Physical Security Attacks

- Inertial (WALNUT)
- GPS Spoofing (Drone)
- Hacking Pacemakers
- Inaudible Voice Commands
Mobile Security
Inaudible Voice Commands
BackDoor: Making Microphones Hear Inaudible Sounds
Microphones are everywhere
Microphones are everywhere
Microphones record audible sounds.

Speaker

I hear that

Audible sound

I record that

Microphone
Inaudible, but recordable!
Inaudible, but recordable! 

I can’t hear that

I record that
Works with unmodified devices

- Speaker
- Camera
- Smartwatch
- Laptop
- Hearing Aid
It’s not “near-ultrasound”
Exploiting fundamental nonlinearity

Microphone hardware
What can we do with it?
Application: Acoustic jammer
Application: Acoustic communication
Threat: Acoustic DOS attack
Threat: Acoustic DOS attack

Jamming hearing aids
Threat: Acoustic DOS attack

Jamming hearing aids

Blocking 911 calls
Talk outline

1. Microphone Overview
2. System Design
3. Challenges
4. Evaluation
Talk outline

1. Microphone Overview
2. System Design
3. Challenges
4. Evaluation
Microphone working principle

Diaphragm Amplifier → Filter → ADC
Microphone working principle

- Diaphragm
- Amplifier
- Filter
- ADC
Microphone working principle

Amplitude

Frequency

Diaphragm Amplifier  Filter  ADC
Microphone working principle

Frequency

Amplitude

10k 20k 30k 40k 50k 60k 70k 80k 90k 100k

Diaphragm Amplifier Filter ADC
Microphone working principle

Amplitude

Frequency

10k 20k 30k 40k 50k 60k 70k 80k 90k 100k

Diaphragm Amplifier Filter ADC
Microphone working principle

Amplifier
Filter
ADC

Amplitude

Frequency

10k 20k 30k 40k 50k 60k 70k 80k 90k 100k

Diaphragm Amplifier Filter ADC
Microphone working principle

Amplifier
Filter
ADC

Diaphragm

Frequency
10k 20k 30k 40k 50k 60k 70k 80k 90k 100k

Amplitude
Microphone working principle

\[ V_{out} = a_1 V_{in} \]

\[ V_{out} = a_1 V_{in} + a_2 V_{in}^2 + a_3 V_{in}^3 + \ldots \]

Amplifier
Microphone working principle

\[ V_{out} = a_1 V_{in} \]

\[ V_{out} = a_1 V_{in} + a_2 V_{in}^2 \]
Microphone working principle

\[ V_{out} = a_1 V_{in} \]

\[ V_{out} = a_1 V_{in} + a_2 V_{in}^2 \]

Frequency range:
- 10k to 100k
Talk outline

① Microphone Overview
② System Design
③ Challenges
④ Evaluation
Exploiting amplifier non-linearity

Amplitude

Microphone filter

F<sub>1</sub> = 50kHz
F<sub>2</sub> = 40kHz

Frequency

10k 20k 30k 40k 50k 60k 70k 80k 90k 100k
Exploiting amplifier non-linearity

$V_{out} = a_1 V_{in} + a_2 V_{in}^2$

$(\sin F_1 + \sin F_2)^2 = \cos 2F_1 + \cos 2F_2 + \cos (F_1 + F_2) + \cos (F_1 - F_2)$
Exploiting amplifier non-linearity

Amplitude

Frequency

10k 20k 30k 40k 50k 60k 70k 80k 90k 100k

$V_{out} = a_1 V_{in} + a_2 V_{in}^2$

$(\sin F_1 + \sin F_2)^2 = \cos 2F_1 + \cos 2F_2 + \cos (F_1 + F_2) + \cos (F_1 - F_2)$

$F_1 = 50\text{kHz}$

$F_2 = 40\text{kHz}$
Exploiting amplifier non-linearity

Microphone filter

$$V_{out} = a_1 V_{in} + a_2 V_{in}^2$$

$$\left( \sin F_1 + \sin F_2 \right)^2 = \cos 2F_1 + \cos 2F_2 + \cos (F_1 + F_2) + \cos (F_1 - F_2)$$

F_1 = 50\text{kHz}
F_2 = 40\text{kHz}
Exploiting amplifier non-linearity

\[ V_{out} = a_1 V_{in} + a_2 V_{in}^2 \]

\[ (\sin F_1 + \sin F_2)^2 = \cos 2F_1 + \cos 2F_2 + \cos (F_1 + F_2) + \cos (F_1 - F_2) \]

Microphone filter

F_1 = 50kHz
F_2 = 40kHz
Exploiting amplifier non-linearity

Amplitude

(F₁ - F₂)

Microphone filter

F₁ = 50kHz
F₂ = 40kHz

Frequency

10k 20k 30k 40k 50k 60k 70k 80k 90k 100k
Exploiting amplifier non-linearity

(F_1 - F_2)

Microphone filter

F_1 = 50kHz
F_2 = 40kHz

Amplitude

10k  20k  30k  40k  50k  60k  70k  80k  90k  100k

Frequency
Talk outline

① Microphone Overview
② System Design
③ Challenges
④ Evaluation
Reminder of Modulation
Modulation

Analog

\[ \cos(2\pi ft) \]

Amplitude

Frequency

10kHz

20MHz

Amplitude

Frequency

20MHz + 10kHz

shift signal
Why is Modulation useful?

