Designing Secure and Usable Wake Words

(Technical Paper)

Regulating the Gold Rush: Using Analogies to Legislate Data Privacy in Smart Speakers

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree

Bachelor of Science, School of Engineering

Joshua Sahaya Arul

Fall, 2020

Technical Project Team Members

Timothy Han

Andrew Wang

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature _____ Date _____

Joshua Sahaya Arul

Approved _____ Date _____

Yuan Tian, Assistant Professor of Computer Science, Department of Computer Science

Approved _____ Date _____

Kathryn A. Neeley, Associate Professor of STS, Department of Engineering and Society

Introduction: The Rise of Smart Speakers and Their Risks

Over 25% of U.S. households owned at least one smart speaker in 2019, and by 2025, it is projected to increase upwards of 75% (Kinsella, 2019, n.p.). As smart speakers' user base increases, so have their access to their users' private information (Gao et al., 2018, p. 27). Their information, however, is not safeguarded. Anyone who walks up to a speaker is free to place an order on the owner's Amazon account, "drop in" on a conversation in another room, or access the owner's fitness logs. An attacker can use a variety of methods to invoke malicious voice commands from a distance. Furthermore, smart speakers can be misactivated without the user knowing, proceeding to record what should have been a private conversation ("Your Microphone Array", 2020, p. 1; Huang et al., 2020, p. 1). Voice assistants are supposed to activate and record the user's question/command only after a user speaks the wake-up word (e.g. "OK Google", "Alexa"). When a user says a similar phrase (e.g. "okay…to go", "I care about") in everyday conversation, voice assistants can mistake it for being the wake-up and proceed to record possibly sensitive audio (Dubois et al., 2020, p. 14). Today, owning a smart speaker comes with significant privacy risk, which one study found was a major deterrent for smart speaker adoption among consumers (Lau et al., 2018, p. 10). Yet, many consumers are willing to make this tradeoff and smart speakers continue to rise.

One of the reasons why consumers love smart speakers is because of how easy they are to use. Google found that fifty-three percent of people who own a smart speaker said "it feels natural speaking to it, with many saying it feels like talking to a friend" (Kleinberg, 2018, n.p.). But how can privacy issues be stopped from killing an otherwise useful, high-potential tool? In the first section, my research and I propose to design secure and usable wake-up words for smart speakers, which can be used by smart speaker manufacturers. Even if this system is successful,

however, it is critical that consumers feel that their data is safe with smart speakers. Lau et al. remarked that “despite the privacy-intrusive potential of smart speakers’ ability to continuously record voices in intimate spaces such as the home, consumers’ privacy concerns with smart speakers and how these concerns affect smart speaker adoption and use have not been studied in detail (2018, p. 2). In my STS research, I will expand upon existing work to investigate consumers' perceived concerns and their effect on smart speaker adoption. This information should help identify actions corporations and governments can take to assure consumers that smart speakers will protect their privacy.

Technical Topic: Designing Secure and Usable Wake-Up Words

Smart speakers play an integral role in the rapid development of Internet of Things (IoT) devices and smart homes since they are the hub for popular smart home platforms ("Your Microphone Array", 2020, p. 1). Each IoT device comes with a mobile application that, unless another device is made by the same manufacturer, usually lacks integration with any other device/app. Smart speakers fill that gap by integrating with the vast majority of smart devices. With a smart speaker, the owner can use their voice to dim the lights, set the thermostat, speak to the visitor waiting at the door, push their security camera feed to their TV, and much more. As smart speakers and IoT devices are both projected to increase by nearly threefold by 2025 (see Figure 1), it is reasonable to expect the amount of control smart speakers hold will increase as well (Dahlqvist et al., 2019, n.p.; Kinsella, 2019, n.p.).

The rise of smart speakers is concerning because smart speakers can be activated without the owner's knowledge in a variety of ways. Figure 2 illustrates the basic interaction between users and smart speakers, and in addition, how third parties can intrude that interaction. The "inherently broadcast nature of voice unlocks a door for adversaries to conduct spoofing attacks

