Light Commands:

Laser-Based Audio Injection on Voice-Controllable Systems

Takeshi Sugawara, **Benjamin Cyr,**
Sara Rampazzi, Daniel Genkin, Kevin Fu

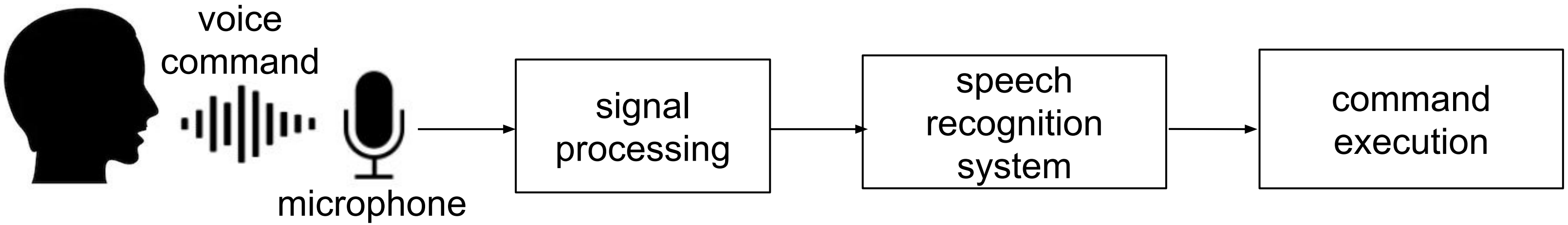
UEC TOKYO

M | CSE COMPUTER SCIENCE AND ENGINEERING UNIVERSITY OF MICHIGAN

Voice Controllable Systems (VCSs)

[Source: pandaily.com]

[Source: developers.google.com]



voice command

microphone

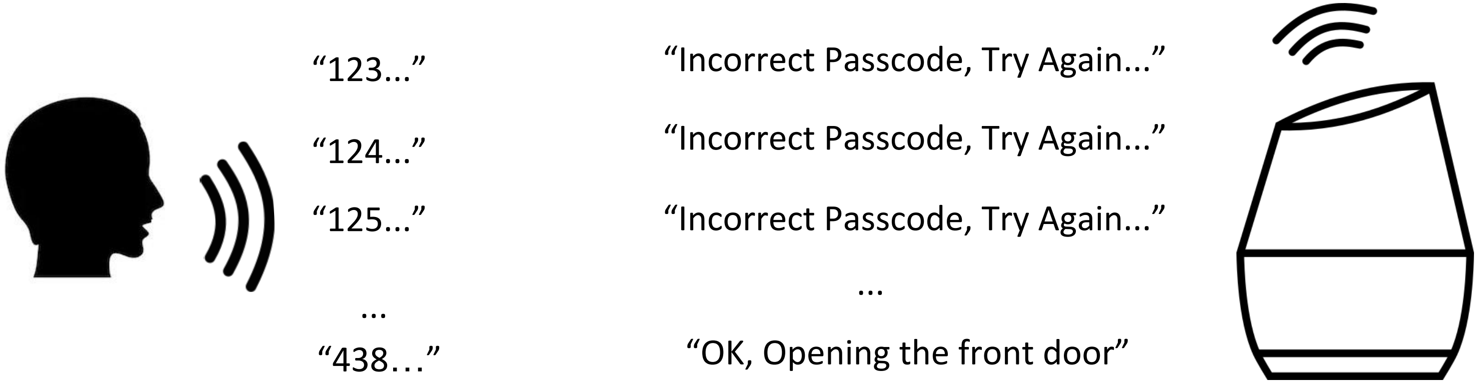
signal processing

speech recognition system

command execution

Security Concerns

- Sacrifice of security to promote usability
- Interfacing with 3rd Party Software
- **Blind trust in microphone readings**



“123...”

“124...”

“125...”

...

“438…”

“Incorrect Passcode, Try Again...”

“Incorrect Passcode, Try Again...”

“Incorrect Passcode, Try Again...”

...

“OK, Opening the front door”

The Problem

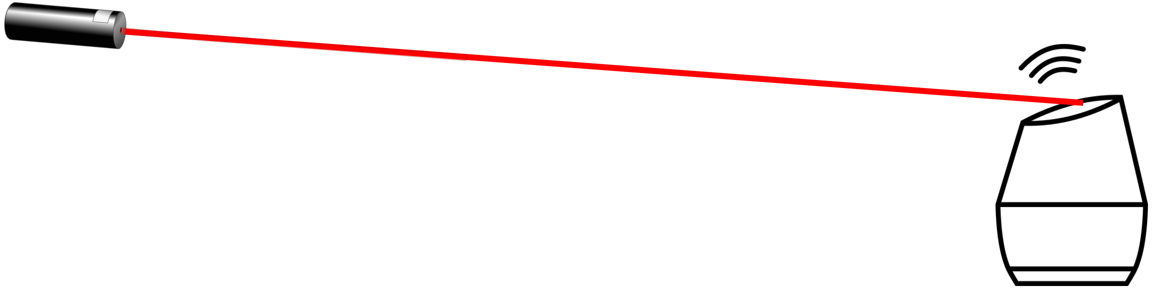
Assumption:

