Using Open Channels to Trigger the Invited, Unintended Consequences of the Internet of Things

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

| Citation         | Siegel, Joshua Eric and Sanjay Sarma. “Using Open Channels to Trigger the Invited, Unintended Consequences of the Internet of Things.” IEEE Security and Privacy 17, 3 (May 2019): 49 - 55 © 2019 IEEE |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| As Published     | http://dx.doi.org/10.1109/msec.2019.2901430                                                                                                                                                                                                                  |
| Publisher        | Institute of Electrical and Electronics Engineers (IEEE)                                                                                                                                                                                                      |
| Version          | Author’s final manuscript                                                                                                                                                                                                                                |
| Citable link     | https://hdl.handle.net/1721.1/128496                                                                                                                                                                                                                       |
| Terms of Use     | Creative Commons Attribution-Noncommercial-Share Alike                                                                                                                                                                                                       |
| Detailed Terms   | http://creativecommons.org/licenses/by-nc-sa/4.0/                                                                                                                                                                                                          |
Using Open Channels to Trigger IoT’s Invited, Unintended Consequences

Joshua Siegel  
Michigan State University Department of Computer Science and Engineering

Sanjay Sarma  
Massachusetts Institute of Technology Department of Mechanical Engineering

We describe bridging IoT’s air gap using device speakers to communicate with voice assistants. We recommend protecting voice channels using two-factor authentication, contextual monitoring, and voice source embedding. These techniques are necessary when connecting critical applications to open channels, and may apply to other exploits resulting from IoT’s growing capabilities.

1 The Internet of Opportunities and Unintended Consequences

Millions of people have Internet of Things (IoT) devices in their homes, workplaces, vehicles, and on their bodies. These devices use sensing, connectivity, inference, and action to improve comfort, health, safety, and efficiency while blending seamlessly into their environments. Changes in these capabilities invalidate assumptions upon which cybersecurity has been built, creating interstitial weakness enabling dire consequences. 1

This paper considers how the “air gap” between a high-criticality physical system and external networks can be eroded by emerging capabilities and unanticipated risks. Specifically, we identify vulnerabilities at the intersection of two convenience- and safety-centric devices (security cameras and smart locks) that, combined, subvert each other’s benefits. We explain how adversaries can leverage vulnerabilities in network webcams, using two-way voice to activate conversational devices trusted to control physical infrastructure, and present risk-mitigating solutions including distance bounding, biometrics, context-aware firewalling, and steganographic source signing.

While this article describes how voice interfaces compromise a planned airgap, we close by exploring how the proposed solutions apply to securing IoT’s emergent technologies.

2 IoT Security Devices

Smart locks enable the use of temporary and revocable digital credentials to allow individuals remote or hands-free access to a property. These locks are valued by property managers and the infirm, who use them to let repair people in remotely or without standing up. 2

Connected locks are hackable, allowing untrusted users remote access. 3 Manufacturers may patch vulnerabilities, but the only way to prevent Internet-enabled exploits is to remove the lock from a wide-area network. Limiting access to short-range protocols like
ZigBee, Z-Wave, or Bluetooth reduces the attack surface, but these can be compromised by man-in-the-middle attacks. The only sure-fire way to prevent distance-based attacks is to eliminate connectivity – negating smart locks’ benefits. Often, users tolerate a degree of insecurity in exchange for convenience, reducing exposure by other means.

One approach is to use networked security cameras, which provide individuals a real-time look at their property. These cameras detect motion or sound, and allow remote audio broadcasting to soothe an agitated pet or demand that an intruder leave. Like smart locks, these cameras are susceptible to attack. Unlike locks, the risks of a compromised webcam are poorly understood – resulting in under-secured cameras, like those the Mirai Botnet employed to disrupt the Internet.4,5

The following sections consider how vulnerable “purely digital” cameras can leverage commonplace intermediate devices to jump the air gap between physical and digital, wresting control of off-line smart locks. The described attack is improbable, but feasible, and representative of an emerging challenge as IoT advances: the same capabilities that make IoT useful also make it vulnerable.

