The Benefits of Standardization for Smart Home Device Privacy Implementations

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

James Tsai

Spring 2023

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

Advisor

M.C. Forelle, Department of Engineering and Society

Introduction

Privacy in the Internet of Things (IoT) age is a commodity wanted by many users, but truly understood by few. IoT is characterized by the usage of smart devices, which are electronic technology connected to other devices through a network; smartphones and computers are widely used examples of smart devices. Smart home devices such as Amazon Echo, Google Home, and Ring Alarm have taken the household appliance market by storm. As smart devices become more prevalent in lower and middle socioeconomic status households, the discussion regarding data collection and invasion of privacy also rises among the smart technology stakeholders. Many find it concerning that home appliances can learn its users' habits and daily routines from the data they collect and store passively. In a survey of smart device consumers, results showed "fairly low levels of trust in IoT, fairly low levels of satisfaction, low risk awareness, and older and less-educated respondents' resistance to IoT" (Cannizzaro et. al., 2020). Many users are unfamiliar with smart home device data collection processes, but the usage is still widespread, almost in 53.9% of American households by 2023 (Georgiev, 2023). Especially with the network of devices connected to the internet, the amount of data collected in a household, if leaked, poses a serious threat to user privacy. As these smart devices grow increasingly invasive, many users may start to question the extent of access and what measures are in place to prevent overreach. Since smart homes just recently grew in popularity, "there are no real consensus standards governing design, manufacture, or performance" of smart home devices, and as for laws and regulations, different agencies have varying views, which results in a lack of consistency and proper legal direction (Embry, 2017).

Smart home device privacy in the United States is negatively impacted by the lack of standards in security designs and privacy policy implementations. The research conducted in this

field covers the types of data being collected and the methods of doing so, how the potential of smart home device security is limited by social barriers, and improvements to device security and privacy policy implementations. I will analyze smart home device data collection transparency, privacy policies, and social barriers of smart device usage with Actor Network Theory (ANT). Through my analysis, I will find that more often than not, users are unaware of third-party access to their collected data. Social and economic barriers such as control, security, and cost concerns contribute to the lack of user trust in device data collection methods. I also find that some companies' privacy policies are more detailed and transparent than others, which brings to light the argument of device overreach concealed by vague privacy policies. From these findings, users are more likely to trust a device with a transparent privacy policy and minimal passive data collection or third party access. By improving standards and regulations for security, social barriers could be reduced and protocols improved, enabling smart home devices to reach their full potential.

Literature Review

Current research discovered that modern smart home devices collect data through two main methods: active and passive listening. Active listening is when the device reacts to an external command, and passive listening is when the device is constantly listening for an external command. Realizing that smart home devices are collecting data passively without voice commands can influence user behavior and trust regarding that device (Malkin et. al., 2019). Working with current active and passive ecological assessment methods, smart home devices can provide "observational multimodal data on individual and relational functioning, specifically within home environments" (Nelson & Allen, 2018). Essentially, smart home devices can record and convert private user actions to data and store them in their databases, although that data may

not be completely accurate. There are various metrics for measuring the accuracy of smart device data collection, and as more control is relinquished, the higher the accuracy (Zainab, Refaat, & Bouhali, 2020). Depending on how much the user agrees to allow the device to listen (strictly active or including passive) and store data, the device can tailor its features to the user. However, barriers exist that prevent complete trust between users and devices, impacting the scope of smart device usage.

Market research has found that the most significant barrier to the adoption of smart devices is upfront cost, followed by lack of awareness and privacy concerns. Other barriers include security, reliability, and interoperability of different technologies (Wilson et. al., 2017). Some barriers stem from the social impact that smart devices unintentionally have.

> "In the long term, though, the fact that society and technology co-evolve in a complex and non-linear manner leading to reshaping of homes and routines by technologies in radical ways (among others) there remains a question mark over whether smart homes can fulfill the promise of reducing energy use and whether other unintended consequences might prevail" (Balta-Ozkan et al., 2013).

Unintended consequences include: speech recognition could discriminate against users with accents; smart security cameras and doorbells could be used as surveillance for the authorities, which could make marginalized communities less safe due to systemic racism (Gilliard, 2019); device sharing data with third parties could result in data breaches due to weak security measures. Standardizing data collection methods and agreements could reduce the impact of those unintended consequences by ensuring users know what how their data is collected and the possible unintended consequences.

Smart home device security can be improved through a more effective security framework and data encryption methods. Current issues with smart device security include: most devices do not have built-in security or privacy controls to protect sensitive data transmissions, listening on smart devices are turned on by default, and online connections can compromise device security (Herold, 2020). Extensive research has been conducted in this field to improve user trust with their devices. A test was conducted to map out trust levels for users of Google Home Mini, Amazon Echo Dot, and Philips Hue Lights, and the levels are (in increasing trust): light and external fixtures, the smart home devices, smartphones, and the user. A trust control flow algorithm was developed from this trust model, which would allow for more trusted connectivity to devices (Ferraris, Bastos, Fernandez-Gago, & El-Moussa, 2021). Current smart home devices also lack user control over privacy settings, especially data collection methods. A machine learning framework can be used to determine for what purpose, to whom, and with what level of details smart home device information should be shared (Keshavarz & Anwar, 2018). The above algorithms and frameworks could be incorporated in a common protocol to improve data collection transparency and users' view of device security.