Microphones capture **acoustic** signals

The Problem

Reality:

Microphones capture **acoustic** signals & LIGHT signals



The Problem

Two Questions:

1. How does laser injection affect VCSs?
2. How can we protect VCSs against laser injection?

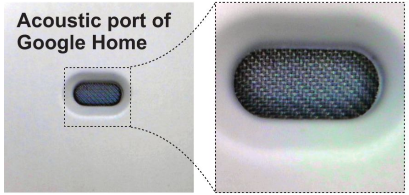
Contributions

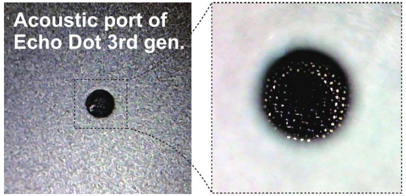
LIGHT
COMMANDS

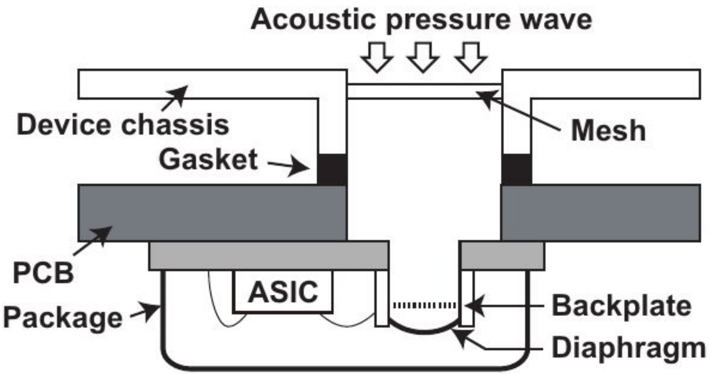
- **Inject light commands** via MEMS microphones
- **Analyze limits** of light-based VCS vulnerabilities
  - Success at 110m with 5mW laser pointer
  - Works through glass windows between buildings
- Demonstrate risks to **smart speakers, phones, smart homes, and cars**
- Suggest HW and SW **countermeasures**

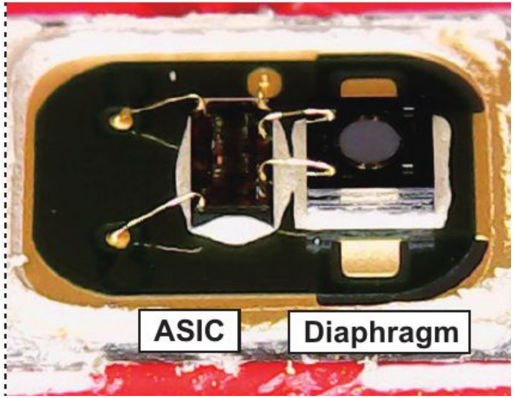
MEMS Microphones

- Used in most Voice Controllable Systems
- The diaphragm and backplate work as a **capacitor**
- When diaphragm moves, causes a change in capacitance
- The ASIC converts the capacitive change to voltage



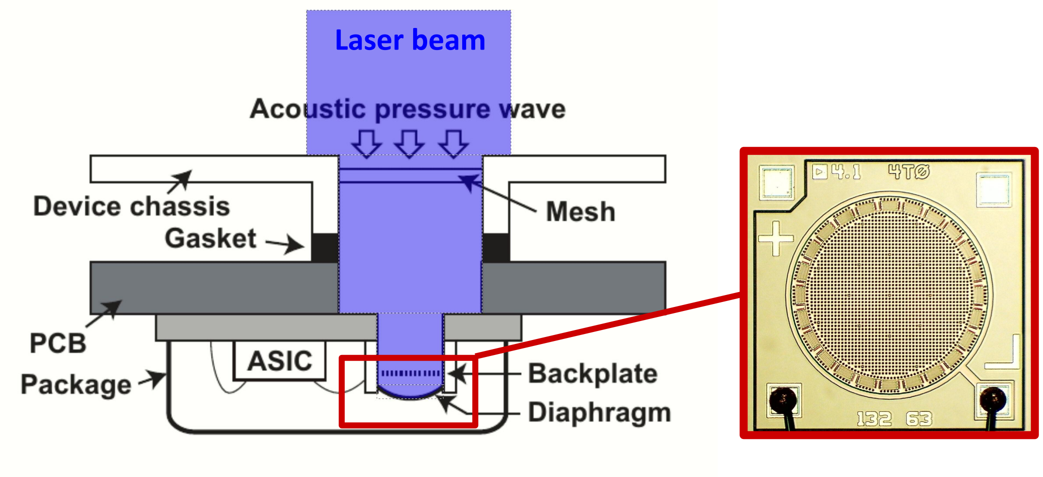






MEMS Microphones

- MEMS microphones exhibit light sensitivity
- Output voltage affected by light **irradiance**
- Inject signal by modulating optical power

