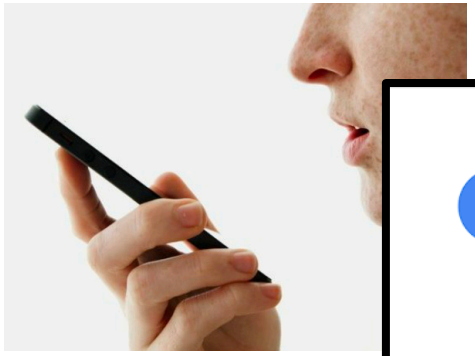
# SurfingAttack: Interactive Hidden Attack on Voice Assistants Using Ultrasonic Guided Waves

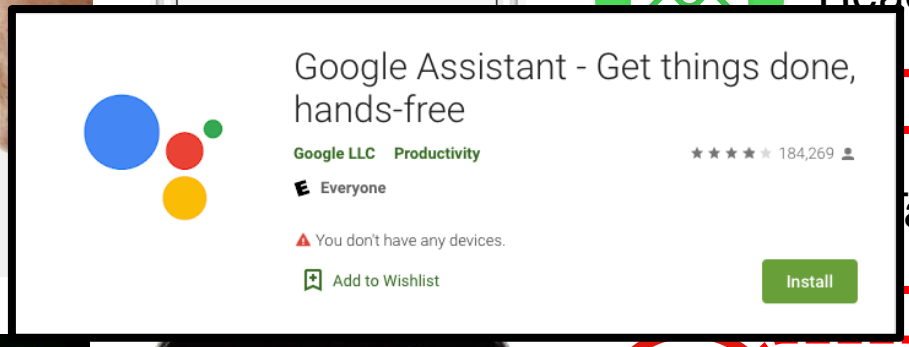
**Qiben Yan**[1], Kehai Liu[2], Qin Zhou[2]
Hanqing Guo[1], Ning Zhang[3]

*[1]Michigan State University, [2]University of Nebraska-Lincoln,*
*[3]Washington University in St. Louis*