3 Home Assistants: Helpful or Harmful?

The Internet of Things relies upon the concept seamless device and service interoperability. This is disadvantageous when such interactions are unexpected. Voice assistants have found uses in homes, cars, and mobile devices, but an unanticipated consequence is that these devices leave open a command channel susceptible to intentional corruption.6

Consider a home with an Amazon Echo, Google Home, or Apple TV. All three devices allow voice control for smart devices over Wi-Fi, Bluetooth, or ZigBee. These devices are an entry point for attackers as they use an open channel to control “trusted” devices. In one possible attack, a compromised camera’s speaker could command the voice assistant to relay a message onto the “protected” local network.

This is not dissimilar from a criminal shouting “Alexa, unlock the door” through an open window – a compromised camera with two-way audio can send the same command to “air-gapped” devices in and around one’s home, allowing an adversary to unlock doors, set HVAC to disable heat causing pipes to burst, or ordering expensive ride share cars to an unknown individual’s home. The more integrated a smart home is, the larger the potential for subversion.

The attack’s elements are already in place – for example, advertisements trigger voice assistants7 while search engines like Shodan.io help adversaries probe for insecure devices (a January 2018 search for “Netwave IP camera,” one of thousands of exploitable devices, turned up 96, 000 results). Manufacturers’ reuse common default passwords, while firmware vulnerable to authentication bypass or high-speed brute force attacks is widespread.8 IoT devices may even leak sensitive location data.9

An adversary can use administrative access to expedite this attack, setting motion- or sound-based triggers.10-12 The same camera used to issue voice commands could capture a
homeowner’s “two-factor” authentication, intercepting an audio passcode, or to identify when the home is empty to facilitate a stealthy break-in. Deep learning could be used to generate audio that sounds like a homeowner to “fuzz” a passcode or biometric voice match, or a social engineering phone call could be used to underpin a replay attack. Going a step further, audio sent through the camera could be played outside the range of human hearing and begin with a “mute” request, keeping all but the most astute homeowners in the dark.

With IoT’s growing scale, the likelihood of a randomly-selected house having a smart lock and voice assistant is increasing. If there is no camera, an adversary might find other entry points – like baby monitors and children’s toys!

4 A Case Study in Villainizing Voice

In this section, we describe an attack disengaging an air-gapped smart lock using a compromised webcam.

First, a hacker uses Shodan or a network scanner to identify unsecured devices with audio capabilities. The attacker may then conduct a port scan to find other devices on the network, selecting a home that responds to ping requests indicative of a voice assistant and a smart lock. With the target identified, the attacker tests common camera entry vectors including default passwords, firmware-specific backdoors, or open Telnet ports. Then, the adversary sets up motion triggers to record video and audio in an attempt to capture the homeowner commanding a passcode-protected door unlock.

Once in possession of the passcode, the attacker watches the camera to determine when the occupants leave for work. He or she then initiates a replay, mimicry, or fuzzing attack to unlock the door. Replay attacks are possible in non-cooperative cases (no target participation required), but mimicry (organic and digital attempts to replicate another individual’s voice) or voice conversion attacks (digital masking of the attacker’s voice as the target’s) could also be used, depending on the home assistant’s use of voice biometrics. Finally, the attacker disables the camera and enters the home with no sign of forced entry. A VPN masks their identity in the event the camera logs access.

To test this attack, we configured a smart home with an August Smart Lock with Connect, an Amazon Echo Dot, an iPhone X and Apple TV 4, and a Foscam C1 Lite webcam. The August lock was chosen due to its compatibility with the Amazon Echo platform and Apple’s HomeKit, which uses biometric voiceprinting to identify users. August also allows remote unlocking (with PIN), whereas some manufacturers disallow voice unlocking. The Foscam C1 Lite was chosen because it was susceptible to Mirai, and the device features two-way audio. The August lock can be configured to use Wi-Fi, Bluetooth, and Z-Wave, and while our attack uses the Wi-Fi connection, this attack targets all protocols.