Another main reason why users distrust their devices is because device manufacturers and vendors have vague privacy notices that give them rights to control user data (Herold, 2020). Many smart device companies are focused on collecting high quality data, while circumnavigating the complexities of user privacy agreements. An analysis of Apple's Siri, Google Assistant, Microsoft's Cortana, Facebook Portal, Amazon's Alexa, and Samsung's Bixby's privacy policy transparency level showed that Samsung's Bixby had the lowest transparency and most privacy/security flaws and Apple's Siri had the highest transparency (Kelly, 2019). A few main differences between the two privacy policies is that Apple requires

account creation in order to collect data, whereas Samsung does not specify account creation in their privacy policy; Apple has multi-factor authentication capabilities, whereas Samsung does not specify; Apple employee access to user information is limited and data in transit is always encrypted and stored in an encrypted format. Samsung has very little details on data encryption. In terms of user access to data, Apple does much better: Apple specifies processes for users to delete or download their data, and user data is deleted when no longer necessary or if their account is terminated. Samsung's privacy policy is unclear in all those aspects. Standardization of privacy policies would likely give users a base level understanding of privacy implications with their device, potentially leading to improving user trust.

To analyze the benefits of implementing standards for smart devices, I will be using Bruno Latour's Actor Network Theory (ANT) framework. ANT focuses on how human and non-human actors aid in the construction of technological systems and how power and influence are generated from the relations between social and natural worlds (Latour, 2007). I will use this framework to understand how smart home devices, users, social barriers, and standards influence the design and social views of smart home technology. By understanding the connections, I can also find ways for researched solutions to be implemented better. Two central concepts of ANT are translation and black boxing. Translation refers to the changes needed for actors to align their interests and form a network together. Black boxing refers to the product of a series of translations resulting in a network that is opaque to outsiders due to technical complexities in the network. I will use these concepts to analyze smart home device privacy policies and the benefits/detriments of their implementations.

Methods

I gathered primary sources in the form of legal texts for information about smart home device data collection and device user agreements/privacy policies. I also gathered secondary sources in the form of academic journals and articles studying the public opinion on smart home data collection and privacy policy improvement strategies. I analyzed privacy policies to find what data is being collected and stored, and how it affects users' perception of the standards/privacy framework of the device. I examined other primary sources to compare/contrast the benefits and detriments to both users and companies, and how the user agreements/privacy policies reflect that. I utilized secondary sources to analyze the network surrounding a smart home device privacy policy, and how device security and privacy policy can be improved to benefit all parties involved. A majority of the data gathered is in the form of official documentation and analysis of official documentation since there are limited academic sources that researched current standards of IoT devices. The secondary sources for security and privacy policy implementation are mainly in the form of proposed protocols and frameworks, since incompatibility between smart devices makes it difficult to analyze current IoT protocols/frameworks (Phan & Kim, 2020). Through these methods, I am able to analyze the actors in this privacy network through my own perspective and understanding of the field.

Analysis

Smart home device data collection methods lack transparency and contain security risks, affecting users' perception and trust level of the device. I use ANT to identify actors in the smart device network and analyze their relationships and contribution to transparency and security risks. Key actors in the smart home network are not limited to just the devices and the user; data storage systems, social barriers, and security considerations also influence the design and social views of smart home technology. These actors set the foundation for the balance between

privacy, security, and functionality of a smart home device (Ferraris et al., 2021). Device standards are a major actor in the network of all smart devices. Such standards are influenced through both technical and social factors including centralized data storages, acceptable end user agreements, and transparent privacy policies (Daivee, n.d.). These factors delegate roles to the standards, which configure both the technology and the users' lifestyles. The IETF has established requirements for secure IoT devices: end device security, local communication security, gateway data security, internet security, cloud data security, and application security (Sadique et. al., 2018). Out of these requirements, only internet and cloud data security have specific data protection standards like the Open Connectivity Foundation's (OCF) Device-to-Cloud Secure Data Access (D2CSDA) specification and Internet Engineering Task Force (IETF)'s Constrained Application Protocol (CoAP) (Open Connectivity Foundation, 2022). These standards exist to regulate how the internet interacts with devices and cloud storage (Shelby et. al., 2014). The lack of standardization in other security fields means different devices have incompatible security implementations and could result in a data breach in the network through the device with the weakest security. However, 48% of IoT device users are unaware of their devices' security risks (Canonical, 2017). This is because a device's trustworthiness is largely based on its privacy policy, as that is the only means regular users have to understand what data is being collected and where it is being shared and stored. The network of users, privacy policies, and device security is constantly evolving with software updates and data leaks, which can impact users' perceptions of how trustworthy their devices are. Implementing policy and security standards could reduce the role of privacy policies in that perception and improve trust level.