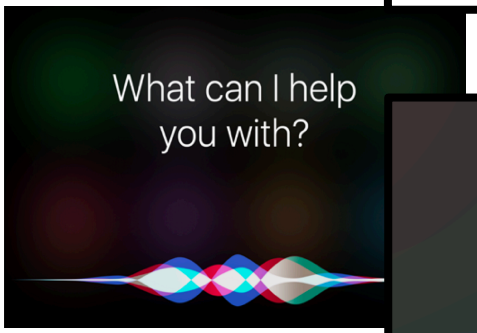
# Voice Assistants

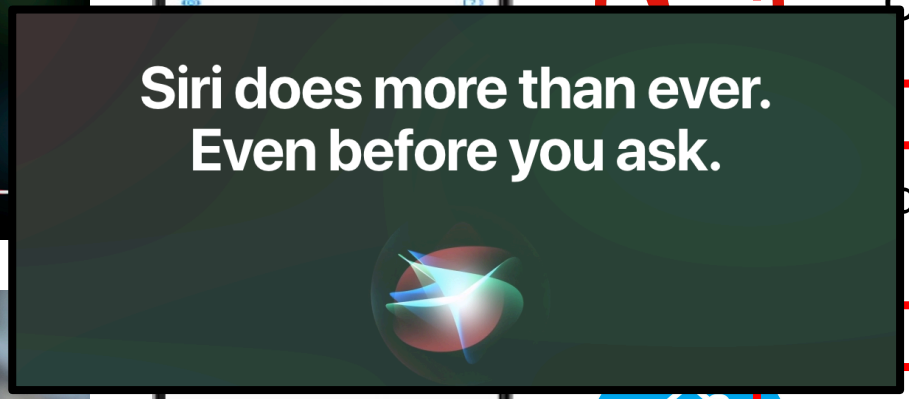





Read my message

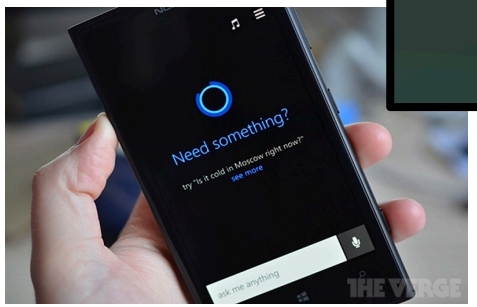
ake a selfie




Calling Sam

d a message to Sam



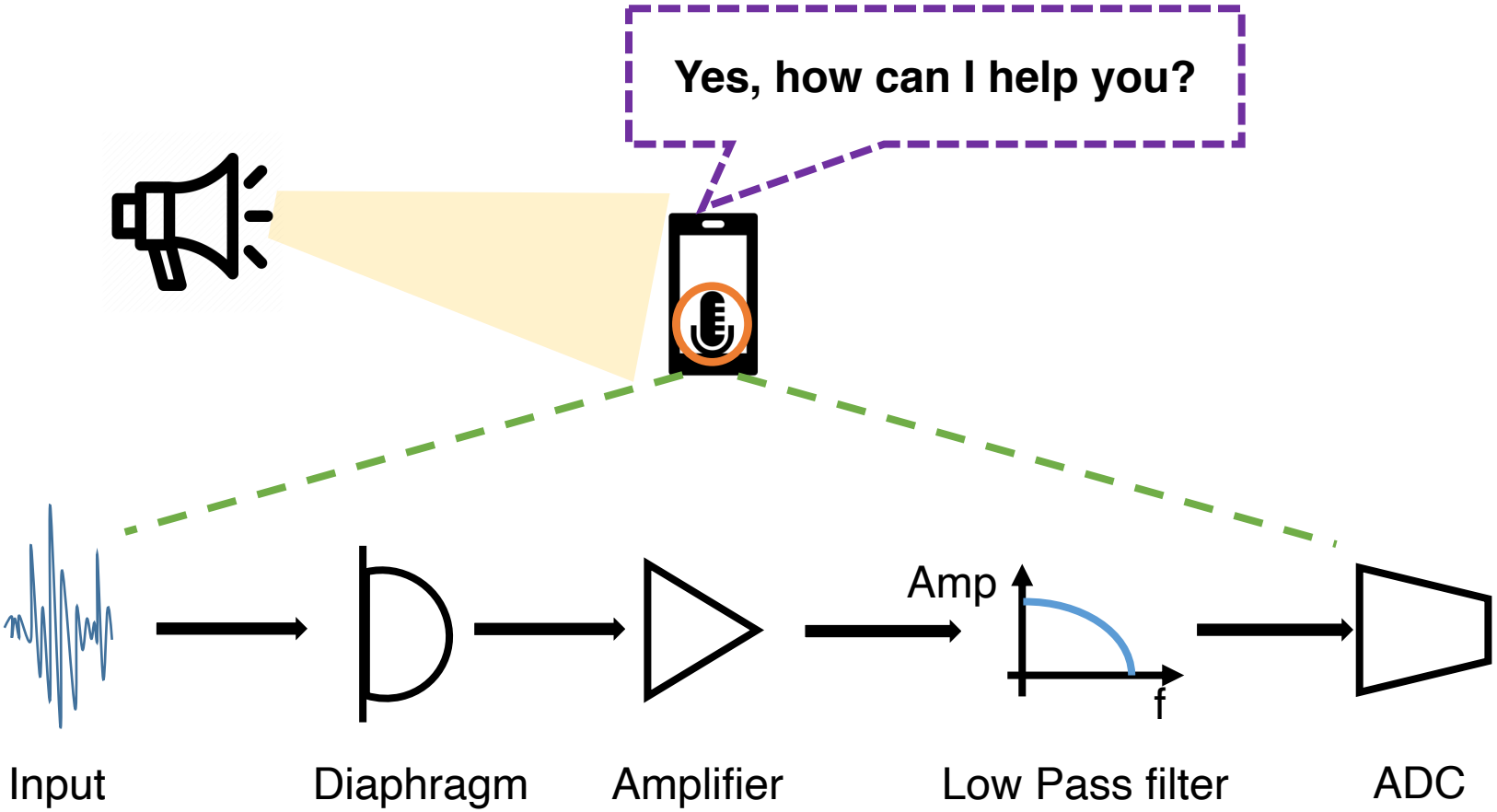




Open my garage door

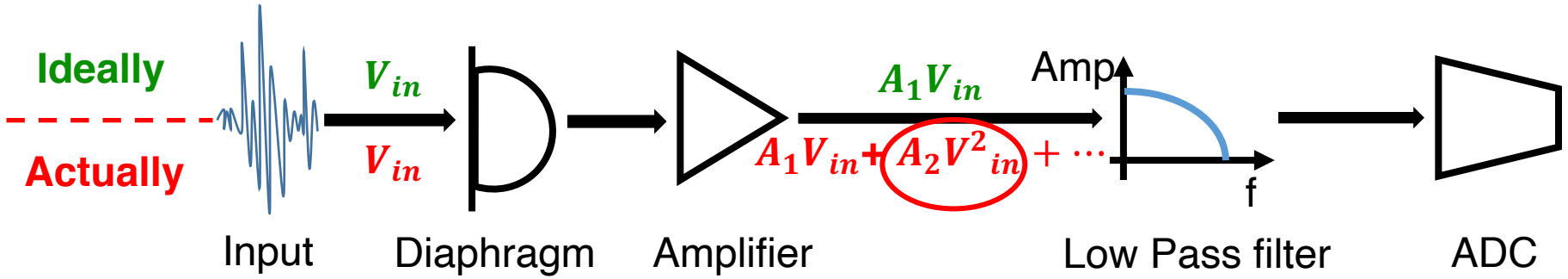
**They are not safe!**

# Over-the-air Inaudible Attack



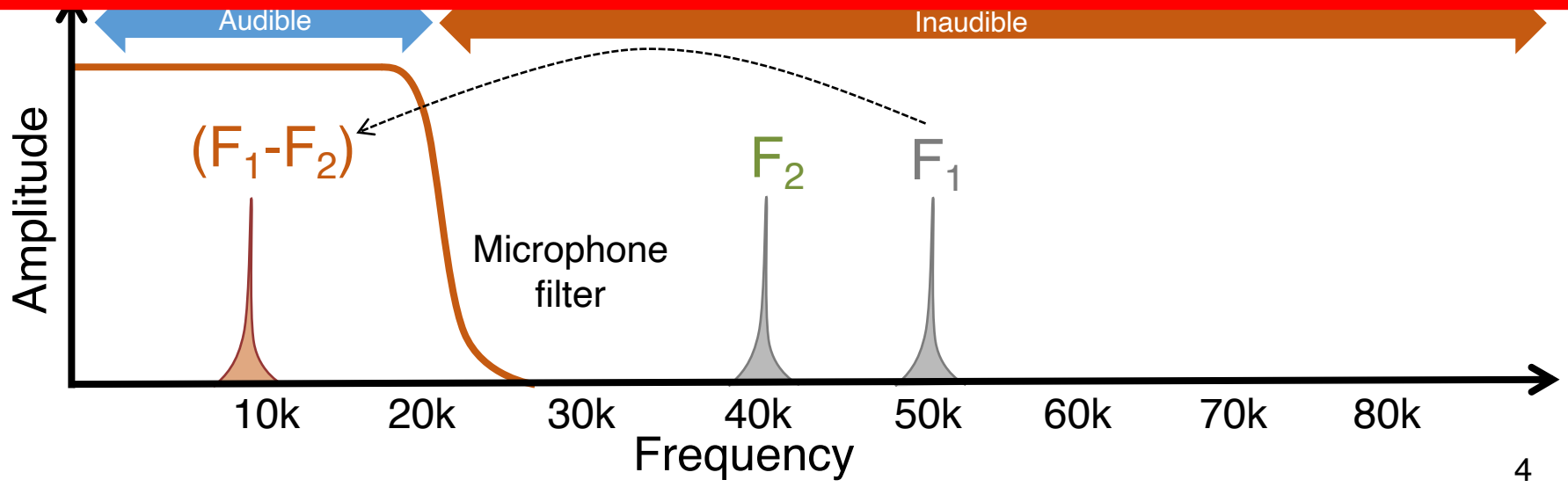

[1] Backdoor: Making microphones hear inaudible sounds. Roy, N. et al., MobiSys 2017.
[2] Dolphinattack: Inaudible voice commands. Zhang, G. et al., CCS 2017.
[3] Inaudible voice commands: The long-range attack and defense. Roy, N., et al. NDSI 2018.