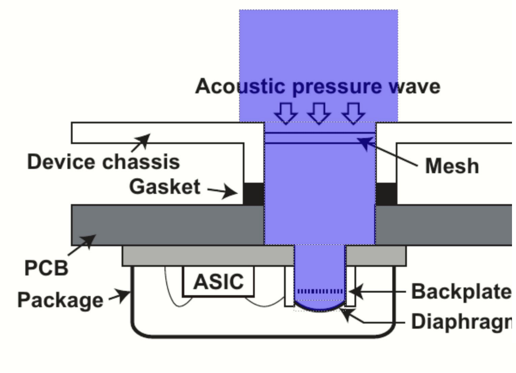


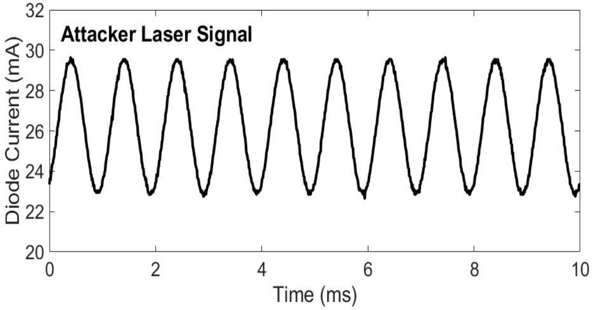
Irradiance:

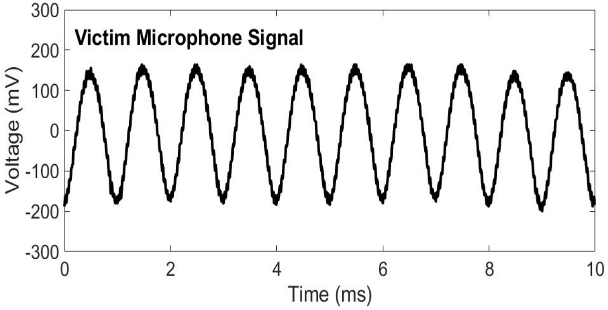
$$I = \frac{Optical\ Power\ (Watts)}{Beam\ Area\ (meters^2)}$$

Key ideas

1. Amplitude modulated light generates a voltage signal on microphone output
2. Higher amplitude light == higher amplitude voltage
3. Very little distortion



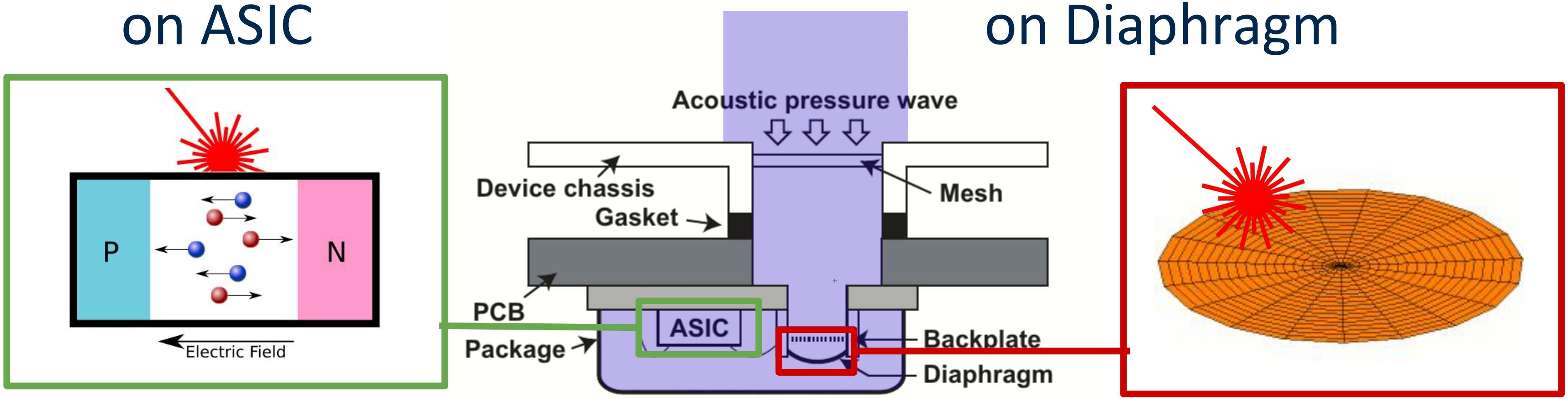




How is this Working?

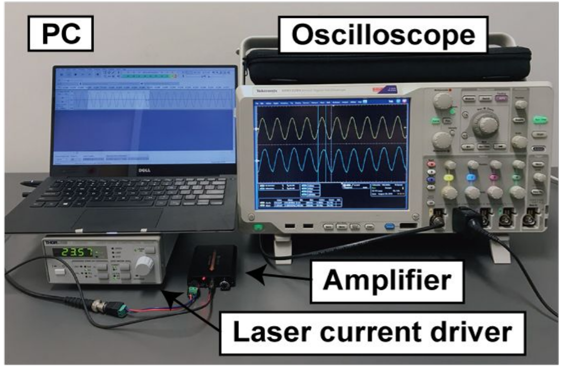
Combination of two physical effects:

