浙江大學
ZHEJIANG UNIVERSITY

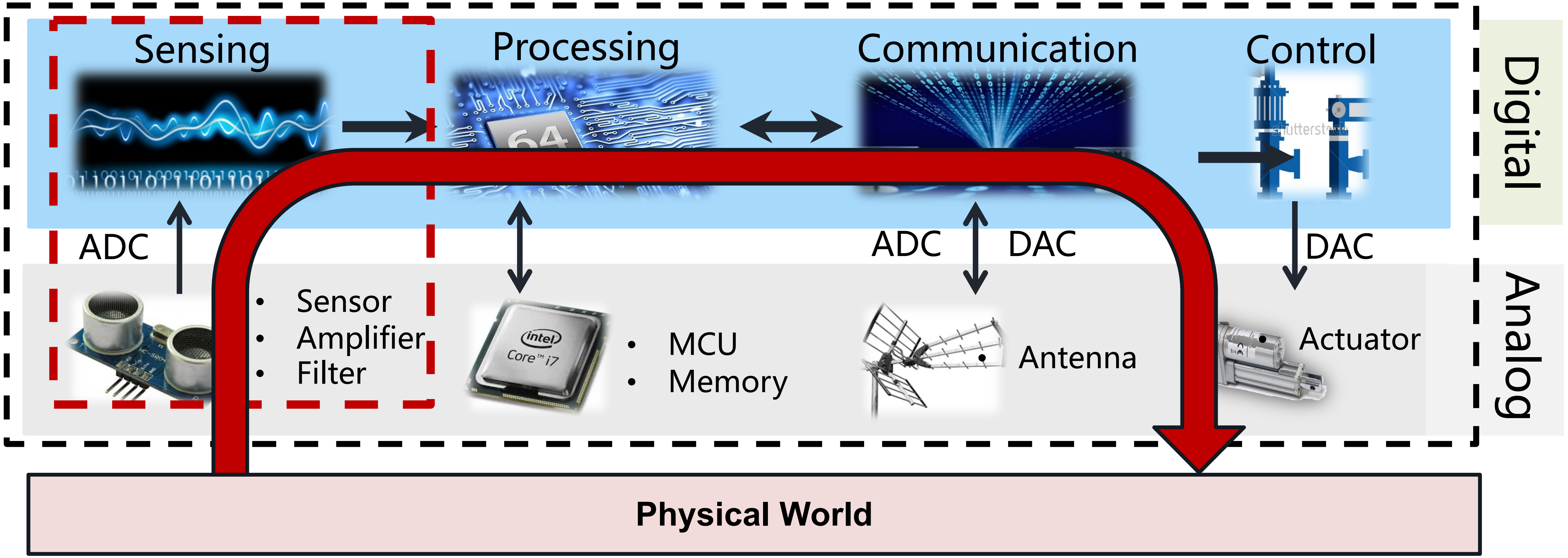
Analog Security of Cyber-Physical Systems: A Sound Story of *Sensors*

Chen Yan

USSLAB, Zhejiang University

智能系统安全实验室
UBIQUITOUS SYSTEM SECURITY LAB.

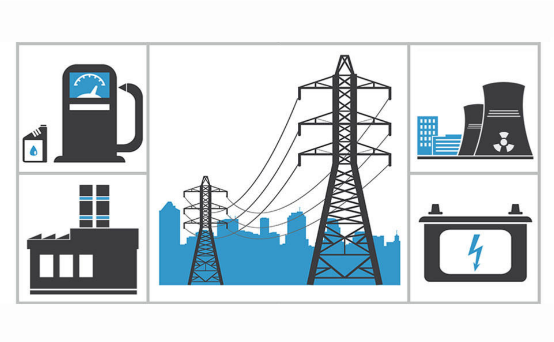
What's new? → Cyber - (Analog) - Physical



It's time to look at the physics of cybersecurity!

Sensors are everywhere

- Smartphone: >14 sensors
- Car: 60-100 sensors now;
  ~200 in the future
- Aircrafts
- Medical devices
- Home and office appliances
- Security
- Power grids
- Industrial plants
- Transportation

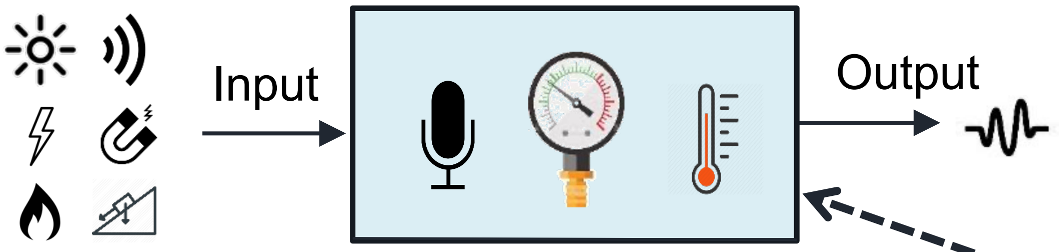


Questioning the trustworthiness of sensors

- Will sensors malfunction under malicious attacks?
- How will the system behave when sensors go wrong?
- How to make sensor measurements trustworthy?

What is inside a sensor module

Transducer



Physical signals -> Transducer -> Electrical signal

- Electromagnetic -> electrical [Electromagnetic induction]
- Mechanical -> electrical [Piezoelectricity]
- Radiant -> electrical [Photoconductivity]
- Magnetic -> electrical [Hall effect]
- Thermal -> electrical [Seebeck effect]
- Chemical -> electrical [Voltaic effect]

Types of Transducer

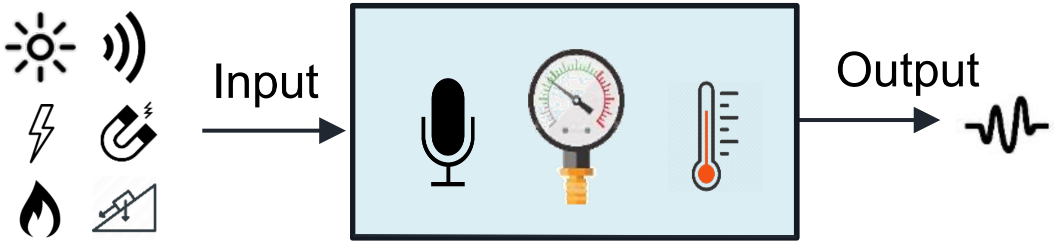
At least **13** types:

Acoustic	Radiation
Electromagnetic	Pressure
Optical	Force
Attitude	Chemical
Thermal	Flow
Humidity	Proximity
Navigation	

Reference: https://en.wikipedia.org/wiki/List_of_sensors

What is inside a sensor module

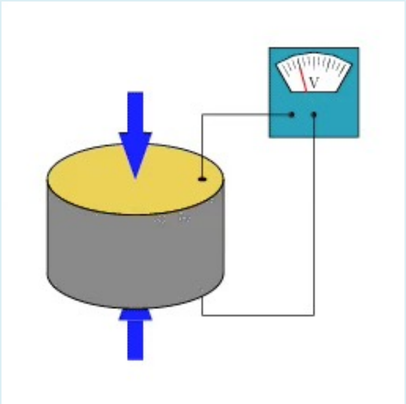
Transducer



Physical signals -> Transducer -> Electrical signal

- Electromagnetic -> electrical [Electromagnetic induction]
- Mechanical -> electrical [Piezoelectricity]
- Radiant -> electrical [Photoconductivity]
- Magnetic -> electrical [Hall effect]
- Thermal -> electrical [Seebeck effect]
- Chemical -> electrical [Voltaic effect]

Piezoelectric Transducer



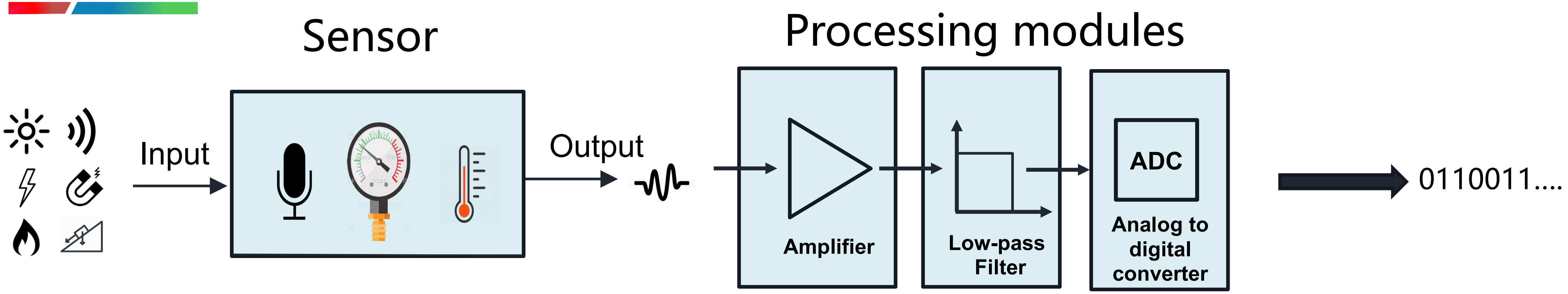
Pressure, acceleration, temperature, strain, or force

Strain-charge equations

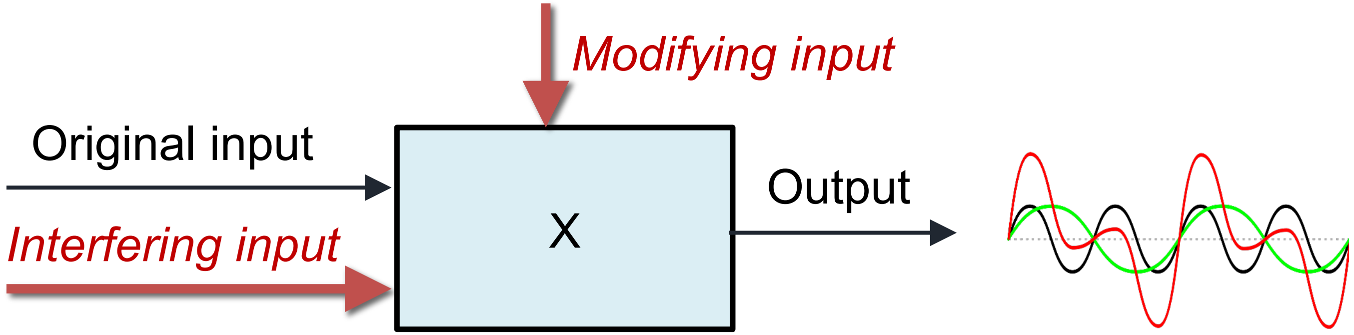
$$S = sT + \delta^t E$$
$$D = \delta T + \varepsilon E$$

Sensors are vulnerable to transduction attack

How do transduction attacks work?



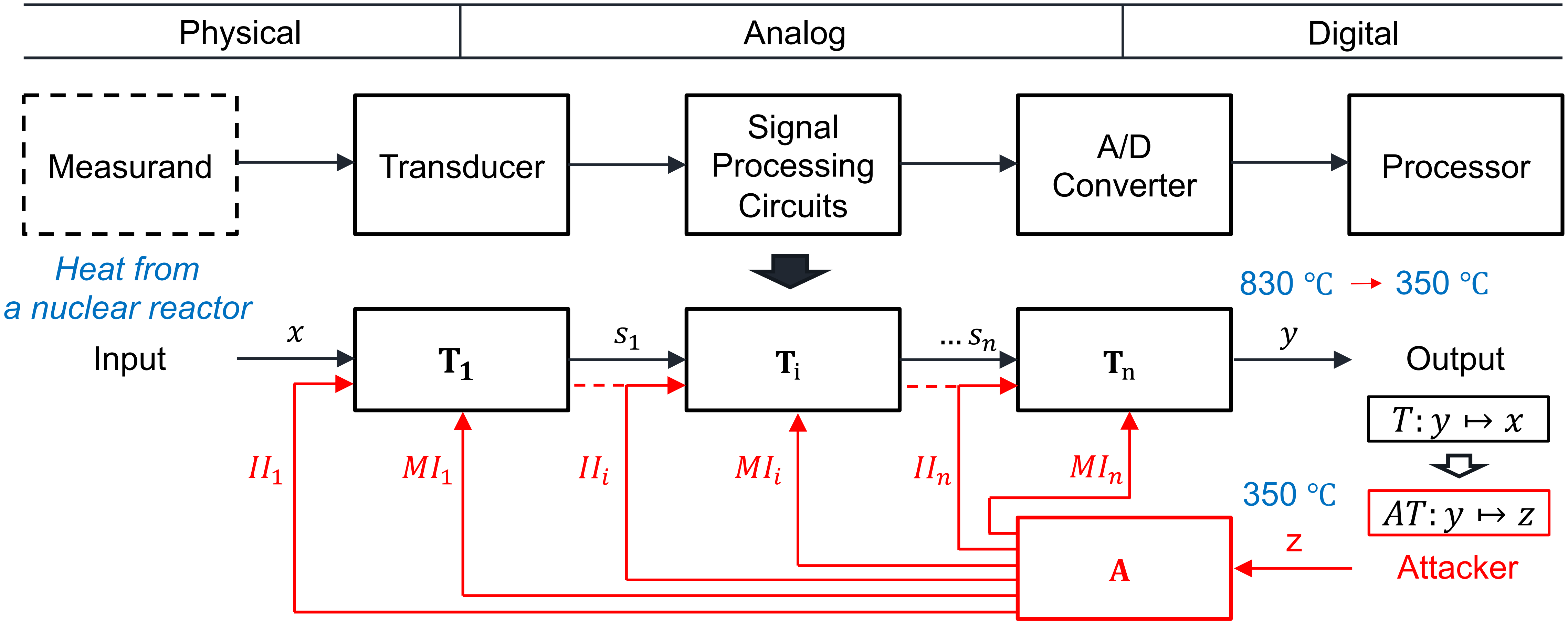
Physical signals -> Sensor -> Electrical signal



Interfering inputs (II): those that are treated as linear superposition with the original inputs, e.g., noises.

Modifying inputs (MI): those that change the transfer functions, e.g., temperature.

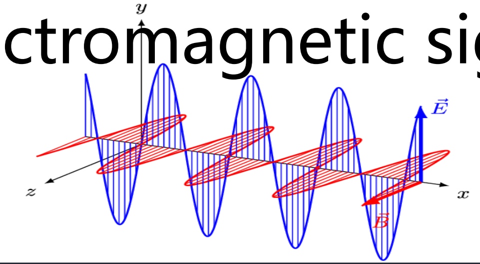
The framework of transduction attacks



Can we trust the sensor readings?

Malicious Signals

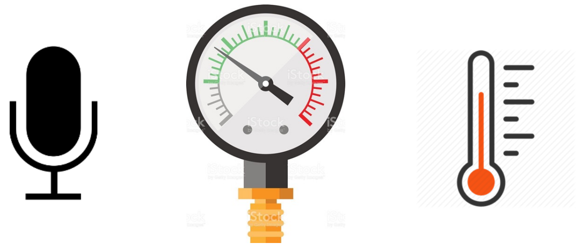
1. Electromagnetic signal

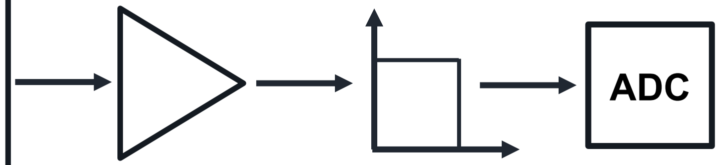


2. Acoustic signals

3. Light
4. Magnet
5. Heat

Sensor Modules

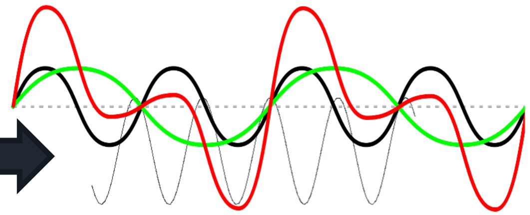




Amplifier

Low-pass Filter

Analog to digital converter



0110011....

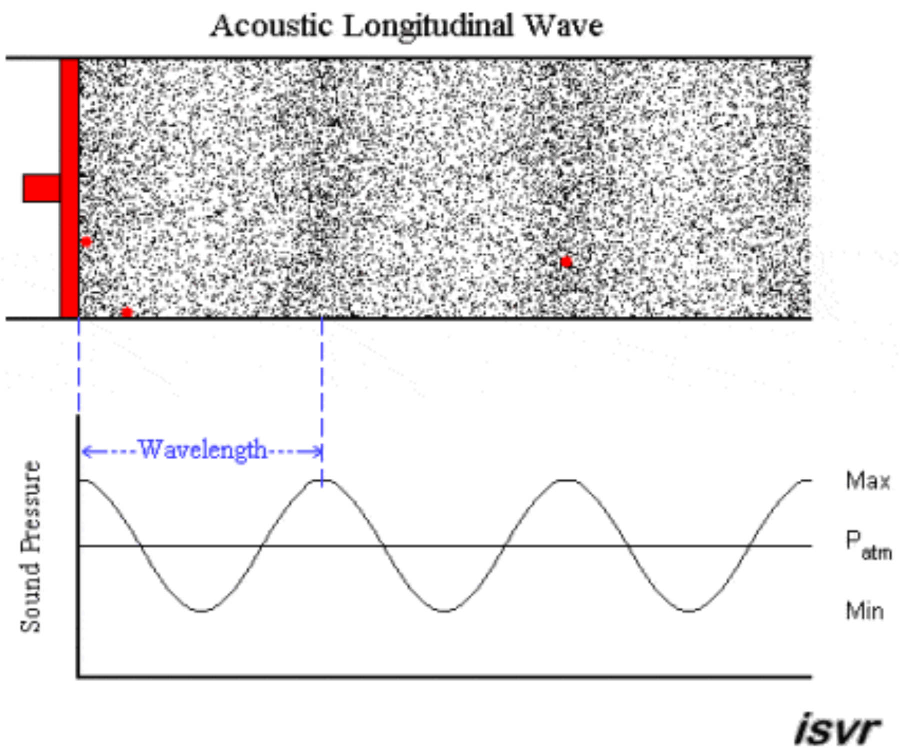
Roadmap: a sound story

- **Sensors on autonomous vehicles**---TeslaHacking
- **Microphone and voice assistants**---DolphinAttack
- **Cameras+AI**---Poltergeist
- **Human ears**---Cuba event

What is a sound wave ?

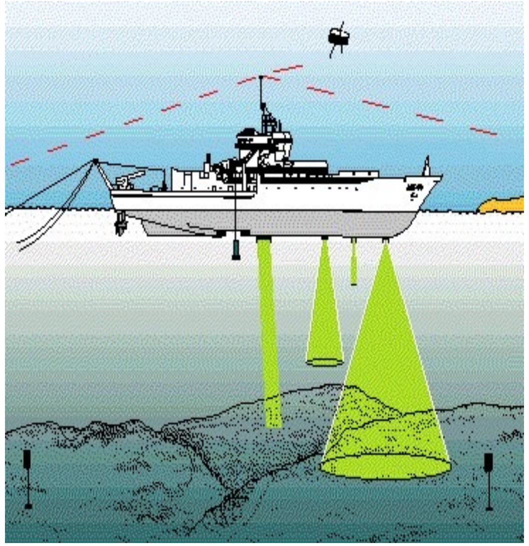
A pressure wave that fluctuates up and down around normal pressure

- **Hz:** the frequency
- **dB:** the intensity



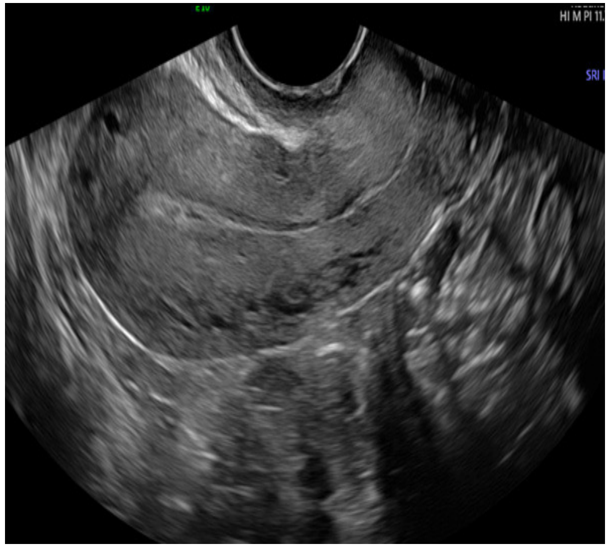
Where are sound waves used ?

Military



Sonar

Medicine



Ultrasonic detector

Industrial processes

Ultrasonic thickness gauge

Daily

Ultrasonic Cleaner

SENSORS ON AUTONOMOUS VEHICLES


ANDY GREENBERG SECURITY 08.04.16 09:00 AM

HACKERS FOOL TESLA S'S AUTOPILOT TO HIDE AND SPOOF OBSTACLES

Analyzing and Enhancing the Security of Ultrasonic Sensors for Autonomous Vehicles

Wenyuan Xu, Chen Yan, Weibin Jia, Xiaoyu Ji, Jianhao Liu

IEEE Internet of Things Journal 5 (6), 5015-5029

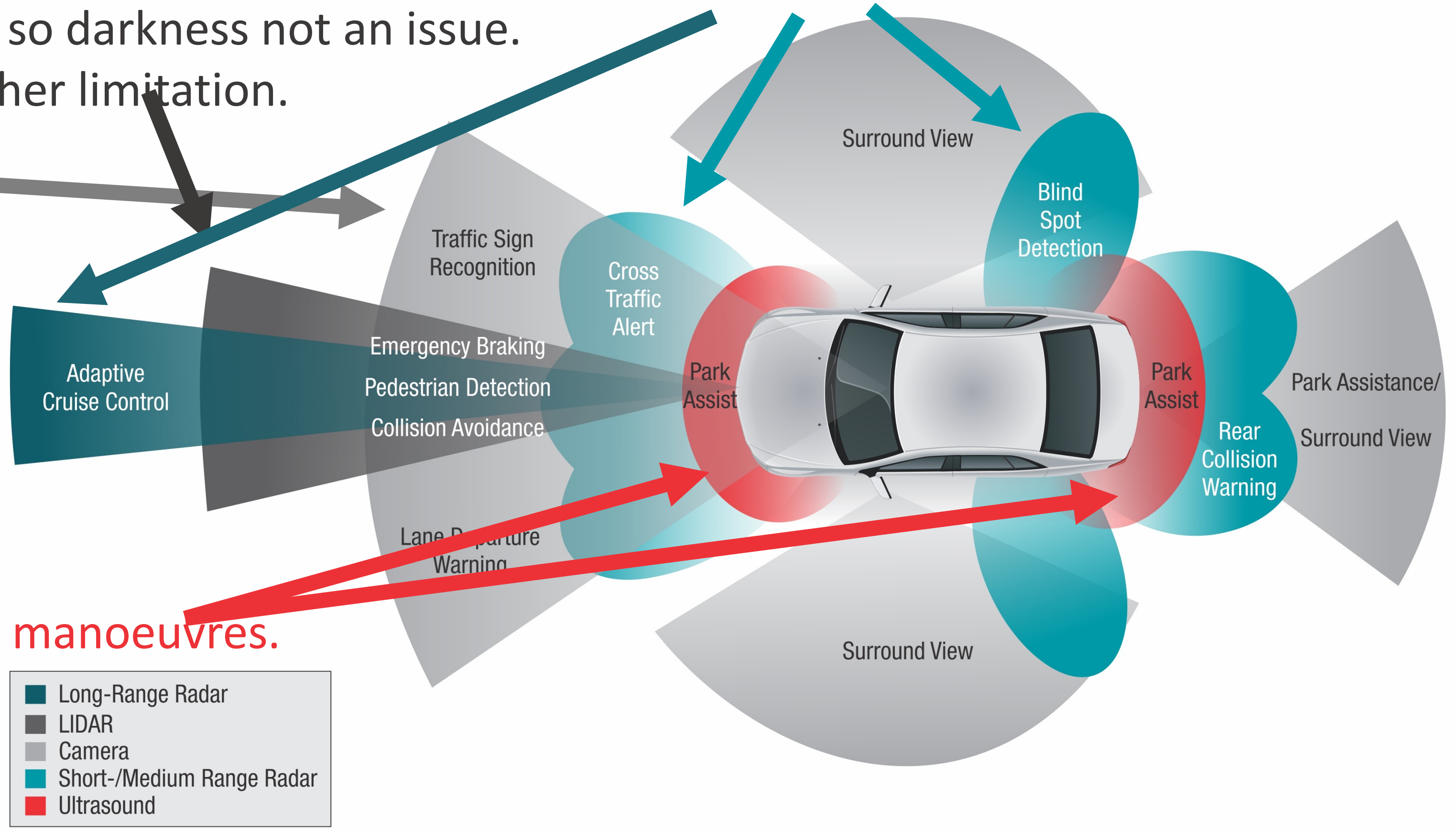
Sensors for Self-Driving

Radar
Works in low light & poor weather, but lower resolution.

LiDAR
Emits light, so darkness not an issue.
Some weather limitation.

Cameras
Senses reflected light, limited when dark.
Sees colour, so can be used to read signs, signals, etc.

Ultrasonic Sensors
Limited to proximity, low speed manoeuvres.



Source: Texas Instruments

What will happen if sensors go wrong?

Sensors

Ultrasonic Sensors

Jamming

Spoofing

MMW Radars

Jamming

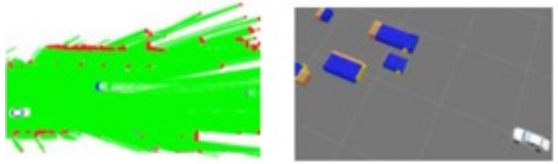
Spoofing

Cameras

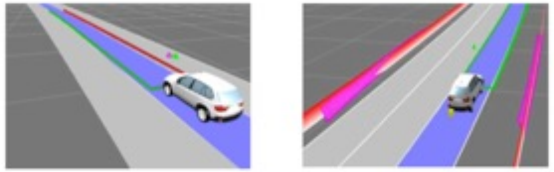
Blinding

Automated System

Representations and Fusion



Road Model and Localization



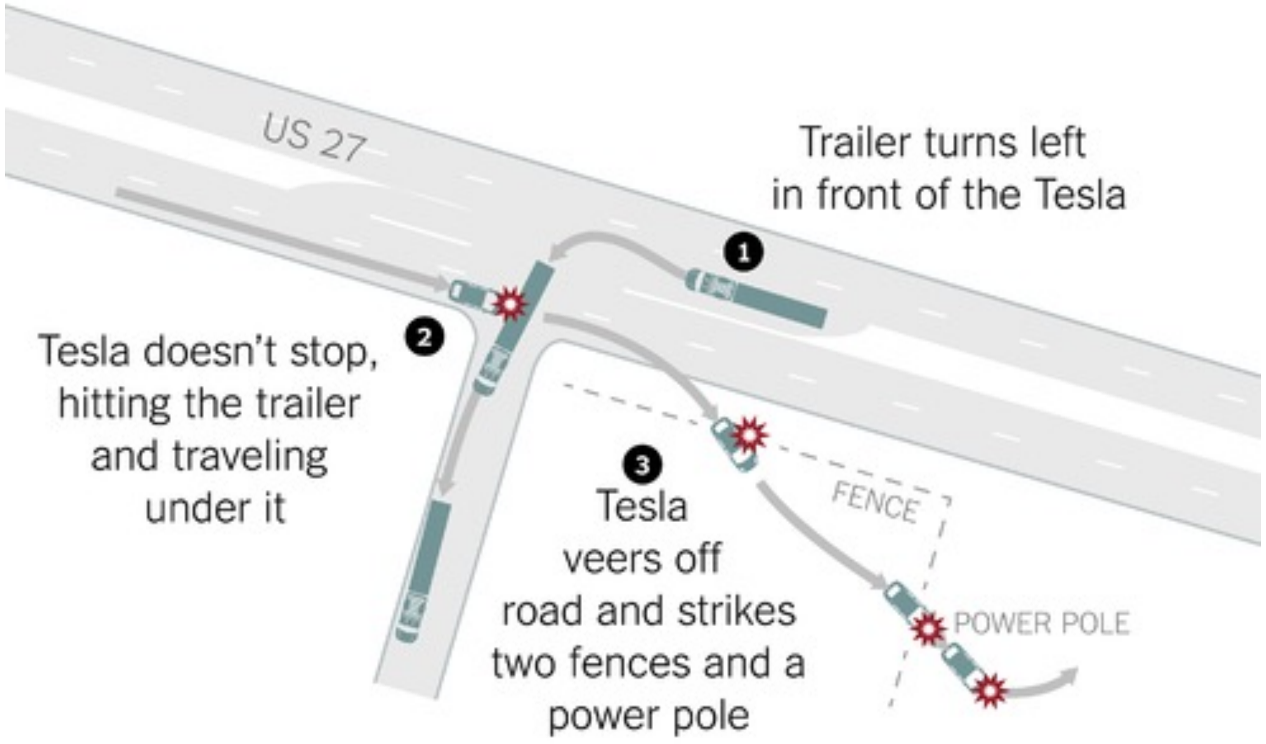
Situation Interpretation

Control

HMI Display

Tesla: A Tragic Loss

- First fatal crash while using Autopilot on May 7, 2016.
- Reliability of sensors.