# Over-the-air Inaudible Attack




$$V_{in} = SinF_1 + SinF_2$$

***How about Inaudible Attack through other media?***

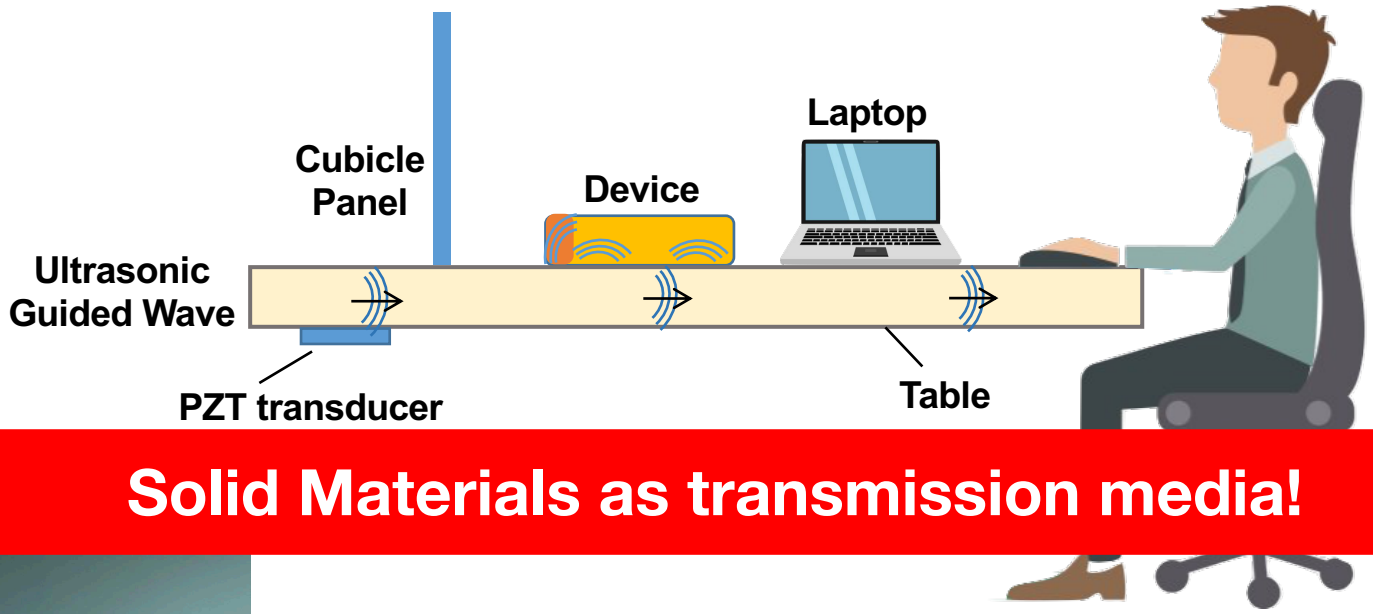





Courtesy: modified image from “Inaudible voice commands: The long-range attack and defense”

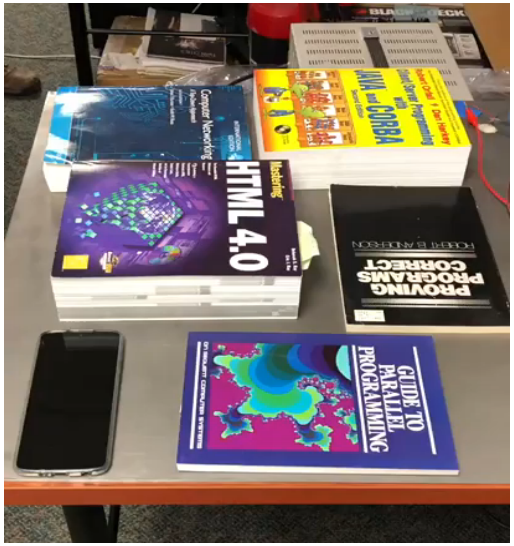
# Inaudible Attack through other media (**a table**)