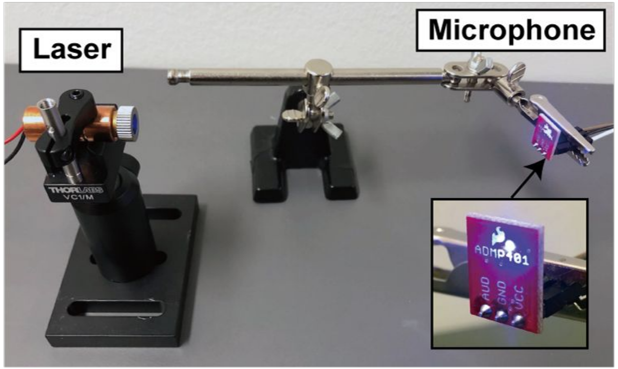
1. **Photoelectric** Effects on ASIC
2. **Photoacoustic** Effects on Diaphragm

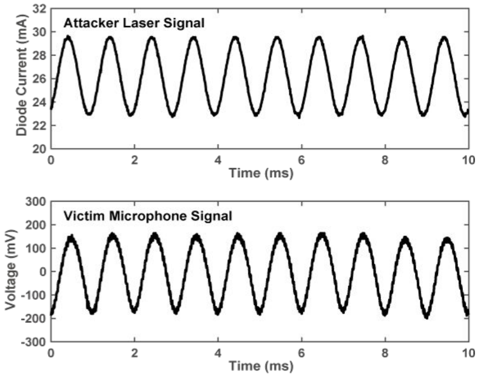


Signal Injection via Laser

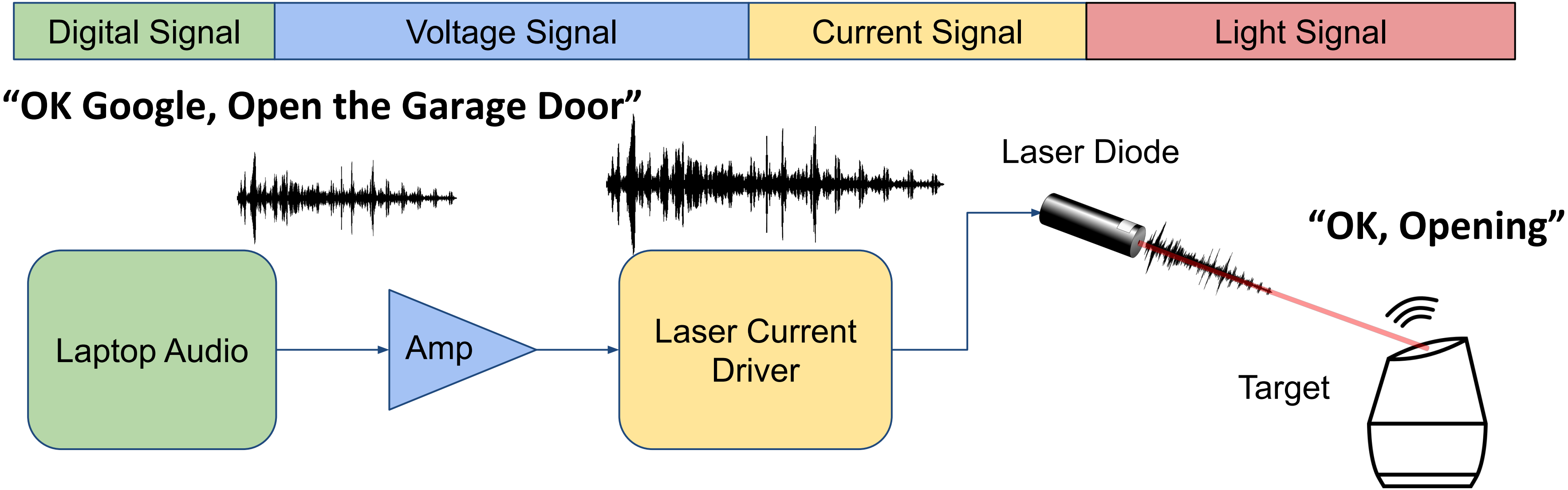
- Audio **voltage** signal from laptop
- Laser current driver converts to **current** signal
  - With DC Bias
- Laser output **power** is proportional to **current**





