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

Qiben Yan¹, Kehai Liu², Qin Zhou²
Hanqing Guo¹, Ning Zhang³

¹Michigan State University, ²University of Nebraska-Lincoln,
³Washington University in St. Louis
They are not safe!
Over-the-air Inaudible Attack

Yes, how can I help you?

Input → Diaphragm → Amplifier → Low Pass filter → ADC

Over-the-air Inaudible Attack

$V_{in} = \sin F_1 + \sin F_2$

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.
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*

Guided Wave Probe

- Sound wave dispersion
- Wave modes
- Material-dependent propagation
Attack Wave Selection

Narrowband input signals

- Low dispersion
- Low attenuation
- Easy excitability
- High attack signal reachability

Circular piezoelectric disc (PZT)

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:

\[ e(t) = (1 + \tau \cos(2\pi f_0 t)) \]

- Depth of modulation: 0.8~1.0

- Central frequency: OK, Google

- Optimize the central frequency, modulation depth, and cosine fraction of Tukey Window by measuring the nonlinearity responses.
Triggering Non-linearity Effect

MEMS microphone

Smartphone

PZT transducer

Table

**Baseband Voice Signal**

**Recorded Voice Signal**

Baseband signal modulated to 25.3 kHz carrier/central frequency.
Attack System Design

Controller

Voice Recording

Voice Commands → TTS Module → Speech Synthesis

Interactive Voice Commands

Attack Device Package

Signal Modulation & Voice Recording

Tapping Device

Voice Response

Signal Modulation & Voice Recording

Waveform generator

Aluminum plate

Smart phone

PZT transducer

Back view
Read my messages
You have one text message. It's from 347268, do you want to hear it?
Sure
It says …, do you want to reply, repeat it and just that for now?
Cancel
Cancelled

OK Google, Turn Volume to 3

Call Sam with speakerphone
OK, calling Sam with speakerphone
Hi, Sam, I forgot the new access code of the lab, can you tell me?
Sure, it is 2501.
OK, thanks.
You are welcome.

Multi-round conversation to steal financial, trade secret, etc.
## Feasibility Across Different Smartphones

<table>
<thead>
<tr>
<th>Manufacture</th>
<th>Model</th>
<th>Assistants</th>
<th>Attack Frequency</th>
<th>Recording</th>
<th>Activation</th>
<th>Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>Pixel 1, 2, 3</td>
<td>Google</td>
<td>27-28 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Moto</td>
<td>G5</td>
<td>Google</td>
<td>27.0 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td></td>
<td>Z4</td>
<td>Google</td>
<td>28.2 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Samsung</td>
<td>Galaxy S8</td>
<td>Google</td>
<td>26.6 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td></td>
<td>Galaxy Note 10+</td>
<td>Google</td>
<td>25.8 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Xiaomi</td>
<td>Mi 5, 8</td>
<td>Google</td>
<td>25-28 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Huawei</td>
<td>Mate 9</td>
<td>Google</td>
<td>27.7 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td></td>
<td>Honor 10</td>
<td>Google</td>
<td>27.3 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Apple</td>
<td>iPhone 5, 5s, 6+, X</td>
<td>Siri</td>
<td>26-27 KHz</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

*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**

**SurfingAttack** attack distance reaches 30ft with 0.8W attack power. In comparison, over-air speaker array reaches 30ft with 6W attack power\(^1\).


GWBP-AMP-X75 Power Amplifier
- Maximum output power of 1.5W (output voltage of 30V)

Reaching 30 feet (900cm+)

Saving attack power by 87%
**Evaluation: Impact of Table Materials**

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

<table>
<thead>
<tr>
<th>Material</th>
<th>Xiaomi Mi 5</th>
<th>Google Pixel</th>
<th>Samsung Galaxy S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminum Metal Sheet (0.3 mm)</td>
<td>910+ cm</td>
<td>910+ cm</td>
<td>910+ cm</td>
</tr>
<tr>
<td>Steel Metal Sheet (0.8 mm)</td>
<td>95+ cm</td>
<td>95+ cm</td>
<td>95+ cm</td>
</tr>
<tr>
<td>Glass (2.54 mm)</td>
<td>85+ cm</td>
<td>85+ cm</td>
<td>85+ cm</td>
</tr>
<tr>
<td>MDF (5 mm)</td>
<td>50 cm</td>
<td>45 cm</td>
<td>48 cm</td>
</tr>
<tr>
<td>Rough polyethylene plastic (5 mm)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

*Impedance mismatch*
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
Thrilling music. Dazzling theater. Stunning spectacle.

No trip to Chicago is complete without a visit to Lyric.
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/](https://surfingattack.github.io/) for more information
We are recruiting graduate students!