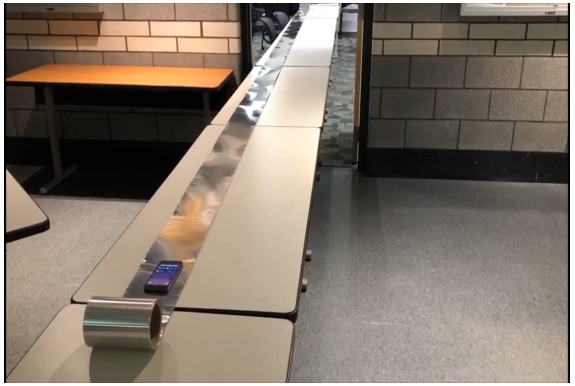
# Typical Attack Setup

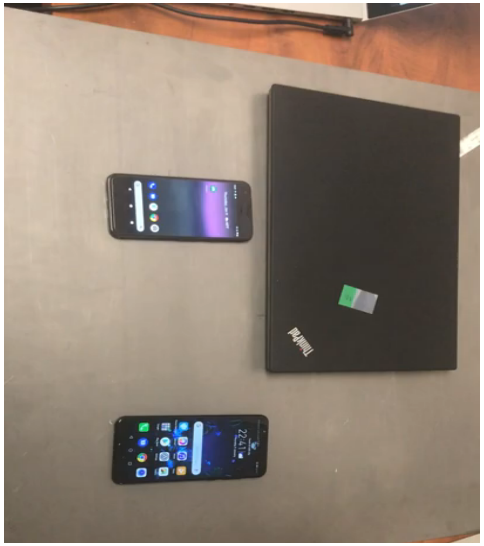



**Solid Materials as transmission media!**

# *SurfingAttack: Surfing Waves in Materials*



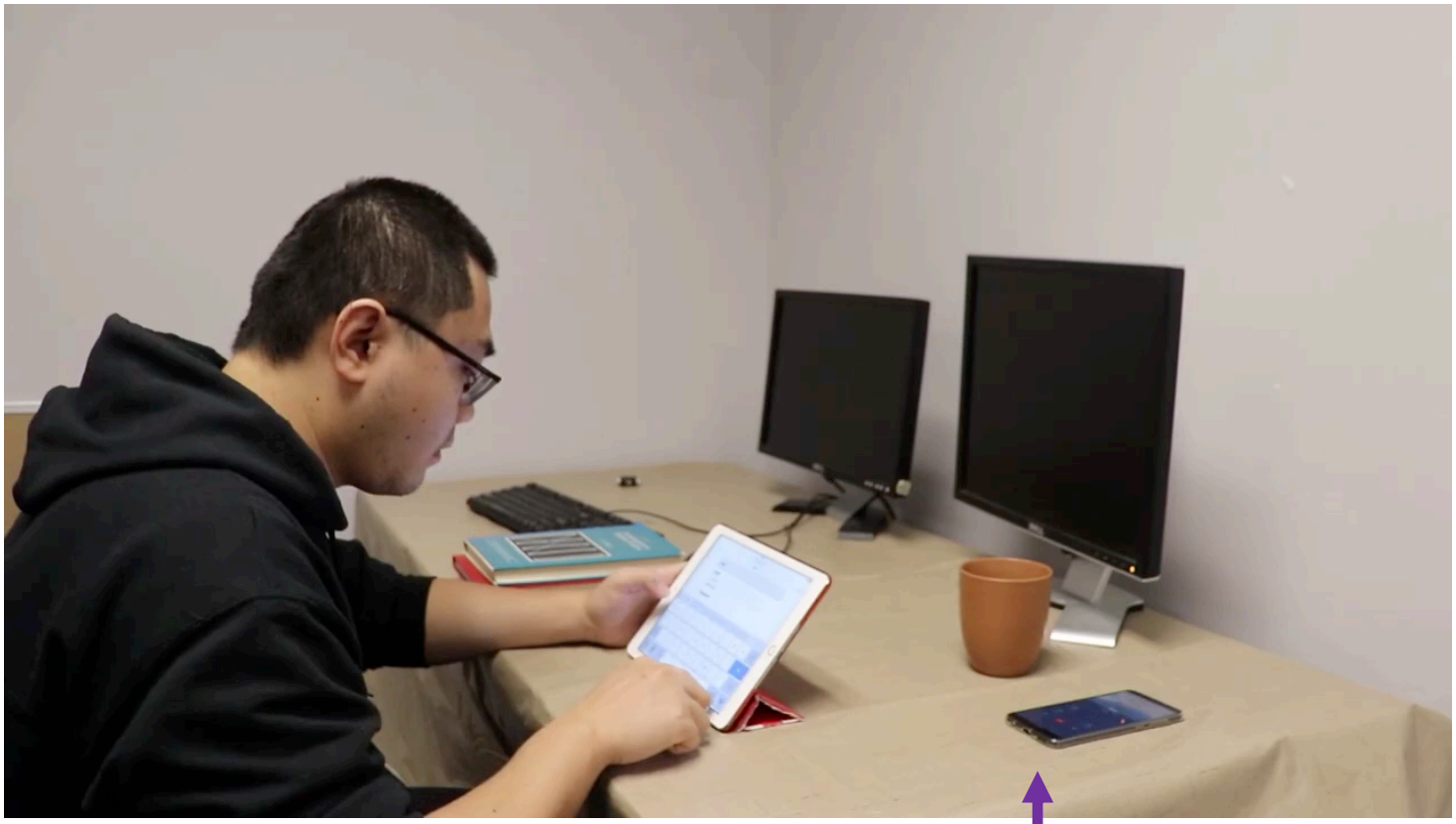


**None Line of Sight & Omni-directional**

**Long Range** Attack

Attack **multiple devices** simultaneously

# SurfingAttack: Hidden Interactive Attack

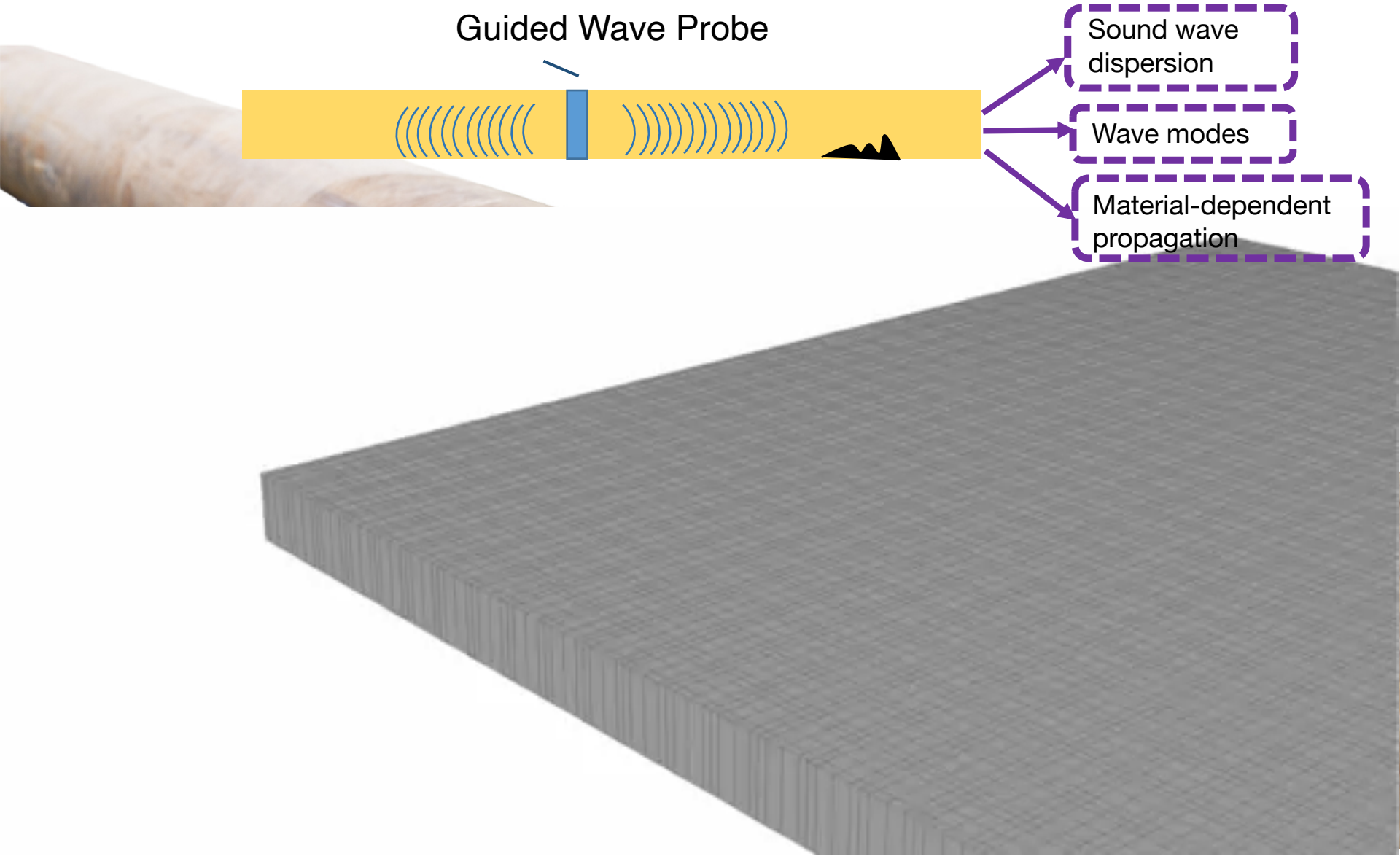


Attack transducer and waveform generator are hidden under the desk


Lyric
Thrilling music. Dazzling theater. Stunning spectacle.
No trip to Chicago is complete without a visit to Lyric.

# How it works?

# Ultrasonic Guided Waves: *Lamb Waves*



# Attack Wave Selection

Narrowband input signals

Circular piezoelectric disc (PZT)