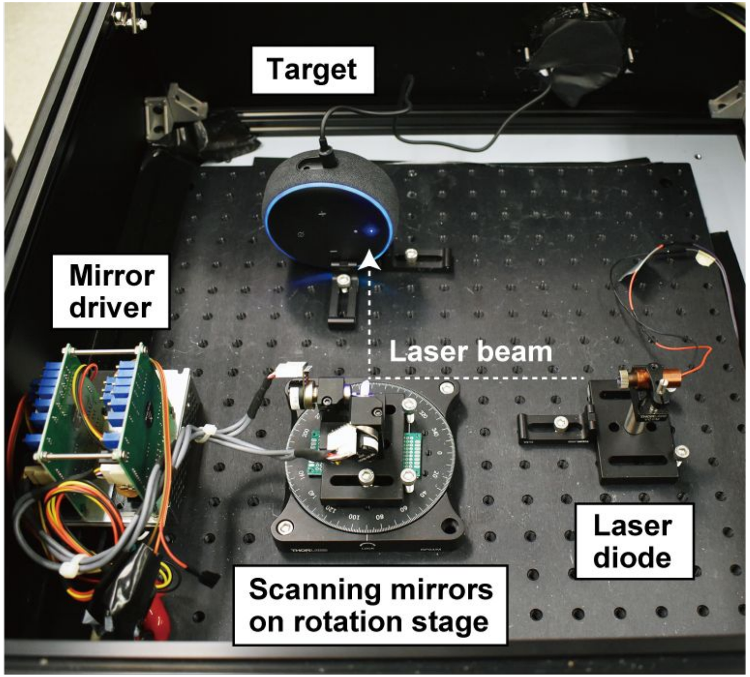


VCS Command Injection via Light



Measuring Vulnerability - Power

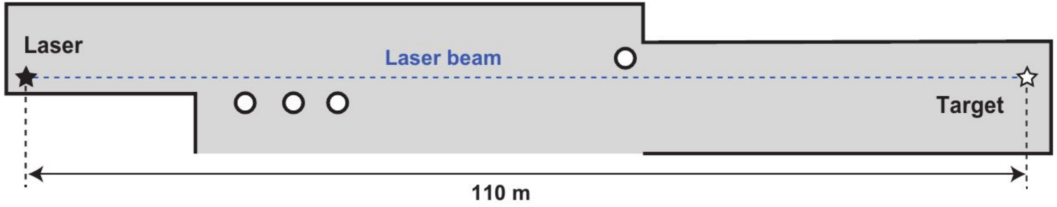
- Investigated 17 devices
- Used scanning mirrors
- Measured minimum optical power to recognize commands



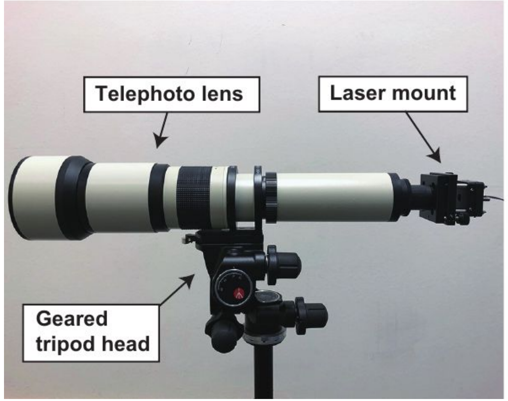
Measuring Vulnerability - Range

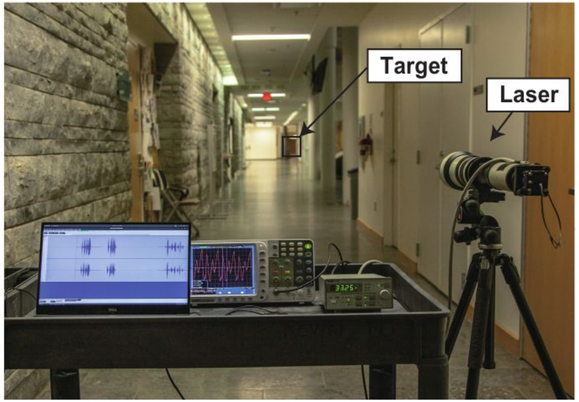
Measuring the maximum range of the attack

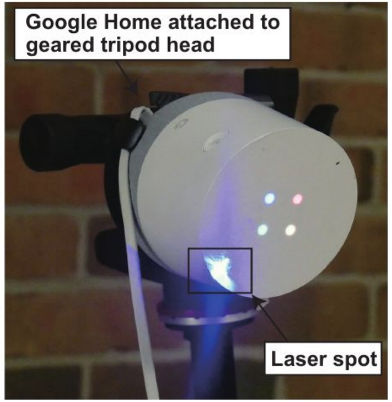
$$\frac{\mathit{Optical\ Power\ (Watts)}}{\mathit{Beam\ Area\ (meters^2)}}$$



Optics!