Source: The New York Times

First Tesla Accident in China Caused by Autopilot

国内发生特斯拉第一起自动驾驶事故

2016-08-05 11:21:06 来源: 盖世汽车(上海)

Existing Sensors on Tesla Model S

One MMW Radar

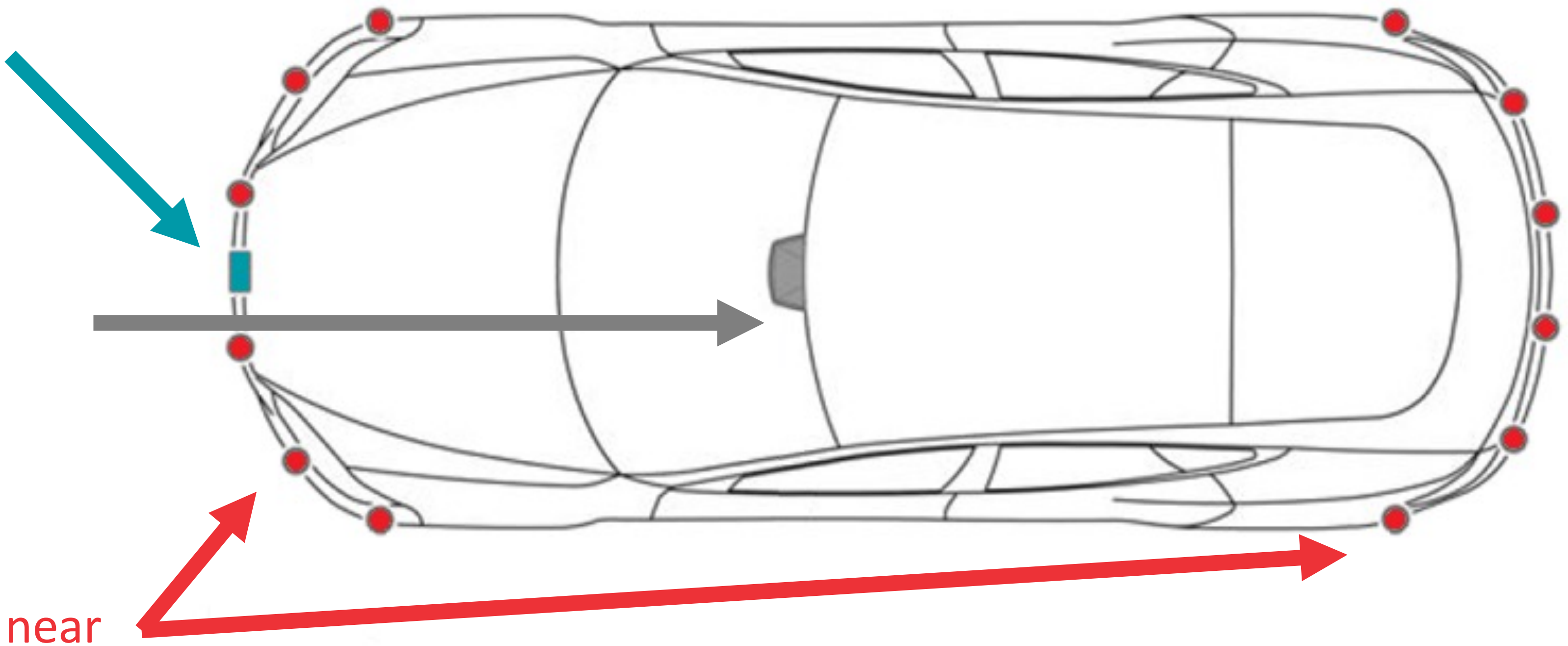
A Medium range Radar is mounted in the front grill.

One camera

A forward looking camera is mounted on the windshield under the rear view mirror.

12 ultrasonic sensors

Ultrasonic sensors are located near the front and rear bumpers.



HMI Display Mistakes – Demo on Tesla

Control Mistakes – Demo on Tesla

Attacking Ultrasonic Sensors

On Tesla, Audi, Volkswagen, and Ford

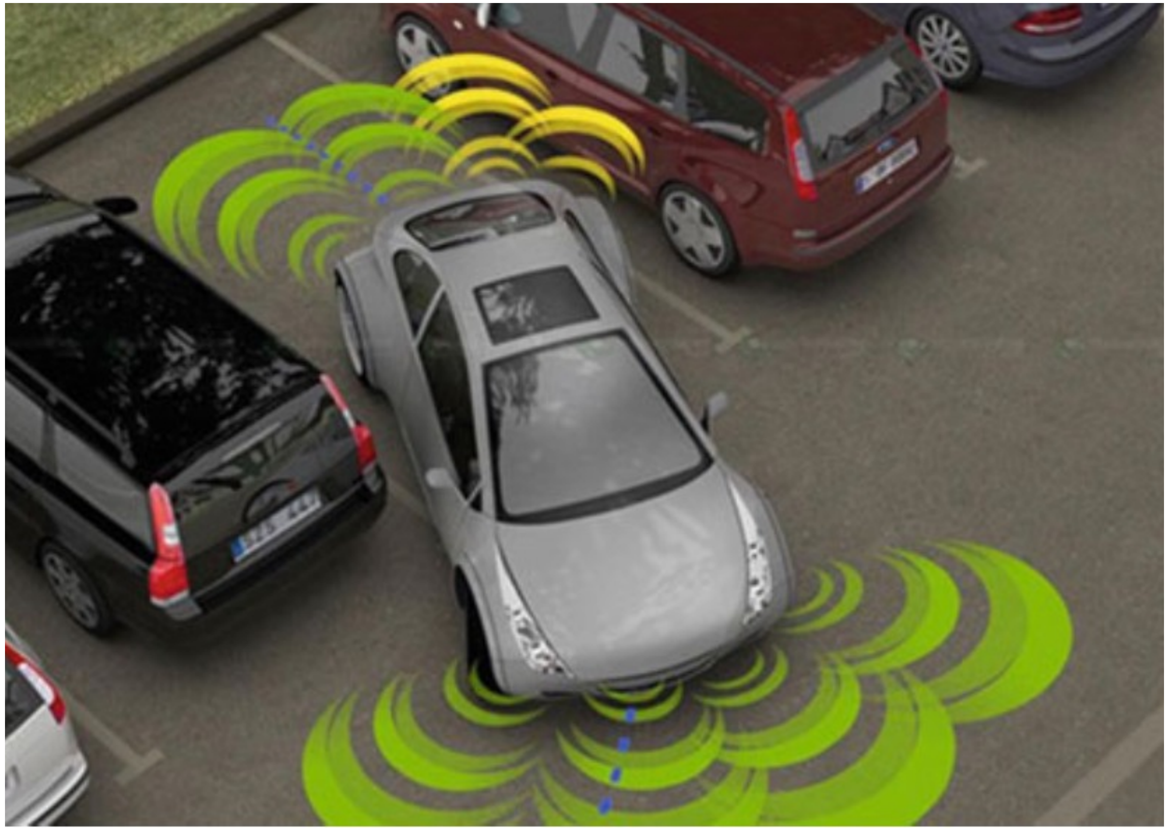
Ultrasonic Sensor

What is ultrasonic sensor?

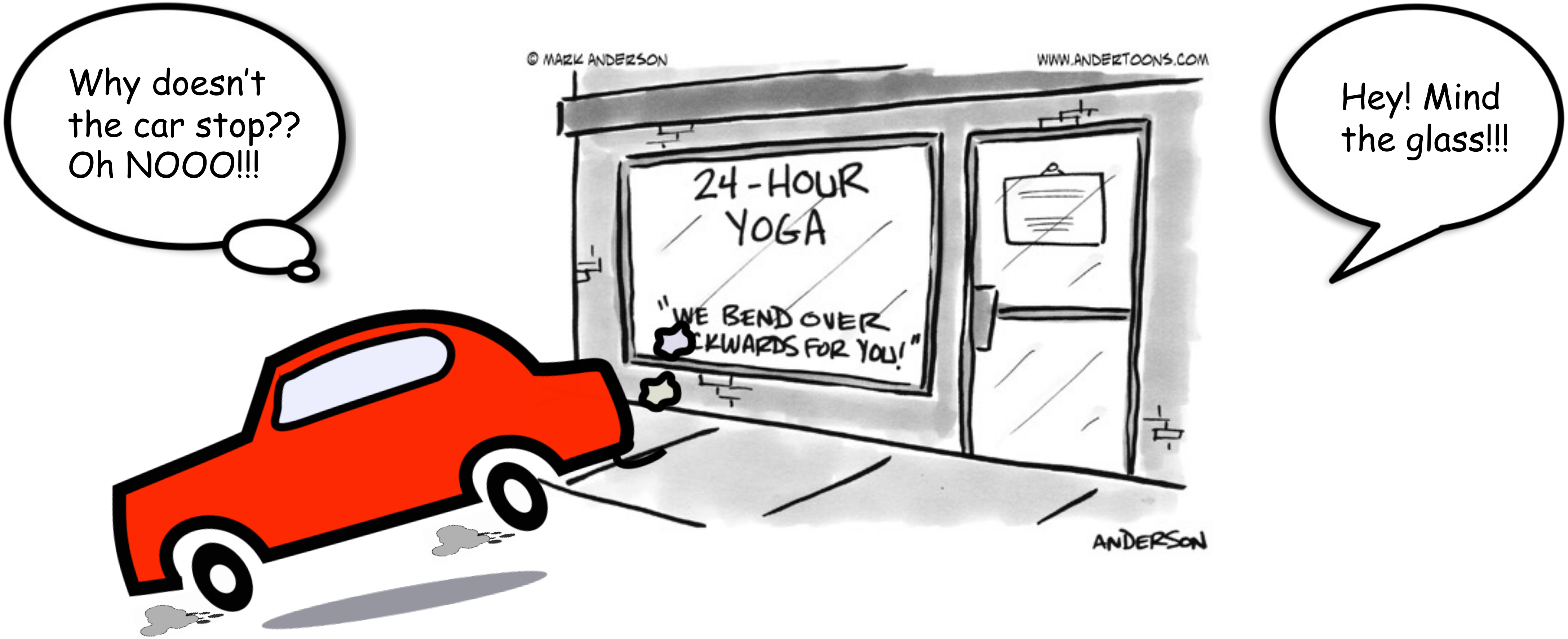
- **Measures distance**
- **Proximity sensor (< 2m)**

- **Applications**
  - Parking assistance
  - Parking space detection
  - Self parking
  - Tesla's summon

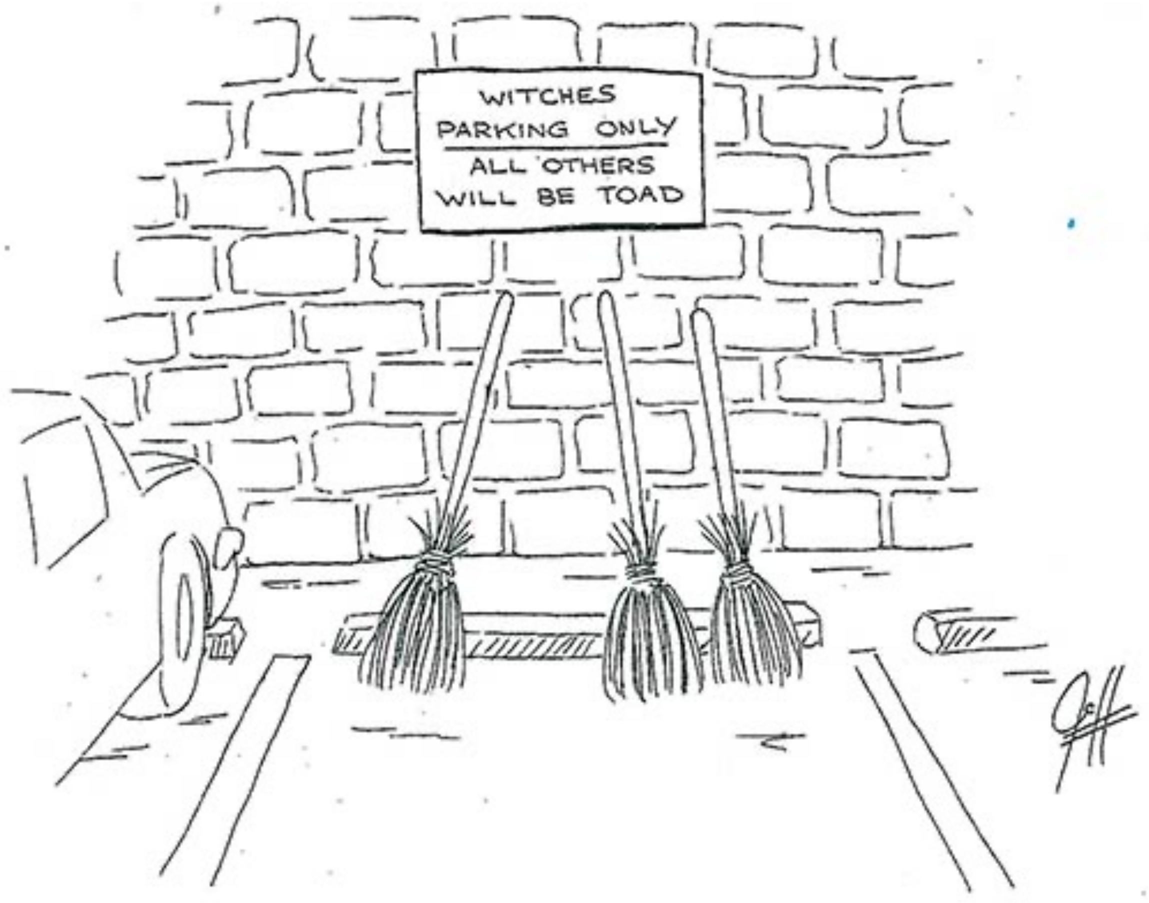
Self-Parking & Distance display



Misuse 1: The car doesn't stop while it should.



Misuse 2: The car stops while it shouldn't.

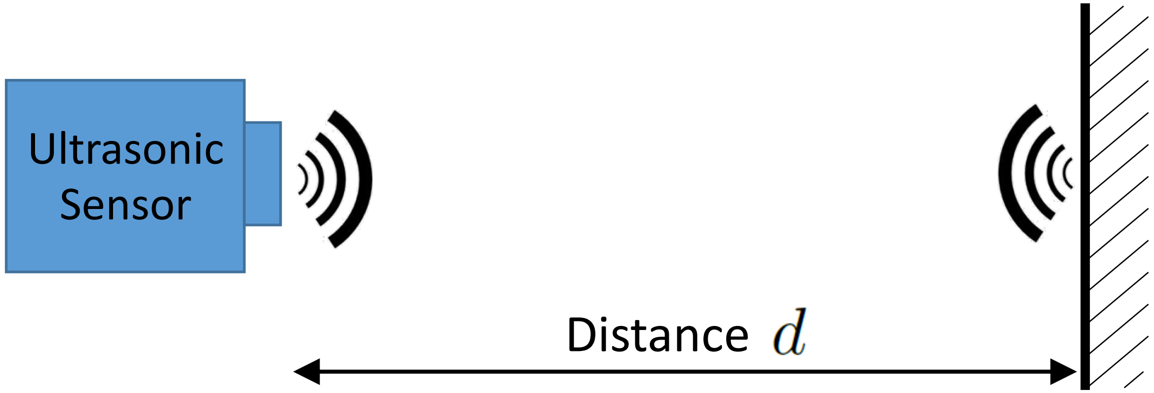


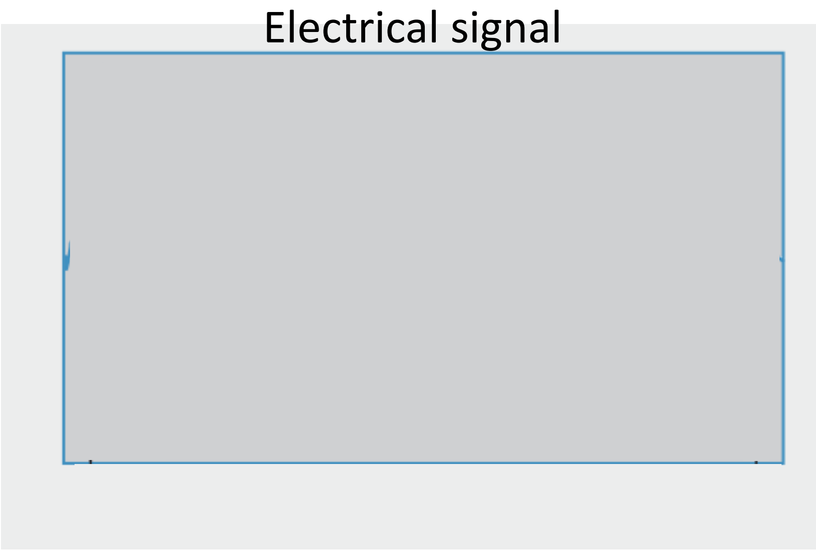
How do ultrasonic sensors work?

- **Emit ultrasound and receive echoes**
- **Piezoelectric Effect**
- **Measure the propagation time (Time of Flight)**
- **Calculate the distance** $d = 0.5 \cdot t_e \cdot c$

t_e : propagation time of echoes

c : velocity of sound in air





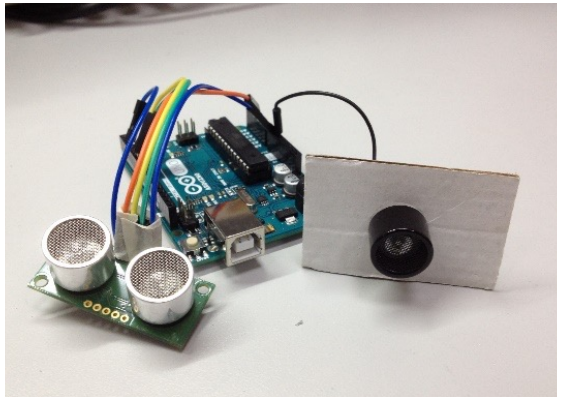
Attacking ultrasonic sensors

Attacks:

- **Jamming – generates ultrasonic noises – denial of service**
- **Spoofing – crafts fake ultrasonic echo pulses – alters distance**
- **Quieting – diminishes original ultrasonic echoes – hides obstacles**

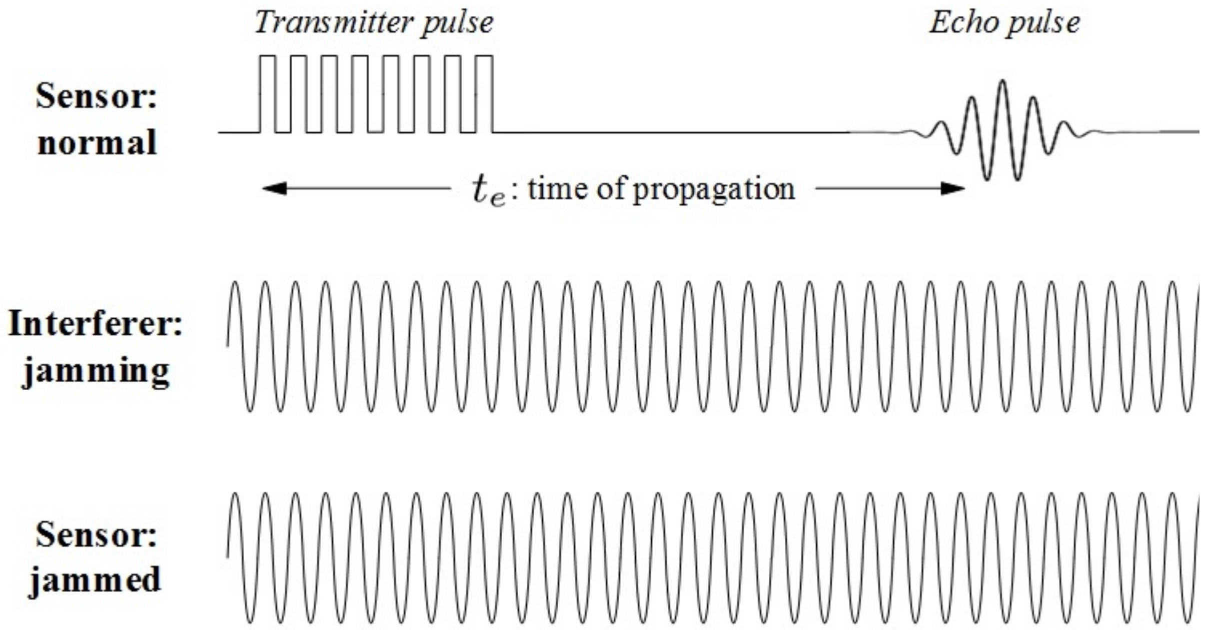
Equipment:

- **Ultrasonic transducers ($0.4) – emit ultrasound**
- **Signal suppliers – generate excitation signals**
  - Arduino ($24.95)
  - Signal generator (~$20)



Jamming Attack

- **Basic Idea:**
  - Injecting ultrasonic noises
  - At resonant frequency (40 – 50 kHz)
  - Causing Denial of Service



- **Tested ultrasonic sensors:**
  - In laboratories: 8 models of stand-alone ultrasonic sensors
  - Outdoors: Tesla, Audi, Volkswagen, Ford

Jamming Attack – in lab

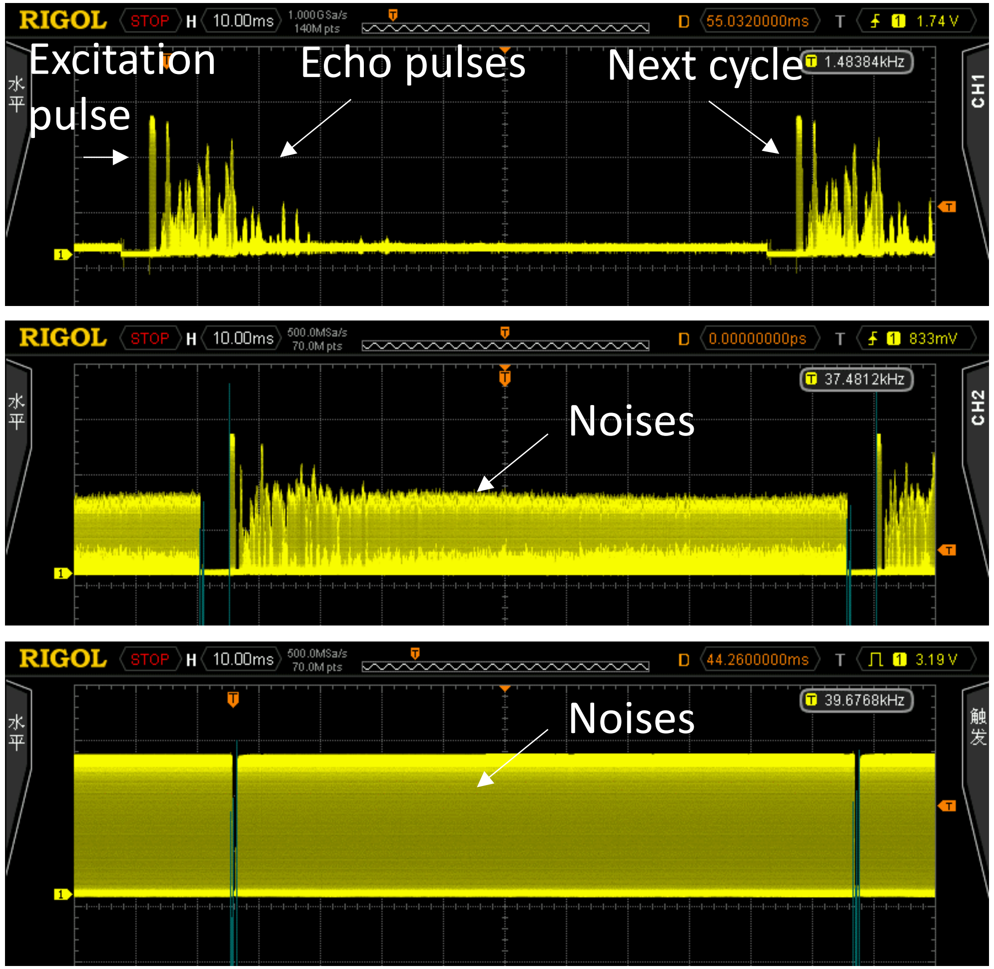
- **8 models of ultrasonic sensors**
  - HC-SR04
  - SRF01
  - SRF05
  - MaxSonar MB1200
  - JSN-SR04T
  - FreeCars V4
  - Grove ultrasonic ranger
  - Audi Q3 sensors
- **Sensor reading**
  - **Zero** distance
  - **Maximum** distance

Received electrical signals at the sensor

No jamming

Weak Jamming

Strong Jamming



How should cars behave to jamming?

Zero distance?

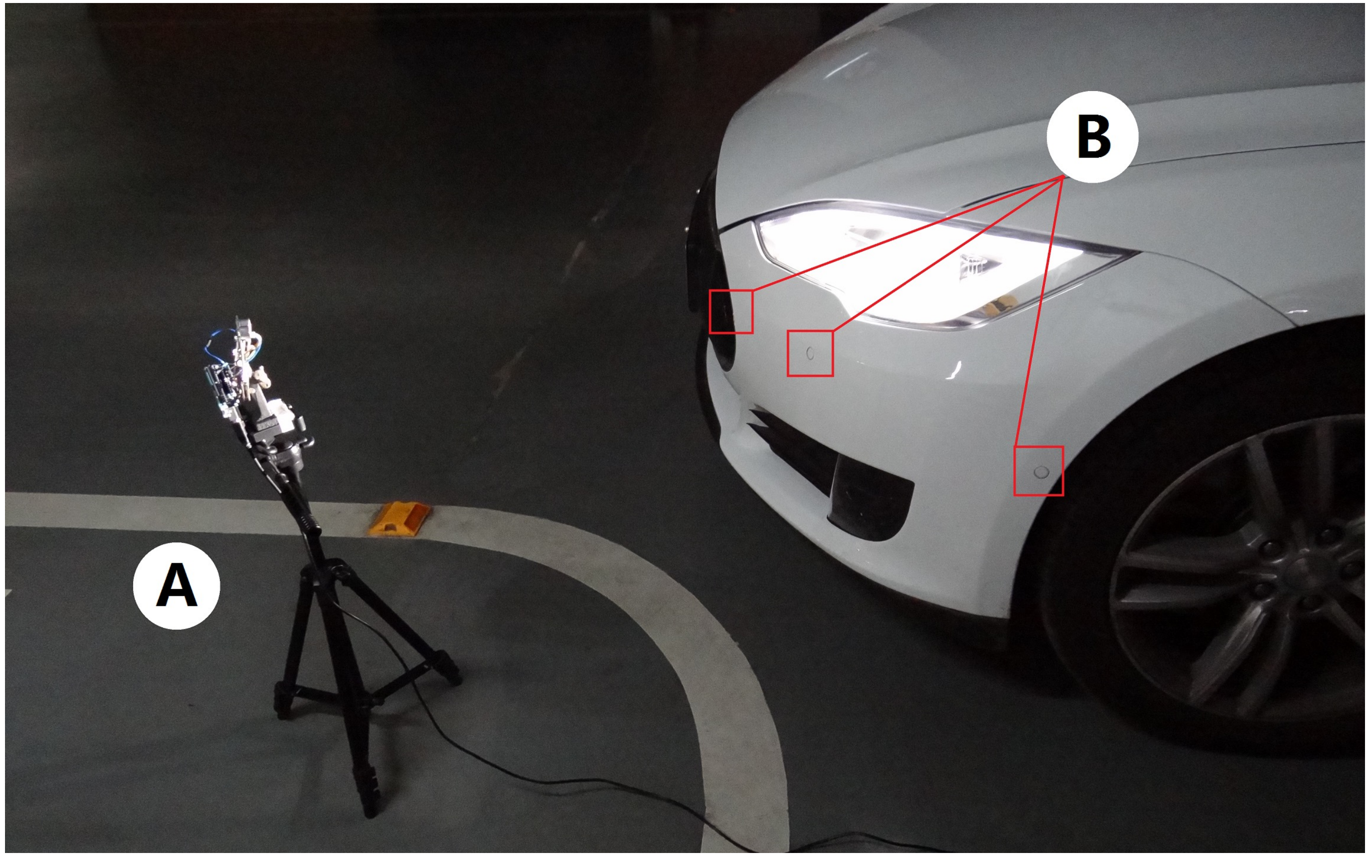
or

Maximum distance?

Jamming Attack – on vehicles

- **4 different vehicles**
  - Audi Q3
  - Volkswagen Tiguan
  - Ford Fiesta
  - Tesla Model S
    - Self parking
    - Summon

- **Results**
  - **Maximum** distance



Experiment setup on Tesla Model S

Jamming Attack – Demo on Audi

Jamming Attack – Results

- **On ultrasonic sensors**
  - Zero or maximum distance
- **On vehicles with parking assistance**
  - Maximum distance
- **On self-parking and summon?**

Note: If a sensor is unable to provide feedback, the instrument panel displays an alert message.

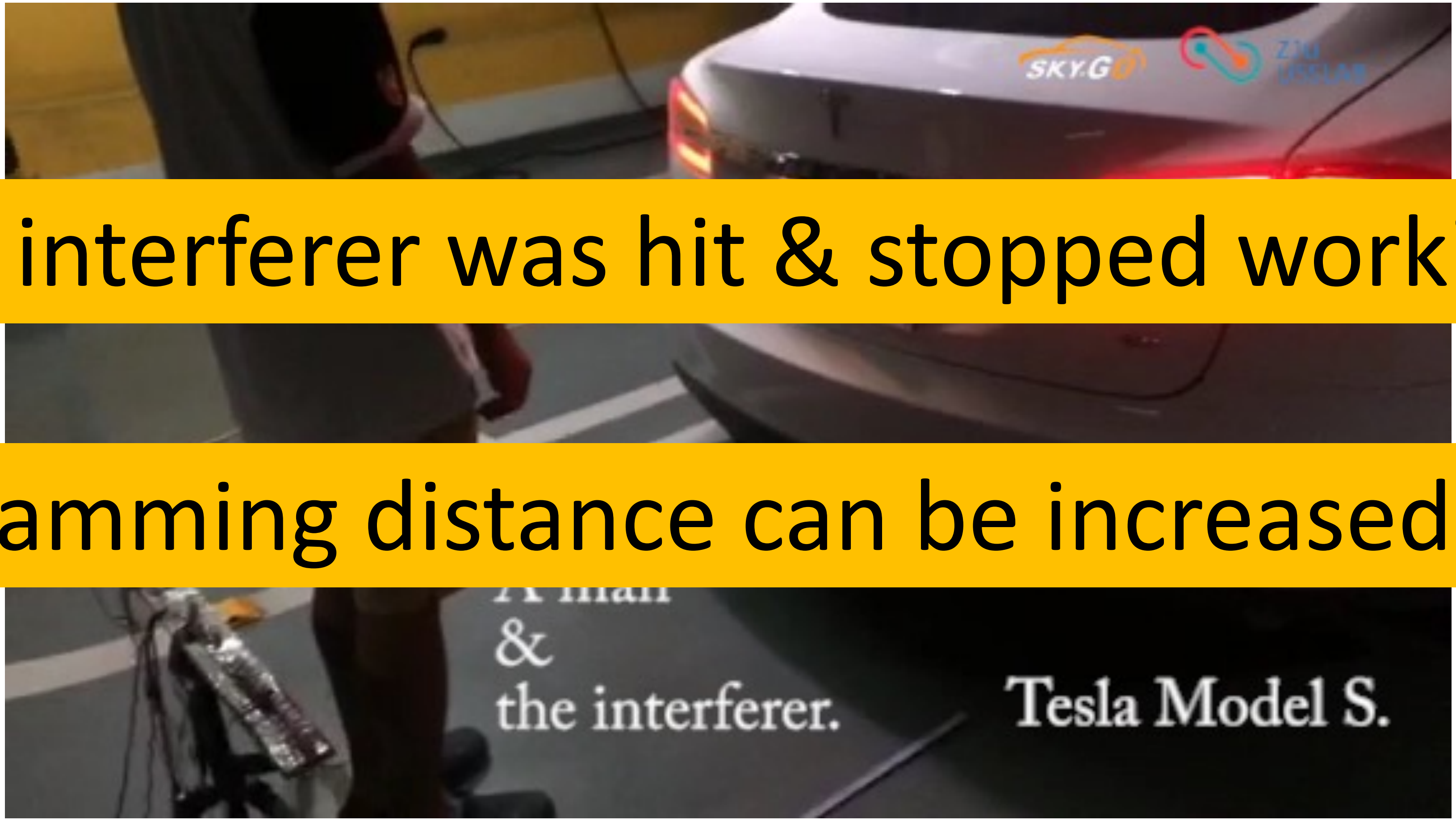
Audi Normal

Audi Jammed

Tesla Normal

Tesla Jammed

Jamming Attack – Demo on Tesla Summon



The interferer was hit & stopped working.