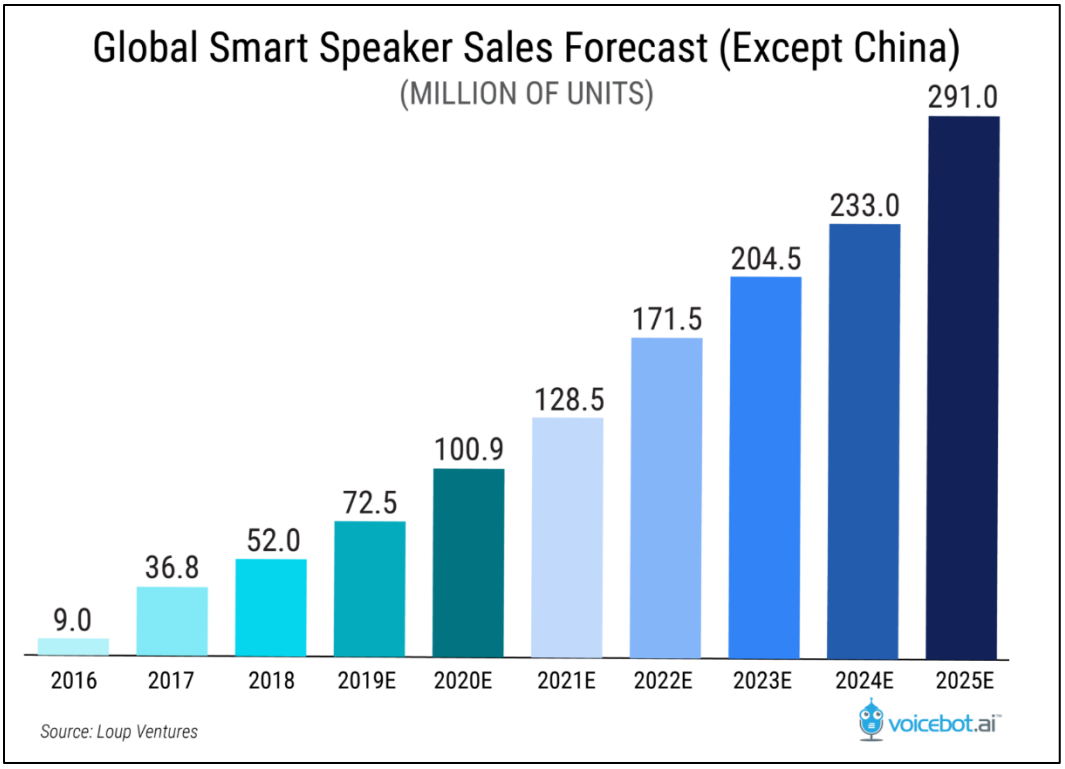

Figure 1: Bar graph of global smart speaker sales in millions of units from 2016 to 2025. The data points from 2019 to 2025 are the projected sales. Market trends suggest that smart speakers will continue to grow over the next five years (Kinsella, 2019, n.p.).

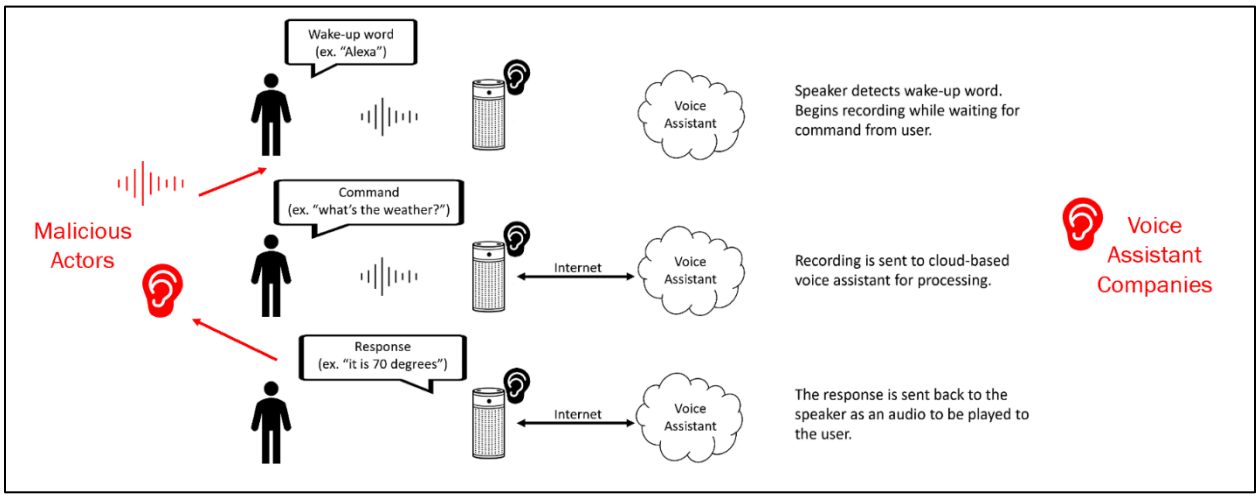

Figure 2: Simplified workflow of user-smart speaker interaction. A malicious actor within range can use a type of spoofing attack to execute a command whether the user is present and/or eavesdrop on sensitive information played aloud by the speaker. Also note that the smart speaker is always listening for the wake-up word, and if detected incorrectly, can proceed to record and send the audio to the voice assistant company (created by author).

to inject malicious commands" into smart speakers (Gao et al., 2018, p. 27; "Your Microphone Array", 2020, p. 1). Imperfections in hardware manufacturing also allow for (inaudible) ultrasound attacks, such as Dolphin attack and BackDoor attack ("Your Microphone Array", 2020, p. 1). Smart speakers can also be triggered accidentally, even in the absence of malicious actors. One study tested ten unique voice assistant/wake word combinations and identified more than 1,000 sequences that incorrectly activated speakers (Schönherr et al., 2020, p. 1).