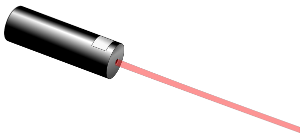


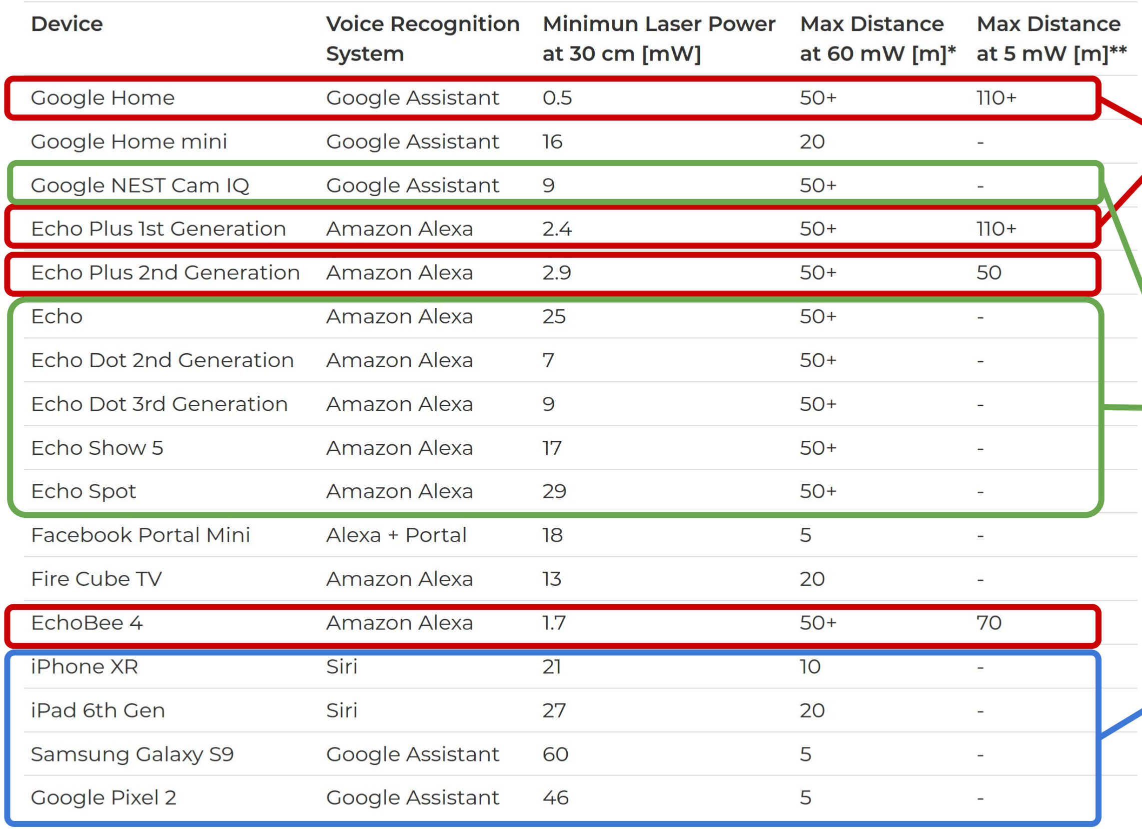




Attack Results

Laser pointer power!





Device	Voice Recognition System	Minimun Laser Power at 30 cm [mW]	Max Distance at 60 mW [m]*	Max Distance at 5 mW [m]**
Google Home	Google Assistant	0.5	50+	110+
Google Home mini	Google Assistant	16	20	-
Google NEST Cam IQ	Google Assistant	9	50+	-
Echo Plus 1st Generation	Amazon Alexa	2.4	50+	110+
Echo Plus 2nd Generation	Amazon Alexa	2.9	50+	50
Echo	Amazon Alexa	25	50+	-
Echo Dot 2nd Generation	Amazon Alexa	7	50+	-
Echo Dot 3rd Generation	Amazon Alexa	9	50+	-
Echo Show 5	Amazon Alexa	17	50+	-
Echo Spot	Amazon Alexa	29	50+	-
Facebook Portal Mini	Alexa + Portal	18	5	-
Fire Cube TV	Amazon Alexa	13	20	-
EchoBee 4	Amazon Alexa	1.7	50+	70
iPhone XR	Siri	21	10	-
iPad 6th Gen	Siri	27	20	-
Samsung Galaxy S9	Google Assistant	60	5	-
Google Pixel 2	Google Assistant	46	5	-