Jamming distance can be increased.

Jamming Attack – Results

- On ultrasonic sensors
  - Zero or maximum distance
- On vehicles with parking assistance
  - Maximum distance
- **On self-parking and summon**
  - Car does not stop under strong jamming

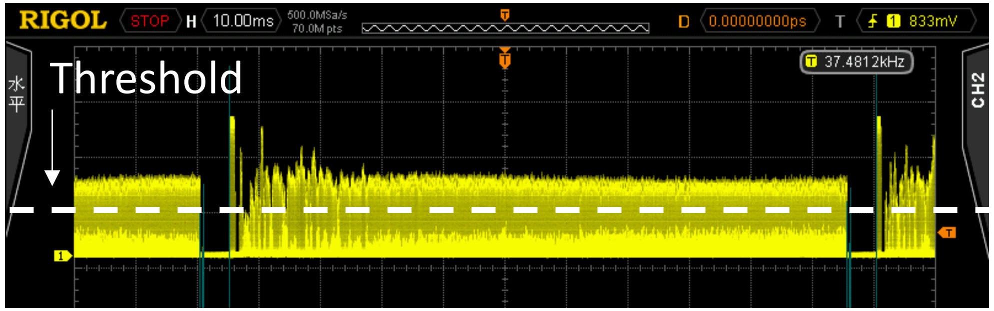
Audi Normal

Audi Jammed

Tesla Normal

Tesla Jammed

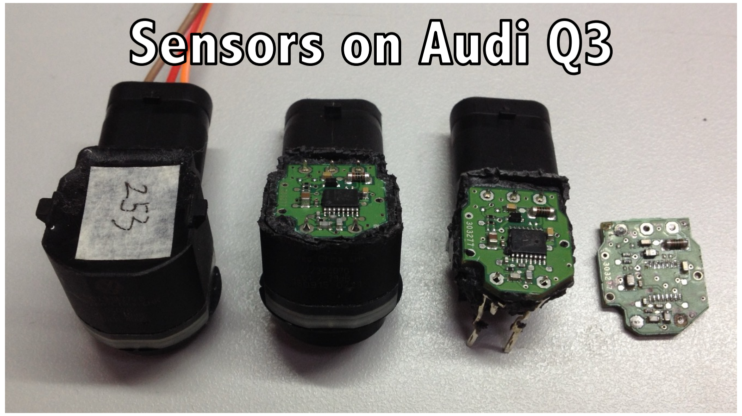
Why Zero or Max distance?

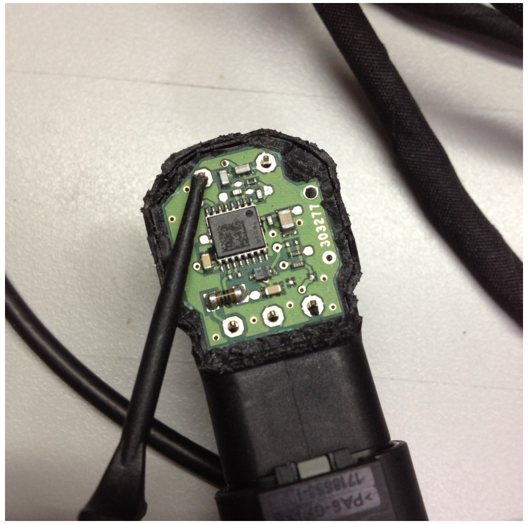


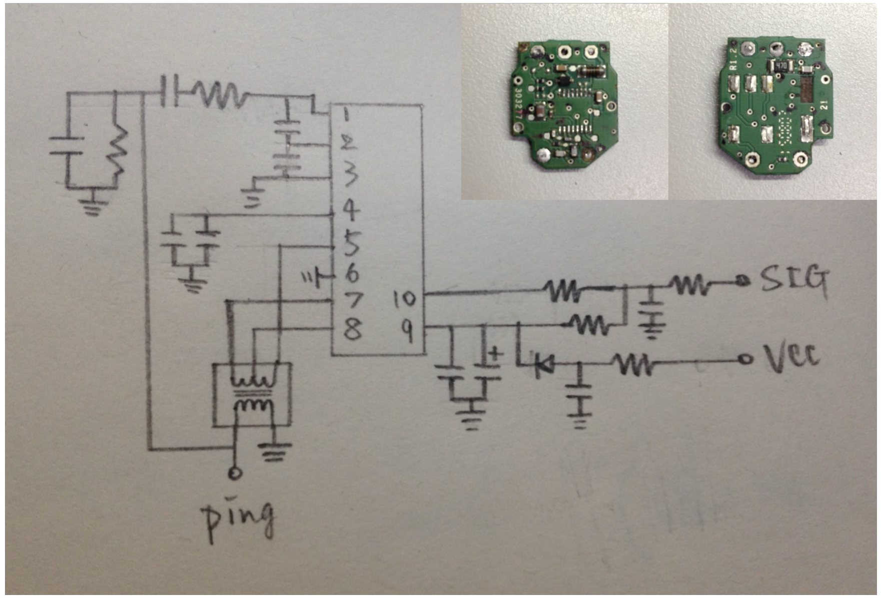
Different sensor designs

- **Zero distance**
  - Compare with a fixed threshold
- **Maximum distance**

  Application Specific IC!



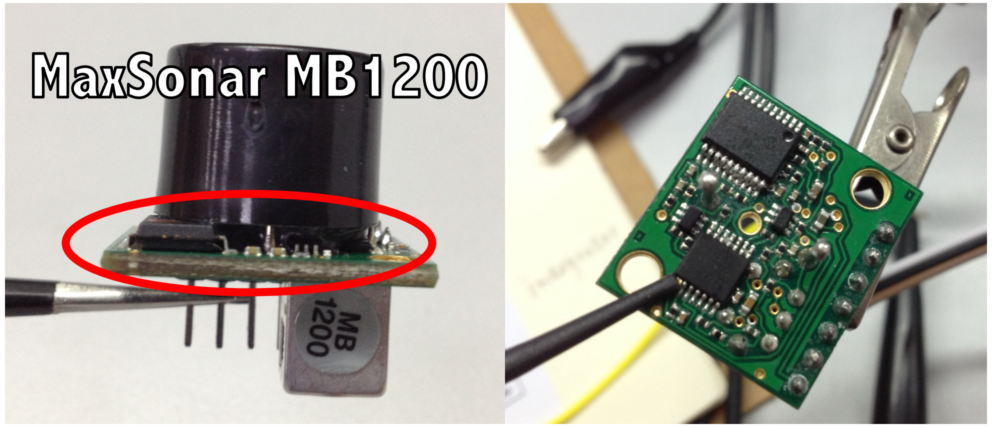




Why Zero or Max distance?

Different sensor designs

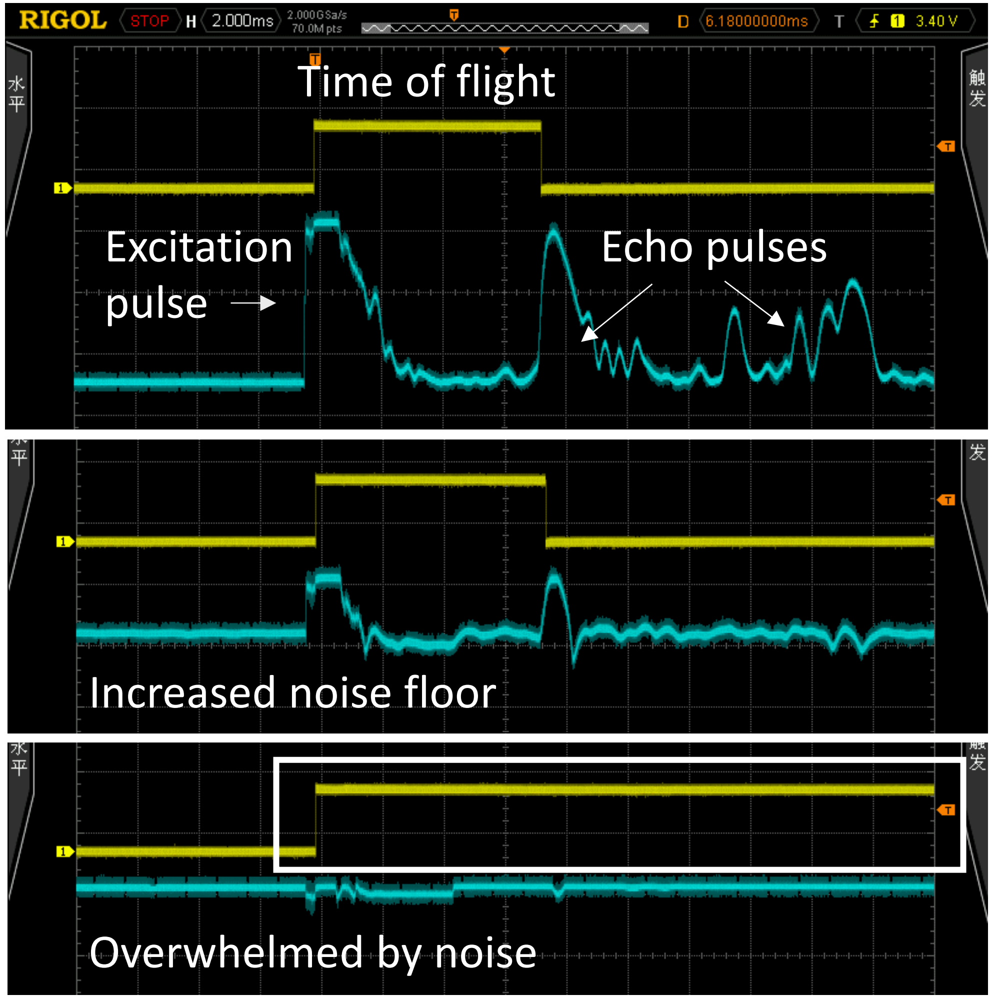
- **Zero distance**
  - Compare with a fixed threshold
- **Maximum distance**
  - Adaptive threshold (Noise Suppression)



No jamming

Weak Jamming

Strong Jamming



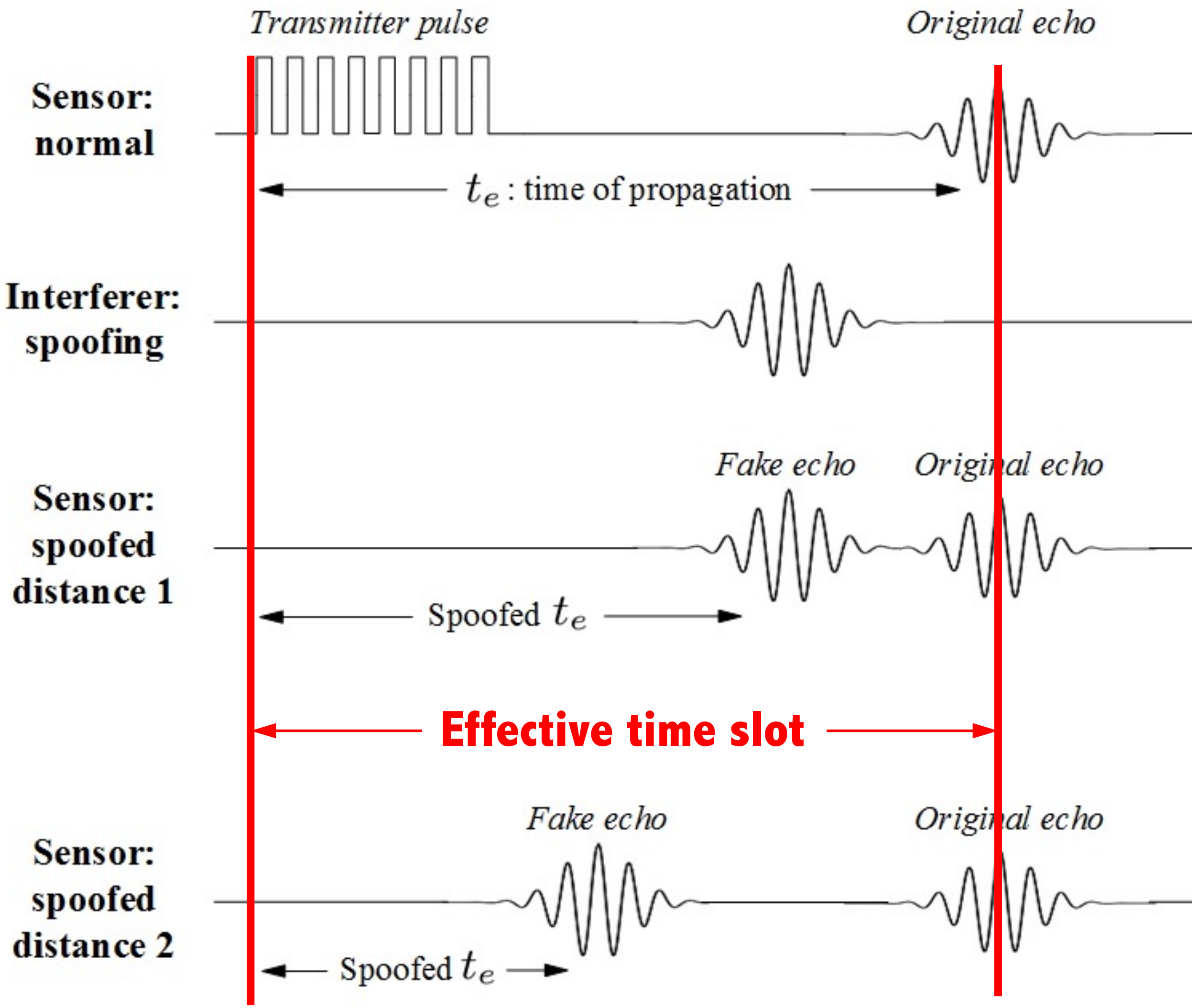
Spoofing Attack

Basic Idea

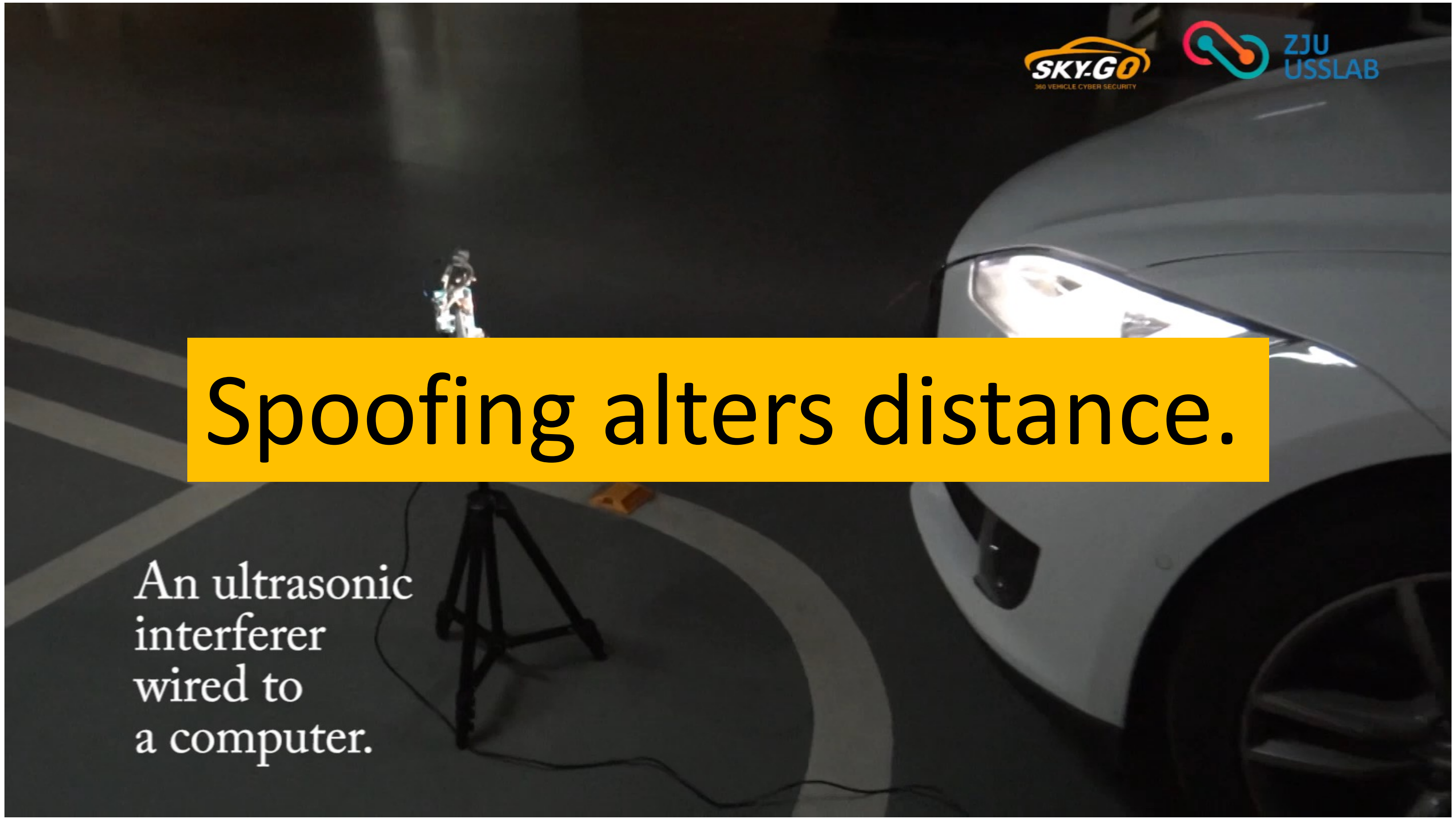
- Injecting ultrasonic pulses
- At certain time

Non-trivial

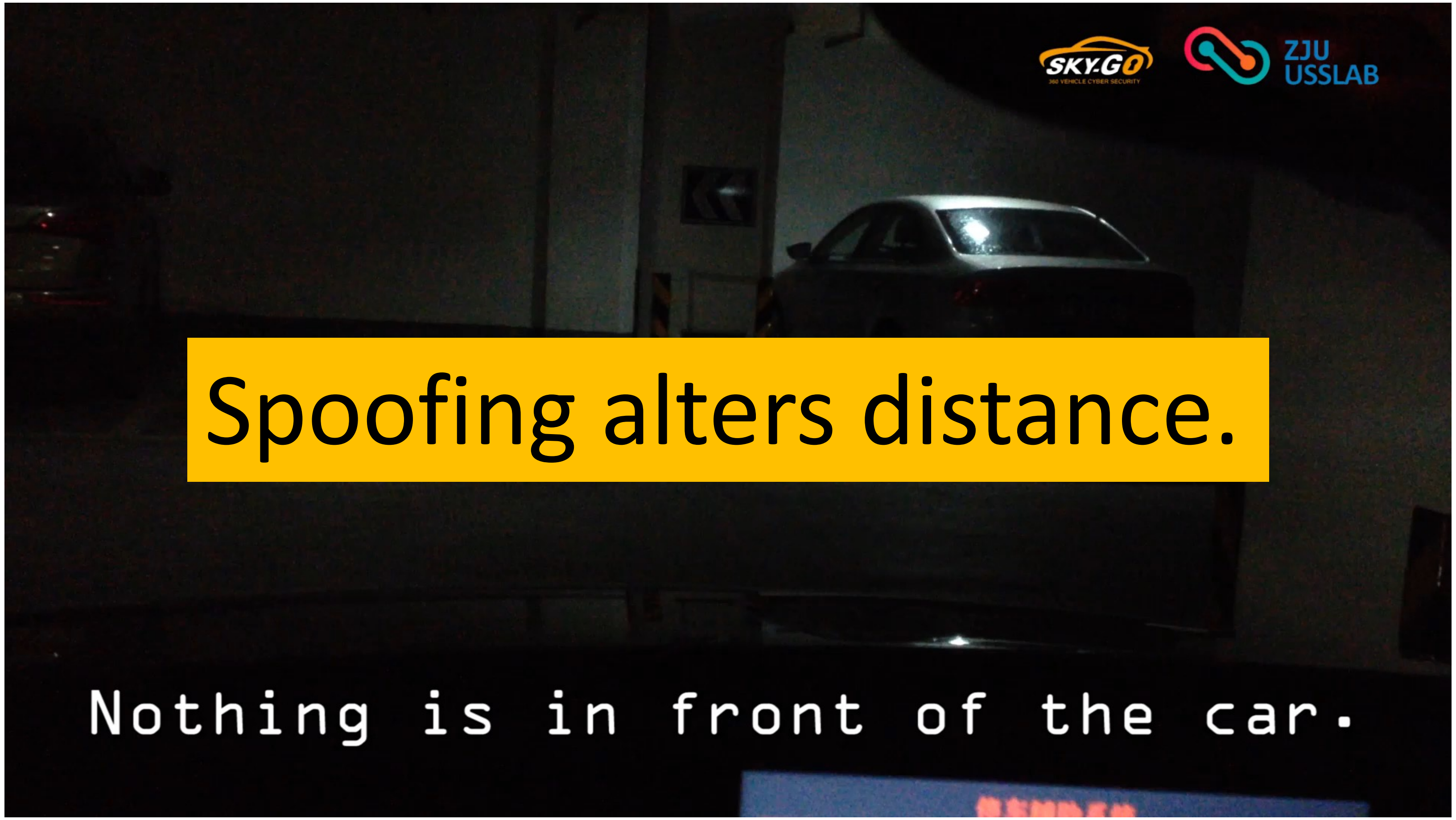
- Only the first justifiable echo will be processed
- **Effective time slot**



Spoofing Attack – Demo on Tesla



Spoofing Attack – Demo on Audi



Spoofing Attack – Results

- **Manipulate sensor readings**
  - On stand-alone ultrasonic sensors
  - On cars

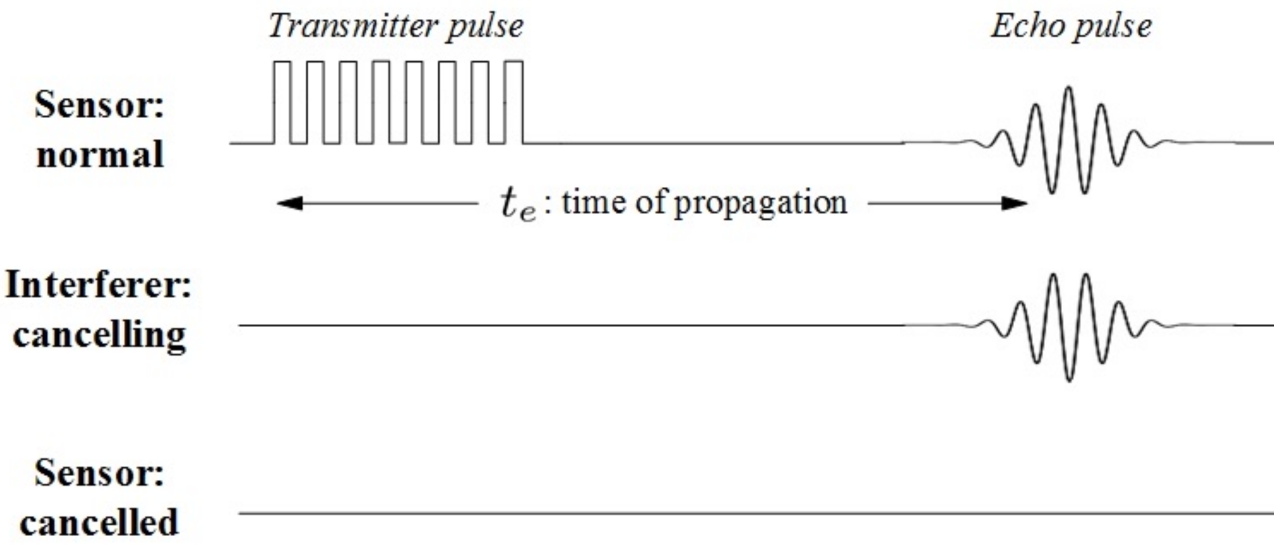
Tesla Normal

Tesla Spoofed

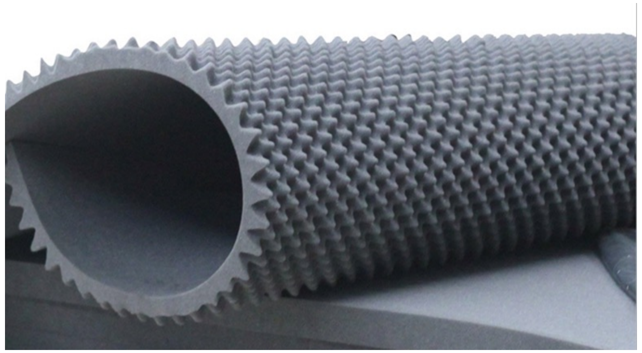
Audi Spoofed

Acoustic Quieting

- **Acoustic Cancellation**
  - Cancel original sound with ones of reversed phase
  - Minor phase and amplitude adjustment



- **Cloaking**
  - Sound absorbing materials (e.g., damping foams ($3/m^2))
  - Same effect as jamming!



Cloaking Car – Demo

Cloaking Human – Demo

Invisible car! Invisible man! Invisible glass! Whee!

浙江大学
ZHEJIANG UNIVERSITY

智能系统安全实验室
UBIQUITOUS SYSTEM SECURITY LAB.

The New York Times

Alexa and Siri Can Hear This Hidden Command. You Can't.

Researchers can now send secret audio instructions undetectable to the human ear to Apple's Siri, Amazon's Alexa and Google's Assistant.

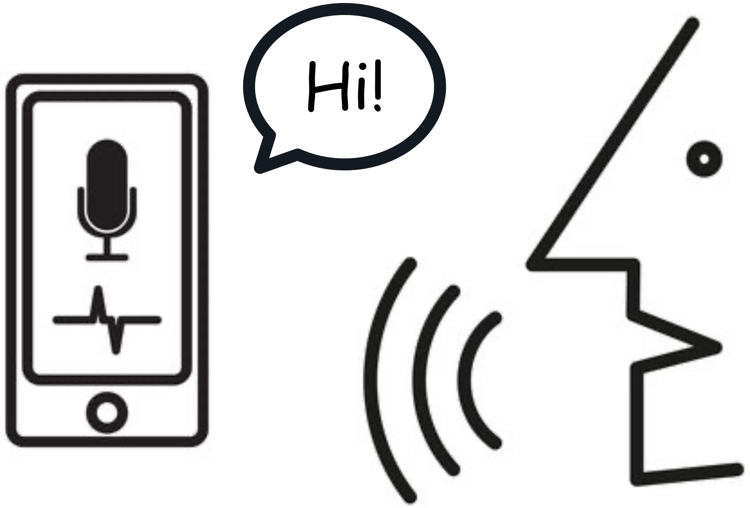
MICROPHONES & VOICE ASSISTANTS

DolphinAttack: Inaudible Voice Commands

Guoming Zhang, Chen Yan, Tiancheng Zhang, Taiming Zhang, Xiaoyu Ji, Wenyuan Xu

Best Paper at ACM CCS 2017

Voice becomes an increasingly important interface



Siri Google Now Alexa Cortana S Voice Hi Voice

Smartphone

Smart speaker

PC & Tablet

Smart watch

How do voice assistants work?

Sound *Electrical signal* *Command*

Voice Command Input

Speech Recognition

Siri
Goolge Now
Alexa
Cortana
S Voice
Hi Voice

Command Execution

What can voice assistants do?

What's on my calendar today?

Open google.com

Tell my wife I love her

Send an email to my boss

Open the front door

Buy something on Amazon

Transfer $100 to Alice

Call 1234567890

Facetime Bob

Drive me to Dallas

What can a malicious user achieve?

What's on my calendar today?

