ADVERSARIAL MACHINE LEARNING AGAINST VOICE ASSISTANT SYSTEMS

Catherine Mathews, David Man, Matt Kokolus, Raymond Huang
Advisor: Dr. Yingying Chen

OUR TEAM

MATT
KOKOLUS

CATHERINE
MATHEWS

DAVID
MAN

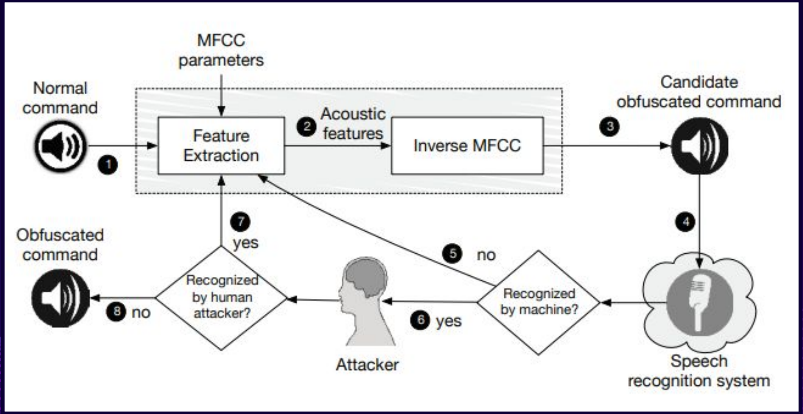
RAYMOND
HUANG

PROJECT OBJECTIVE

1. To study the security of voice assistant systems (e.g. Google Home, iPhone Siri, Amazon Alexa) under adversarial machine learning
2. To develop a system to generate hidden voice commands to attack voice assistants
3. To explore options to use a drone to carry a loudspeaker and attack voice assistant systems

HIDDEN VOICE COMMANDS

- Audio samples that have been slightly altered to fool speech recognition systems
  - Unintelligible to human listeners
  - Interpretable by voice assistant systems
- Generation of Commands
  - Noise is generated through the use of eight autonomously optimized parameters

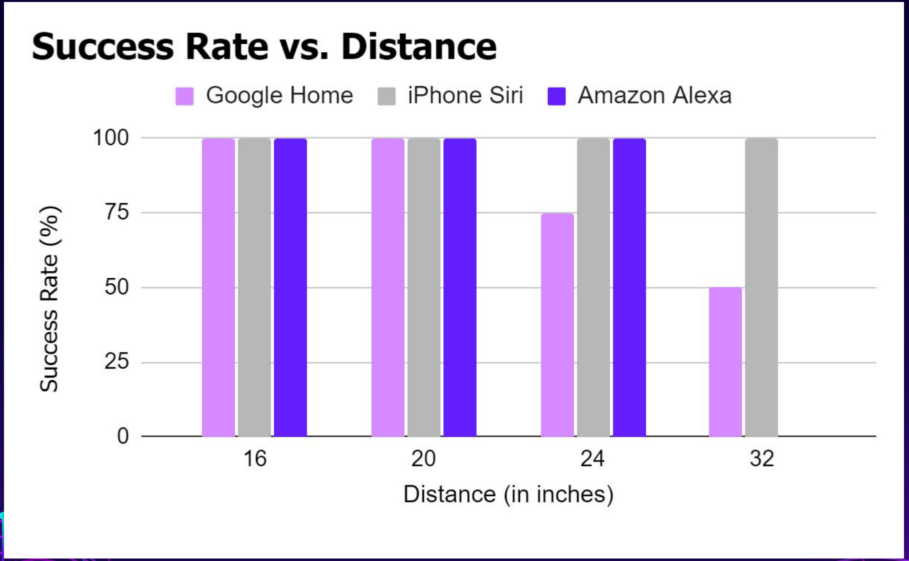


HIDDEN VOICE COMMANDS EXPERIMENT

- Experiment Procedure:
    - Recorded voice commands & obfuscated them
        - Example: 🔊 → 🔊
    - Played obfuscated commands through speaker facing voice assistant and measured success at varying distances
    - Gradually increased distance between speaker & device

EXPERIMENT RESULTS

- iPhone Siri
  - Recognized all commands at <11 ft.
- Google Home
  - Recognized all commands at <22 in.
- Amazon Alexa
  - Recognized all commands at <30 in.



* Amazon Alexa was not tested past 30 inches *

DRONE PROGRESS

- Able to pilot and fly drone: Yuneec H920 drone
- Set up procedure in future to use smaller Holy Stone HS700 drone to carry out attacks

Holy Stone HS700

Yuneec H920 Pro

FUTURE WORK

- Generate commands less recognizable to humans
  - Allows for a more realistic scenario
- Utilize reinforcement learning for further hidden command generation
- Attach loudspeaker to drone to carry out attacks over the air

THANK YOU

Any questions?