5mW:
110+ meters

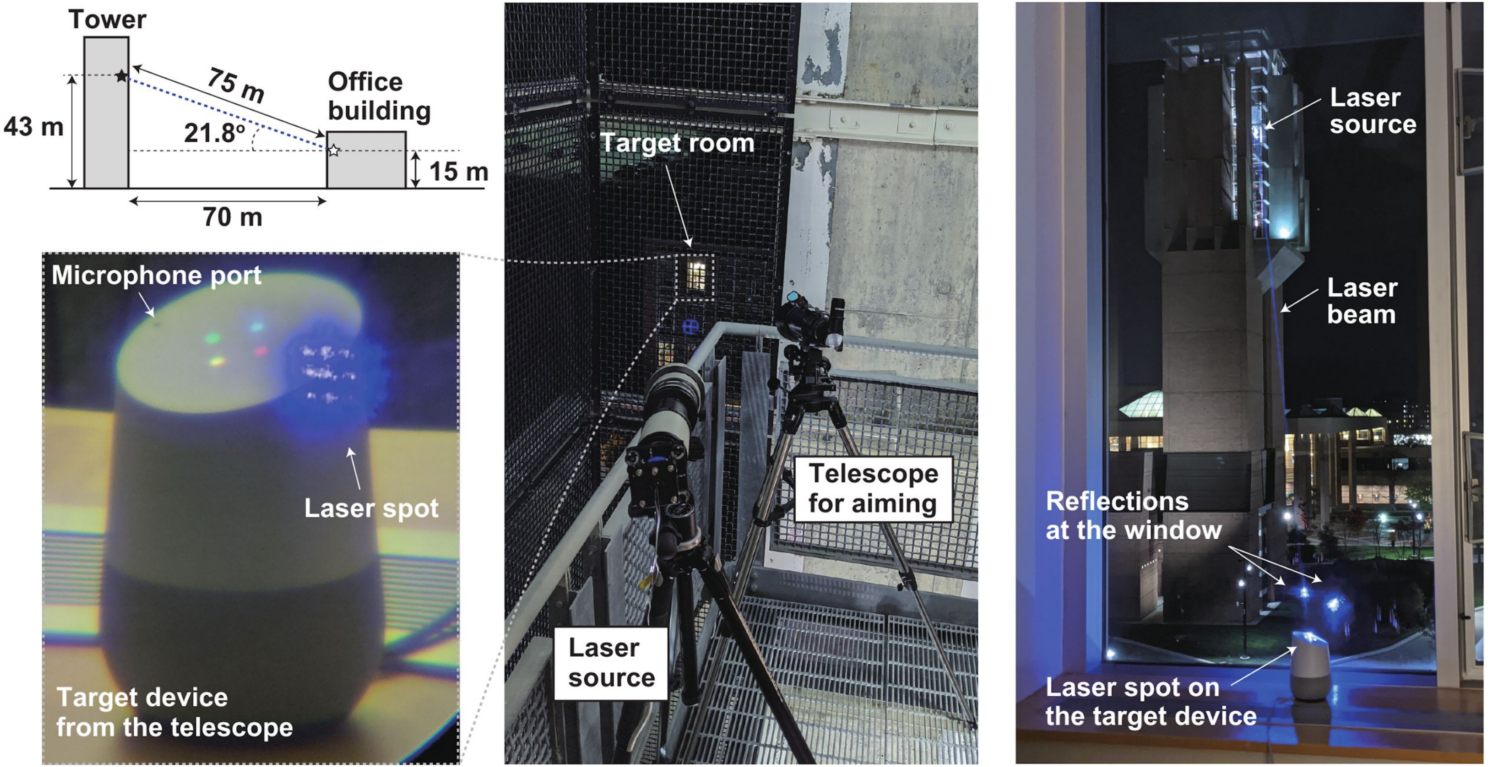
60mW:
50+ meters

Phones/Tablets

60mW:
5-20 meters

* Limited to a 50 m long corridor.
** Limited to a 110 m long corridor.

Cross-Building Attack Scenario



Attack Demonstration

DEMO #1 OPENING GARAGE DOOR

Injecting “OK Google, open the garage door” to a Google Home by shining a cheap laser pointer

DEMO #2 LONG DISTANCE

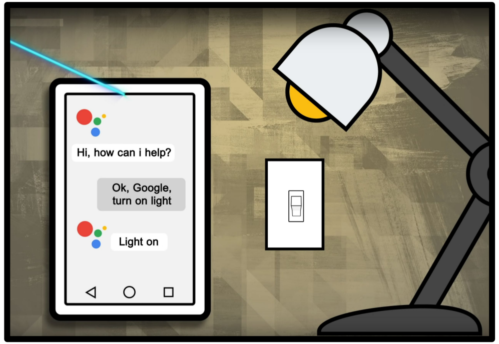
Injecting “OK Google, what time is it?” to a Google Home by shining a laser from 110 meters away

Demos available at lightcommands.com

DEMO #3
THROUGH A WINDOW

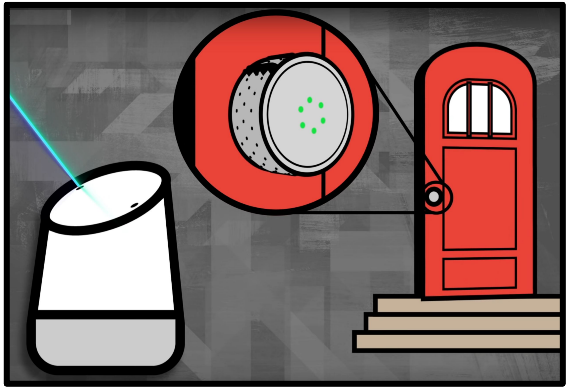
Injecting “OK Google, open the garage door” to a Google Home by shining a laser from another building

Consequences

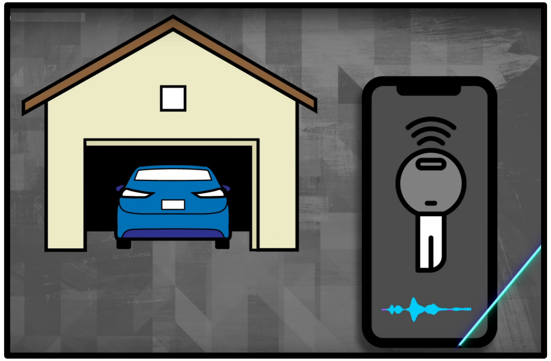


Turn on/off
Enable/Disable

Brute force unlock door



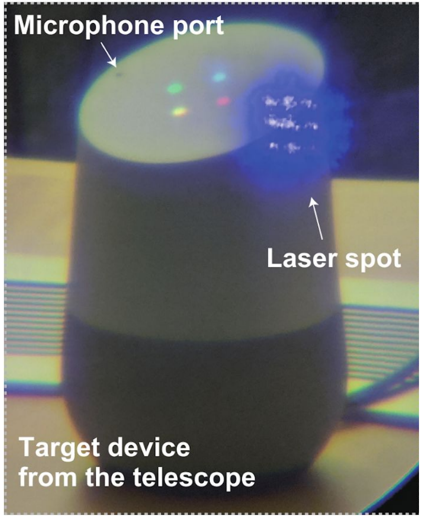
Unauthorized purchases



Open garage doors
Unlock car
Start engine

Limitations

- Dependence on Focusing, Aiming, Acoustic Noise, and Audio Quality
- Requires **Line of Sight**
  - Very little diffraction
  - Difficult to target top microphones
- Limited **Feedback**