Sensitive information

Open evil.com

Malicious website

Tell my wife I love her

Fake message

Send an email to my boss

Social engineering

Open the front door

Break-in

Buy something on Amazon

Lose money

Transfer $100 to Eve

Steal money

Call 1234567890

Eavesdrop

Facetime Eve

Spy

Drive me to Austin

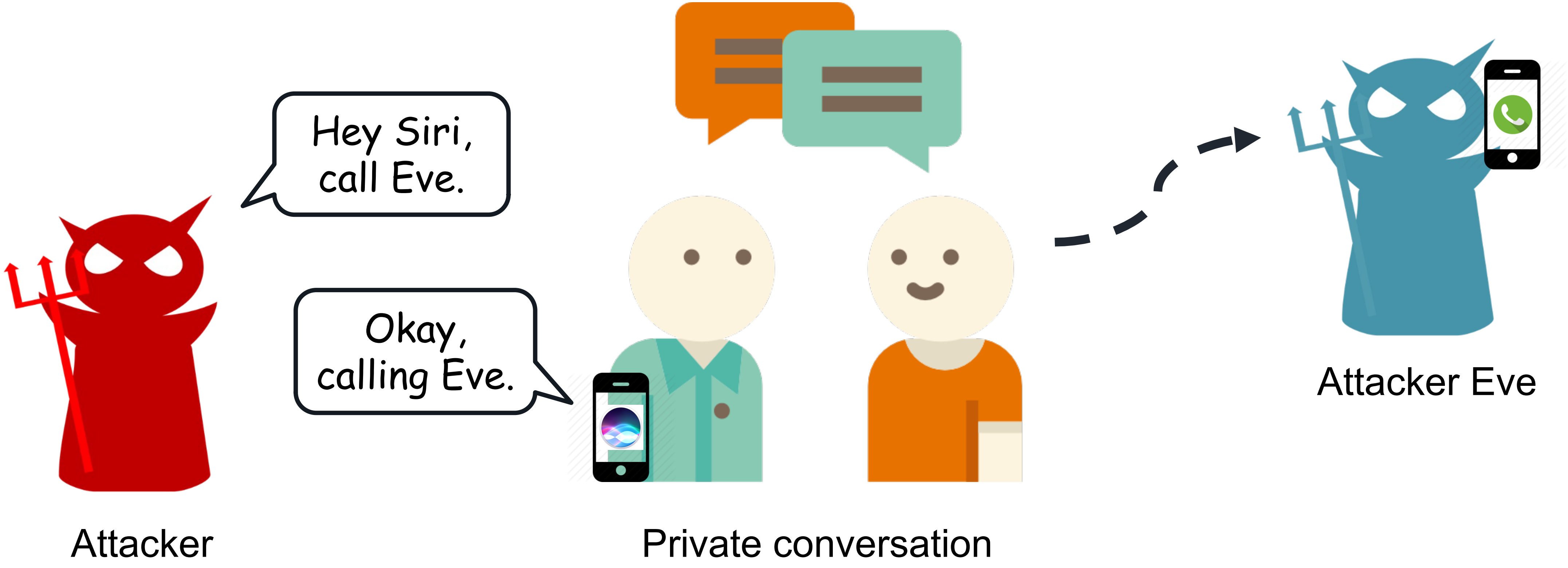
Mislead

Attack Scenario 1: fake online orders

Attacker

Amazon Echo

Attack Scenario 2: spying phone/video calls



Attack Scenario 3: exposing user privacy

Related Work

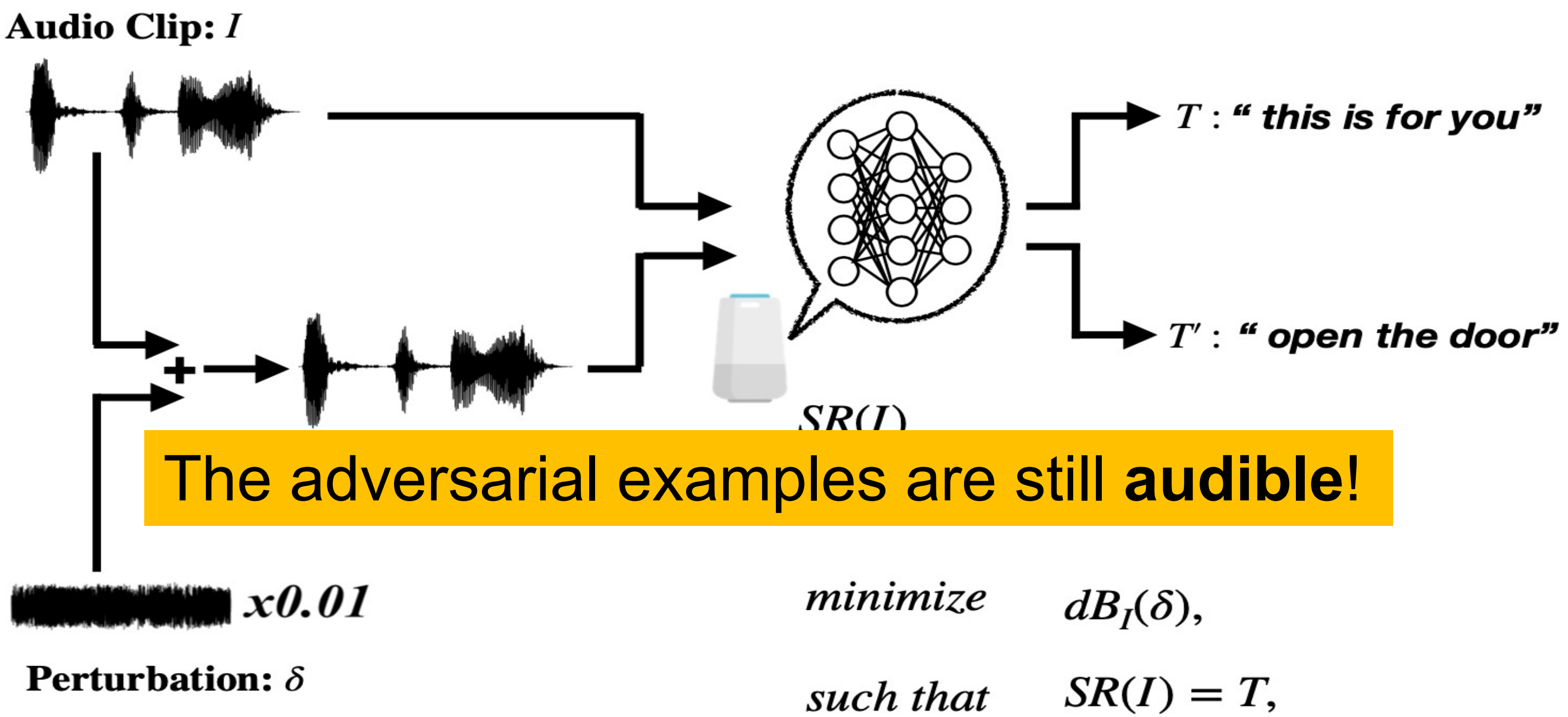
Vaidya et al., **Cocaine Noodles** (WOOT 2015)
Carlin et al., **Hidden Voice Commands** (Usenix Security 2016)

The attacking commands are **audible**, and can be **noticed**!

The attacking commands are still **audible.**

Related work: adversarial examples (AE)

- Generating an audio adversarial example via optimization



The adversarial examples are still audible!

$$\begin{aligned} \textit{minimize} \quad & dB_I(\delta), \\ \textit{such that} \quad & SR(I) = T, \\ & SR(I+\delta) = T' \end{aligned}$$

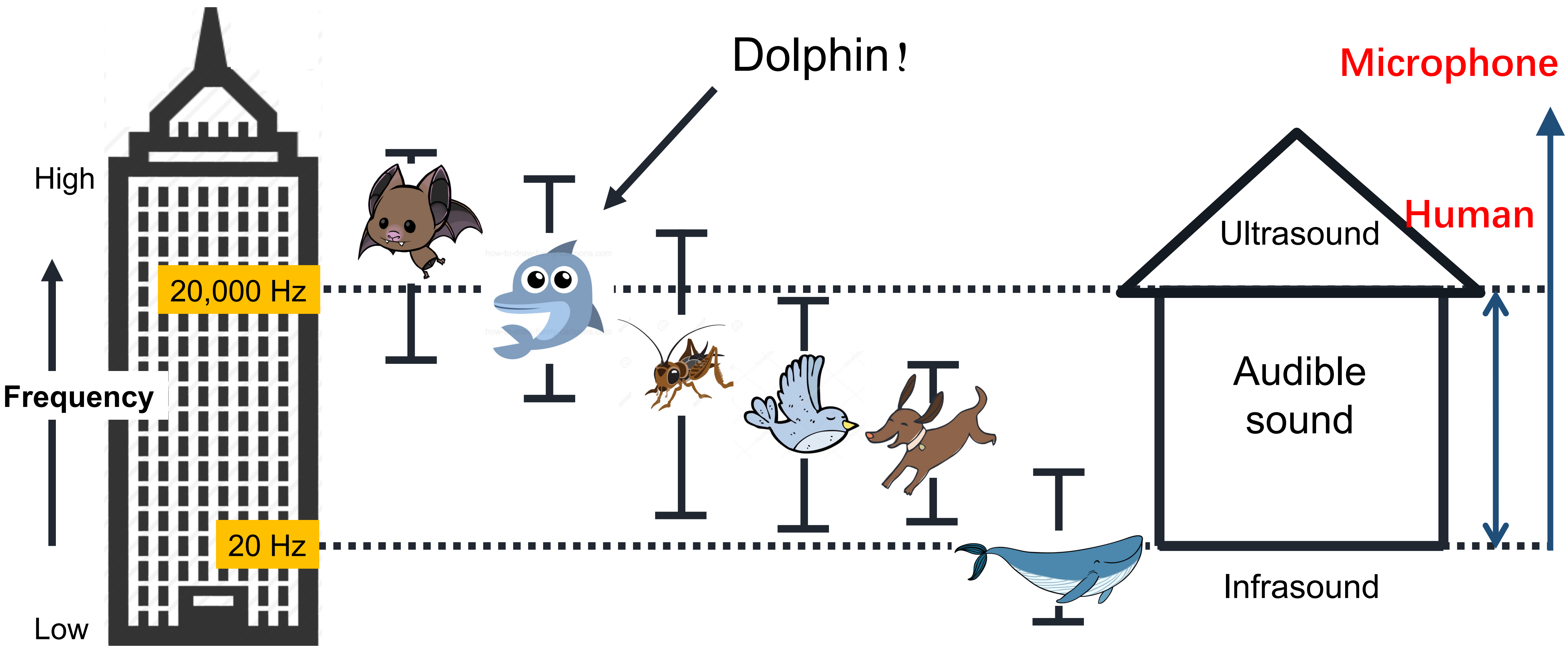
DolphinAttack

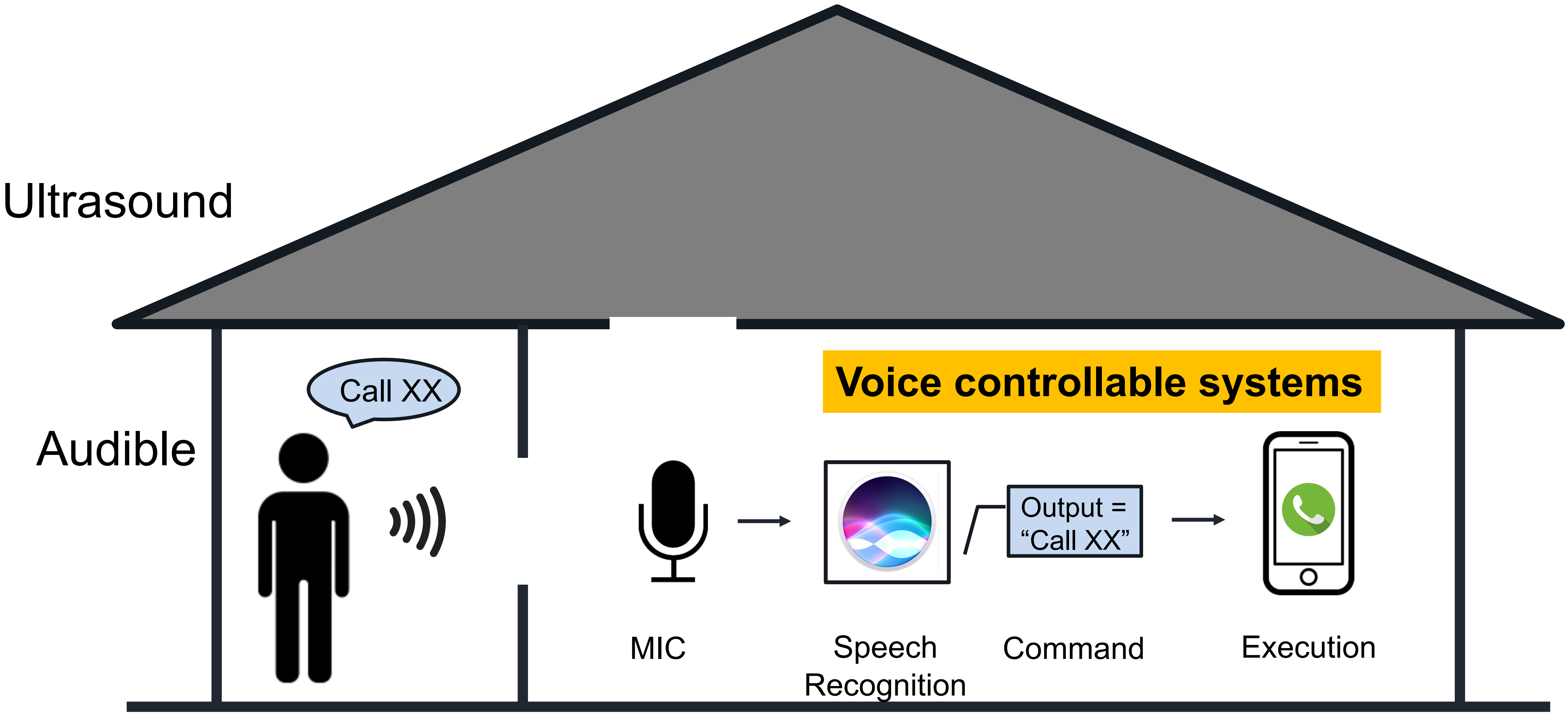
ATTACKED DEVICE : AMAZON ECHO

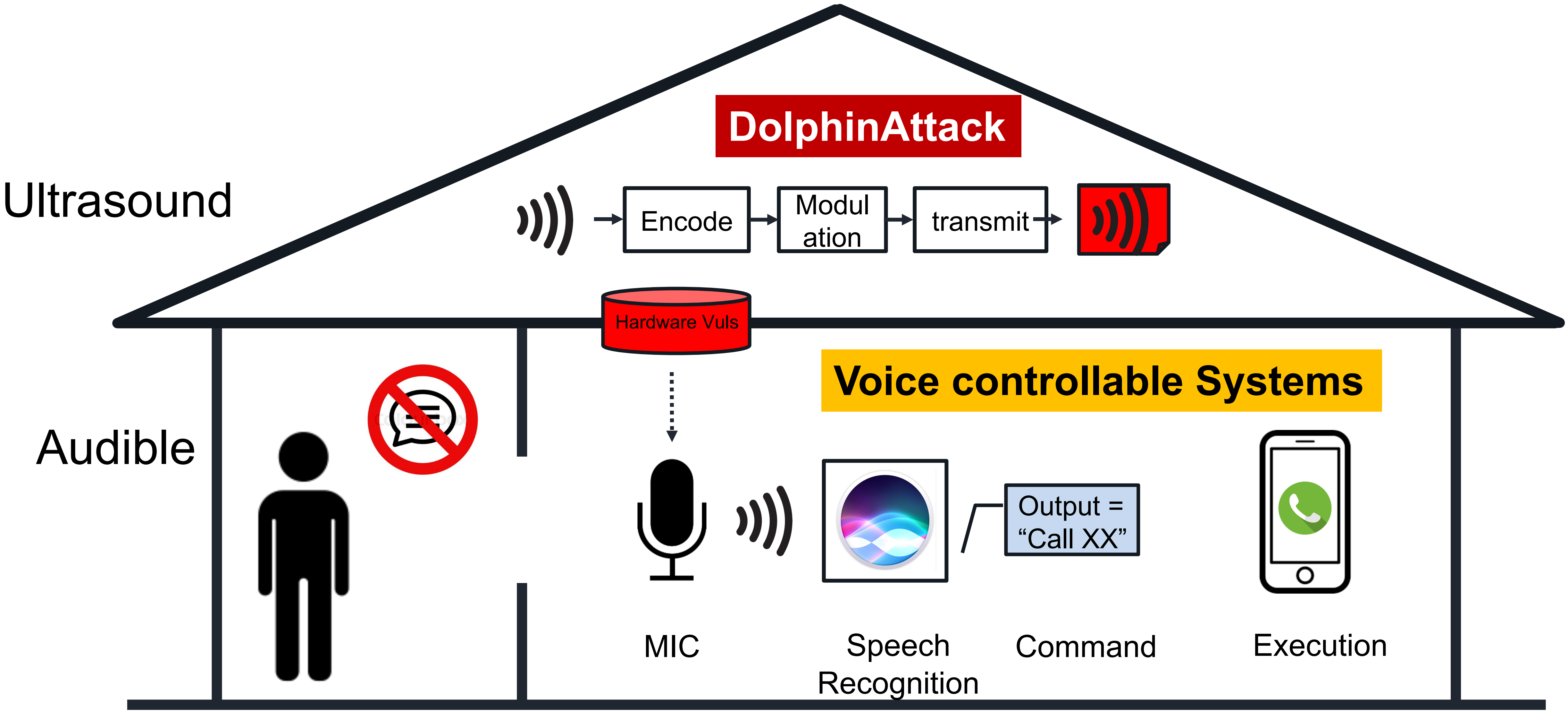
Attack Scenario

- Order stuff
- Make a call
- Read to-do list
- Open the door


Scenario 1: Shopping

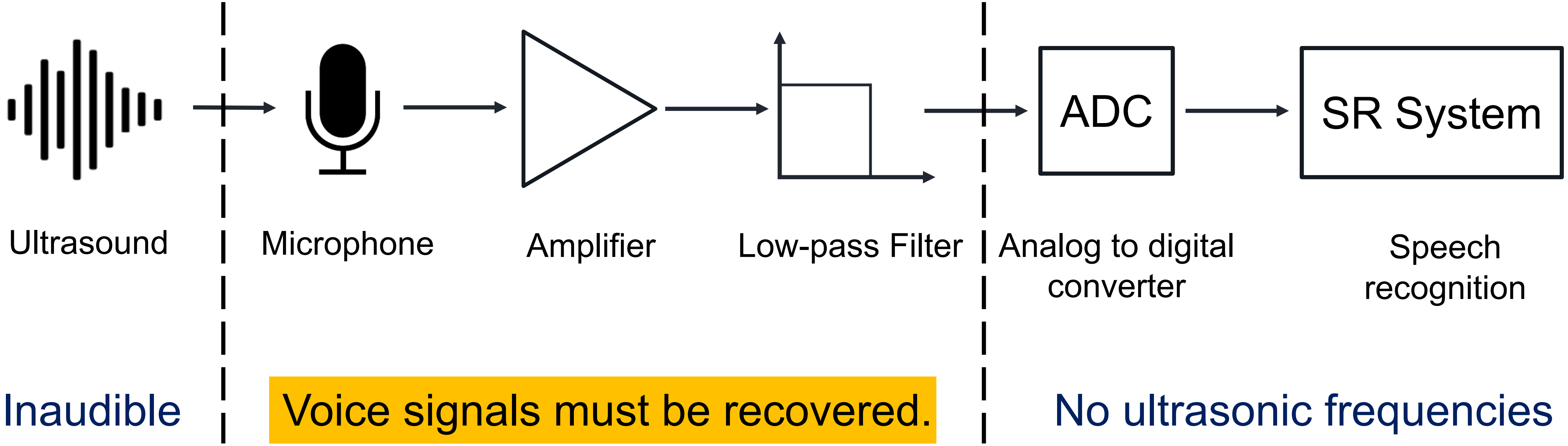

Dolphin！
Microphone
Human
High
20,000 Hz
Frequency
20 Hz
Low
Ultrasound
Audible sound
Infrasound


Ultrasound
Audible
Call XX
Voice controllable systems
Output =
"Call XX"
MIC
Speech
Recognition
Command
Execution

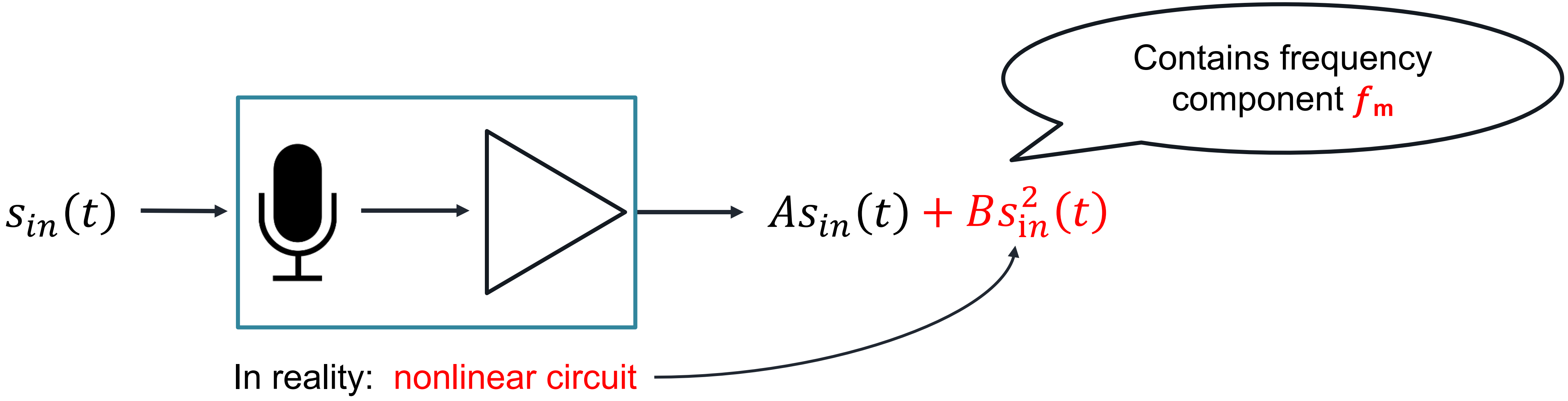

DolphinAttack
Ultrasound
Encode
Modul ation
transmit
Hardware Vuls
Voice controllable Systems
Audible
Output = “Call XX”
MIC
Speech Recognition
Command
Execution

How can voice assistants accept ultrasound?

- The low-pass filter will remove ultrasonic frequencies to avoid aliasing.



Exploiting the Nonlinearity of Microphone

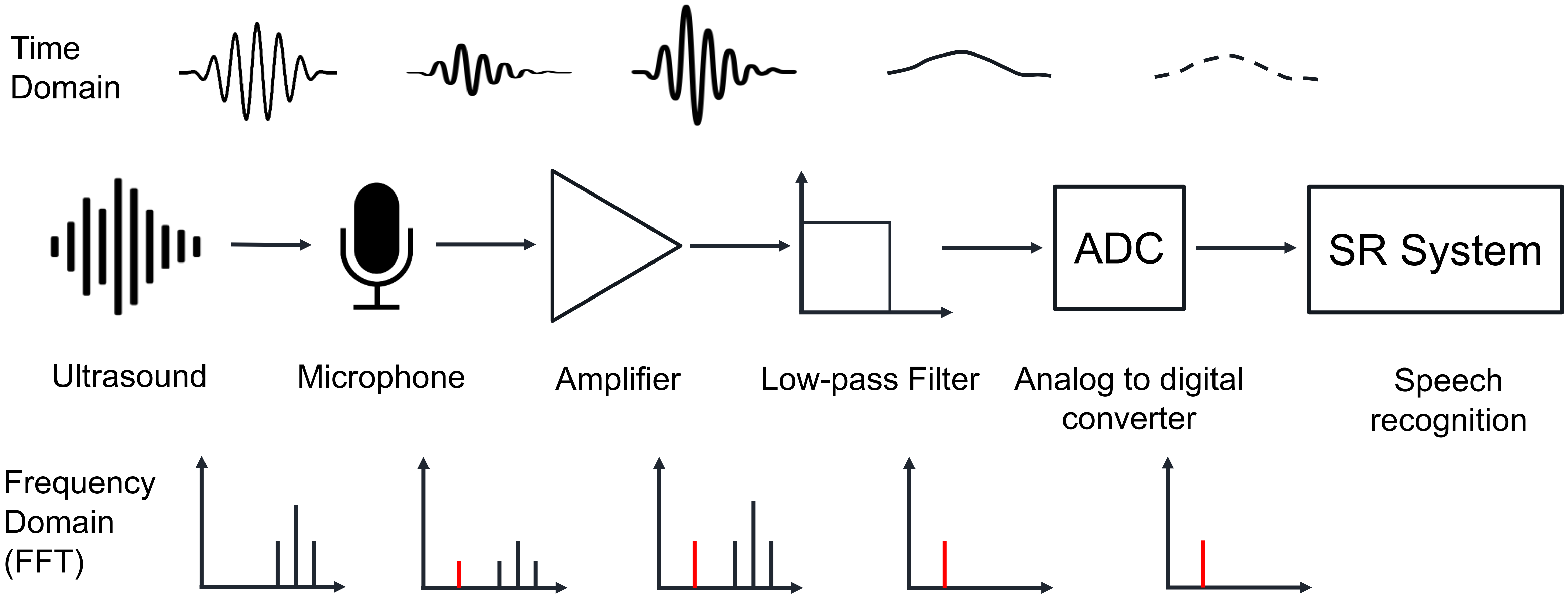


Let input be $s_{in}(t) = m(t)\cos(2\pi f_c t) + \cos(2\pi f_c t)$

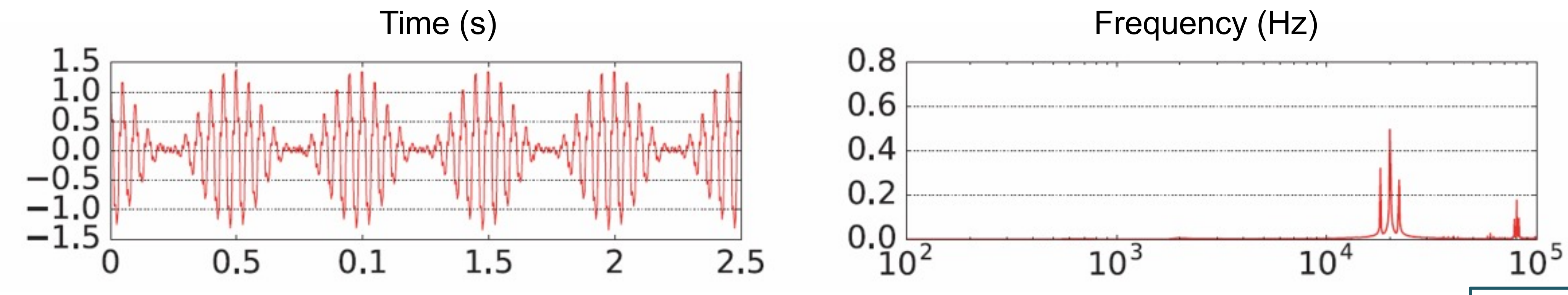
Where $m(t)$ is a baseband voice signal, $m(t) = \cos(2\pi f_m t)$

The baseband voice signals can be demodulated by microphones.

Signal Flow of DolphinAttack



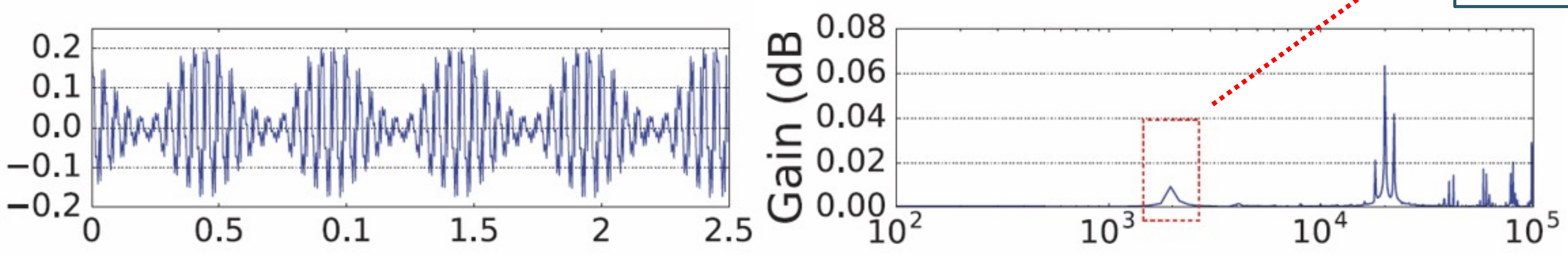
Nonlinearity Effect Validation



$f_c = 22$ kHz, $f_m = 2$ kHz

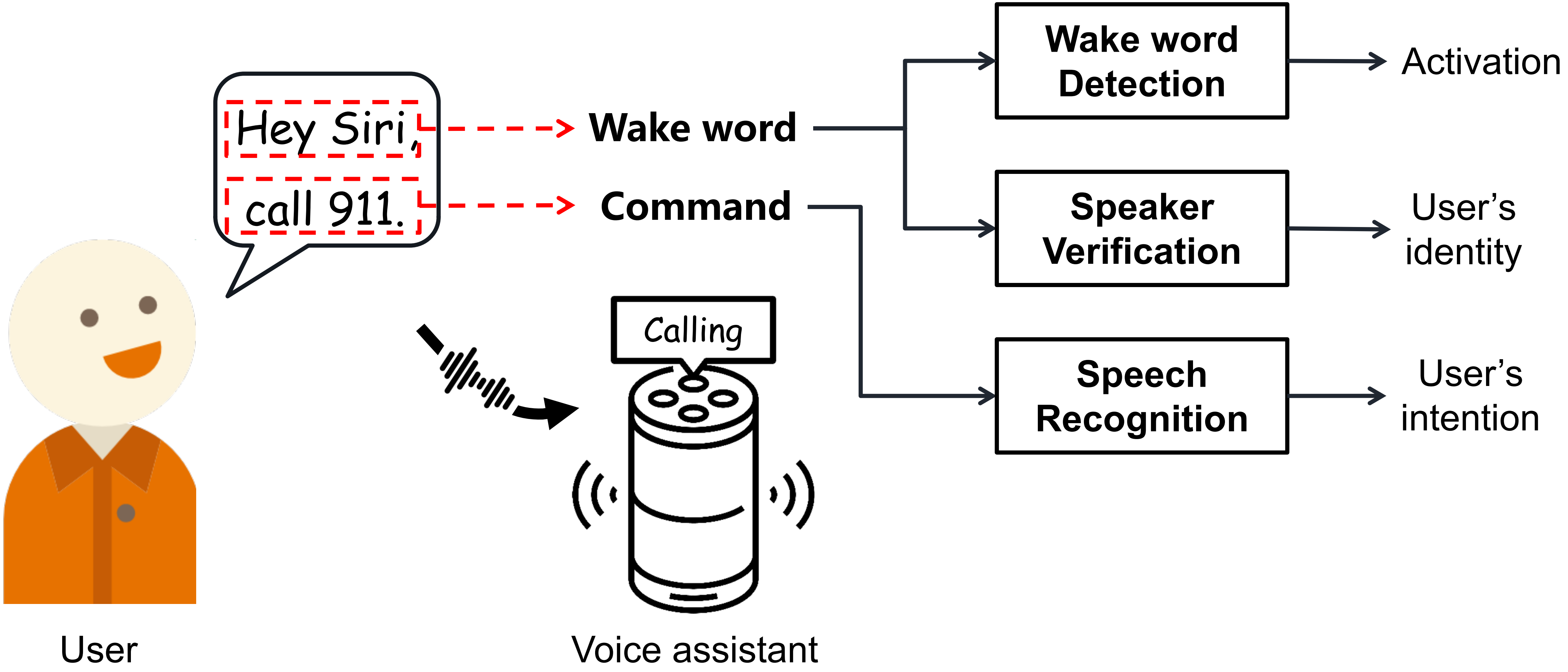
Signals of DolphinAttack

Nonlinearity Effect

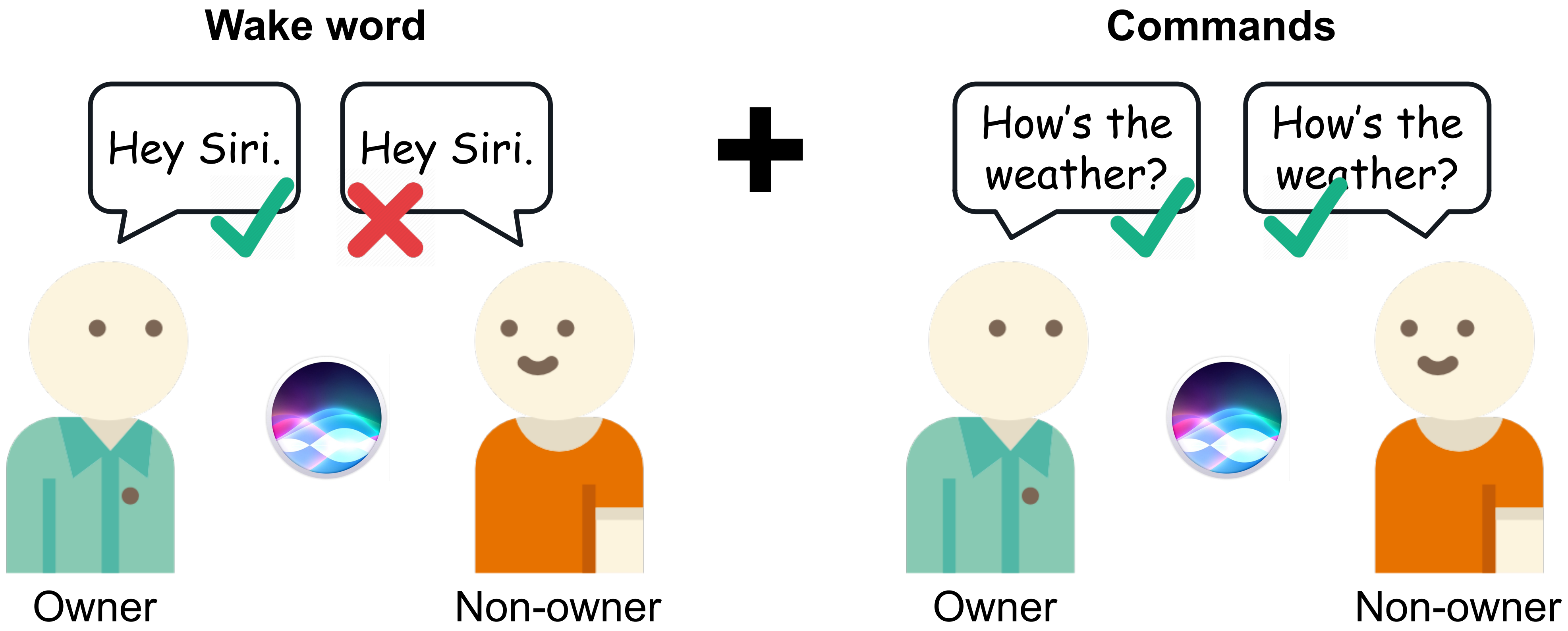


Signals received by a MEMS microphone

How does a voice assistant work?