A representative system and attack schematic is shown in Figure 1.

Testing the in-home attack, we commanded Apple HomeKit and the August Home Alexa skill to unlock the door. Live audio relayed through the C1 speaker successfully unlocked the door on both assistants, while the Echo was more susceptible to replayed audio. Poor
connectivity and compression artifacts caused commands to occasionally fail (particularly against Apple’s biometric filtering), though it was clear that webcam audio is generally treated as authentic voice, allowing a compromised camera to jump the airgap.

We then estimated how long it would take an adversary to “fuzz” a passcode for devices with no request throttling (the August Smart Lock Alexa app has a rate limit and disables voice unlocking after three incorrect attempts, but high-rate brute forcing or bypass attacks are not uncommon 8). For this calculation, we assumed a four-digit passcode (10,000 possibilities) and no artificial rate limiting.

A command-and-pin request takes 4.5 seconds, with a 5 second unsuccessful unlock response and a 2.5 second gap required between commands, leading to a 12.5s cycle time. This allows for 2, 400 attempts in an 8-hour work day. To accelerate brute-forcing, an adversary could work longer hours by timing the attack for when a homeowner is out of town, or could fuzz the passcode using ultrasonic requests while the target sleeps. In any case, a four-digit pin could be broken within a work week and accelerated playback or prioritized passcode dictionaries could reduce this.

This approach is convoluted, but only uses known exploits. It is reasonable to assume that other zero-day vulnerabilities exist to simplify the attack (rate-limit bypasses, device scanning techniques, etc.), and that a well-resourced adversary could exploit these. The scale of the IoT also presents a unique challenge – elements of this attack are scriptable and simplified by advances in deep learning, and many homeowners shop based on price rather than capabilities, leading to an installed base of poorly-secured smart locks (the August implements several best practices).

5 Addressing Open-Channel Vulnerabilities

There are several technical solutions for mitigating the risk of adversaries jumping the voice air gap.

First, IoT systems must incorporate improved logging, log retention, and administrator notification to ensure authentication attempts are recorded immutably. We must also elevate device capabilities from speech recognition to voice recognition. Biometrics will improve security in mission-critical voice systems, e.g. banking and aspects of home automation. 16 Today, Apple uses this approach to authorize access to Siri while Google uses “Trusted Voice” and banks use Nuance’s “Free Speech.” 19 It may even be possible to infer a voice’s emotional state 20 to detect whether a voice command was issued by humans under duress.

These techniques, however, remain vulnerable to impersonation attacks – Google requires users to acknowledge that voice is less secure than passwords due to feature dimension reduction and increased collision likelihood. 6 Improvements are necessary to prevent replay attacks and to address political challenges (the characteristics used for individual recognition constitute a biometric personal record in the Privacy Act, FERPA, and HIPAA). 2 Allowing users to choose unique wake-words and device names will further reduce adversaries’ abilities to successfully conduct a brute force attack. As device
capabilities improve, it will be imperative to tailor models to individual users – lest domain squatting play a role in conversational interfaces.  

Mandatory two-factor authentication would also mitigate risk by using data from webcams, passive infrared motion sensors, or Bluetooth beacons to ensure that someone is near the house or approves each request from a trusted device. Typed passcodes or fingerprints can provide this second degree of protection. Wearable devices can record vibration to continuously authenticate voice commands, and acoustic device proximity bounding is also feasible. These second-factor inputs can become part of a context-aware “Cognitive Firewall” considering data from multiple devices and using AI to learn typical behaviors and request human-in-the-loop approval when an atypical request is received. Similarly, a Cognitive Supervisor could monitor the home’s network to detect anomalous traffic, alerting the homeowner to the presence of an observer before their PIN has been intercepted. The more inputs there are to such a model, the more an adversary must compromise in order to stage an effective attack.

