Smart Speaker Security
Voice – An Emerging Interaction Channel

- Smart speakers grow 48% annually and hit 90M in US
- While driving, cooking, exercising, etc...
- Companies introduce many different voice assistants
  - Google: Google Now
  - Apple: Siri
  - Microsoft: Cortana
  - Amazon: Alexa
Your Voice Assistant Listens to Everyone

No Authentication for Voice Assistants
Voice attacks

• Set alarm for 2:00 a.m.
• Set a timer for 1 hour
• Install embarrassing apps (skills/actions)
• Send text to SMS short code
• Call someone, e.g. “Call Home”
• Play with smart home devices by, for example, saying “Light off” or “TV off”
• Saying “Restart”, which restarts the most recently played music
• Full volume, to turn up your speaker to full volume
• Buy something, which could trigger the device to make a purchase
• Repeat, which repeats the last command
• Simon says / Repeat after me, which will cause the device to repeat whatever is said afterwards
• Enable the “do not disturb” mode, which will ignore incoming calls and notifications
• Use the “where is my phone?” function to make your smartphone ring at full volume
Voice attacks (Cont.)

• Burger King launched a TV ad to trigger google home to read Wikipedia about whopper burger

• Google’s super bowl ad

• Sound park episode triggered google home, alexa, siri to say “big hairy balls”

• Parrot ordered a gift box 😊
Ultrasonic attacks

• Exploit non-linearity in the microphones to hear the commands that are inaudible to human
WIFi attacks

• Any device on the same network as smart speaker can change settings of the smart speaker
Software attacks

• Hackers gain access to the root shell of the Amazon echo speaker to steal customer authentication tokens and stream live mic data

• Reported 8 Bluetooth vulnerabilities to change echo response
Countermeasures

- Aware of the security issues with smart speakers
- Disallow smart speakers to control critical IoT (e.g., front doors)
- Protect WiFi and use WPA2
- Provide backend control (e.g., Google’s fix)
- Research on authenticating users
  - Voice features (vulnerable to replay and mimic)
  - Biometrics
  - Vauth