1. Interference, Technology Co-existence
2. Spectrum Access (Legal)
3. Antenna size (wavelength/4)

WiFi? LTE? 5G?
Challenges

Problem: speaker has non-linearities => Audible sound

\[ S_{AM} = a \cdot \sin(\omega_m t) \cdot \sin(\omega_c t) \]

Amplitude modulation

\[ S_{out, AM} = A_2 \{a \sin(\omega_m t) \cdot \sin(\omega_c t)\}^2 \]

\[ = -A_2 \frac{a^2}{4} \left\{ \cos(\omega_c t - \omega_m t) - \cos(\omega_c t + \omega_m t) \right\}^2 \]

\[ = -A_2 \frac{a^2}{4} \cos(2\omega_m t) + (\text{terms with frequencies above } \omega_c \text{ and DC}) \]
Challenges

Frequency modulation

\[ S_{FM} = \sin(\omega_c t + \beta \sin(\omega_m t)) \]

Ultrasonic speaker
Challenges

Frequency modulation

\[ S_{FM} = \sin(\omega_c t + \beta \sin(\omega_m t)) \]

Ultrasonic speaker
Challenges

Frequency modulation

\[ S_{FM} = \sin(\omega_c t + \beta \sin(\omega_m t)) \]

Ultrasonic speaker

\[ S_{FM}^2 \sim 1 + \cos(2\omega_c t + \text{other terms}) \]

Problem: microphone can’t measure inaudible sound
Add another speaker
How do we structure its signal?

\[ S_{FM} = \sin(\omega_c t + \beta \sin(\omega_m t)) \]
Talk outline

1. Microphone Overview
2. System Design
3. Challenges
4. Evaluation
Hardware generalizability

![Diagram showing BackDoor Signal (dB) for different devices]

- Hearing aids
- Camera
- iPhone
- Android phone
- Smart-watch
- Laptop

BackDoor Signal (dB)

- 40 kHz
- 50 kHz

Devices
Implementation

Communication prototype

Jammer prototype
Communication performance

FM data packets

4kbps up to 1 meter

More power can increase the distance
Jamming performance

BackDoor jammer
Jamming performance

BackDoor jammer

Spy microphone
Jamming performance

BackDoor jammer

Spy

microphone
Jamming performance

BackDoor jammer

Spy microphone
Jamming performance

BackDoor jammer

Spy microphone
Jamming performance

2000 spoken words

BackDoor jammer

Jammed recording
Jamming performance

2000 spoken words

BackDoor jammer

Jammed recording

Human listener

Speech recognition
Jamming performance

2000 spoken words

BackDoor jammer

% of legible words

Human listener

Speech recognition

Jammed recording
#### Jamming performance

- **Human users**
- **Automatic speech recognition**

<table>
<thead>
<tr>
<th>Jamming distance</th>
<th>Legibility of words (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0m</td>
<td></td>
</tr>
<tr>
<td>1.5m</td>
<td></td>
</tr>
<tr>
<td>2.0m</td>
<td></td>
</tr>
<tr>
<td>2.5m</td>
<td></td>
</tr>
<tr>
<td>3.0m</td>
<td></td>
</tr>
<tr>
<td>3.5m</td>
<td></td>
</tr>
<tr>
<td>4.0m</td>
<td></td>
</tr>
<tr>
<td>4.5m</td>
<td></td>
</tr>
<tr>
<td>5.0m</td>
<td></td>
</tr>
<tr>
<td>No Jam</td>
<td></td>
</tr>
</tbody>
</table>

- **Legibility of words (%)**
  - The graph shows the legibility of words (%) against jamming distance for both human users and automatic speech recognition systems. The legibility decreases as the jamming distance increases.
How would you design a system to secure against this attack?
Summary

• IoT Security: both digital and analog
• “Sensor” security & attacks:
  - Mobile acoustic attacks (inaudible voice commands)
  - Analog Sensor attacks (on MEMS accelerometers)
  - Drone Security (Spoofing GPS)
  - Medical Security (Hacking Pacemakers)
• Modulation schemes
  - AM
  - FM
  - Inter-modulation
• Fundamentals have implications beyond IoT (e.g., Cuban “acoustic attack”)
MUTE: Bringing IoT to Noise Cancellation

Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, Romit Roy Choudhury
University of Illinois at Urbana-Champaign

ACM SIGCOMM 2018