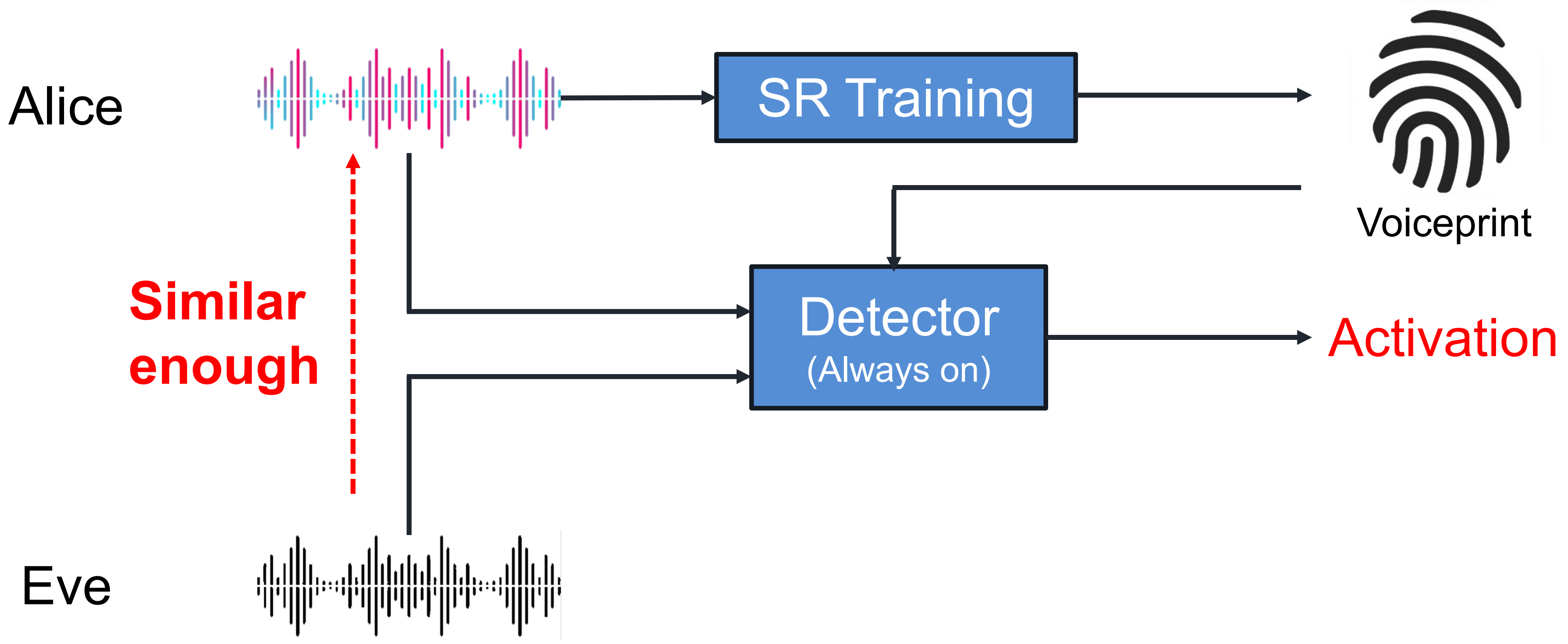


Speaker Dependent vs Speaker Independent

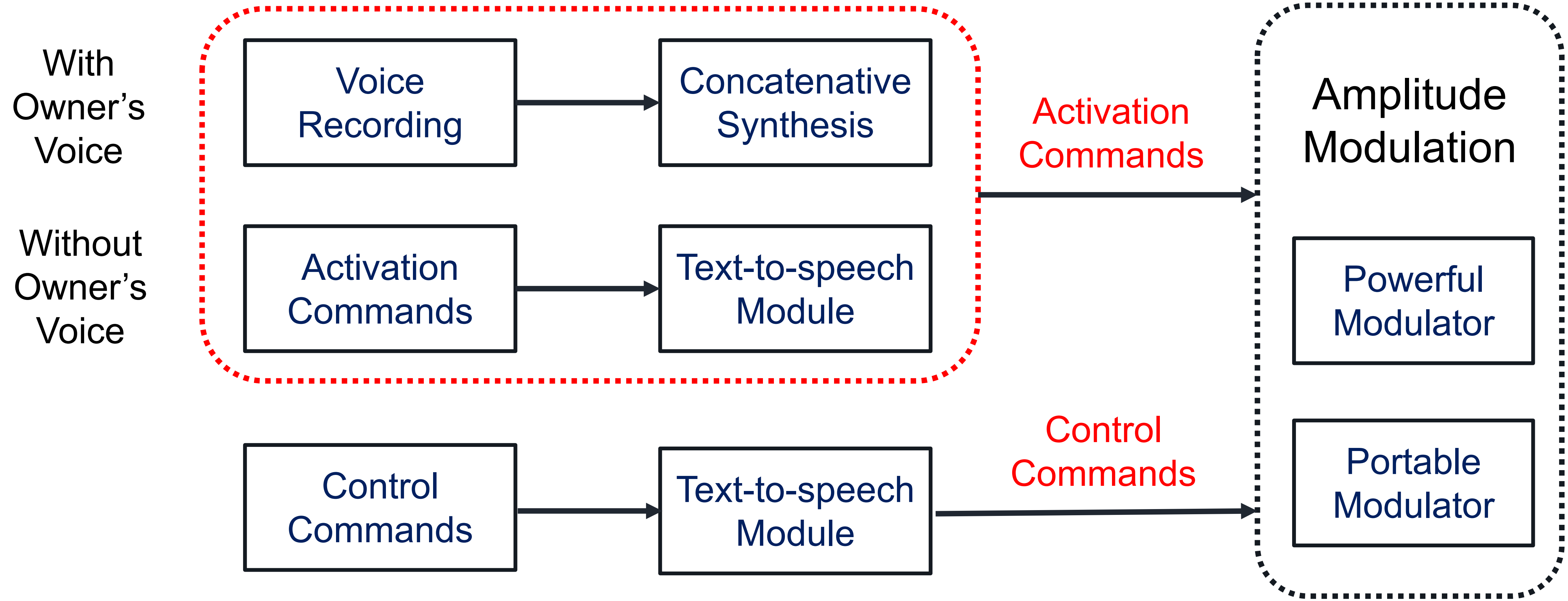


Both **activation** and **control** commands are required for DolphinAttack.

Speaker Dependent SR – Activation



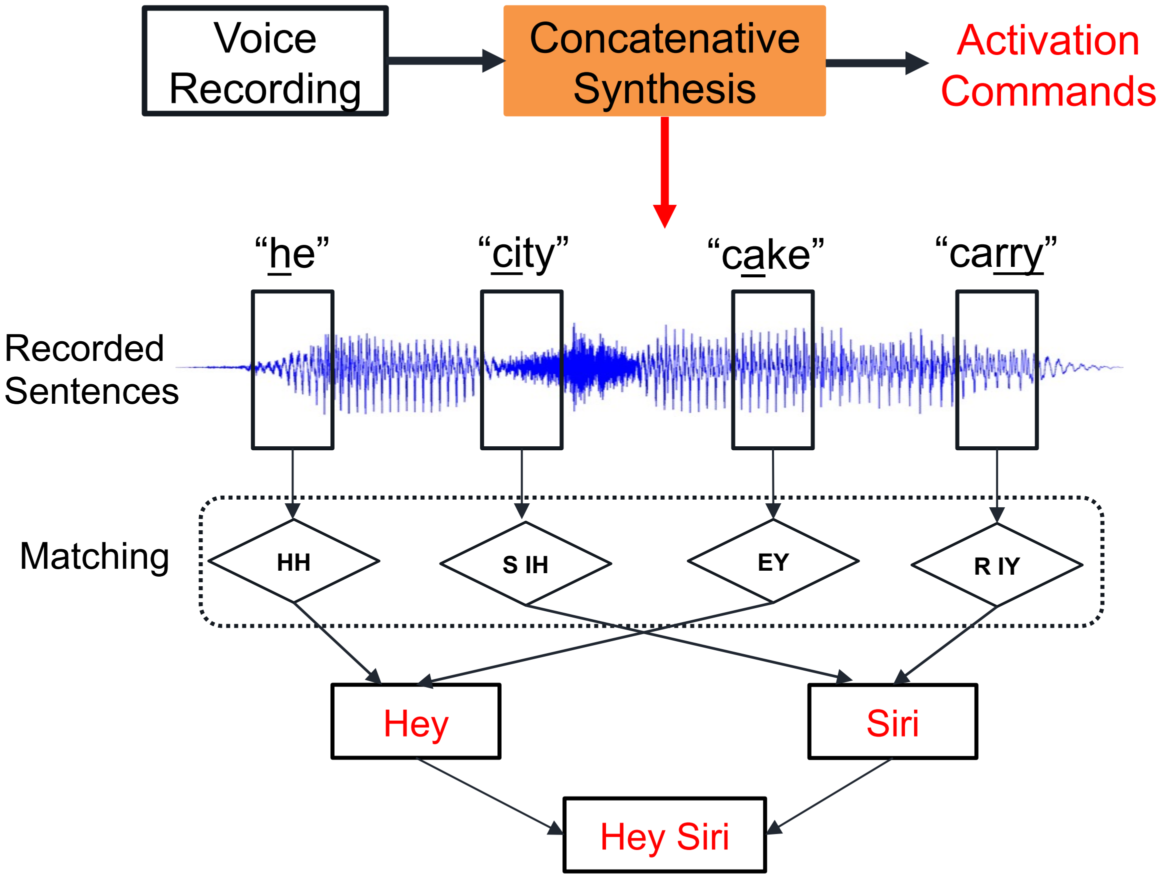
Design of DolphinAttack



1. Concatenative Synthesis – with owner's voice

- 44 phonemes in English.
- "Hey Siri" includes 6 of them

  (i.e., HH, EY, S, IH, R, IY).
- Synthesize a desired activation command by searching for relevant phonemes from other words in available recordings.



Concatenative synthesis of "Hey Siri"

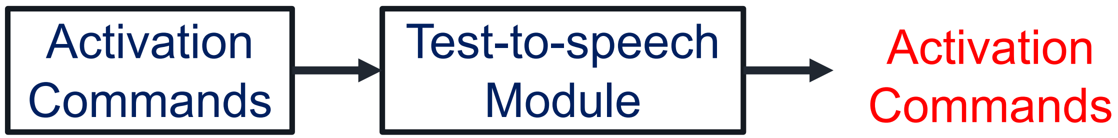
2. TTS-based Approach – without owner's voice

TTS: Text to Speech

Observation

- Two users with similar vocal tones can activate the other's Siri.

Method



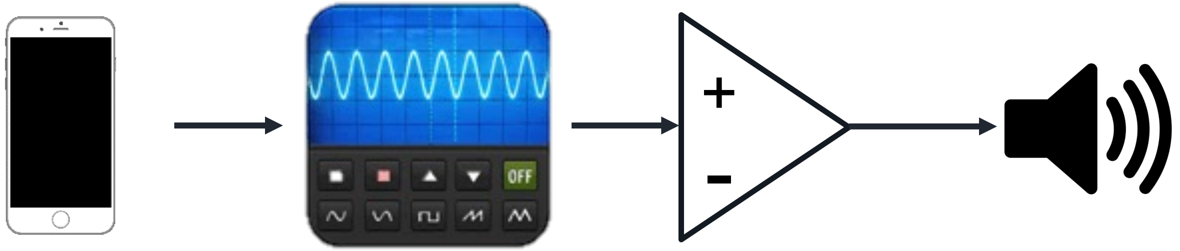
- 35 out of 89 TTS systems can successfully activate a trained Siri.

TTS Systems	voice type #	# of successful types	
		Call 12..90	Hey Siri
Selvy Speech [51]	4	4	2
Baidu [8]	1	1	0
Sestek [45]	7	7	2
NeoSpeech [39]	8	8	2
Innoetics [59]	12	12	7
Vocalware [63]	15	15	8
CereProc [12]	22	22	9
Acapela [22]	13	13	1
Fromtexttospeech [58]	7	7	4

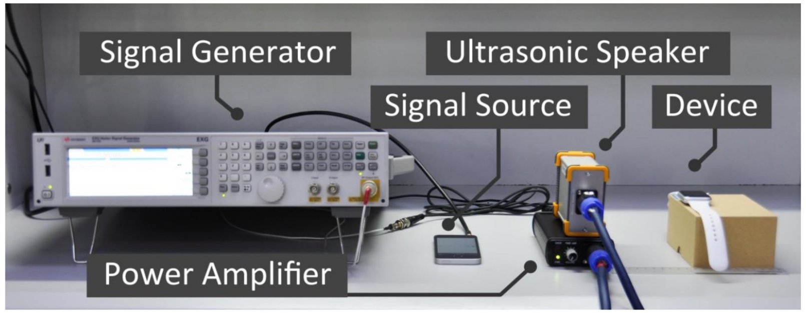
The list of TTS systems used for attacking the Siri trained by the Google TTS system, and the evaluation results on activation and control commands.

Inaudible Voice Commands Transmitter

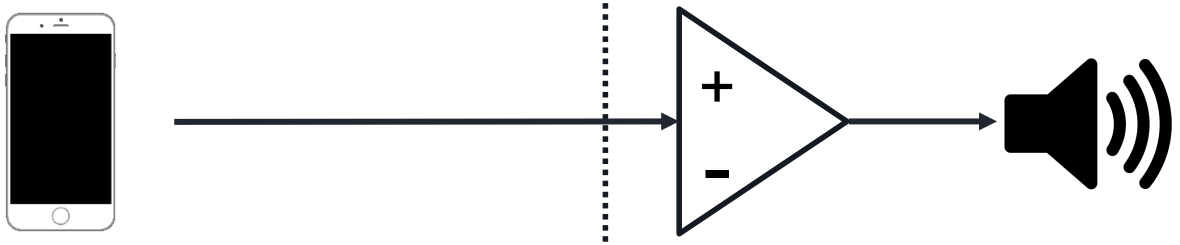
Rich man solution



Signal source — Signal generator (Hardware Modulator) — Amplifier — Professional speaker

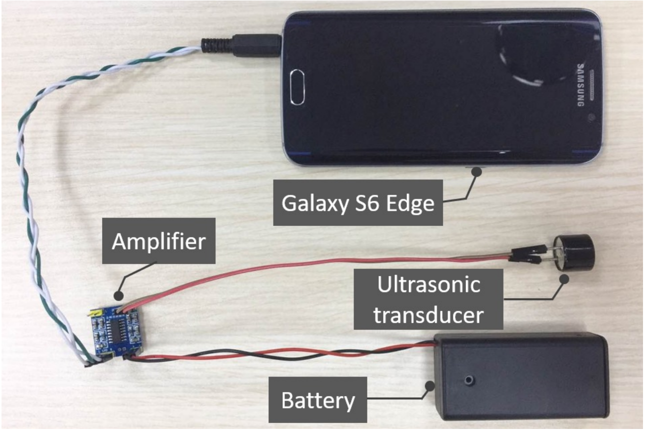


Poor man solution



Signal source & Software Modulator — Amplifier — Low-cost speaker

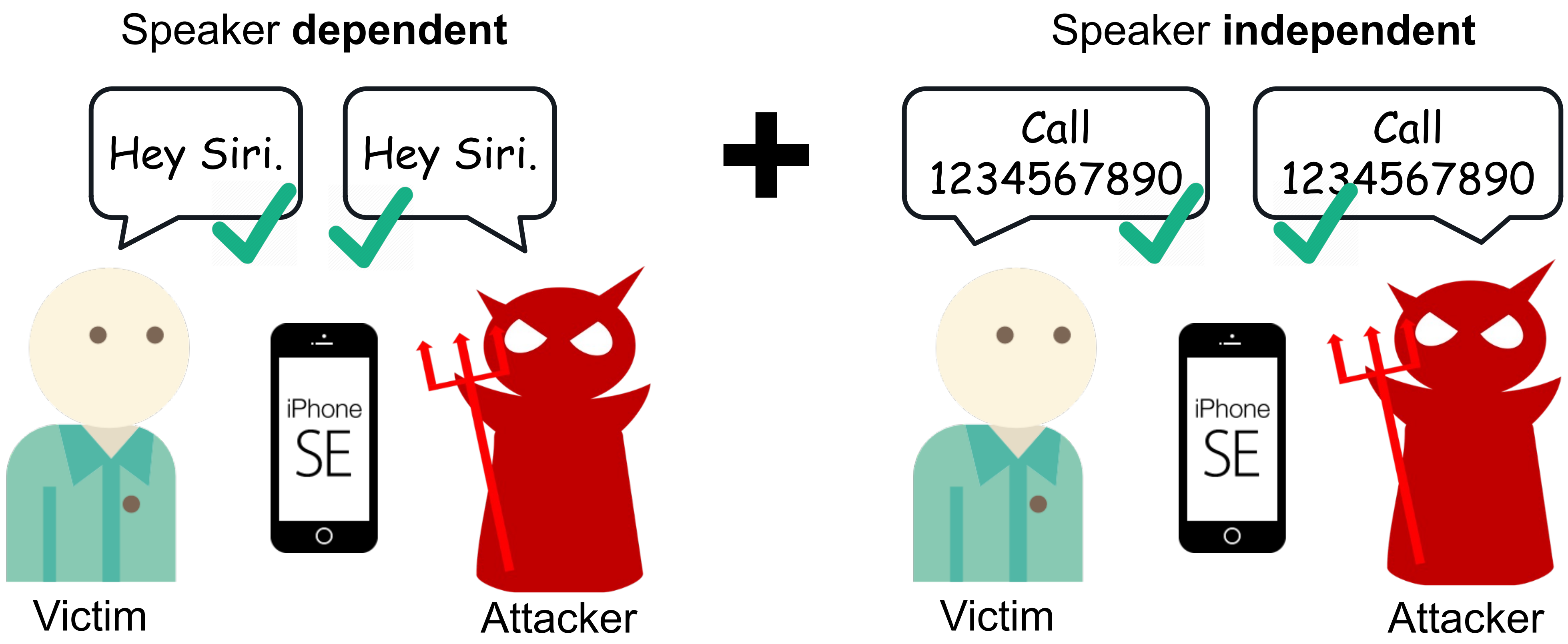
Less than $3



DolphinAttack

ATTACKED DEVICE : IPHONE

Attack Scenario: Make Spying Phone Call

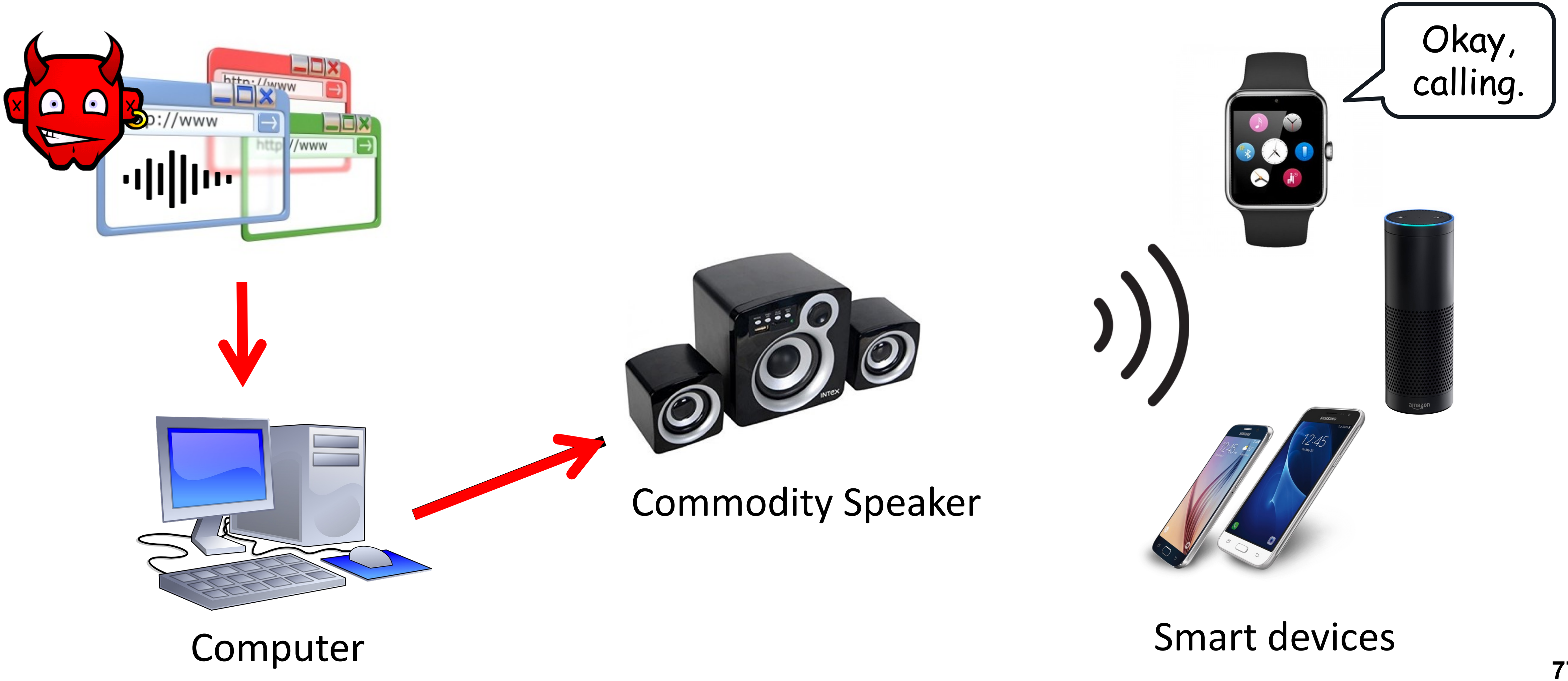



55.5
Activate Siri and make a phone call with a normal voice.

DolphinAttack

ATTACKED DEVICE: APPLE WATCH

Attack Scenario: Remote Attack




"Facetime 1551072xxxx"
Under attack

DolphinAttack

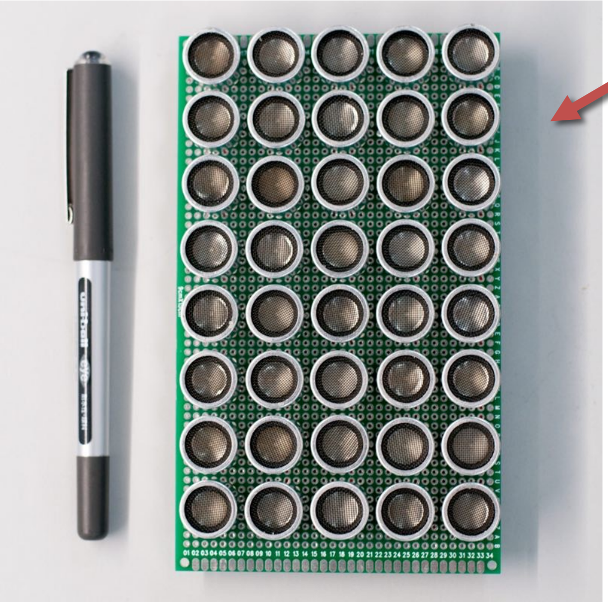
COMPROMISED DEVICES

Manuf.	Model	OS/Ver.	SR System	Attacks		Modulation Parameters		Max Dist. (cm)	
				Recog.	Activ.	f_c (kHz) & [Prime f_c] ‡	Depth	Recog.	Activ.
Apple	iPhone 4s	iOS 9.3.5	Siri	√	√	20–42 [27.9]	≥ 9%	175	110
Apple	iPhone 5s	iOS 10.0.2	Siri	√	√	24.1 26.2 27 29.3 [24.1]	100%	7.5	10
Apple	iPhone SE	iOS 10.3.1	Siri	√	√	22–28 33 [22.6]	≥ 47%	30	25
			Chrome	√	N/A	22–26 28 [22.6]	≥ 37%	16	N/A
Apple	iPhone SE †	iOS 10.3.2	Siri	√	√	21–29 31 33 [22.4]	≥ 43%	21	24
Apple	iPhone 6s *	iOS 10.2.1	Siri	√	√	26 [26]	100%	4	12
Apple	iPhone 6 Plus *	iOS 10.3.1	Siri	×	√	— [24]	—	—	2
Apple	iPhone 7 Plus *	iOS 10.3.1	Siri	√	√	21 24-29 [25.3]	≥ 50%	18	12
Apple	watch	watchOS 3.1	Siri	√	√	20–37 [22.3]	≥ 5%	111	164
Apple	iPad mini 4	iOS 10.2.1	Siri	√	√	22–40 [28.8]	≥ 25%	91.6	50.5
Apple	MacBook	macOS Sierra	Siri	√	N/A	20-22 24-25 27-37 39 [22.8]	≥ 76%	31	N/A
LG	Nexus 5X	Android 7.1.1	Google Now	√	√	30.7 [30.7]	100%	6	11
Asus	Nexus 7	Android 6.0.1	Google Now	√	√	24–39 [24.1]	≥ 5%	88	87
Samsung	Galaxy S6 edge	Android 6.0.1	S Voice	√	√	20–38 [28.4]	≥ 17%	36.1	56.2
Huawei	Honor 7	Android 6.0	HiVoice	√	√	29–37 [29.5]	≥ 17%	13	14
Lenovo	ThinkPad T440p	Windows 10	Cortana	√	√	23.4–29 [23.6]	≥ 35%	58	8
Amazon	Echo *	5589	Alexa	√	√	20-21 23-31 33-34 [24]	≥ 20%	165	165
Audi	Q3	N/A	N/A	√	N/A	21–23 [22]	100%	10	N/A

‡ Prime f_c is the carrier wave frequency that exhibits highest baseband amplitude after demodulation. — No result
† Another iPhone SE with identical technical spec.
* Experimented with the front/top microphones on devices.

