March 2020

Protecting Voice-Activated Devices from Laser or Electromagnetic Attacks

Zhiping Yang

Protecting Voice-Activated Devices from Laser or Electromagnetic Attacks

ABSTRACT

Researchers have demonstrated that a laser beam or a high-frequency electromagnetic (EM) wave can be used to remotely and inaudibly gain control of a voice-based virtual assistant. This disclosure describes techniques to protect a voice assistant device from such laser or EM attacks. The techniques leverage the differing sensitivities of the multiple microphones of a virtual assistant device to a laser or to EM waves to flag a received voice command as authentic or suspicious. The techniques also determine a user's proximity to the virtual assistant device to determine if a received command is authentic or suspicious.

KEYWORDS

- Laser hack
- Electromagnetic hack
- Remote activation
- Smart speaker
- Smart display
- Voice-activated device
- Demodulation efficiency
- Voice command
- Virtual assistant
- Computer security
- Intrusion detection

BACKGROUND

Researchers have demonstrated that a laser beam or a high-frequency electromagnetic wave can be used to remotely and inaudibly gain control of a voice-based virtual assistant device. In one demonstration, a laser was used to remotely (from a 50-100 meter distance) and photoacoustically excite the diaphragm of the MEMS microphone of a voice-based virtual assistant device such that the microphone captured false commands. Although the laser attack required line-of-sight access to the virtual assistant, in another demonstration, researchers were able to gain control of a voice-based virtual assistant device using high-frequency electromagnetic waves capable of penetrating thick walls. In the case of the electromagnetic (EM) attack, a false command was modulated atop an EM carrier, and naturally occurring non-linearities in the components of the virtual assistant device enabled accurate demodulation of the false command. In both demonstrations, the virtual assistant executed the false commands, e.g., “open the garage door,” “purchase a product,” etc.

Although some virtual assistant devices offer authentication protections that might foil a laser or EM attack, these typically involve additional steps or inconveniences, e.g., requiring a user to prove identity using fingerprinting or face recognition. For most virtual assistant devices, a wake word that begins a voice command is spoken in the voice of the owner of the device for the command to be accepted as genuine. However, a determined attacker can obtain or fabricate the target user's own voice.

A laser attack can possibly be foiled by redesigning the virtual assistant device, e.g. by incorporating a light shield around the microphone; however, this can have performance, cost, and aesthetic implications. Also, this does not address the problems for devices that do not include such a light shield, which includes millions of voice-activated devices that are currently

in use. In any case, a light shield does not protect against an electromagnetic attack, as such shields are transparent to electromagnetic waves.

DESCRIPTION

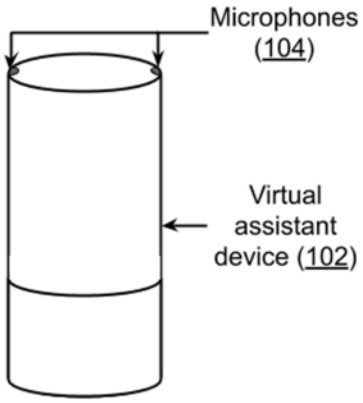

Fig. 1: Virtual assistant with multiple microphones

As illustrated in Fig. 1, voice-based virtual assistant devices (102) typically have two or more microphones (104) to better capture user commands. The microphones are designed to have a similar response to acoustic waves, but are likely to have very different responses to light or EM waves. For example, one microphone might have a peak electromagnetic demodulation efficiency at 4 GHz while the other might have a peak electromagnetic demodulation efficiency at 8 GHz. An electromagnetic attack mounted at 4 GHz might be demodulated by the two microphones at a power-level difference ratio of one-hundred-to-one, e.g., 20 dB.

Under normal usage, the distinct microphones pick up voices at similar, or marginally different, levels depending on the speakers' voice volume, distance, direction, environment settings, etc. Under laser or EM attack, even if the laser or EM waves reach the microphones at similar energy levels, due to the differing sensitivities of the microphones to laser or EM waves, the command transmitted over the laser or EM wave can be received at the two microphones at

widely different power levels, e.g., by 20 dB or more. A volume difference for a command that is substantially greater than the volume difference under regular-usage mode can thus be indicative of a laser or EM attack.