Countermeasures

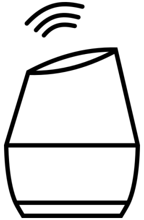
Software Approaches

- Stronger **Authentication**
- **Liveness** Tests
- **Sensor Fusion**: Compare Multiple Microphones

"Please give the passcode to unlock the garage door"

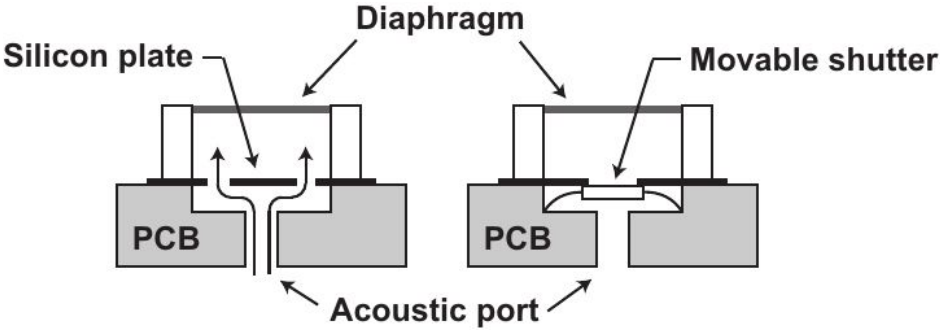
...

"Please confirm by repeating the second digit of your passcode"



Hardware Approaches

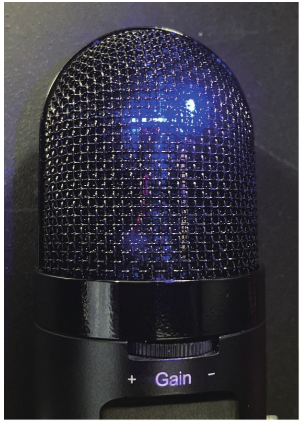
- Light-Blocking **Covers**
  - On the VCS (fabric)
  - Inside the MEMS Mic

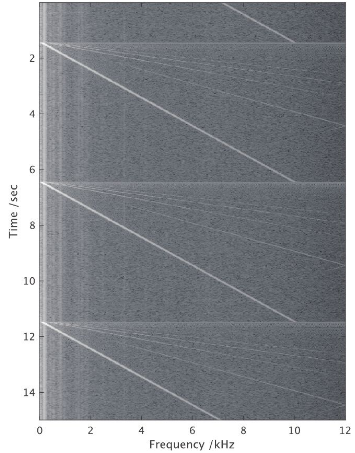


Future Work

- Deep exploration of physical causality
  - Lead to future defenses

- Other Vulnerabilities:
  - Non-MEMs Microphones
  - Other Motion Sensors





Related Work

Attacks on VCS Speech Recognition:

- Vaidya et al., “Cocaine noodles: exploiting the gap between human and machine speech recognition,” USENIX WOOT, 2015.
- Carlini et al., “Hidden voice commands.” in USENIX 2016.
- Yuan et al., “CommanderSong: A systematic approach for practical adversarial voice recognition,” in USENIX 2018
- Kumar et al., “Skill squatting attacks on Amazon Alexa,” in USENIX 2018.

Acoustic Injection on VCS via Ultrasound:

- Roy et al., “Backdoor: Making microphones hear inaudible sounds,” in ACM MobiSys 2017.
- L. Song and P. Mittal, “Inaudible voice commands,” arXiv preprint arXiv:1708.07238, 2017
- Zhang et al., “DolphinAttack: Inaudible voice commands,” in ACM CCS 2017.
- Roy et al., “Inaudible voice commands: The long-range attack and defense,” in USENIX NSDI 2018.

Conclusion

- Lasers can inject commands into VCSs
- **Long range** with **low optical power**
- Physical vulnerability in MEMS microphones
- Highlights security flaws in VCSs
- **Blind trust** of any input often points to vulnerabilities

Thank You!

Authors:

Takeshi Sugawara, **Benjamin Cyr,** Sara Rampazzi, Daniel Genkin, Kevin Fu

Questions?

Website: lightcommands.com

My Email: bencyr@umich.edu

Team: LightCommandsTeam@gmail.com

This research was funded by JSPS KAKENHI Grant #JP18K18047 and #JP18KK0312, by DARPA and AFRL under contracts FA8750-19-C-0531 and HR001120C0087, by NSF under grants CNS-1954712 and CNS-2031077, gifts from Intel, AMD, and Analog Devices, and an award from MCity at the University of Michigan.