Improvement to long-range attacks

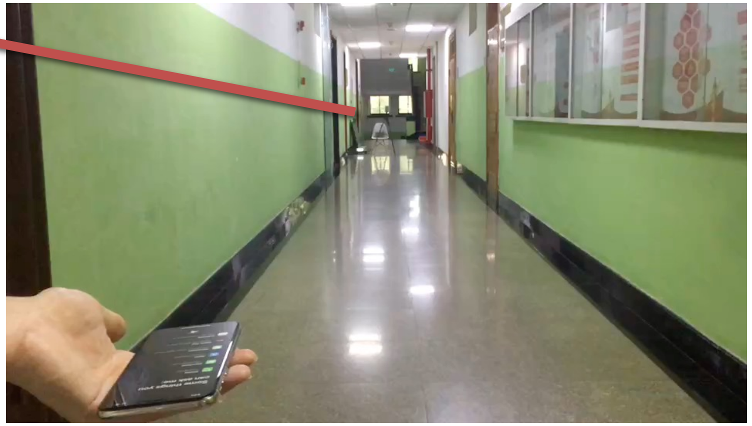
Ultrasonic transducer array



Long-range setup

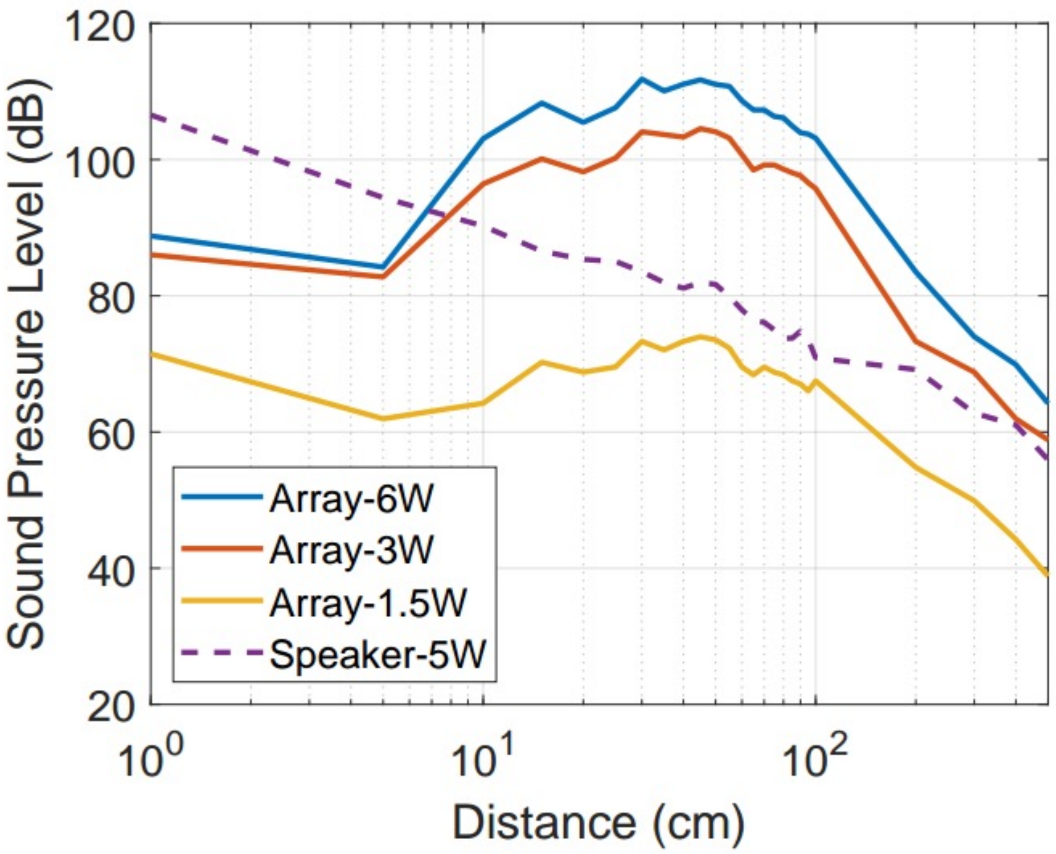
10 meters

20 meters



Can we infinitely boost the attack range with more power?

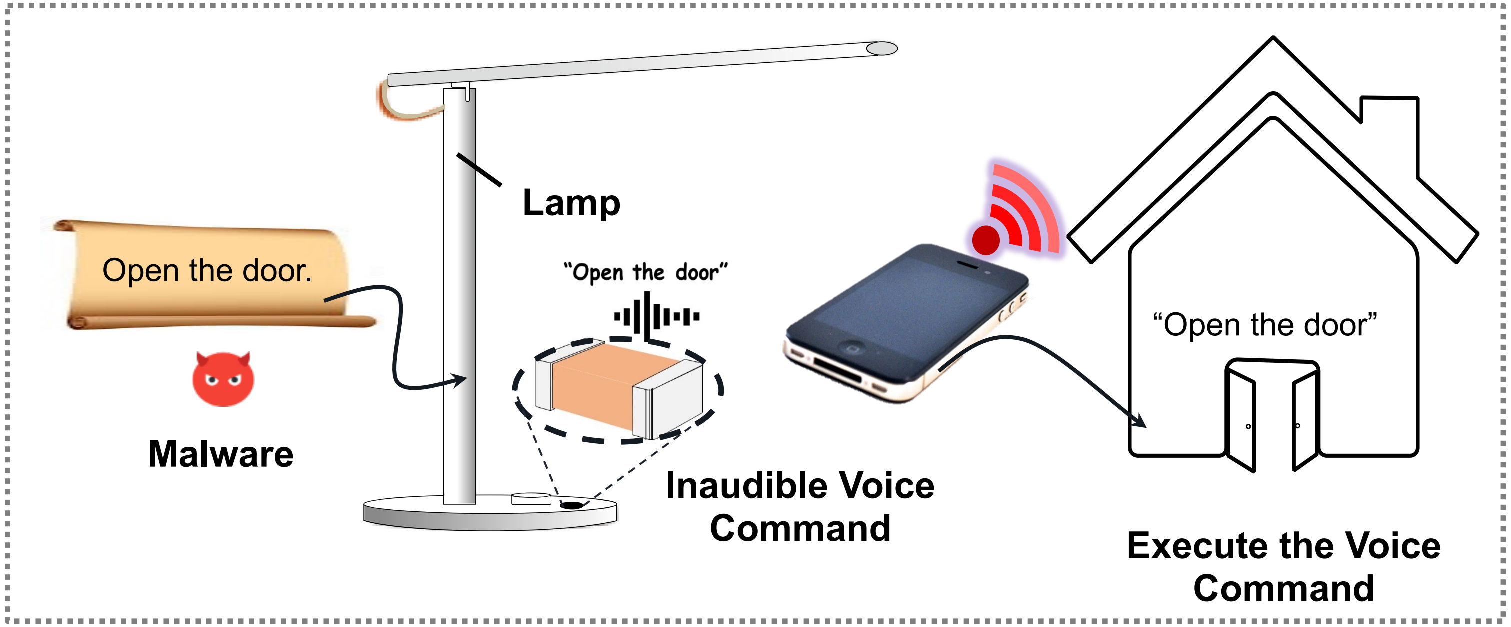
- Inaudible voice commands become audible when the transmission power is high!
- **Nonlinear acoustics** happens when sounds have sufficiently large amplitudes



What if there is no speaker at all

CapSpeaker: Turn capacitors in to speakers!

- Multi-layer Ceramic (MLC) Capacitors can be used as a speaker to attack voice systems due to **inverse piezoelectric effect**

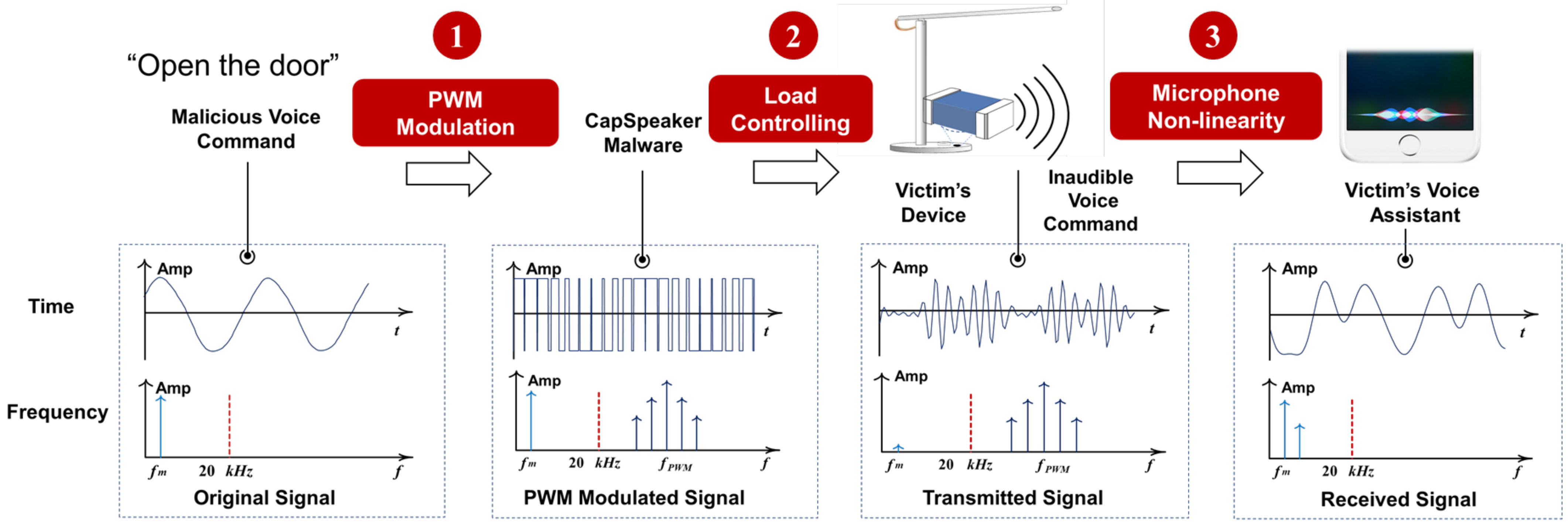


Xiaoyu Ji, Juchuan Zhang, Shui Jiang, Jishen Li, Wenyuan Xu. CapSpeaker: Injecting Voices to Microphones via Capacitors, ACM CCS 2021

What if there is no speaker at all

CapSpeaker: Turn capacitors in to speakers!

- Using **PWM** and exploiting **microphone nonlinearity** to inject voices to a microphone



Xiaoyu Ji, Juchuan Zhang, Shui Jiang, Jishen Li, Wenyuan Xu. CapSpeaker: Injecting Voices to Microphones via Capacitors, ACM CCS 2021

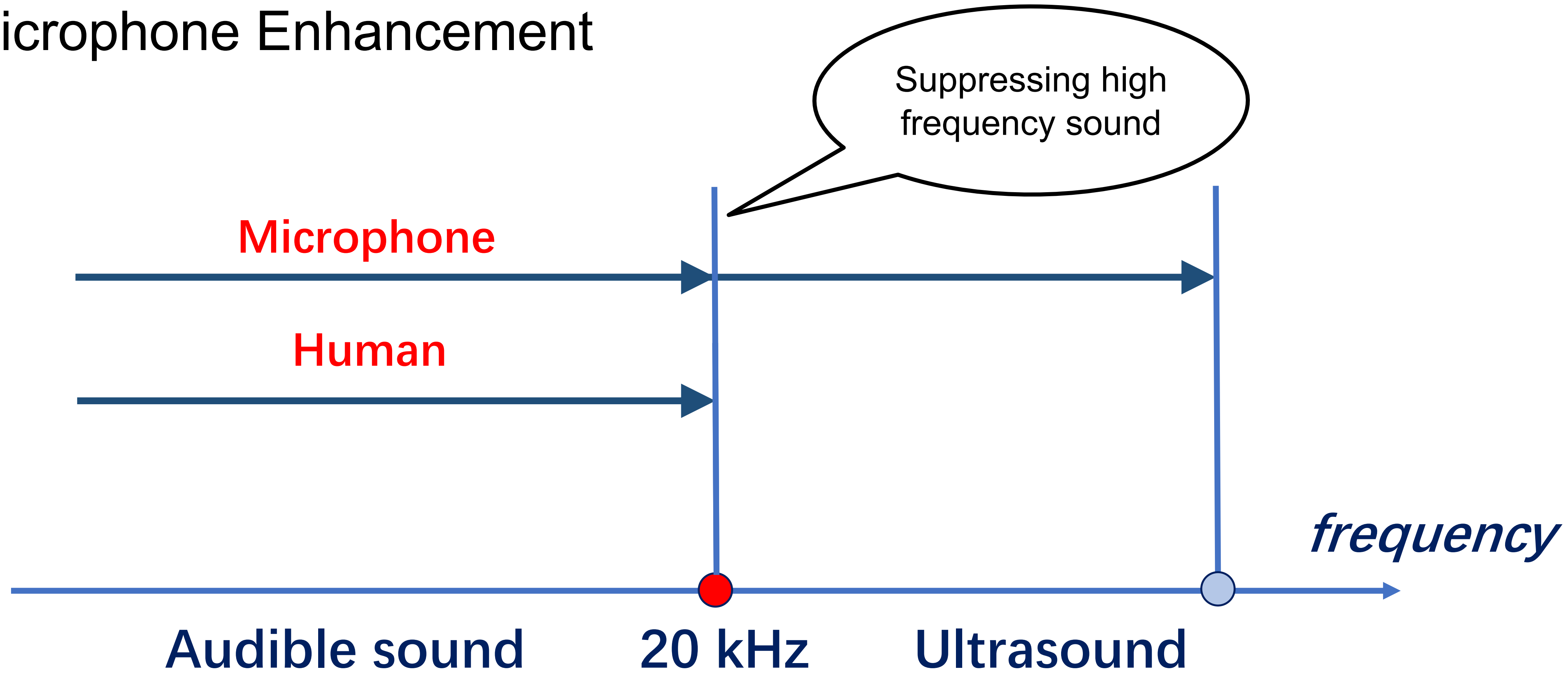
Defense


T-Mobile Wi-Fi
8:21 AM
44%
Settings
Do Not Disturb
General
Display & Brightness
Wallpaper
Sounds
Siri
Touch ID & Passcode
Battery
Privacy
iTunes & App Store
Wallet & Apple Pay


T-Mobile Wi-Fi
8:22 AM
44%
Settings
Siri
Siri
Press and hold the home button to start speaking to Siri, then release the button when you are done. About Siri and Privacy...
Access When Locked
Allow access to Siri when locked.
Allow "Hey Siri"
You can speak to Siri without pressing the home button by saying "Hey Siri" anytime.
Language
English (United States)
Siri Voice
American (Female)
Voice Feedback
With Switch
My Info
Chris Hauk
App Support

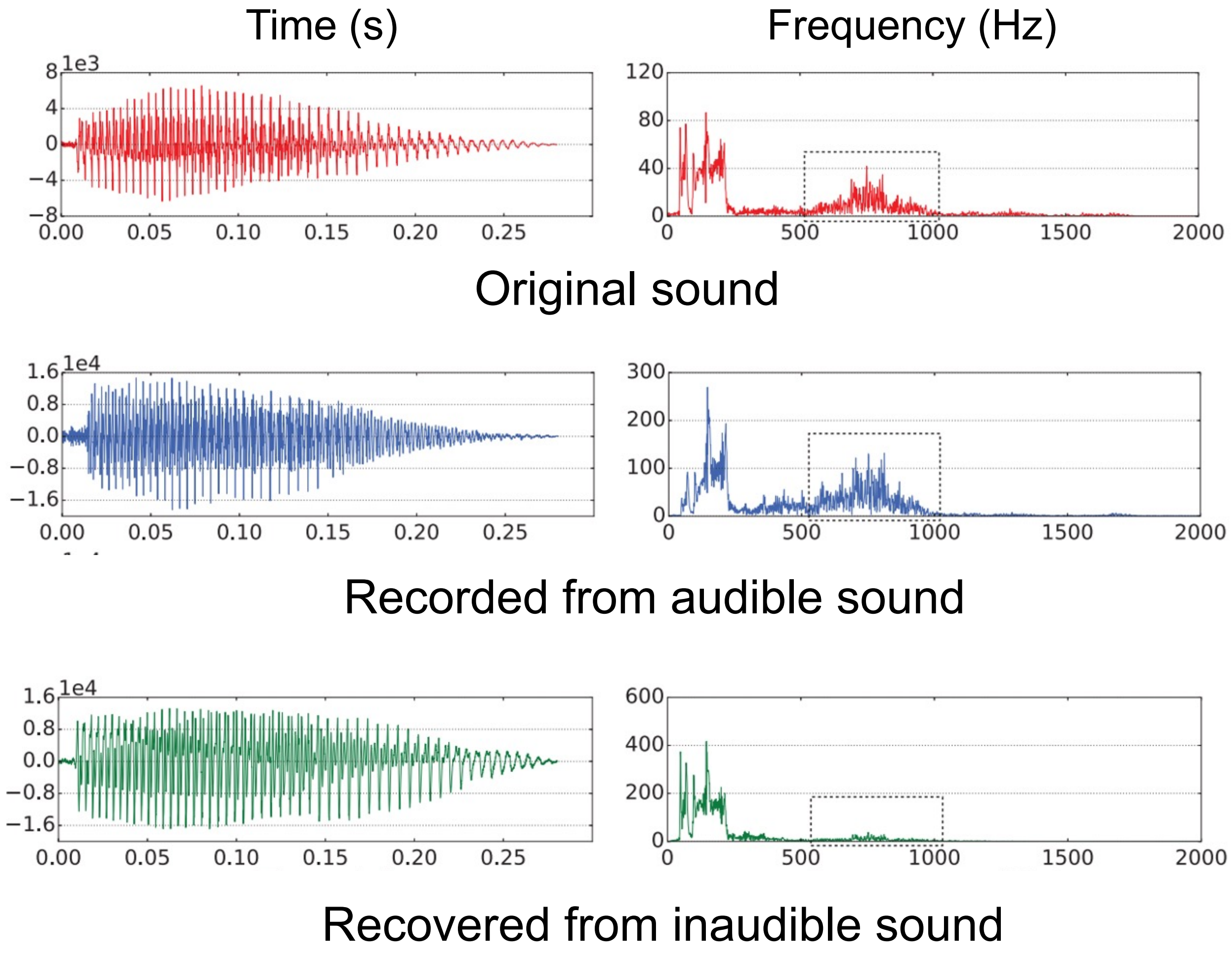
Hardware-Based Defense

- Microphone Enhancement



Software-Based Defense

- Modulated voice commands are distinctive from genuine ones.
- Supported vector machine (SVM) as the classifier to detect the malicious command from the normal command.
- Result: 100% true positive rate (7/7) and 100% true negative rate (7/7).



浙江大学 ZHEJIANG UNIVERSITY

智能系统安全实验室 UBIQUITOUS SYSTEM SECURITY LAB.

CAMERAS + AI

Poltergeist: Acoustic Adversarial Machine Learning against Cameras and Computer Vision

Xiaoyu Ji; Yushi Cheng; Yuepeng Zhang; Kai Wang; Chen Yan; Wenyuan Xu; Kevin Fu

IEEE S&P 21

IoT vs. AIoT

IoT

- Connecting **everything**
- Connecting **the physical world**

Sensor —human→ Control

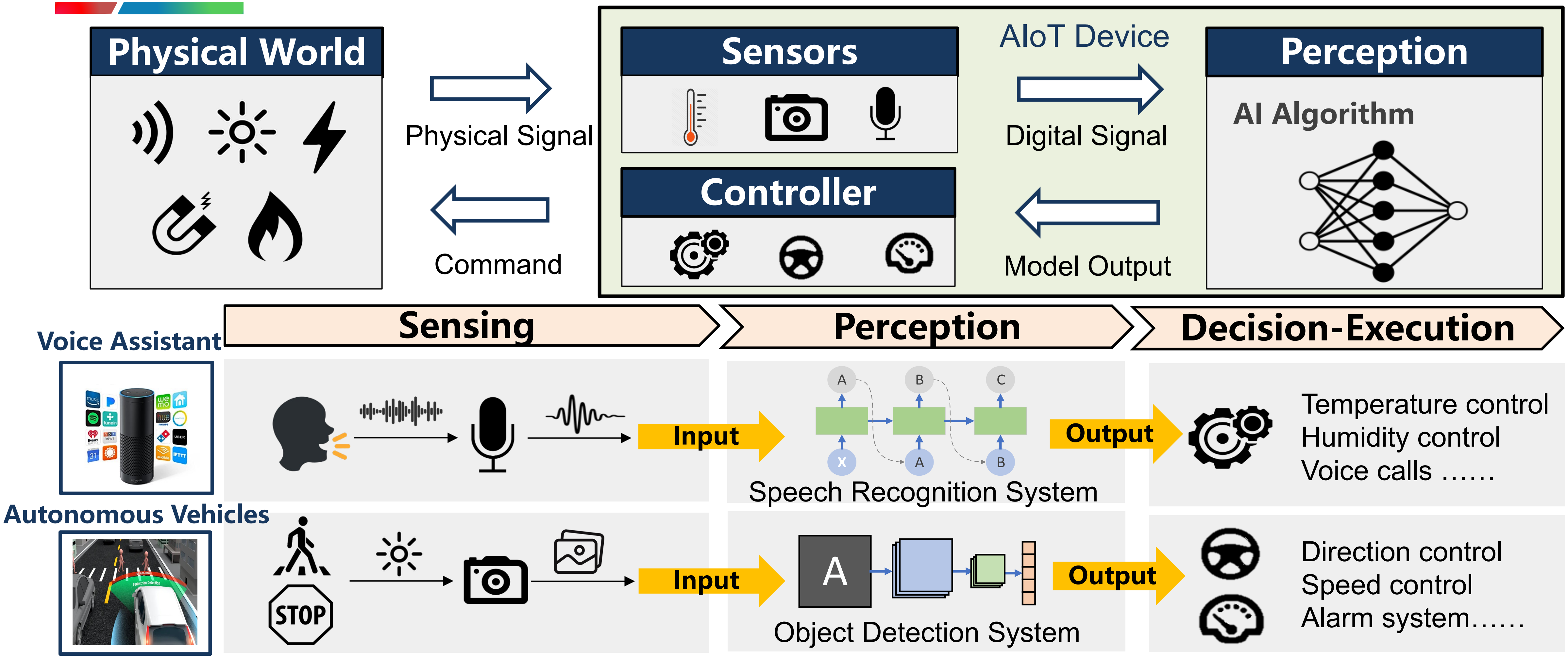
VS

AIoT

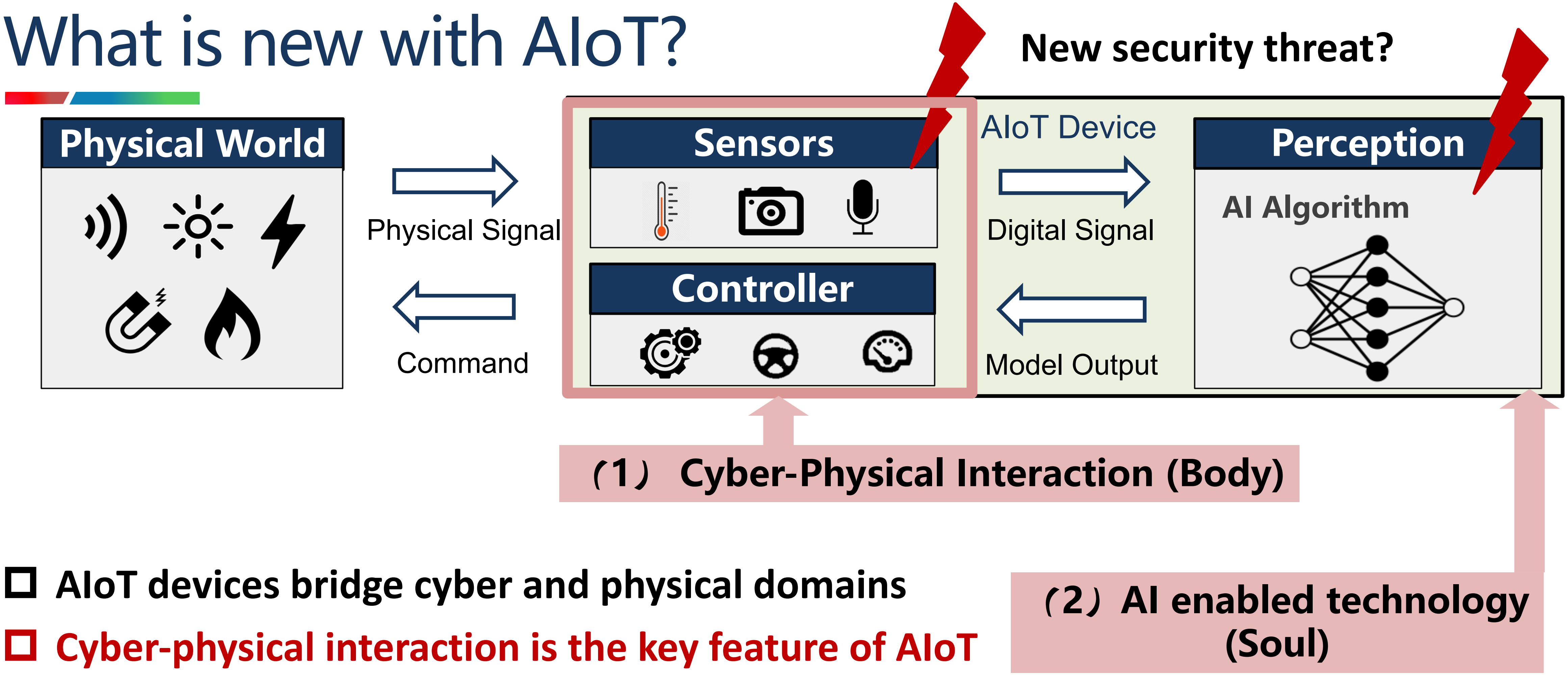
- **Data-Driven** & **AI-Empowered**
- **Faster** Communication Tech. (e.g. 5G)

Sensor → AI Computing → Control

How AIoT devices works

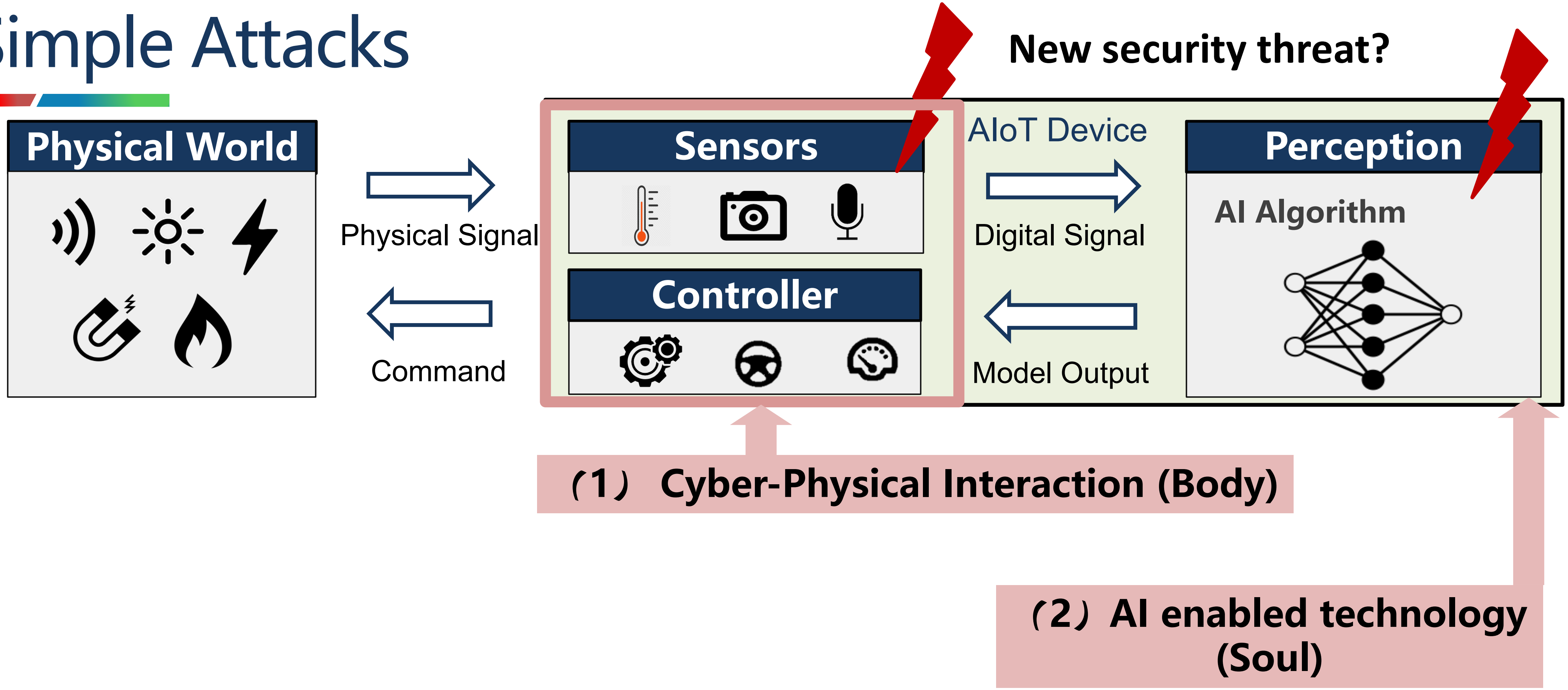


What is new with AIoT?



- AIoT devices bridge cyber and physical domains
- Cyber-physical interaction is the key feature of AIoT
- Traditional vulnerabilities do not cover….