Low dispersion

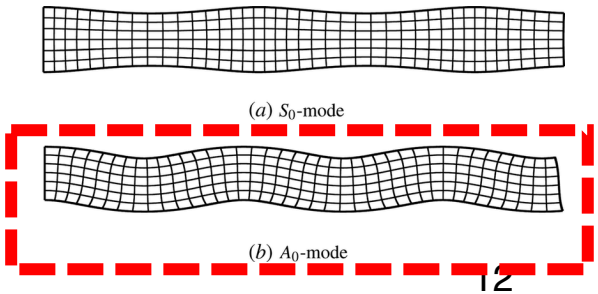
Low attenuation

Easy excitability

High attack signal reachability

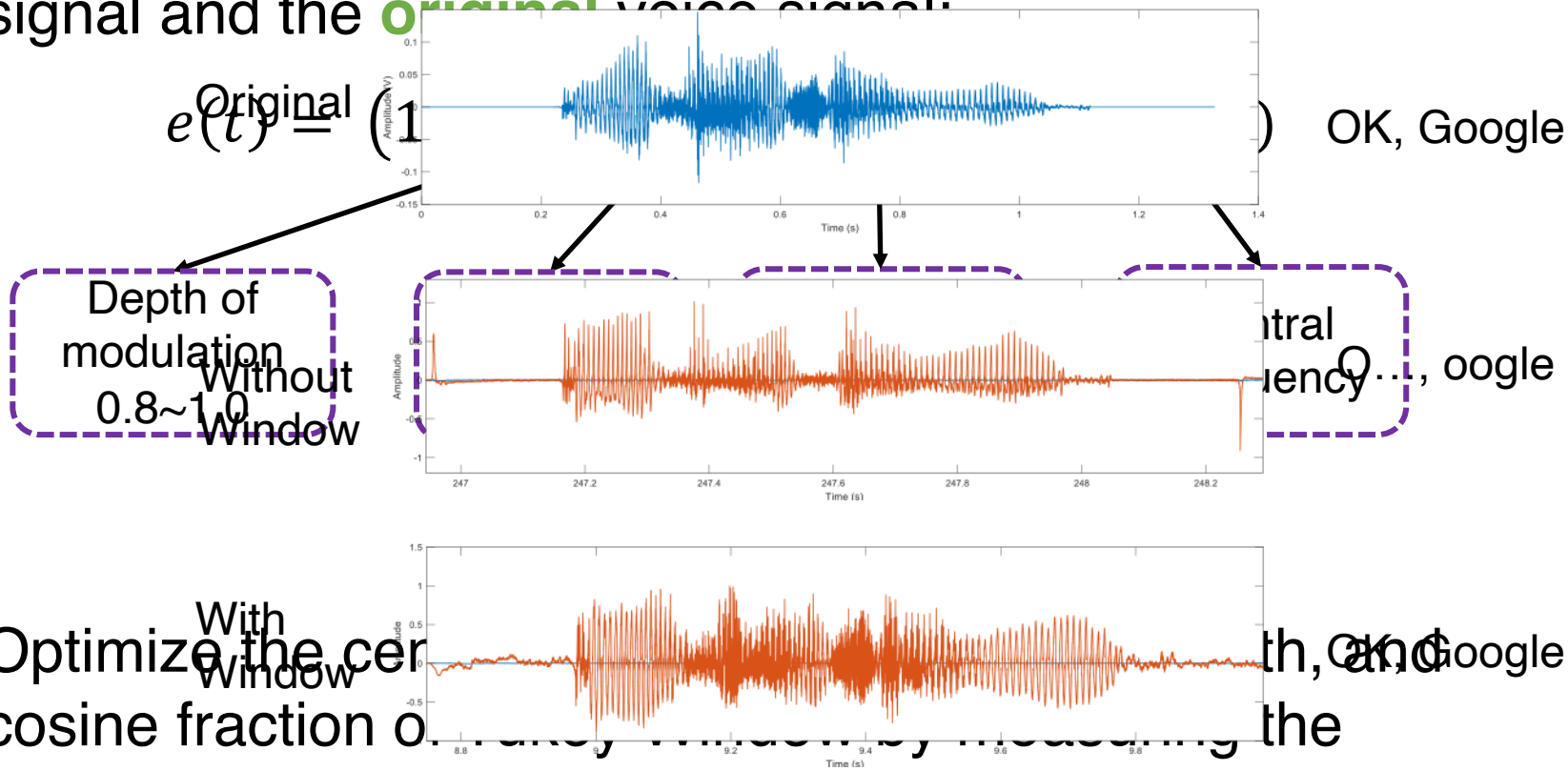
Ultrasonic guided wave

Lower-order Lamb wave modes ($A_0$)



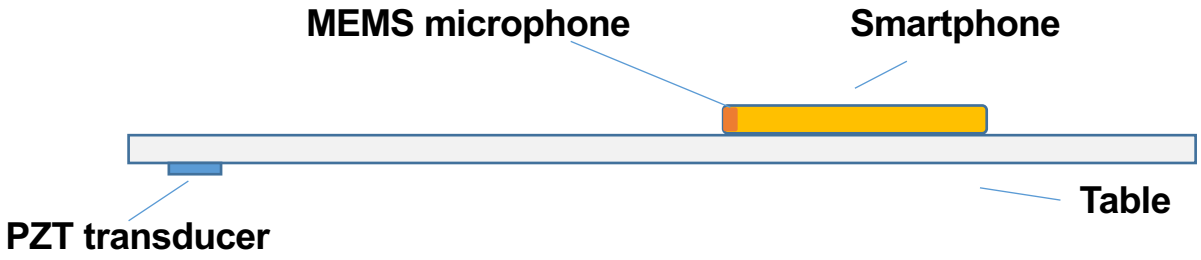
# Attack Wave Generation

- Goal: Preserve the similarity between the **recovered** voice signal and the **original** voice signal:

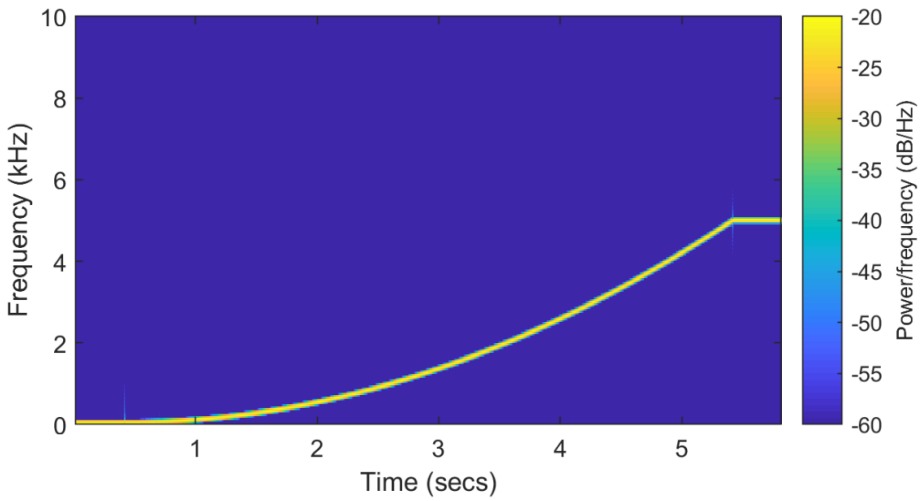



- Optimize the cer… th, and cosine fraction o… the nonlinearity responses.