To add another point of resilience, we propose that future IoT devices use convolved signals to encode source metadata into outbound audio (or visual) signals. This signature injection would take place in hardware – so that software vulnerabilities could not bypass its insertion – and would include a unique identifier comprised of information about the steganography protocol (standard) used, device manufacturer, device type, unique serial, and software revision. This unique encoded information would be multiplexed with outbound media and comprise a “Thing Identifying Number,” or TIN, identifying device provenance and akin to a vehicle’s VIN.

Implemented properly, steganographically modified media would appear to be unedited to humans, but voice assistants, webcams, and other IoT devices could extract metadata in order to identify audio or video as coming from a particular (trusted or untrusted) source. Devices could then act based upon specific rules taking device provenance into consideration. This would also allow voice assistants to ignore replayed commands, such as the earlier-mentioned TV ad, while high-priority, trusted messages could be signed differently to ensure their receipt. This is shown in Figure 2. Note that trust could be established using a Certificate Authority model, which would allow device certificates to be revoked when corruption is detected. This same approach can be used for visual signals, such as validating that a QR-code ticket to a sporting event hasn’t been scanned from multiple devices, or that a video was captured by a particular IP camera. These modifications could use sub-block pixel modification, phase shift, or amplitude modulation to embed new information on top of de-facto media standards.

As a last means of security, seed/key pairs, rolling codes, or other handshaking could be used to ensure that an adversary could not simply inject high-priority metadata into their messages. Open channel broadcasting and receiving devices could belong to a consortium or otherwise be certified as “safe” when implementing standardized versions of this technology. The ability to verify open-channel inputs will become increasingly important as IoT engages with critical infrastructure, and as conversational commerce and voice banking grow. There would be significant challenges in developing and enforcing appropriate standards designed for a growing, global market, and in hiding the metadata sufficiently well to avoid the perception of signal corruption. However, the value could be immense to IoT practitioners from cybersecurity and other perspectives (e.g. using
metadata to communicate a sensors’ noise characteristics, so that an application can appropriately bias data weights depending on trustworthiness).

While this section focuses on addressing voice-related side channel attacks, other threats target the hardware, software, and people associated with IoT. In the next section, we consider approaches to securing IoT against existent and emerging threat models.

6 Protecting Other Channels

The IoT has brought about opportunities, capabilities, and vulnerabilities in equal measure, though consumers and developers are often ignorant to the risks of hyperconnectivity. We must recognize that innovations are “broken” as quickly as breakthrough innovations happen.

IoT’s growth has increased the number of attackable endpoints. Individuals and organizations put these devices on sensitive internal networks where they can intercept data or take down other devices from behind firewalls. Problems are often worsened by deep-rooted design issues like hard-coded passwords or constraints imposed by outsourcing hardware to the lowest bidder. Beyond voice attacks, we must think broadly about how to protect the IoT using a blend of policy, technology, and education.

With regard to policy, NIST, CIS, IEEE, the FTC and DoD have made recommendations for IoT security, but these guidelines are not necessarily future-proof and enforcement is lacking. Unified guidelines cognizant of an evolving technology landscape and improved compliance will benefit outcomes across the board. Of course, IoT is used outside the United States – and other countries have an equal right to participate in defining guidelines working towards the global good.

On the technology side, it may be possible to reduce the number and severity of attacks by implementing industry best practices, including end-to-end encryption, requiring changes from default passwords and regular updates thereafter, and including provisions for secure over-the-air firmware updates. The use of intelligent and adaptive context-aware firewalls can further reduce the risk of a malicious command being passed onto an end device. We must begin to build devices and services that put at least as much emphasis on security as cost-effectiveness, and consider both the challenges and affordances of the IoT (increased scale and potential for impact, but equal potential for disruption) throughout the design process. For manufacturers, this means over-provisioning hardware to leave computational overhead for security, developing firmware with security in mind, implementing best practices for over-the-air updates, and planning for a long service life (as many IoT devices are installed in durable goods, rather than “disposable” consumer goods).