Misactivation is an issue because voice-assistant companies, such as Amazon and Google, have employees that listen to a random sample of recordings for quality assurance. That means companies will, unbeknownst to the user, be listening and discovering unsavory things. Dubois et al. found that smart speakers are misactivated nearly once per hour, with 10% of them lasting at least 10 seconds (2020, p. 1). This high rate explains how, as one Apple whistleblower revealed, "'there have been countless instances of recordings featuring private discussions between doctors and patients, business deals, seemingly criminal dealings, sexual encounters and so on'" (Hern, 2019, n.p.). Smart speaker activation needs to be improved because otherwise, consumers' private lives will be vulnerable to malicious intruders and eavesdropping smart speaker companies.

The technical research paper will focus on designing safe and usable wake-up words. Schönherr et al. and Dubois et al. have studied patterns within phrases that misactivated voice assistants, and Schönherr et al. developed a method for crafting artificial triggers (2020, pp. 8-11; 2020, p. 14). Their findings can be a starting point for implementing a machine learning algorithm that will score the security of a given phrase. Then, the secure phrases can be evaluated for usability to select the best wake-up words. The biggest foreseeable challenges will be to generalize the security of a phrase for any accent and to develop an automated method to

quantify how usable a word is. This project should yield a process for producing secure and easy-to-pronounce wake words so that users can safely bring smart speakers into the confinements of their home.

STS Topic: Developing Trust Between Consumers and Smart Speakers

While many consumers are embracing smart speakers because of their convenience, others avoid them because of privacy and security concerns. Participants in one study shared that their primary concerns were: "device hacking; recording of private conversations; 24/7 listening activities; the collection, sharing, and storage of private data; and the 'creepy' nature of the devices" (Huang et al., 2020, p. 1). Their apprehension is well-justified; Amazon and Google have confessed listening to user conversations, and "smart devices have been shown to be able to collect a wide range of user information, including sleeping patterns, exercise routines, child behavior, and medical information" (Gao et al., 2018, p. 27). Furthermore, researchers have discovered that the data that Alexa stores in the cloud can easily be used to extrapolate additional information, such as driving routes, dinner time, and general interests (Huang et al., 2020, p. 2). Smart speakers are responsible for storing numerous sensitive data points, and many do not feel that their data is safe. In the research study conducted by Lau et al., most of non-user participants "did not trust the smart speaker company to abide by the [terms of service] in perpetuity (2018, p. 10). A couple non-users also "posited that since major companies like Yahoo and other IoT devices like baby monitors have been hacked, it is highly unlikely that smart speaker companies can guarantee safety from hackers (Lau et al., 2018, p. 10). These are the same issues raised and illustrated in the technical section (see Figure 2). The problem is not necessarily that smart speakers store sensitive information – credit card companies, mobile apps, and even car

dealerships store sensitive information as well. It is that users do not trust smart speaker companies to protect them from eavesdroppers or to not eavesdrop themselves.

Consumers rightfully should value privacy and feel confident that outsiders are not listening in or accessing their data unbeknownst to them before buying a smart speaker. Figure 2 depicts a simplified diagram flow of the data from the user to the voice assistant and back, and how smart speakers are always listening in anticipation for the wake-up word. Huang et al. found that while "non-users were certain that they would never have smart speakers in their homes, others mentioned that if these concerns were addressed, they might consider getting one in the future" (2020, p. 10). They also found that many who do use smart speakers are consciously making a compromise between utility and privacy but are "[frustrated] about such an all-or-nothing trade-off" (Huang et al., 2020, p. 9). If consumers cannot adequately trust their data in the hands of smart speakers, consumers will abstain from using them. These limited interviews also suggest, however, that if privacy concerns were addressed, consumers would more readily adopt smart speaker technology. Therefore, it is important to identify what can be done to resolve privacy issues and advance smart speaker technology.

The goal of this paper is to identify what actions can be taken, particularly by smart speaker manufacturers and governments, to build consumers' trust for smart speakers. Huang et al. has observed users' mental models of smart speakers and of coping risks, which gives insight into users' thought process in using smart speakers (2020). Lau et al. conducted a research study on consumers to investigate privacy concerns with smart speakers and how they affect smart speaker adoption, concluding with a list of features, policies, and controls smart speaker companies should provide to users to fit their security needs (2018). Their paper, however, did not consider other players in the smart speaker ecosystem, particularly the government. In this

study, actor-network theory will be used to better identify the actors at play, the conflicts between them, and how to resolve them. The main challenge is in ensuring proposed governmental policies are within the scope of the federal government according to legal precedent. Still, this research will hope to enumerate practical steps that can be taken by key players in the smart speaker landscape to bring smart speakers that will honor consumer privacy.

