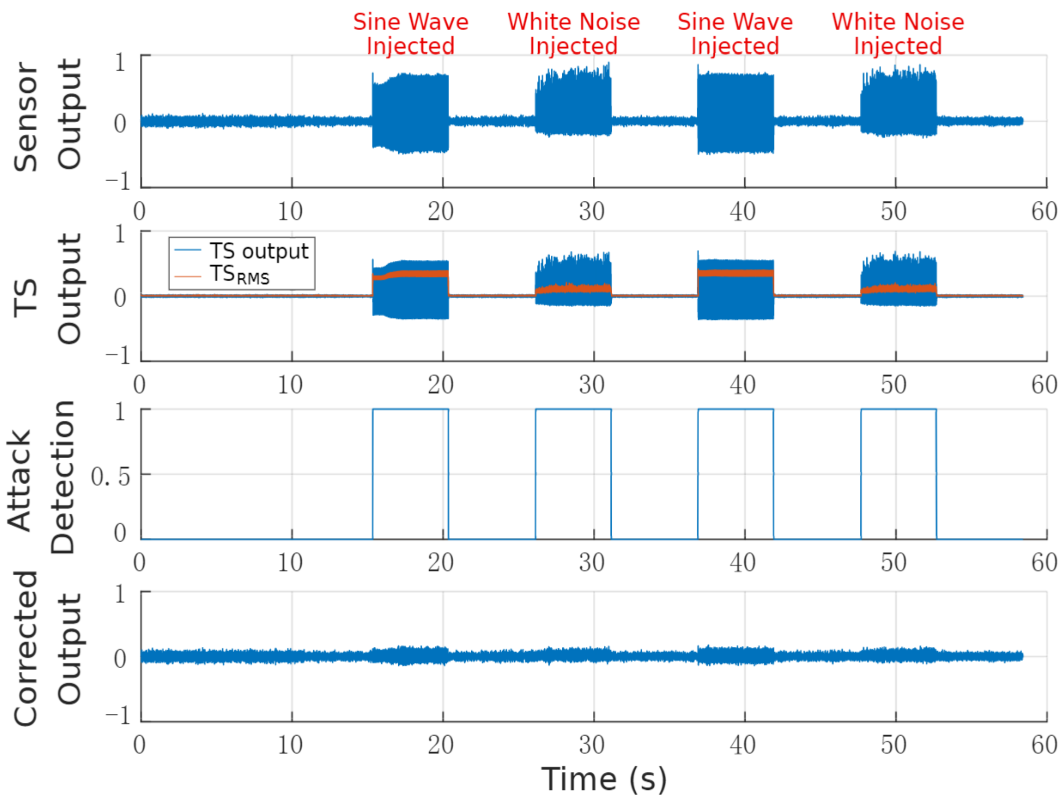


- Detect/correct corrupted sensor measurements in real-time

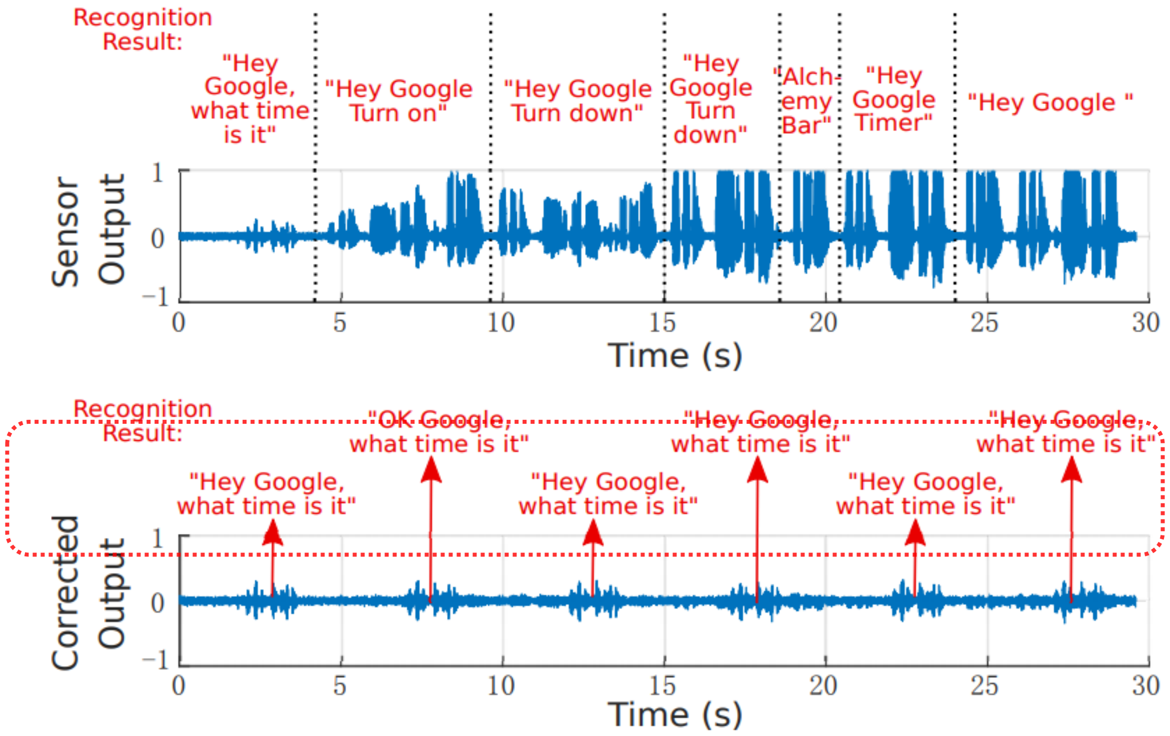


Results

Sine wave and white noise injection



Malicious voice injection



Legitimate signal: Hey Google, what time is it?

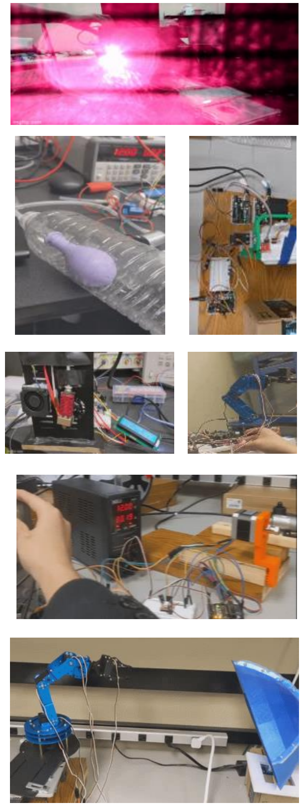
Acknowledgment

- Collaborators:
  - **Zhiqiang Lin** (Ohio State University)
  - **Insup Lee** (University of Pennsylvania)
  - **Kevin Fu** (Northeastern University)
  - **Zhongqi Pan** (Electrical Engineering, University of Louisiana at Lafayette)
  - **Sara Rampazzi, Kevin Butler** (University of Florida)

- Students and visiting scholars:
  - **Yazhou Tu, Md Imran Hossen**, **Bin Hao**, **Vijay Srinivas Tida**, **Liqun Shan**, **Md Fazle Rabby**, **Yuan Ping**, **Jianyi Zhang**
- **NSF**

Summary

- **Threat models** and **computations** of CPS/IoT should NOT assume an ideal, imaginary physical world (**Illusion**)
  - Side channels and out-of-band signal injections (EMI, Acoustic, Light)
  - Unexpected security, privacy, and safety risks (**Assumptions** broken)
- Identifying, detecting and **correcting** the attacks
  - Considering the underlying physics
- **Validating** security and reliability
  - Testbeds, prototype CPS systems, education



Testbeds for **CPS security research** and **education**

Demos: https://youtu.be/Y1LLiyhCn9I
https://youtu.be/8BjvIbu4aJM

Email: xiali.hei@louisiana.edu