Education is equally important. People must become “informed digital citizens,” with an awareness of the capabilities and risks associated with their devices, and their own responsibilities for ensuring safe operation. This includes users implementing best computing practices, like regularly changing passwords, avoiding password reuse, enabling two-factor authentication, and keeping software up-to-date (as-is, one report showed that 96% of devices configured with their default passwords remained configured
Perhaps most importantly, we can minimize risk exposure in the first place by being thoughtful about the devices we install and services we use. This includes mapping out their full capabilities (intentional and incidental), permissions (implicit or explicit), configuration, and contemporary and future attack vectors, and evaluating the benefit of a device or system relative to its potential risks (e.g. a personal home lock is likely safe to put on the IoT, but the lock on a Fortune 100 data center is a riskier proposition). Developers and installers must take an active role in this education process.

While the benefits of each individual improvement may be small, these changes are critical to the long-term success of the IoT. Increasing the percentage of educated users and resilient devices will go a long way toward creating a “heard immunity” for the IoT, limiting the potential for adversaries to build massive botnets or to build long attack chains extending well inside a firewall’s defensive perimeter.

7 Conclusion

This article considered a realistic but improbable dystopian future where webcams go rogue, using advances in voice interfaces to undermine smart lock security. We posed possible solutions to this problem, including increased use of biometrics, distance bounding, and media source signing. These solutions target a specific problem, but the larger challenge is that advances in technical capabilities can easily invalidate the assumptions underpinning cybersecurity.

With increased awareness of system capabilities, proactive configuration, and active monitoring, the IoT can improve all our lives without undue risk. This requires constant education, use of industry best practices, and working with the understanding that new problems require new thinking to solve.

As practitioners, our ultimate goal should be to ensure that the IoT is something users trust, rather than fear. The suggestions laid out in this document, from encoding multimedia metadata to improved education, standards, and enforcement, will go a long way to achieving that goal. And, while there’s always another vulnerability to be found (a homeowner’s keys could be pickpocketed even if their smart lock is secured), these protections can make the IoT at least as secure as “business as usual” and lead to incredible societal benefit.

References

1. B. Duncan, A. Happe, and A. Bratterud, “Enterprise IoT security and scalability: how unikernels can improve the status quo,” in 2016 IEEE/ACM 9th International Conference on Utility and Cloud Computing (UCC). IEEE, 2016, pp. 292–297.
2. S. Gray. (2016, April) Always on: Privacy implications of microphone-enabled devices. [Online]. Available: https://web.archive.org/web/20180511155208/https://fpf.org/wp-content/uploads/2016/04/FPF_Always_On_WP.pdf
3. Jmaxxz, “Backdooring the frontdoor,” 11 2016, dEF CON 24. [Online]. Available: https://media.defcon.org/DEF%20CON%2024/DEF%20CON%2024%20presentations/DEF CON-24-Jmaxxz-Backdooring-the-Frontdoor-UPDATED.pdf

4. S. Sarma and J. Siegel. (2016, November) Bad (internet of) things. [Online]. Available: https://web.archive.org/web/20180325135538/https://www.computerworld.com/article/3146128/Internet-of-things/bad-internet-of-things.html

5. U. Lindqvist and P. G. Neumann, “The future of the internet of things,” Communications of ACM, vol. 60, no. 2, pp. 26–30, 1 2017. [Online]. Available: http://doi.acm.org/10.1145/3029589

6. H. Feng, K. Fawaz, and K. G. Shin, “Continuous authentication for voice assistants,” arXiv preprint arXiv:1701.04507, 2017.