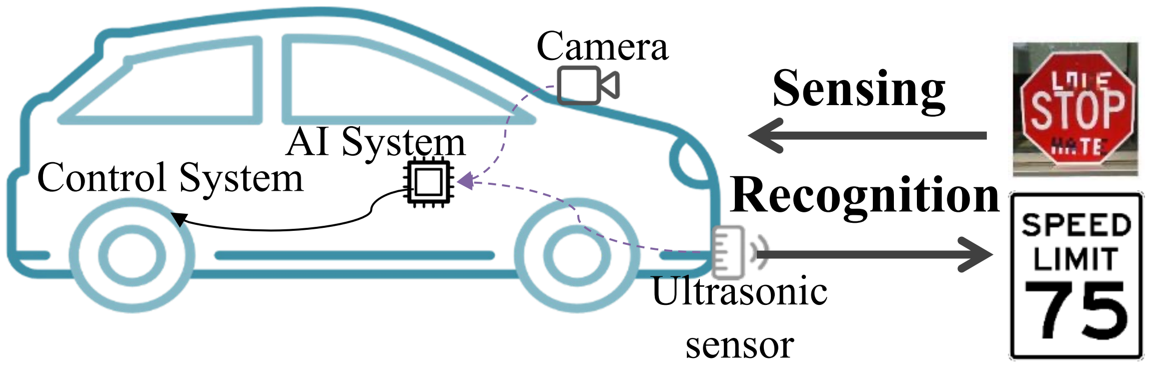
Simple Attacks



- **Simple Attacks utilize vulnerability of each component: Sensors OR AI Algorithms**

New AIoT Security Issues: AI + Sensing errors

Simple Attacks		Ample Attacks
Sensor vulnerabilities *or* AI algorithm vulnerabilities	*Combine* →	Sensor vulnerabilities *and* AI algorithm vulnerabilities

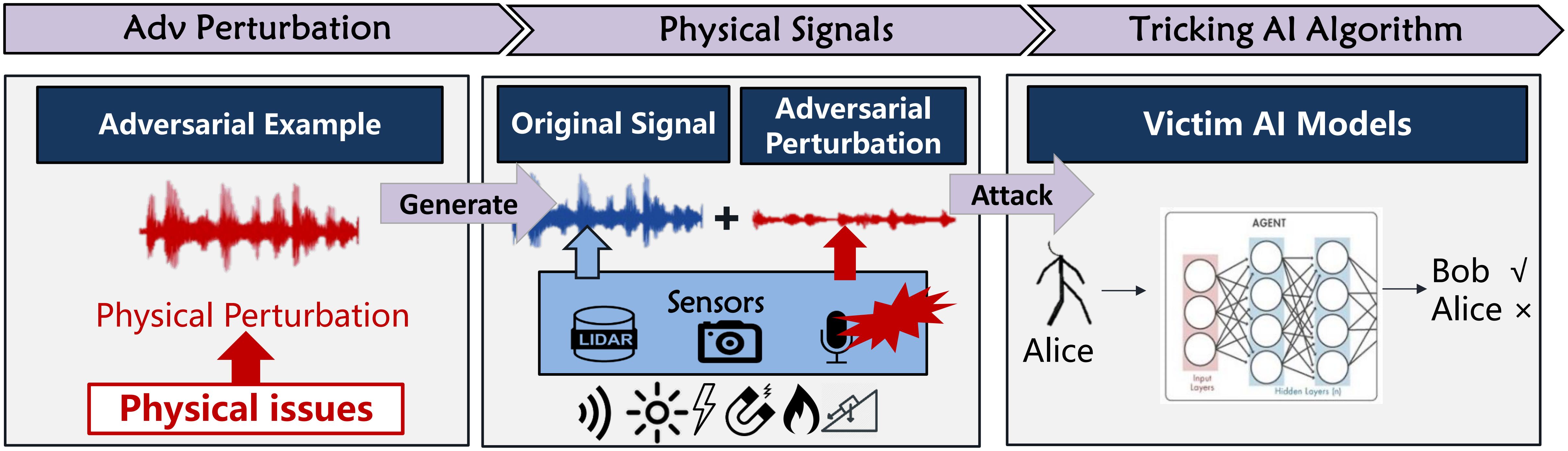


Advantages:

- High concealment
- Physically achievable

Injecting **p**hysics into **A**dversarial **M**achine **Le**arning

AMpLe Attack

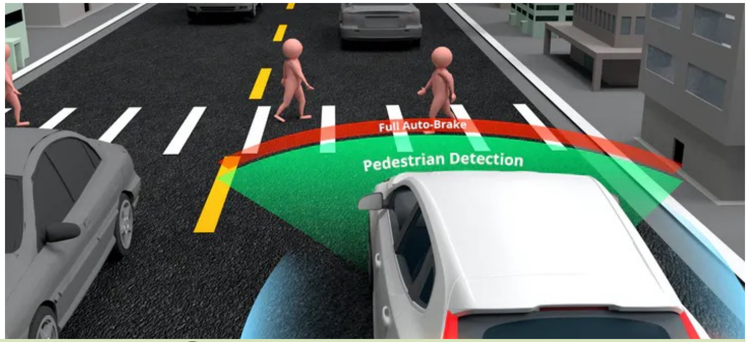


Ample Attacks Case Study: Computer Vision in AV

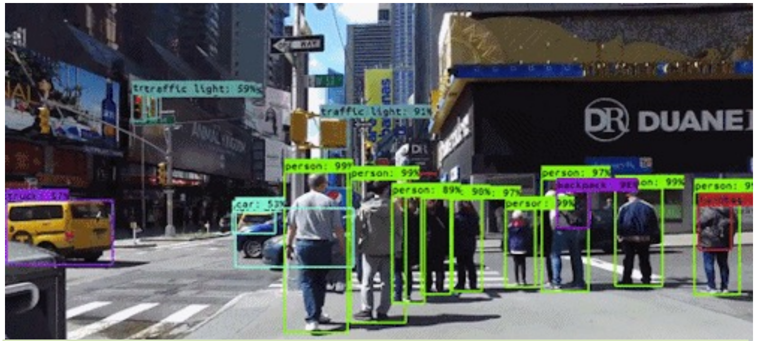
Object | *Light* **Input** → **Sensor**: Camera System → **CV Algorithm**: Object Detector → "Car" "Person" … **Output** → **Decision Algorithm**: Decision System

STOP

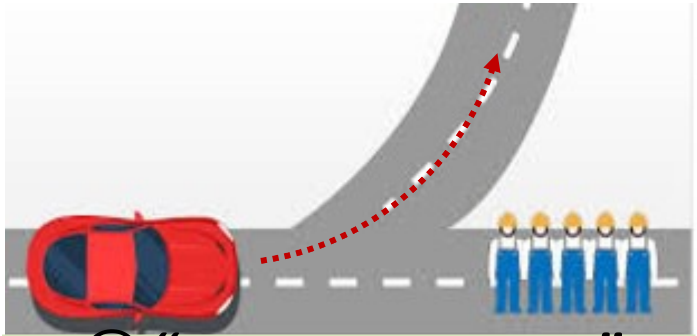
Sensing → Perception → Decision-making



① ***Camera sensing***



② ***Pedestrian detected***



③ ***"Take a turn" or "Stop"***

Poltergeist Attack

Objects

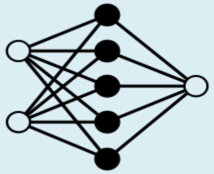
Light →

Sensors

Camera System

Digital Images

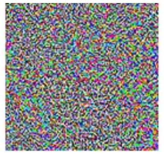
CV Algorithms



Object Detector

① **Physical** stickers or patterns on the objects
[K. Eykholt et al, CVPR’18]
[Y. Zhao et al, CCS’19]
…

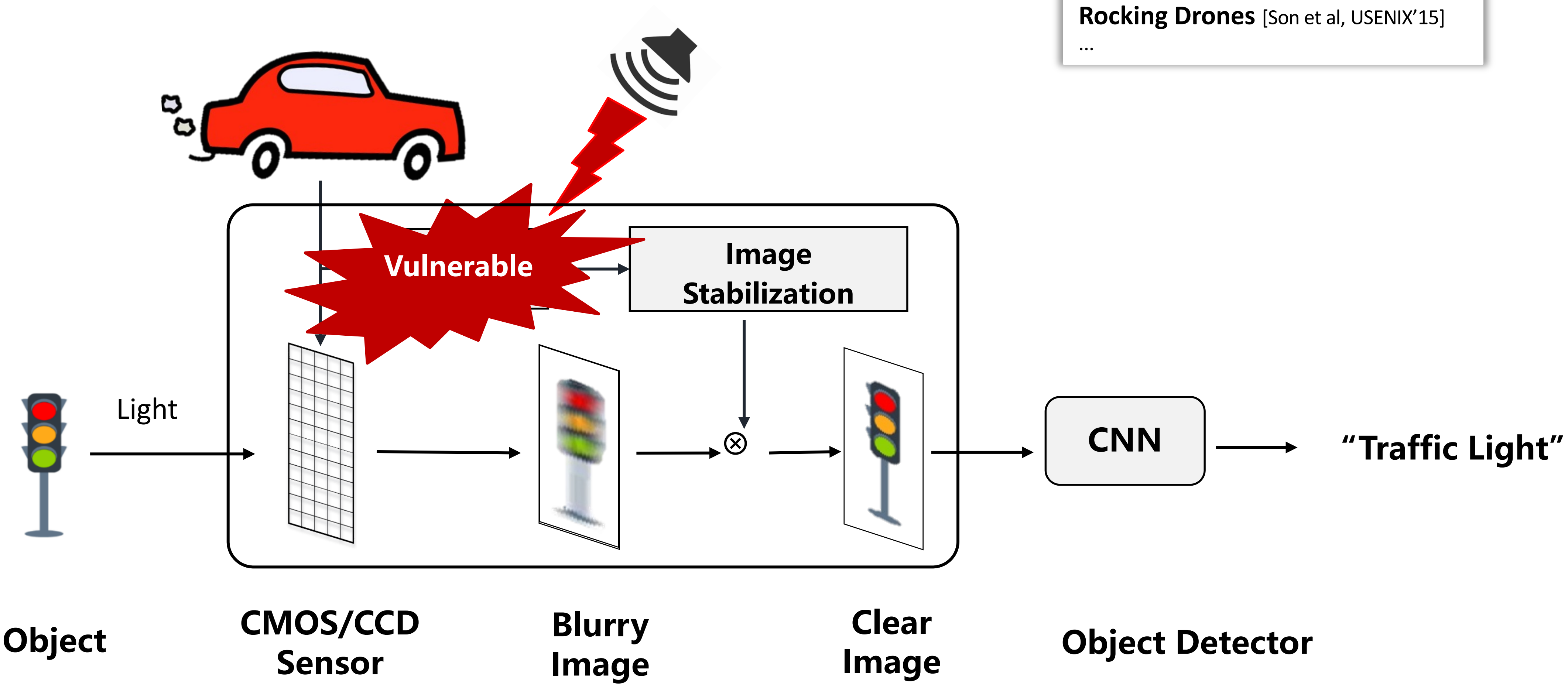
② **Digital** pixel perturbation on the images
[S.-M. Moosavi-Dezfooli et al, CVPR’16]
[N. Carlini and D. Wagner, S&P’17]
…



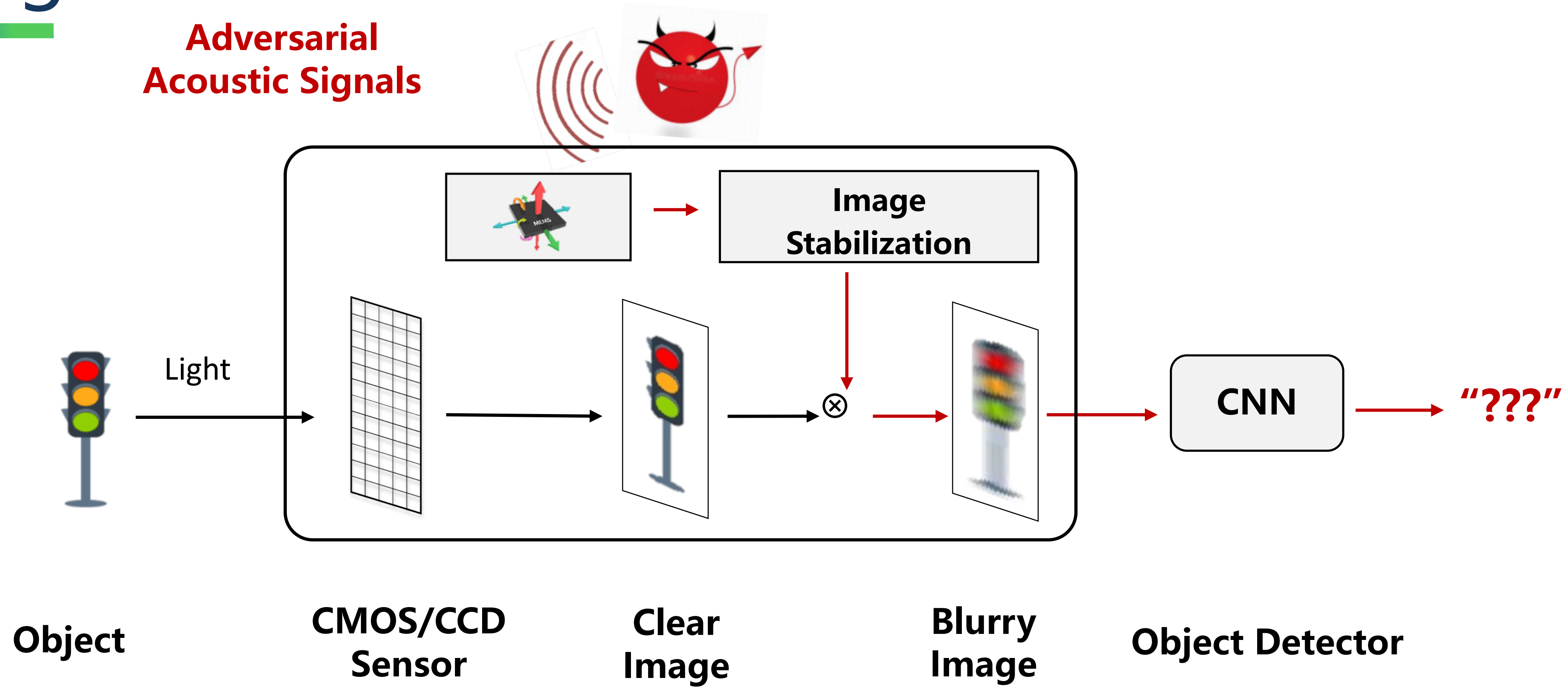
Can we achieve adversarial examples by attacking sensors?

Poltergeist Attack

Walnut [T. Trippel et al, EuroS&P'17]
Rocking Drones [Son et al, USENIX'15]
...



Poltergeist Attack



Adversarial examples by injecting acoustic signals

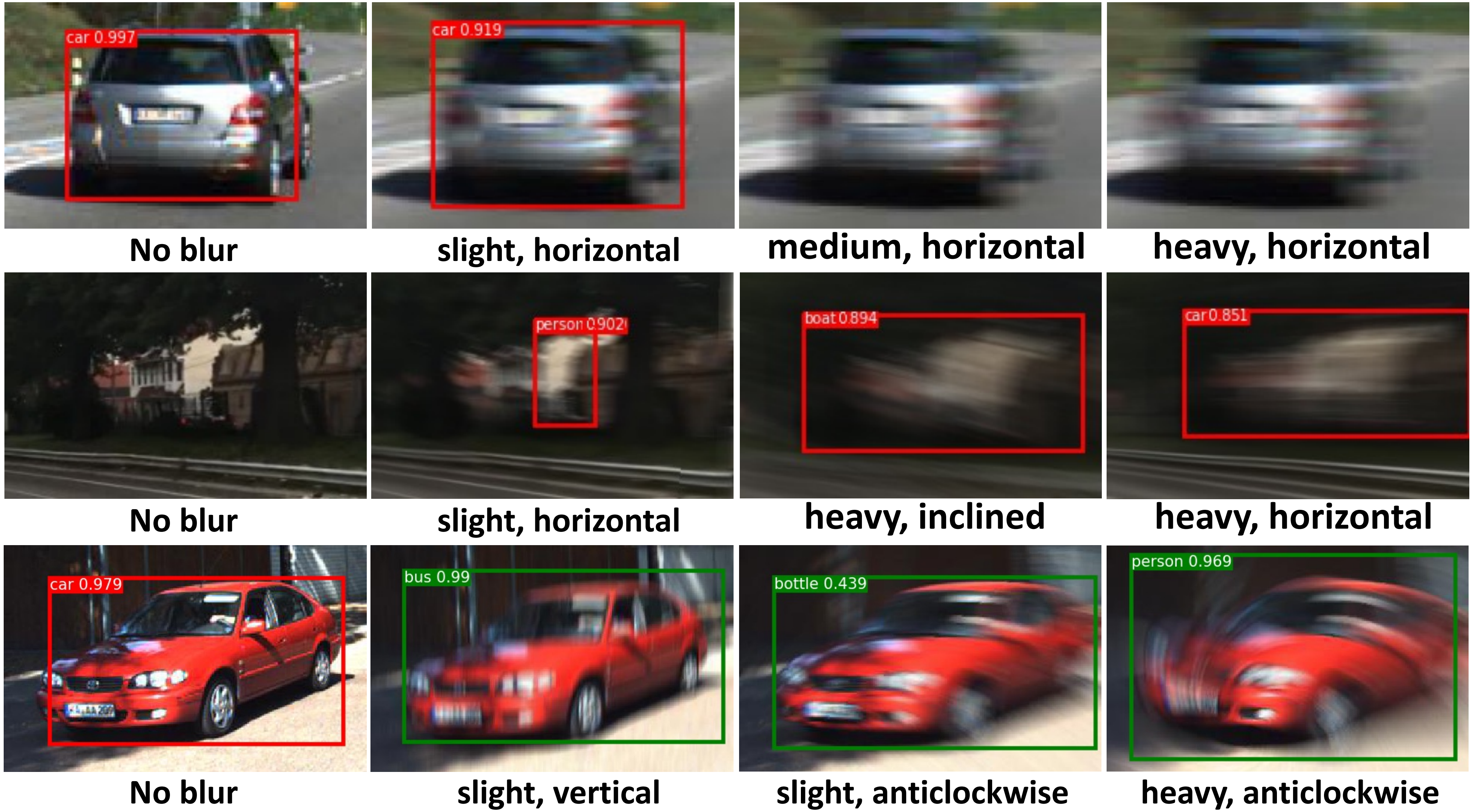
Xiaoyu Ji; Yushi Cheng; Yuepeng Zhang; Kai Wang; Chen Yan; Wenyuan Xu; Kevin Fu. Poltergeist: Acoustic Adversarial Machine Learning against Cameras and Computer Vision IEEE S&P21'

Poltergeist Attack

Hiding
“A” → None

Creating
None → “A”

Altering
“A” → “B”



Real-world Evaluation

- **Target:** Samsung S20 smartphone in a moving vehicle
- **Attack device:** Ultrasonic Speaker
- **Scenes:**
  - City Lane
  - City Crossroad
  - Tunnel
  - Campus Road

Real-world Attack Videos

Altering car into person

Creating a truck

Hiding the car

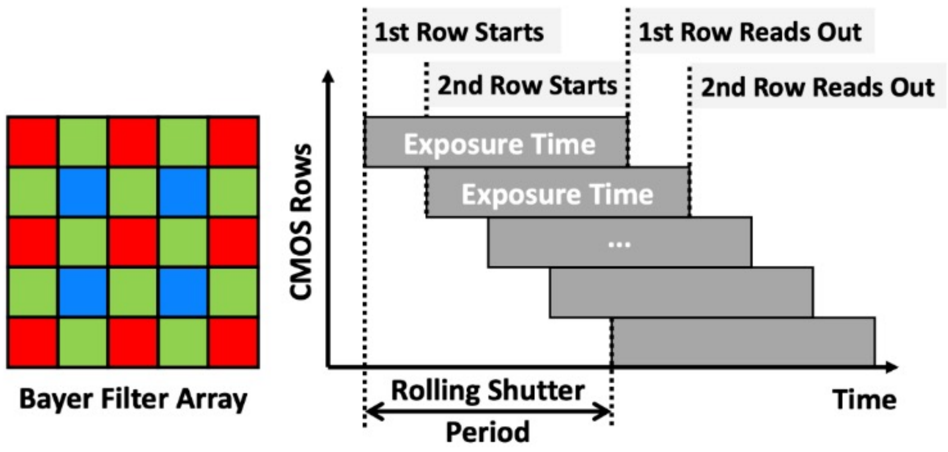
https://github.com/USSLab/PoltergeistAttack

Ample in CV: Rolling Color

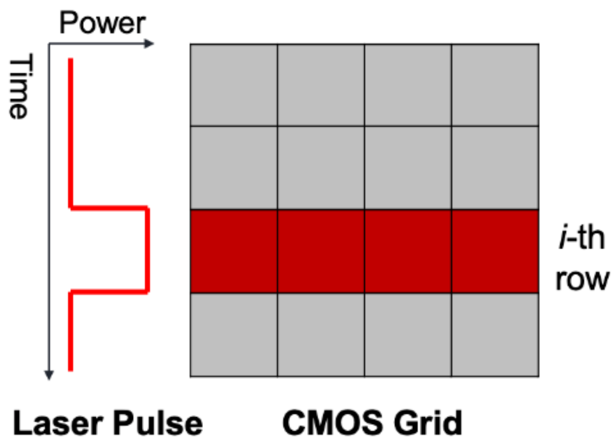
Rolling Colors: using laser to fool traffic light recognition

- The **rolling shutter** mechanism in CMOS cameras can be exploited to inject color stripes into the captured image using modulated laser
- An elaborate color stripe can **fool traffic light recognition** (recognize red as green or vice versa)

Rolling shutter mechanism



Attack design



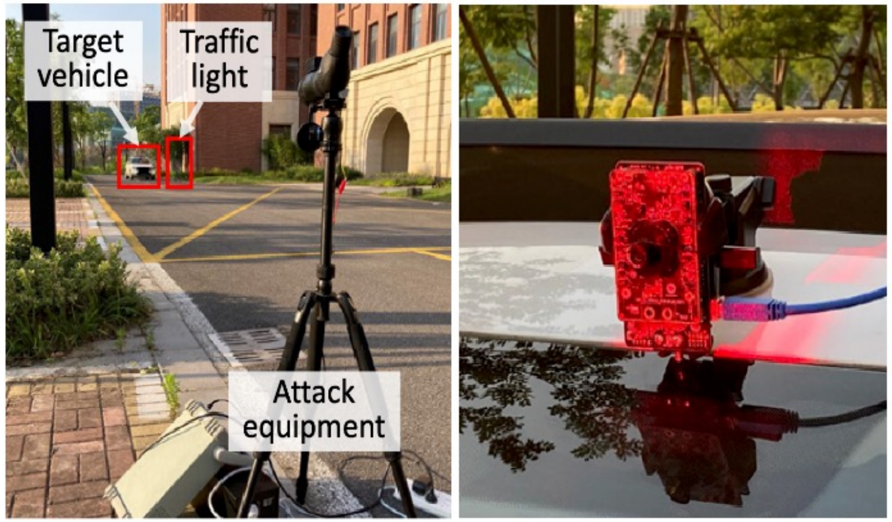
Injected color stripe

Chen Yan, Zhijian Xu, Zhanyuan Yin, Xiaoyu Ji and Wenyuan Xu, Rolling Colors: Adversarial Laser Exploits against Traffic Light Recognition", Usenix Security 2022

Ample in CV: Rolling Color

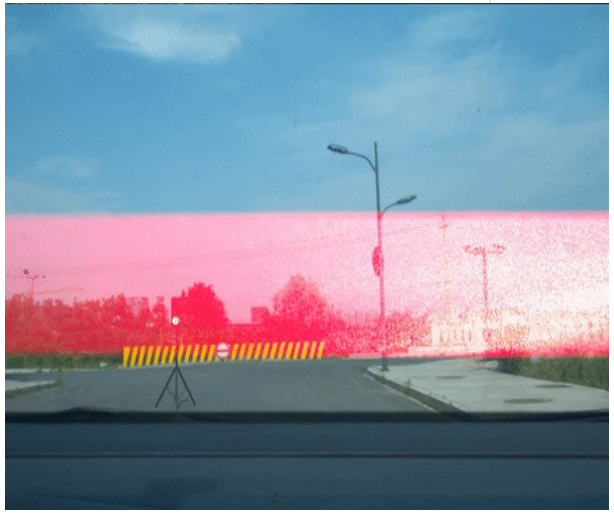
Rolling Colors: using laser to fool traffic light recognition

- Real-world attack evaluation on a moving vehicle using self-made attack equipment



Red → Green

Green → Red

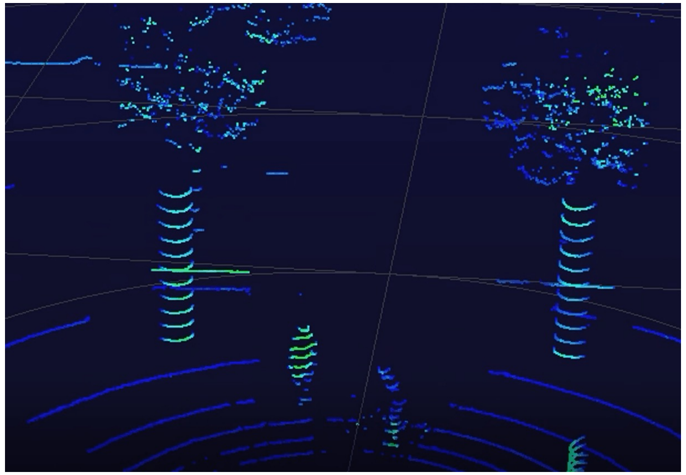


Chen Yan,; Zhijian Xu,; Zhanyuan Yin; Xiaoyu Ji and Wenyuan Xu, Rolling Colors: Adversarial Laser Exploits against Traffic Light Recognition", Usenix Security 2022

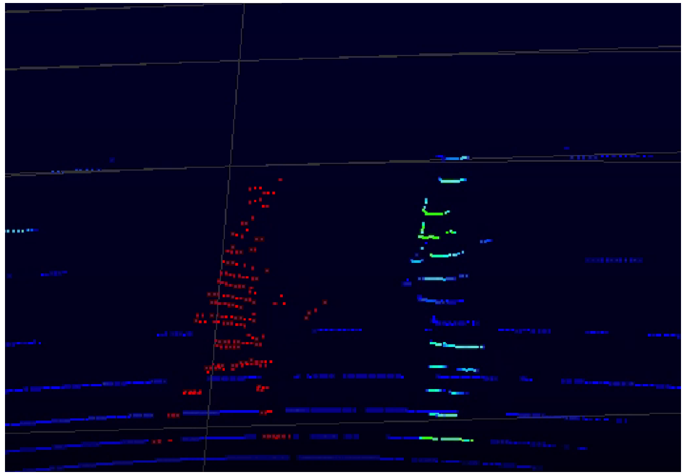
AMple in Lidar

PLA-LiDAR: using laser to spoof LiDAR-based 3D Object Detection!

- LiDAR can be spoofed due to its **periodic work cycle** and **lack of echo verification mechanism**.

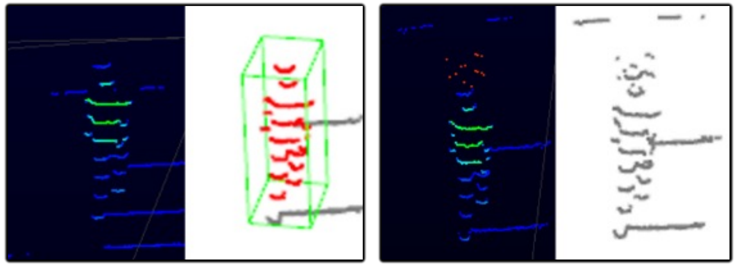


Hiding. Hide the point cloud of pedestrians and cyclists at 5 meters.



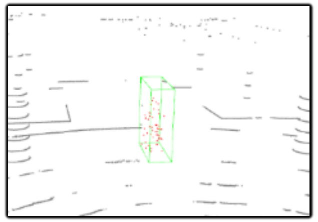
Creating. Create a fake pedestrian (left) next to the real pedestrian (right).

Optimization-based Attack



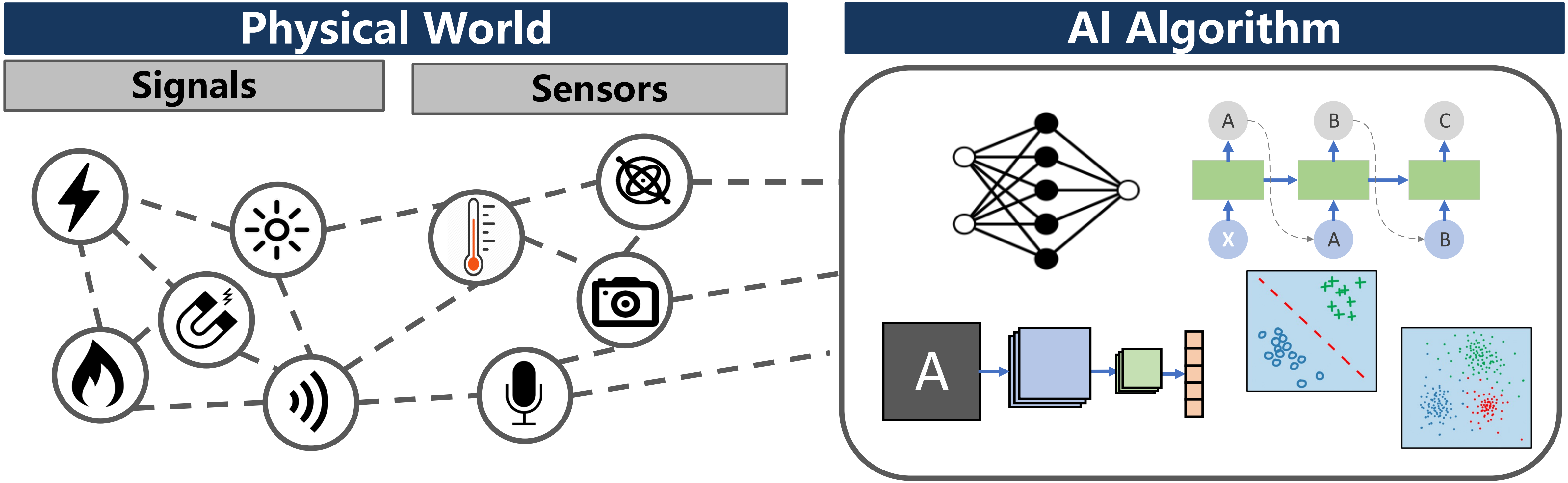
benign spoofing

Hiding. Only a dozen adversarial points can keep the pedestrian undetected.



Creating. Just a few dozen adversarial points can be detected as a pedestrian.

Other types of AMple attacks?



Inject alternative signals covertly and affecting systems

Countermeasures

Simple Attack	AMpLe Attack
Solutions for Sensor vulnerabilities ➢ Passive vs. Active ➢ Microprocessors should not blindly trust sensors ➢ Rethink ICs and hardware-software APIs **Solutions for AI vulnerabilities** ➢ Model: ● Adversary training & Gradient hiding ➢ Input: ● Detection & Rectification & Input denoising ……	**MEMS Inertial Sensors Safeguarding** ➢ Acoustic Isolation ➢ Secure Low-pass Filter, amplifier **Image Stabilization Techniques** ➢ Other types of Digital Image Stabilization **Object Detection Algorithms** ➢ Input Image De-blur ➢ Detection Model Improvement **Sensor Fusion Techniques** ➢ LiDARs, radars combined with cameras ……

Testing is important!