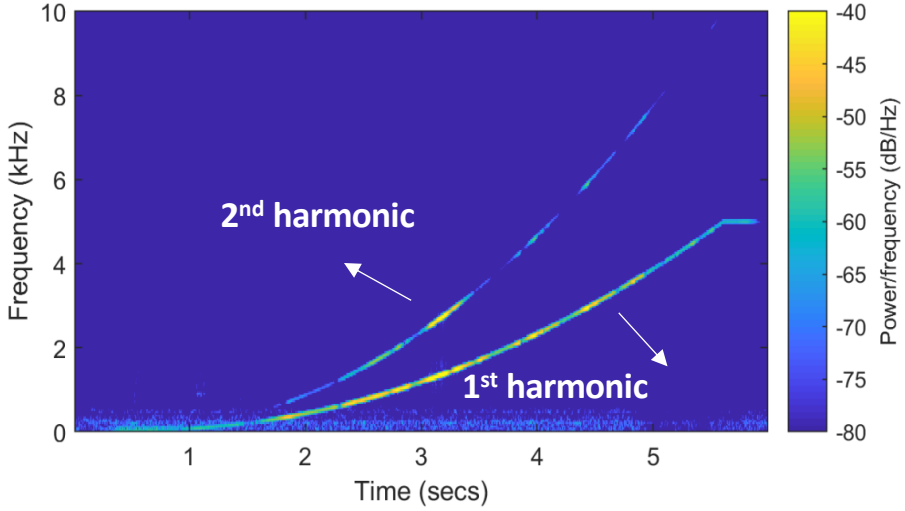
# Triggering Non-linearity Effect




Baseband Voice Signal

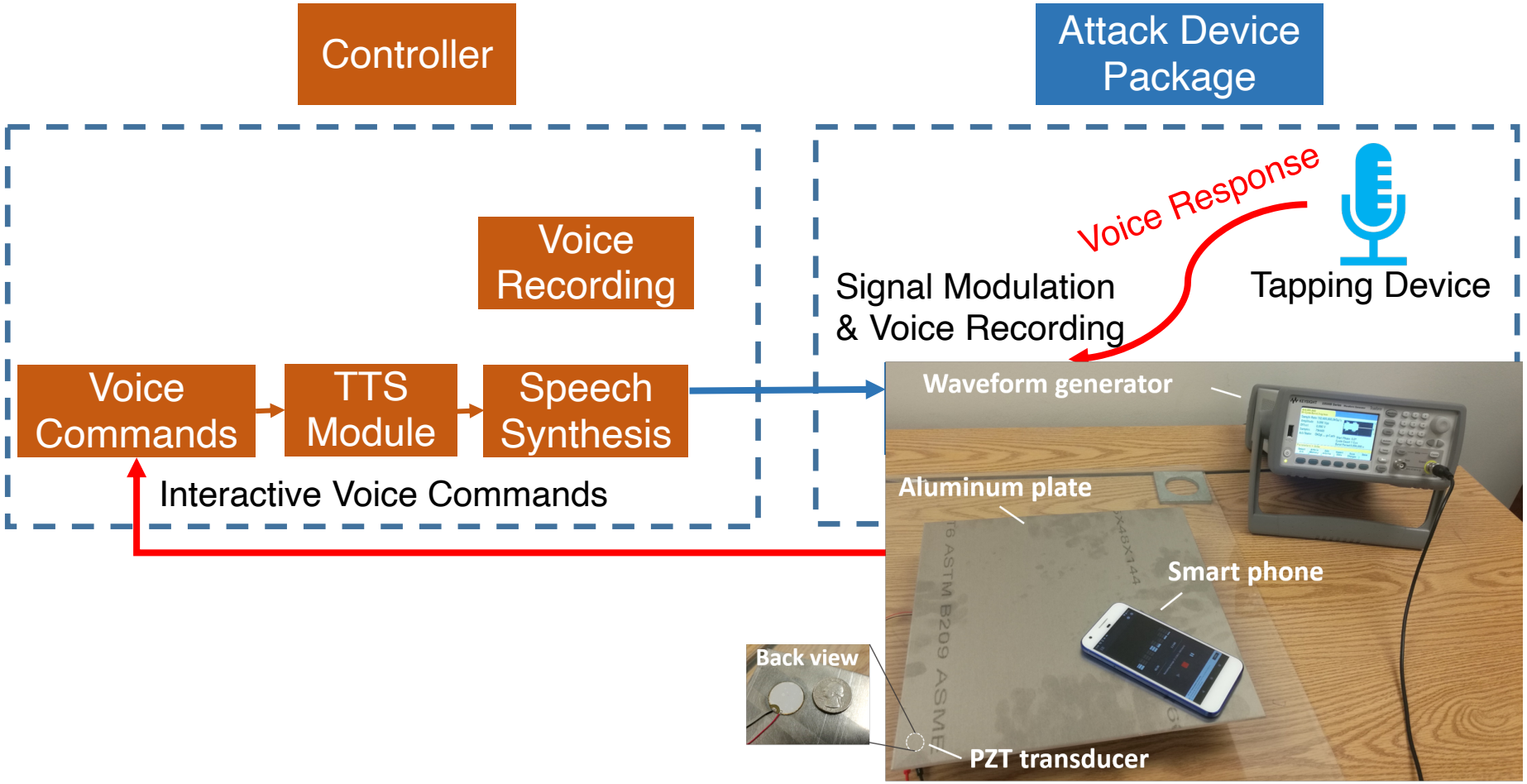


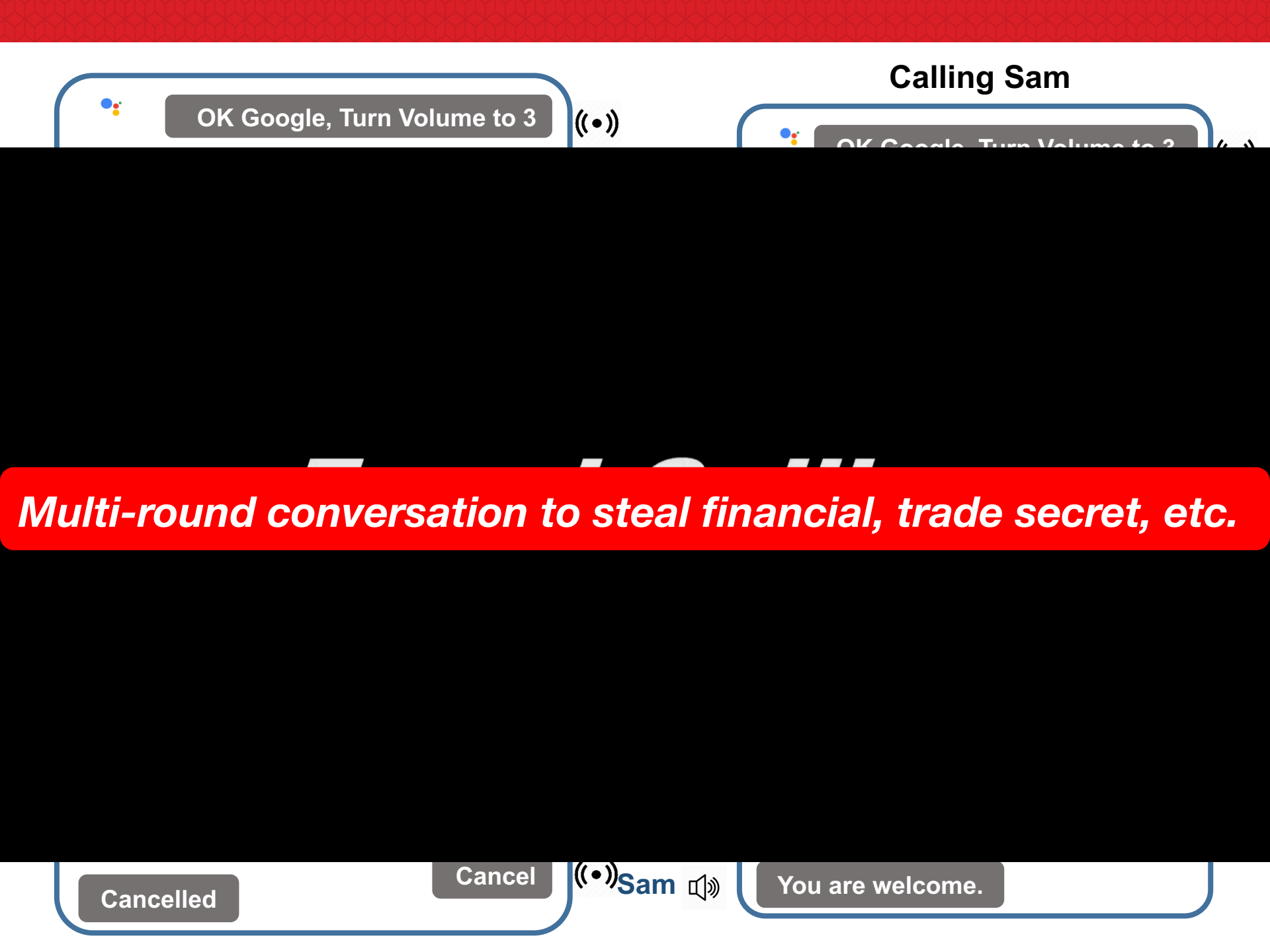

Recorded Voice Signal




**Baseband signal modulated to 25.3 kHz carrier/central frequency.**

# Attack System Design




Calling Sam
OK Google, Turn Volume to 3
Multi-round conversation to steal financial, trade secret, etc.
Cancelled
Cancel
Sam
You are welcome.

# Feasibility Across Different Smartphones

| Manufacture | Model | Assistants | Attack Frequency | Attacks | | |
|---|---|---|---|---|---|---|
| | | | | Recording | Activation | Recognition |
| Google | Pixel 1, 2, 3 | Google | 27-28 KHz | ✓ | ✓ | ✓ |
| Moto | G5 | Google | 27.0 KHz | ✓ | ✓ | ✓ |
| | Z4 | Google | 28.2 KHz | ✓ | ✓ | ✓ |
| Samsung | Galaxy Note 10+ | | | | | ✗ |