7. J. Kastrenakes. (2017, April) Burger king's new ad forces google home to advertise the whopper. [Online]. Available: https://web.archive.org/web/20181026182003/https://www.yalehome.com/en/yale/yale-Engineering-of-Science,-Engineering-and-Technology,-International-Journal-of-Biomedical-and-Biological-practices-for-presentation-at-J.-Sanchez-Communications-and-Network-Security-security-and-S.-Notra,-M.-Siddiqi,-H.-H.-Gharakheili,-V.-Sivaraman,-and-R.-Boreli,-"An-experimental-study-of-IoT-vulnerabilities,"

8. S. Shekyan and A. Harutyunyan, “To watch or to be watched: Turning your surveillance camera against you,” April 2013, Hack In the Box 2013. [Online]. Available: https://web.archive.org/web/20180910113323/http://conference.hitb.org/hitbsecconf2013ams/materials/D2T1%20Sergey%20Shekyan%20and%20Artem%20Harutyunyan%20-%20Turning%20Your%20Surveillance%20Camera%20Against%20You.pdf

9. B. Krebs. (2018, June) Google to fix location data leak in Google Home, Chromecast. http://web.archive.org/web/20180916095055/https://krebsonsecurity.com/2018/06/google-to-fix-location-data-leak-in-google-home-chromecast/

10. M. Capellupo, J. Liranzo, M. Z. A. Bhuiyan, T. Hayajneh, and G. Wang, “Security and attack vector analysis of IoT devices,” in International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage. Springer, 2017, pp. 593–606.

11. T. Lackorzynski and S. Koepsell, “’Hello Barbie’-hacker toys in a world of linked devices,” in Proceedings of Broadband Coverage in Germany; 11. ITG-Symposium. Berlin: VDE Verlag GMBH, March 2017, pp. 41–47.

12. C. van den Bogaard, “Security analysis of cloud-based video cameras,” in Proceedings of the Twente Student Conference on IT, July 2017.

13. G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, “DolphinAttack: Inaudible voice commands,” in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2017, pp. 103–117.

14. M. Stanislav and T. Beardsley, “Hacking IoT: A case study on baby monitor exposures and vulnerabilities,” Rapid7, 2015. [Online]. Available: https://web.archive.org/web/20180520171300/https://www.rapid7.com/docs/Hacking-IoT-A-Case-Study-on-Baby-Monitor-Exposures-and-Vulnerabilities.pdf

15. S. Notra, M. Siddiqi, H. H. Gharakheili, V. Sivaraman, and R. Boreli, “An experimental study of security and privacy risks with emerging household appliances,” in 2014 IEEE Conference on Communications and Network Security, Oct 2014, pp. 79–84.

16. J. Sanchez-Casanova, I. Goicoechea-Telleria, J. Liu-Jimenez, and R. Sanchez-Reillo, “Best practices for presentation attack detection evaluation in voice biometrics,” World Academy of Science, Engineering and Technology, International Journal of Biomedical and Biological Engineering, vol. 5, no. 5, 2018.

17. (2017, March) Yale smart locks now offer amazon Alexa compatibility via SmartThings and Wink. [Online]. Available: https://web.archive.org/web/20181026182003/https://www.yalehome.com/en/yale/yale-home/news/news/2017/yale-announces-amazon-alexacompatibility/

18. L. Tung. (2017, June) Hardcoded password in Foscam’s C1 gives remote access to stored video, audio. [Online]. Available:
19. I. Agadakos, C.-Y. Chen, M. Campanelli, P. Anantharaman, M. Hasan, B. Copos, T. Lepoint, M. Locasto, G. F. Ciocarlie, and U. Lindqvist, “Jumping the air gap: Modeling cyber-physical attack paths in the internet-of-things,” in Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and Privacy, ser. CPS ’17. New York, NY, USA: ACM, 2017, pp. 37–48. [Online]. Available: http://doi.acm.org/10.1145/3140241.3140252

20. H. Jin and S. Wang, “Us patent application 15/457,846 - voice-based determination of physical and emotional characteristics of users,” October 2018.