Conclusion

This paper will strive to improve the status quo of smart speakers by proposing solutions to increase security and trust. The technical component will yield an algorithm for evaluating the security and usability of a word, which can then be used to generate safer wake-up words for smart speakers. The STS component will help reveal what consumers need from manufacturers, governments, and other actors to reassure privacy and feel safe. With these two deliverables, consumers' daily, private conversations at home will less likely be recorded, protecting their privacy. If players within the smart speaker industry take actions accordingly, consumers should feel comfortable bringing smart speakers to their households while still being able to control and limit their privacy exposure. As a result, consumers will be more satisfied and will embrace smart speakers, encouraging further innovation in the industry.

Bibliography

5 Ways Consumers Interact With Smart Speakers. (2019, November 18). *Mindstream Media Group*. https://mindstreammediagroup.com/introduction-smart-speakers-voice-search-brand-advertisers/

Bentley, F., Luvogt, C., Silverman, M., Wirasinghe, R., White, B., & Lottridge, D. (2018). Understanding the Long-Term Use of Smart Speaker Assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 91:1–91:24. https://doi.org/10.1145/3264901

Cavill, S. (2020, February 13). *Consumers Are Reluctant To Use Smart Speakers For Purchases*. https://insights.digitalmediasolutions.com/paid-search/purchases-using-smart-speakers

Dahlqvist, F., Patel, M., Rajko, A., & Shulman, J. (2019, July 22). Growing Opportunities in the Internet of Things. McKinsey & Company. https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/growing-opportunities-in-the-internet-of-things#

Dubois, D. J., Kolcun, R., Mandalari, A. M., Paracha, M. T., Choffnes, D., & Haddadi, H. (2020). When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers. *Proceedings on Privacy Enhancing Technologies*, *2020*(4), 255–276. https://doi.org/10.2478/popets-2020-0072

Gao, C., Chandrasekaran, V., Fawaz, K., & Banerjee, S. (2018). Traversing the Quagmire that is Privacy in your Smart Home. *Proceedings of the 2018 Workshop on IoT Security and Privacy*, 22–28. https://doi.org/10.1145/3229565.3229573

Hern, A. (2019, July 26). Apple contractors "regularly hear confidential details" on Siri recordings. The Guardian. http://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings

Huang, Y., Obada-Obieh, B., & Beznosov, K. (Kosta). (2020). Amazon vs. My Brother: How Users of Shared Smart Speakers Perceive and Cope with Privacy Risks. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376529

Kinsella, B. (2019, June 18). *Loup Ventures Says 75% of U.S. Households Will Have Smart Speakers by 2025, Google to Surpass Amazon in Market Share*. Voicebot.Ai. https://voicebot.ai/2019/06/18/loup-ventures-says-75-of-u-s-households-will-have-smart-speakers-by-2025-google-to-surpass-amazon-in-market-share/

Kleinberg, S. (2018, January). OK, marketers: Here's what people are saying about their voice-activated speakers. Think with Google. https://www.thinkwithgoogle.com/marketing-strategies/app-and-mobile/voice-assistive-speaker-technology/

Lau, J., Zimmerman, B., & Schaub, F. (2018). Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 102:1–102:31. https://doi.org/10.1145/3274371

Schönherr, L., Golla, M., Eisenhofer, T., Wiele, J., Kolossa, D., & Holz, T. (2020). "Unacceptable, where is my privacy?": Exploring Accidental Triggers of Smart Speakers. ArXiv:2008.00508 [Cs]. http://arxiv.org/abs/2008.00508

Smart home privacy: What Amazon, Google and Apple do with your data. (2019, November 8). The Ambient. https://www.the-ambient.com/features/how-amazon-google-apple-use-smart-speaker-data-338

Transparency Market Research. (2018, September 19). *Smart Speaker Market Rising at Impressive 18% CAGR, Thanks to Wide Ranging Products on Sale—TMR*. https://www.prnewswire.com/news-releases/smart-speaker-market-rising-at-impressive-18-cagr-thanks-to-wide-ranging-products-on-sale-tmr-871259047.html

Wueest, C. (2017, November 20). *Everything You Need to Know About the Security of Voice-Activated Smart Speakers*. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/security-voice-activated-smart-speakers

Your Microphone Array Hides Your Identity: Robust Voice Liveness Detection for Smart Speakers. (2020, February 23). Unpublished manuscript. Network and Distributed Systems Security (NDSS) Symposium.