| Xiaomi | Mi 5, 8 | | | | | ✓ |
| Huawei | Mate 9 | | | | | ✗ |
| | Honor 10 | | | | | ✓ |
| Apple | iPhone 5, 5s, 6+, X | | | | | ✓ |

***SurfingAttack* succeeds on 15 out of 17 smartphones!**

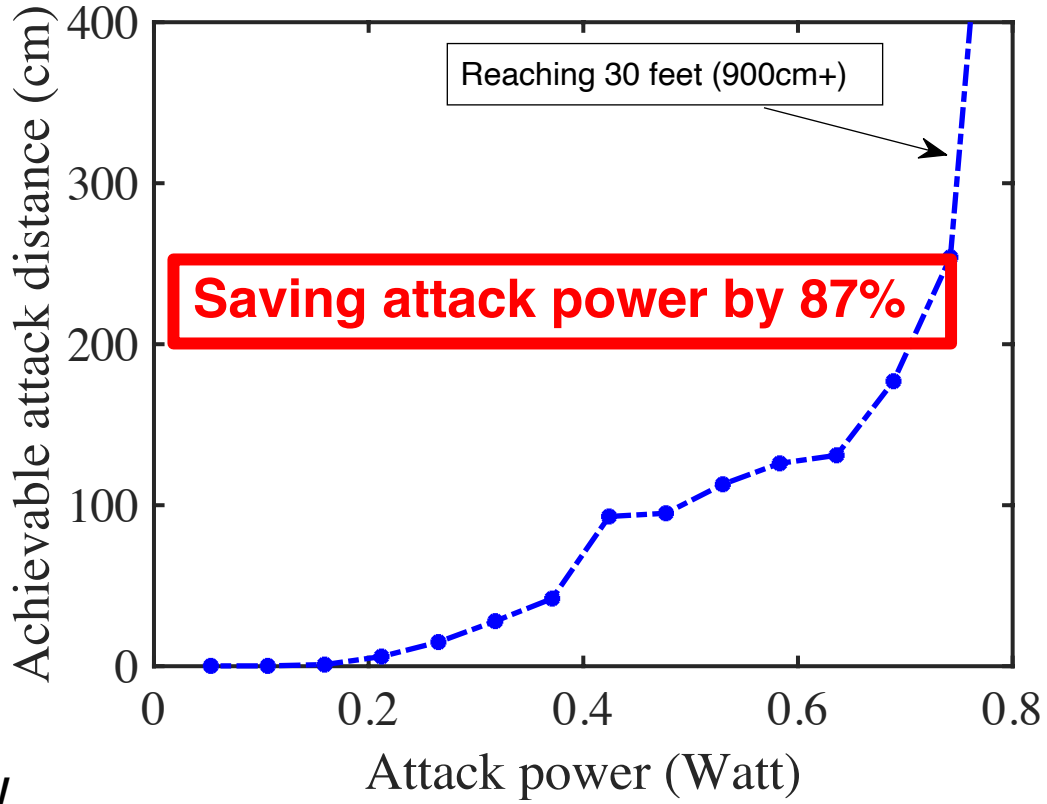

# Evaluation: Impact Analysis of Factors

- Noise and Verbal Conversations
- Directionality
- Attack Distance
- Table Materials
- Lock Screen
- Table Thicknesses
- Interlayers on the Table
- Phone Cases

# Evaluation: Attack Distance

GWBP-AMP-X75 Power Amplifier
- Maximum output power of 1.5W

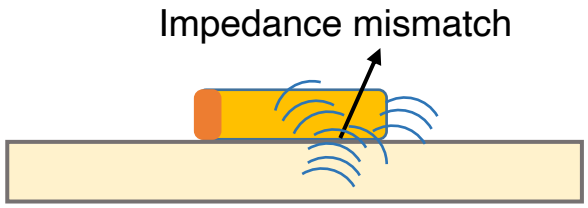
(output voltage of 30V)



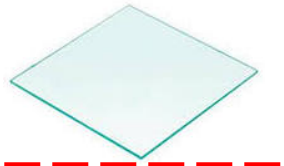

***SurfingAttack*** **attack distance reaches 30ft with 0.8W attack power. In comparison, over-air speaker array reaches 30ft with 6W attack power[1].**