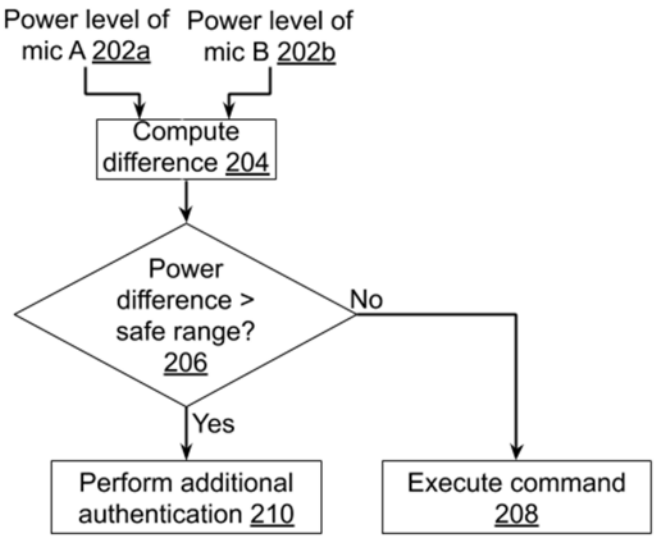

Fig. 2: Detecting a laser or EM attack by computing power differences between microphones

Fig. 2 illustrates detecting a laser or EM attack on a virtual assistant, per techniques of this disclosure. Based on the setup and regular-usage information, e.g., the maximum and the minimum volumes under regular usage, a nominal or safe range for the ratio of the power level (volume) of commands received by the device is defined. The power levels of a command received at two or more microphones of the device are determined (202a-b).

The difference of the power levels is computed (204). If the power-level difference of the received command is outside a safe range (206), then the command is flagged as suspicious, and additional steps are taken to authenticate the command (210) before accepting (or rejecting) it. A received voice command that falls within the safe range of power-level difference is executed (208) without additional authenticating steps. Additionally, a machine learning (ML) model can be used to differentiate authentic from suspicious usage, using, e.g., power-level differences as

one of the features input to the ML model. The use of such an ML model enables adaptation and improved performance over time.

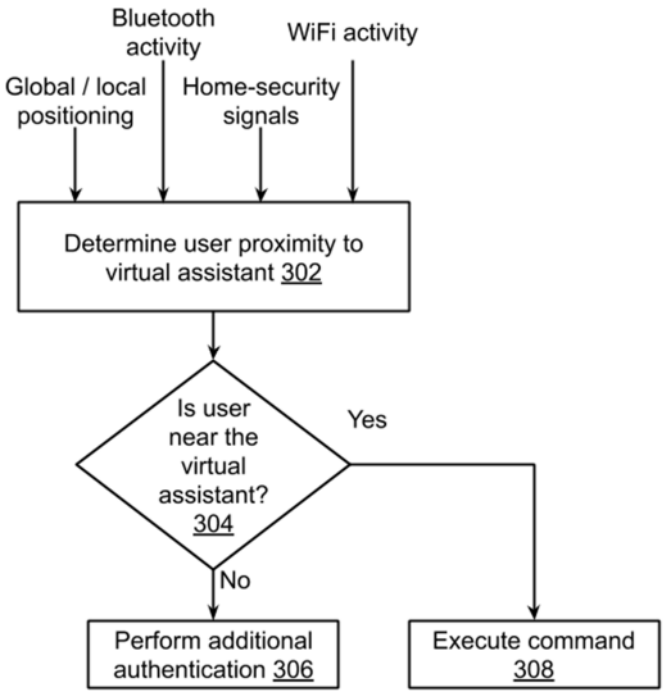

Fig. 3: Detecting a laser or EM attack by determining the proximity of a user to the virtual assistant

Fig. 3 illustrates an additional or alternative technique to detect a laser or EM attack on a virtual assistant. With user permission, the virtual assistant obtains values of various factors such as location (determined by global or local positioning systems), Bluetooth activity, signals from home-security devices, WiFi device activity, etc., to determine if an authentic user is near the virtual assistant device (302). If a command is received while the authentic user is not determined to be proximate to the device (304), the command is treated as suspicious and the user is requested to take additional steps to authenticate the command (306). If the user is determined to be nearby, then the command is executed (308).

Further to the descriptions above, a user is provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may

enable collection of user information (e.g., information about a user's current location or wireless environment). In addition, certain data is treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity is treated so that no personally identifiable information can be determined for the user. Thus, the user has control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques to protect a voice assistant device from laser or EM attacks. The techniques leverage the differing sensitivities of the multiple microphones of a virtual assistant device to a laser or to EM waves to flag a received voice command as authentic or suspicious. The techniques also determine a user's proximity to the virtual assistant device to determine if a received command is authentic or suspicious.

REFERENCES

[1] "Researchers used a laser to hack Alexa and other voice assistants" https://www.cnn.com/2019/11/04/tech/alexa-siri-laser-attack-research/index.html accessed Feb. 19, 2020.

[2] "With a Laser, Researchers Say They Can Hack Alexa, Google Home or Siri" https://www.nytimes.com/2019/11/04/technology/digital-assistant-laser-hack.html accessed Feb. 19, 2020.

[3] Hackers Can Use Lasers to 'Speak' to Your Amazon Echo or Google Home https://www.wired.com/story/lasers-hack-amazon-echo-google-home/ accessed Feb. 19, 2020.

[4] “Lasers can seemingly hack Alexa, Google Home and Siri”

https://www.cnet.com/news/lasers-can-seemingly-hack-alexa-google-home-and-siri/ accessed Feb. 19, 2020.