21. D. Kumar, R. Paccagnella, P. Murley, E. Hennenfent, J. Mason, A. Bates, and M. Bailey, “Skill squatting attacks on Amazon Alexa,” in 27th USENIX Security Symposium (USENIX Security 18). Baltimore, MD: USENIX Association, 2018, pp. 33–47. [Online]. Available: https://www.usenix.org/conference/usenixsecurity18/presentation/kumar

22. M. Mitrevski, “Conversational interface challenges,” in Developing Conversational Interfaces for iOS: Add Responsive Voice Control to Your Apps. Berkeley, CA: Apress, 2018, pp. 217–228. [Online]. Available: https://doi.org/10.1007/978-1-4842-3396-2_8

23. N. Z. Gong, A. Ozen, Y. Wu, X. Cao, R. Shin, D. Song, H. Jin, and X. Bao, “PIANO: Proximity-based user authentication on voice-powered internet-of-things devices,” arXiv preprint arXiv:1704.03118, 2017.

24. J. E. Siegel, S. Kumar, and S. E. Sarma, “The future internet of things: Secure, efficient, and model-based,” IEEE Internet of Things Journal, vol. 5, no. 4, pp. 2386–2398, August 2018. [Online]. Available: https://doi.org/10.1109/JIOT.2017.2755620

25. G. Hua, J. Huang, Y. Q. Shi, J. Goh, and V. L. Thing, “Twenty years of digital audio watermarking - a comprehensive review,” Signal Processing, vol. 128, pp. 222 – 242, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168416300263
Figure 1: This figure shows a proposed camera-to-smart lock attack. An attacker identifies homes with cameras, smart locks, and voice assistants and then compromises the camera to record audio and observe the homeowners. When the homeowners are out, the attacker transmits audio to the voice assistant and commands the door to unlock. Finally, the attacker gains physical access.

Figure 2: This figure shows the last stages of our attack scenario, but this time, the camera digitally watermarks its outbound audio. The voice assistant is able to detect the command is being transmitted through another devices and is therefore not authentic. The assistant disregards the command, logging it and optionally sending an alert to the homeowner about the thwarted attack.

At the time of submission, Josh Siegel (jsiegel@msu.edu) was a Research Scientist at the Massachusetts Institute of Technology (MIT). He is now an Assistant Professor of Computer Science and Engineering at Michigan State University (MSU) and the lead instructor for MIT's IoT and DeepTech Bootcamps. He received Ph.D., S.M. and S.B. degrees in Mechanical Engineering from MIT. Josh and his automotive companies have been recognized with accolades including the Lemelson-MIT Student Prize and the MassIT Government Innovation Prize. He has multiple issued patents, published in top scholarly venues, and been featured in popular media. Dr. Siegel's ongoing research develops architectures for secure and efficient...
CONNECTIVITY, APPLICATIONS FOR PERVERSIVE SENSING TO VEHICLE DIAGNOSTICS, AND NEW APPROACHES TO AUTONOMOUS DRIVING.

SANJAY SARMA (sesarma@mit.edu) RECEIVED A BACHELOR’S DEGREE FROM THE INDIAN INSTITUTE OF TECHNOLOGY, A MASTER’S DEGREE FROM CARNEGIE MELLON UNIVERSITY AND A PH.D. DEGREE FROM THE UNIVERSITY OF CALIFORNIA AT BERKELEY. HE IS THE FRED FORT FLOWERS (1941) AND DANIEL FORT FLOWERS (1941) PROFESSOR OF MECHANICAL ENGINEERING AND THE VICE PRESIDENT FOR OPEN LEARNING AT MIT. HE CO-FOUNDED THE AUTO-ID CENTER AT MIT AND DEVELOPED MANY OF THE KEY TECHNOLOGIES BEHIND THE EPC SUITE OF RFID STANDARDS NOW USED WORLDWIDE. HE WAS ALSO THE FOUNDER AND CTO OF OATSYSTEMS. HE HAS AUTHORED ACADEMIC PAPERS IN COMPUTATIONAL GEOMETRY, SENSING, RFID, AUTOMATION, AND CAD, AND IS A RECIPIENT OF NUMEROUS AWARDS FOR TEACHING AND RESEARCH.