[1] Roy, N., Shen, S., Hassanieh, H., & Choudhury, R. R. (2018). Inaudible voice commands: The long-range attack and defense. In *15th USENIX Symposium on Networked Systems Design and Implementation (NSDI 18).*

# Evaluation: Impact of Table Materials



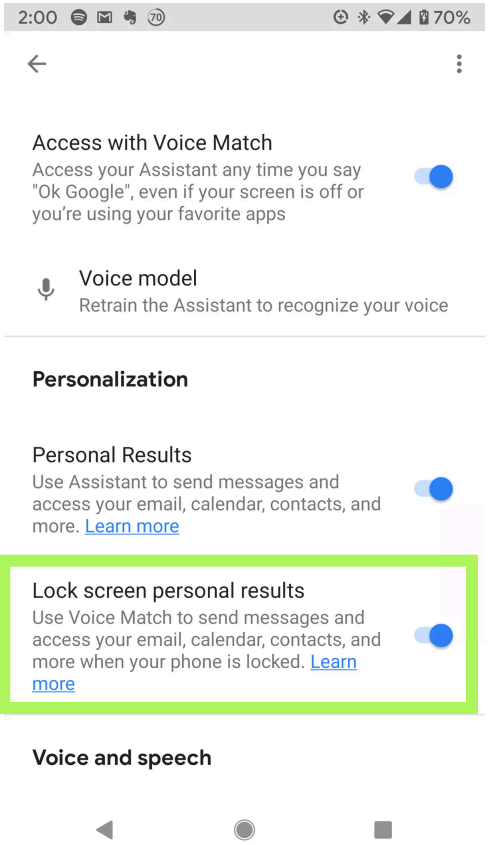


| | Aluminum Metal Sheet (0.3 mm) | Steel Metal Sheet (0.8 mm) | Glass (2.54 mm) | MDF (5 mm) | Rough polyethylene plastic (5 mm) |
|---|---|---|---|---|---|
| Xiaomi Mi 5 | 910+ cm | 95+ cm | 85+ cm | 50cm | X |
| Google Pixel | 910+ cm | 95+ cm | 85+ cm | 45cm | X |
| Samsung Galaxy S7 | 910+ cm | 95+ cm | 85+ cm | 48cm | X |

**The best energy delivery can be achieved when the table material is the same as the device body material. Porous structure absorbs ultrasound.**

# Evaluation: Lock Screen


**The attack works on Voice Assistants even if the device is locked, if we enable voice assistants on the lock screen.**

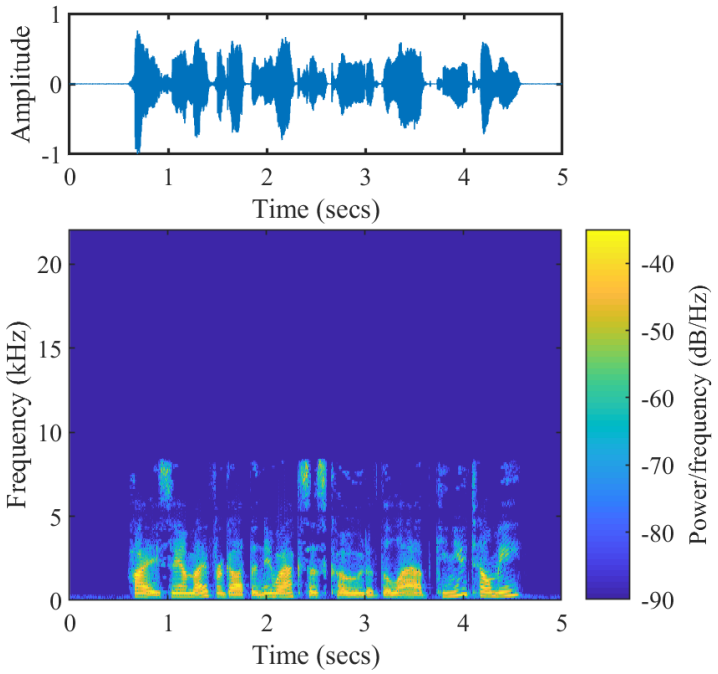
# How to defend?

# Countermeasure I

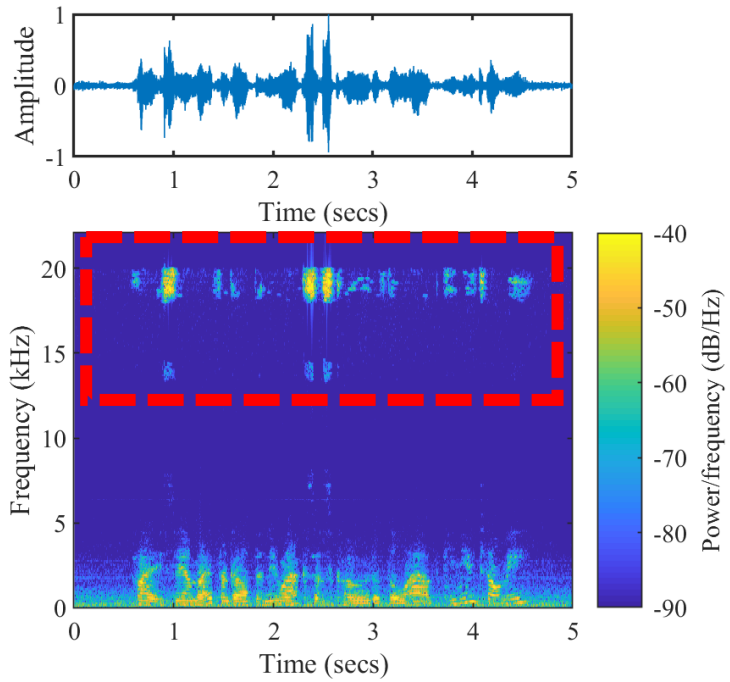
- Keep an eye on your devices.
- Reduce the touching surface area of your phones with the table.
- Place the device on a soft woven fabric before touching the tabletops.
- Use thicker phone cases made of uncommon materials such as wood.
- Disable your Voice Assistant on lock screen and lock your device.

# Countermeasure II

- Software-based Defense
  - Difference between recovered signal and the baseband signal in spectrogram (10 – 20 kHz)




**Recorded Normal Voice**




**Recorded Attack Signal**


Lyric
Thrilling music. Dazzling theater. Stunning spectacle.
No trip to Chicago is complete without a visit to Lyric.
DON'T

# Can We Attack Standing Voice Assistants?


**Further increasing the power of ultrasound signals: the guided waves can be converted into in-air ultrasound signals.**

# Conclusion

1. Explore the feasibility of launching inaudible ultrasonic attack leveraging **ultrasonic guided waves through solid materials**

2. Enable **conversations** between the adversary and the voice controllable device

3. *SurfingAttack* successfully attacks **15** popular smartphones on different solid materials and achieves **30ft** long-range attack through a metal table with a low power profile.

**Visit https://surfingattack.github.io/ for more information**

# We are recruiting graduate students!


SEIT

MICHIGAN STATE UNIVERSITY