News | Semiconductors | Devices

1 Mar 2018 | 13:00 GMT

Finally, a Likely Explanation for the "Sonic Weapon" Used at the U.S. Embassy in Cuba

Researchers say bad engineering, not a deliberate attack, may be to blame

By **Jean Kumagai**

HUMAN EARS

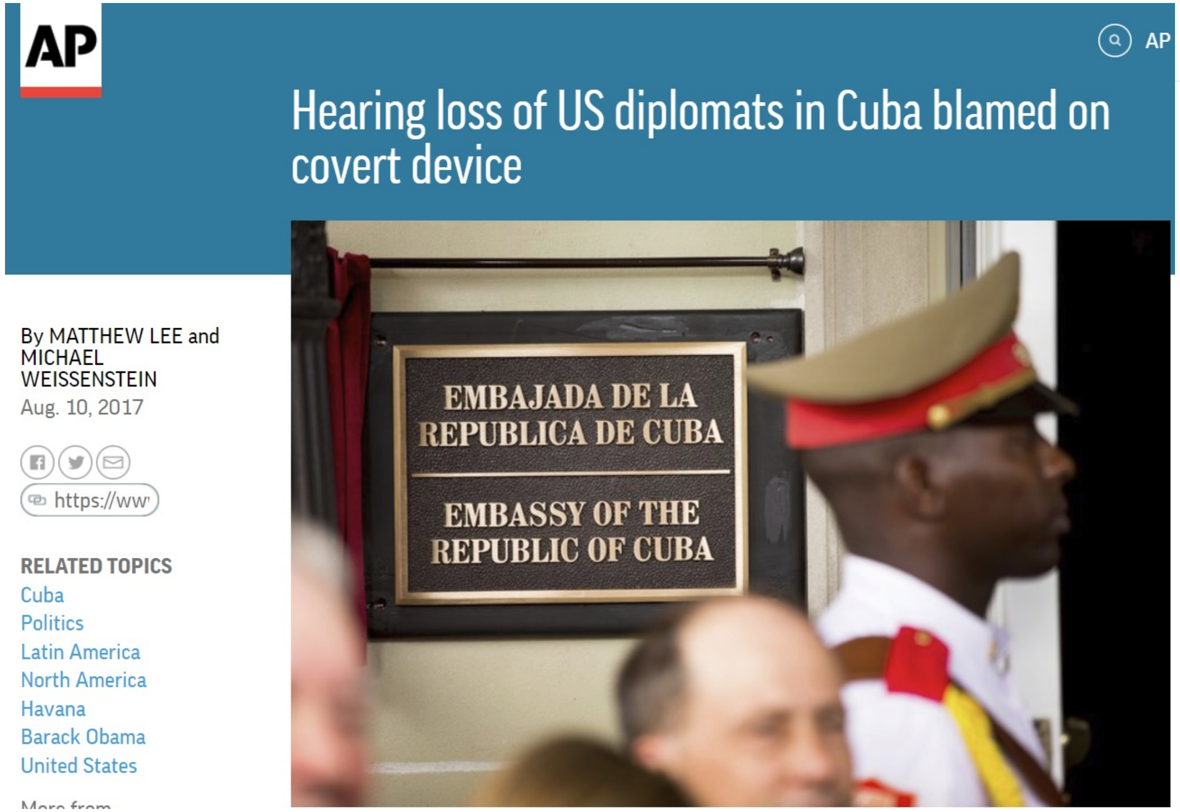
On Cuba, diplomats, ultrasound, and intermodulation distortion

Chen Yan, Kevin Fu, Wenyuan Xu

Computers in Biology and Medicine 104, 250-266

The most dangerous sound?

Two dozen US embassy workers in Cuba suffered headaches, hearing loss, and brain swelling—but no one knows why


AP

Hearing loss of US diplomats in Cuba blamed on covert device

By MATTHEW LEE and MICHAEL WEISSENSTEIN
Aug. 10, 2017

RELATED TOPICS
Cuba
Politics
Latin America
North America
Havana
Barack Obama
United States


CNN politics
45 CONGRESS SECURITY THE NINE TRUMPMERICA 2018

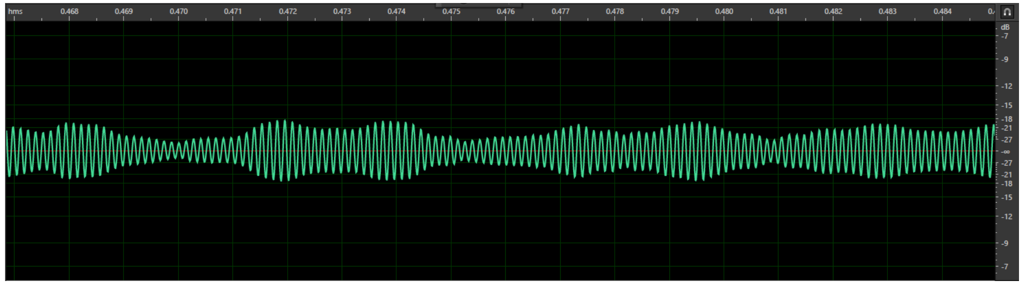
Sonic attacks in Cuba hit more diplomats than earlier reported, officials say

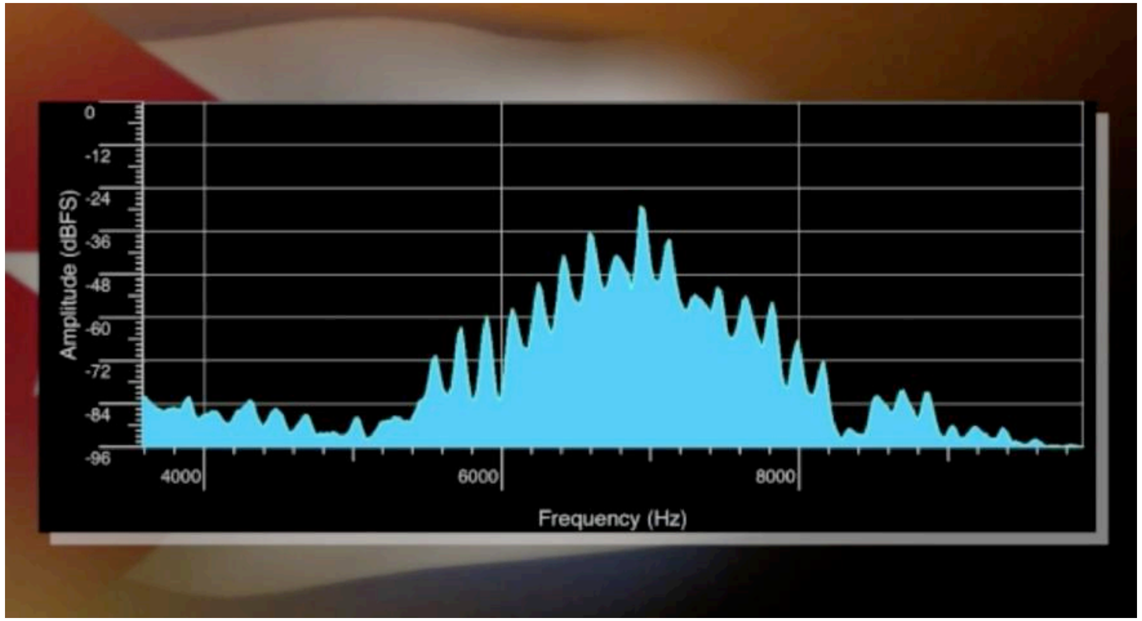
By Patrick Oppmann and Elise Labott, CNN
Updated 5:36 PM ET, Sun August 20, 2017

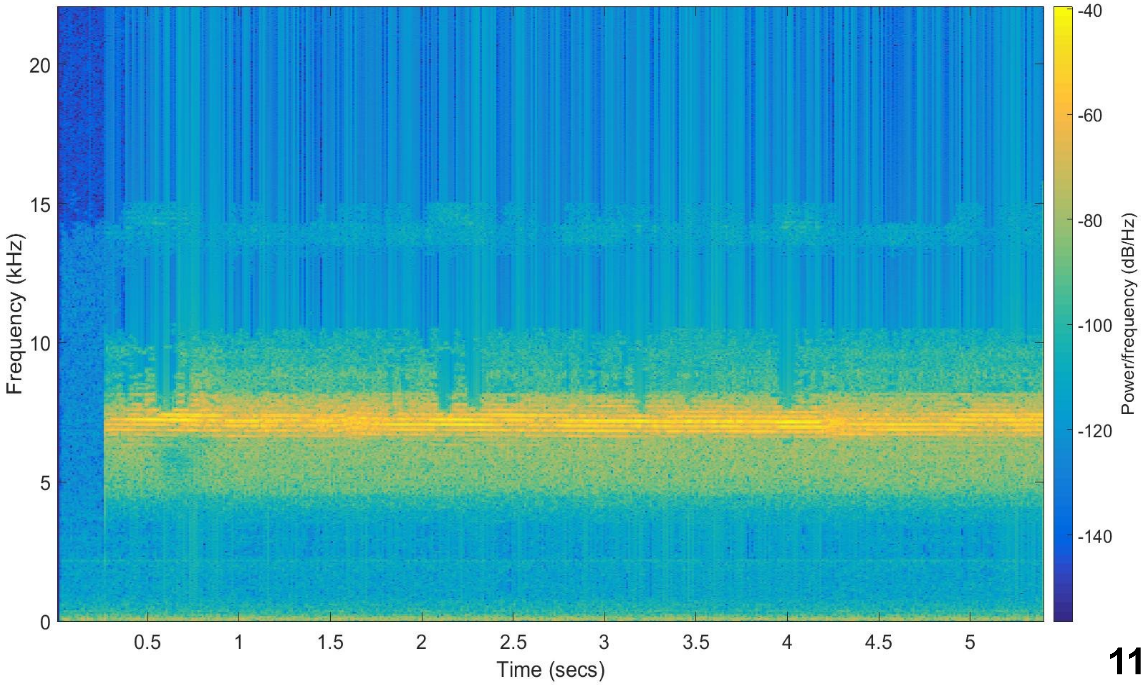
The most dangerous sound?

Reverse engineering of the recording

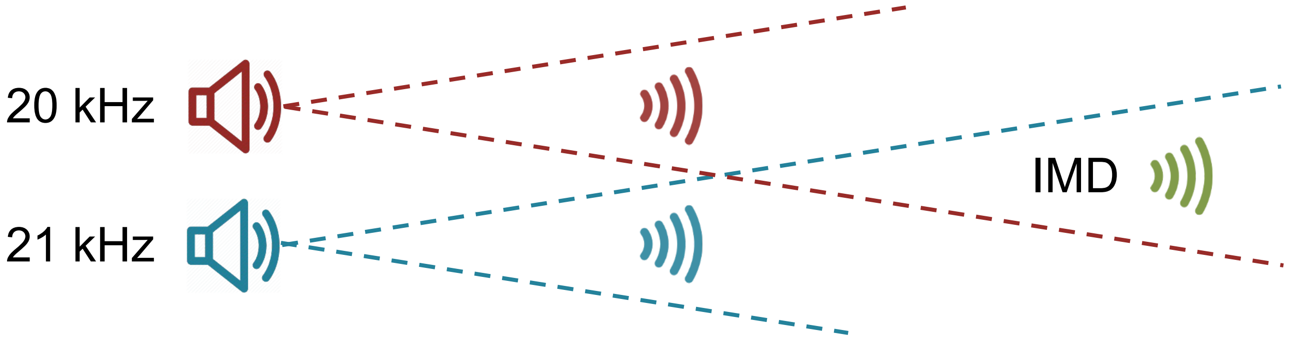
- Analysis in the time and frequency domains
- Frequency peaks separated by 180 Hz around 7kHz

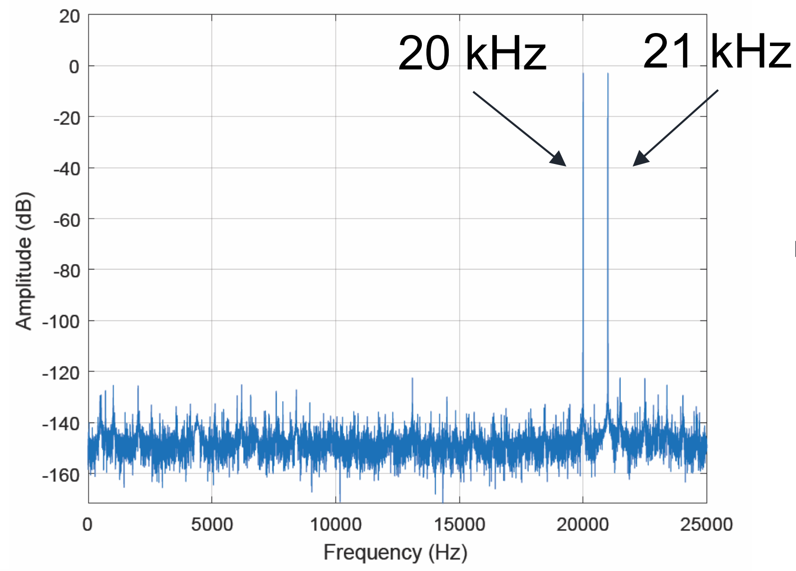


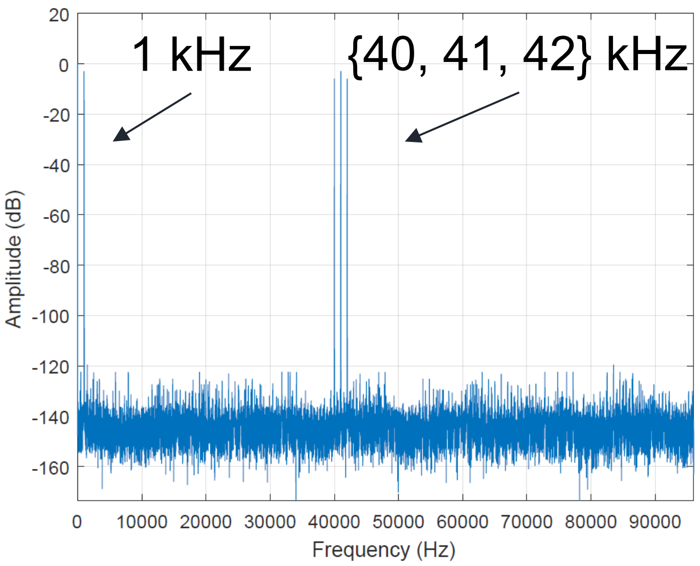




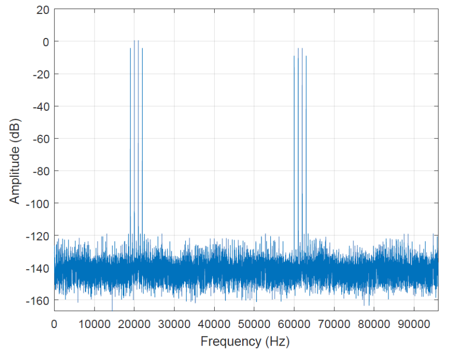
IMD of two tones



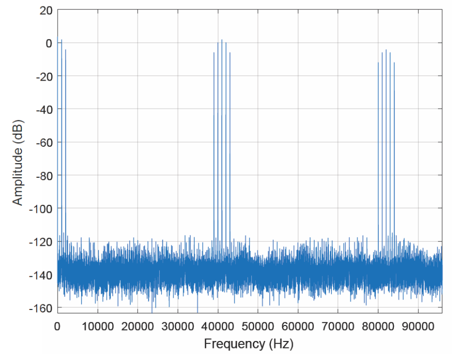




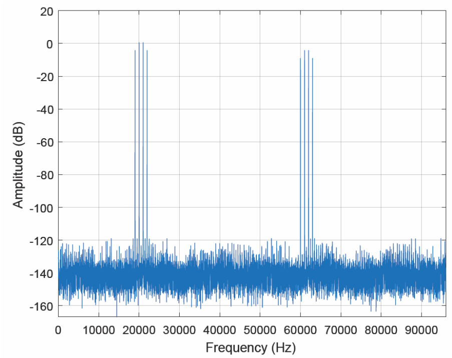
+



3th



4th

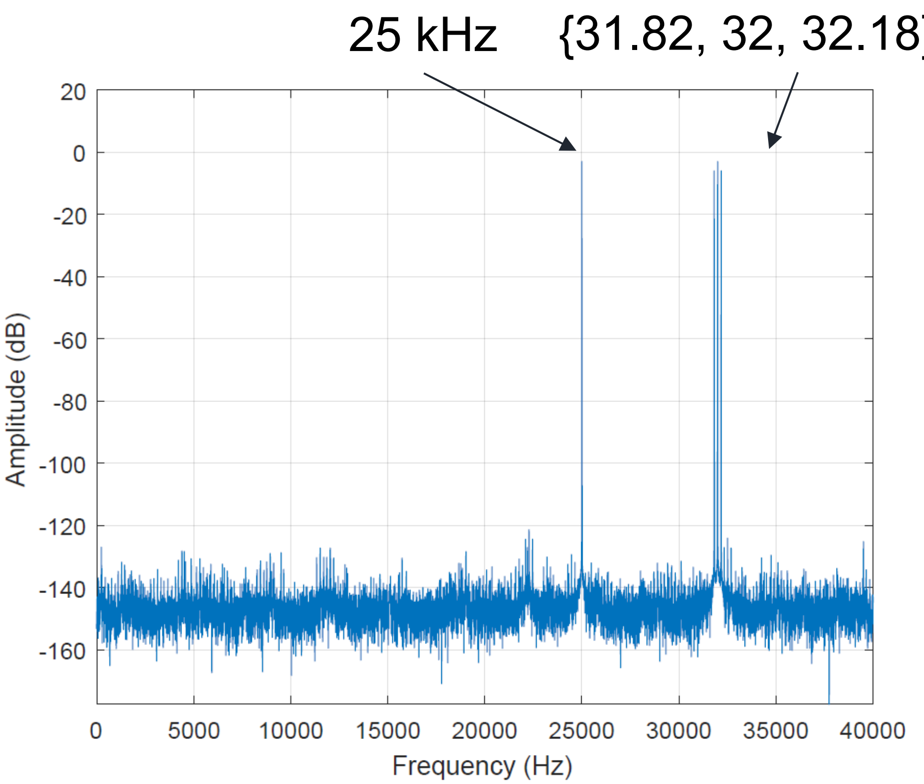


5th

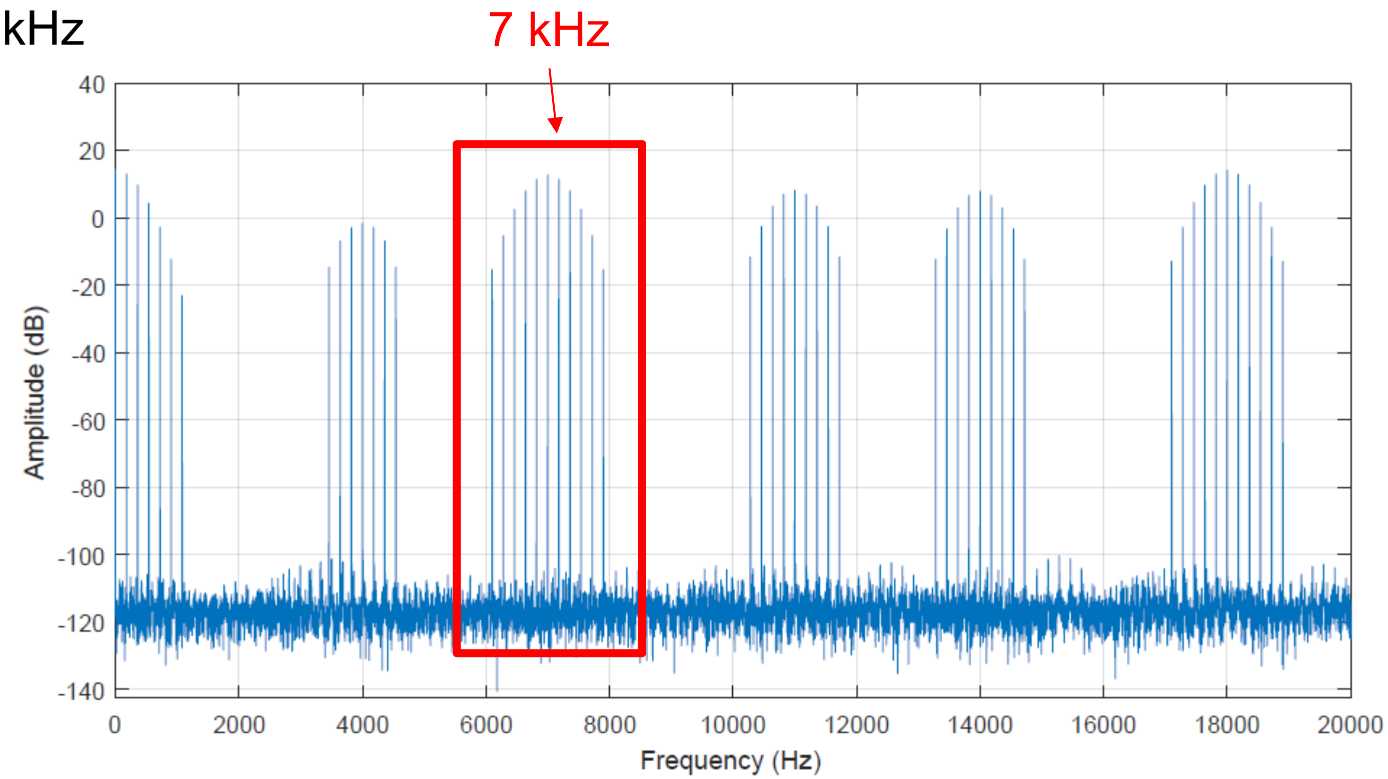
two tones $s_{out} = a_1 s_{in} + a_2 s_{in}^2 + \dots$ second order higher orders

IMD of an AM signal

- 25 kHz tone + AM signal (32 kHz carrier， 180 Hz baseband)

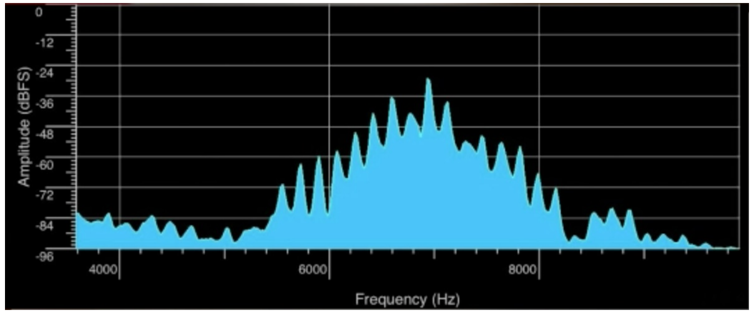


AM signal

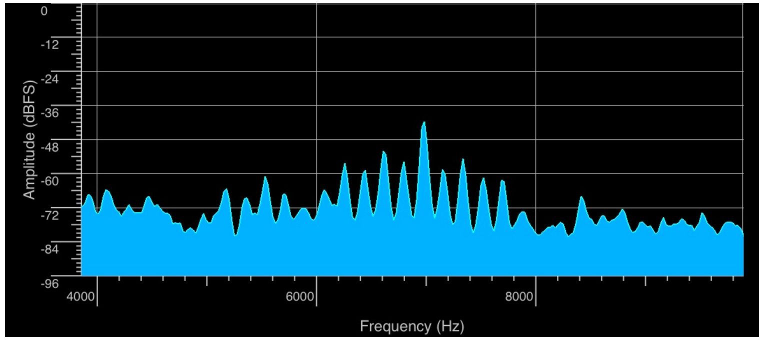


IMD under 20 kHz

Reproducing the “dangerous sound”

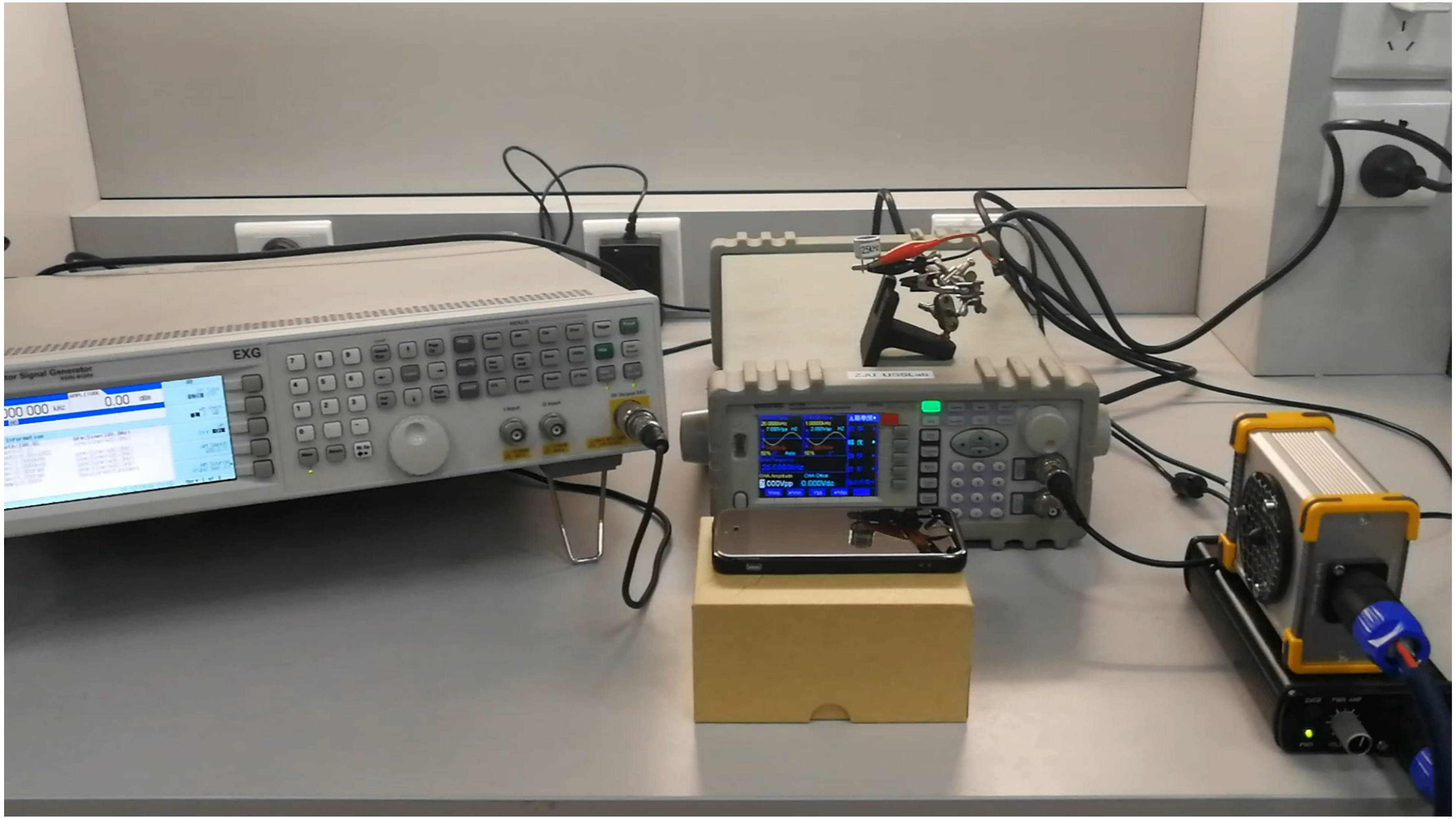


AP news

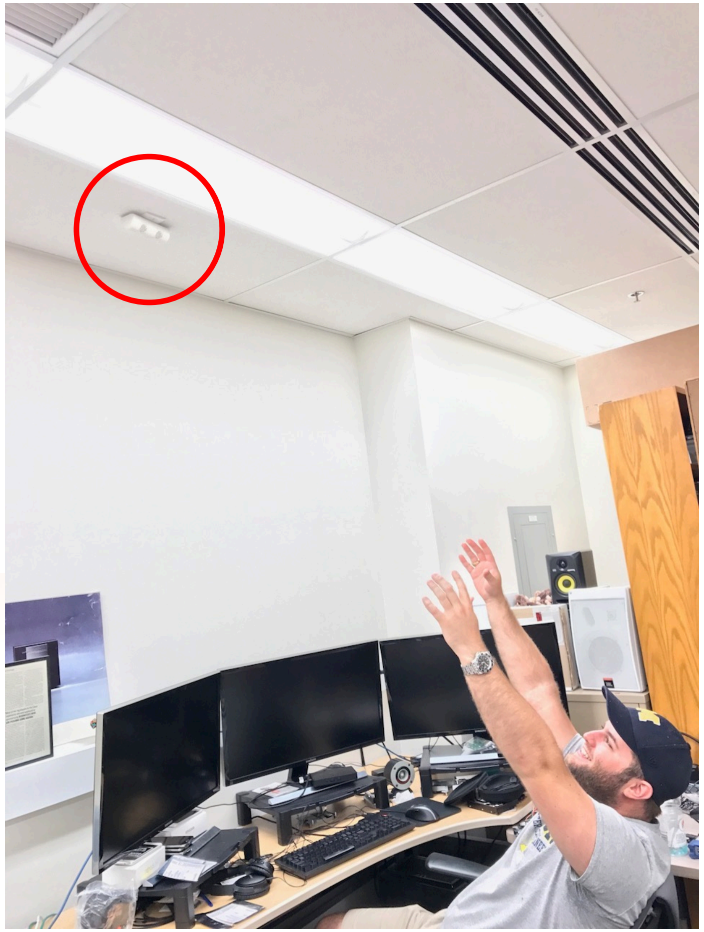


Reproduced

Reproducing the “dangerous sound”



Ultrasound sources in daily life



Conclusions: Analog is the new digital

- Analog security risks
  - Analog Sensors --- RF
  - MEMS Sensors --- Acoustic
  - Active Sensors --- Sensing principle

- Solutions
  - Microprocessors should not blindly trust sensors
  - Rethink ICs and hardware-software APIs

Questions and Answers

yanchen@zju.edu.cn

USSLab Homepage: usslab.org