To tackle the lack-of-trust problem, users should understand why functional and economic barriers restrict the potential of devices and how standardization could improve stability in the actor network. Barriers include control, security, and tradeoff concerns, which prevent smart home devices from collecting more data. I analyze these barriers through their relationships with different actors in the network. Although relinquishing more control to smart home devices improves their performance, the costs of maintaining a more connected system mostly outweigh the benefits, for now (Zainab, Refaat, & Bouhali, 2020). The additional computing power and data storage needed will increase the price of those devices, which can impact security. Manufacturers may see production cost and ease of use as a tradeoff for security since security is often overlooked by users. Manufacturer actions, like trading lower production cost and ease of use for security, interact with the security protocols to create openings for attackers.

User actions interact with security protocols to also create potential security risks. Actions such as sharing passwords, not updating software, or failing to configure device security settings properly can create vulnerabilities. Standardizing smart device security software could reduce the amount of user-created security risks by providing a common interface for users to configure security settings properly (Moustafa et. al., 2021). Users, manufacturers, and smart device protocols all have a role in shaping the security aspect of the actor network by preventing introduction of security risks. A good foundation for the network can be created through device manufacturers implementing “more security capabilities and applications with secure and easy-to-configure user interfaces”, and governmental authorities’ contribution of “legal support, security standards, and law enforcement policies” (Ali & Awad, 2018). Manufacturers and governmental support can develop standards to mitigate security risks introduced by actors’

actions. The standardization can ease the tension created by barriers, but until such measures are established and recognized by users, data privacy will likely remain privatized by smart device companies.

The lack of federal standardization for privacy policies and data protection measures also contributes to smart device companies establishing vague privacy policies and compromising device security in the United States. The European Union has a federal General Data Protection Regulation (GDPR), which provides safeguards and education on the data collected through any technological means (Piasecki & Chen, 2021). However, the United States and most other countries have no such federal protection, allowing smart devices like speakers and TVs to collect sensitive information and send that data to a third party. Main legislation includes the Consumer Online Privacy Rights Act (COPRA), which essentially gives users access and control of their personal data upon request (S.3195, 2021). Although it can be argued that the COPRA establishes requirements for user privacy and security, there is very little enforcement and the requirements are very general, which enables manufacturers to shift their investments from security implementations to ease of access and new features. Manufacturers take advantage of the lack of regulation by black boxing their user agreements and incorporating them in the network for their users. Privacy policies and user agreements are the device security protocols translated in user-readable word form for user accessibility. However, 36% of Americans do not read privacy policies that are presented to them (Atske, 2019), so manufacturers can capitalize by translating security risks into vague wording in the privacy policy. Standardizing privacy policies to maintain a level of transparency would hold manufacturers accountable for their security practices. In-depth explanations of security protocols highlighted in the policy makes a big difference for the few Americans who read the policy and emphasizes the role of consumer

advocacy groups in the network. Making risks known to all users enables advocacy groups to enact changes in product design and security, protecting users and improving manufacturer protocols.

Conclusion

Ultimately, users will not likely use a device they do not trust. Standards and regulations are one method of improving transparency to establish trust between users and devices. Whether it be standardizing clauses regarding third party access to data or security protocols, commonality between different companies and devices will benefit all parties in the long run through strengthened security and high user trust.

By bringing the issues of smart device data collection and privacy policies to light, device manufacturers and policymakers can find ways to improve user experience. There are many directions to go in terms of future research: implementing standards for privacy policies to guarantee a certain level of transparency in data collection; improved device security frameworks can also help create new standards for device security; socioeconomic analysis of various smart devices can distinguish what social groups use which device and how much they trust that device. Future research should build off the privacy policy analysis referenced in the literature review and have all smart device policies follow general format and specification rules.

It is important to view smart home devices not as a potential security threat, but a technology that improves control and convenience within society. As IoT devices are seeing rampant growth in homes throughout the world, human existence gradually becomes more defined by the data we produce. Therefore, it is the responsibility of all involved - users, manufacturers, and governments - to protect our data as technology continuously evolves.

References

Ali, B., & Awad, A. (2018). Cyber and physical security vulnerability assessment for iot-based smart homes. *Sensors*, *18*(3), 817. https://doi.org/10.3390/s18030817

Atske, S. (2019, November 15). 4. Americans' attitudes and experiences with privacy policies and laws. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2019/11/15/americans-attitudes-and-experiences-with-privacy-policies-and-laws/

Cannizzaro, S., Procter, R., Ma, S., & Maple, C. (2020). Trust in the smart home: Findings from a nationally representative survey in the UK. *PLOS ONE*, *15*(5), e0231615. https://doi.org/10.1371/journal.pone.0231615

Canonical. (2017). *Taking charge of the IoT's security vulnerabilities*. Canonical. https://pages.ubuntu.com/IoTSecurityWhitepaper-Fullreport.html

Daivee. (n.d.). IEEE device app end user agreements. IEEE Brand Experience. Retrieved October 30, 2022, from https://brand-experience.ieee.org/guidelines/digital/mobileapp-andresponsive-design-guidelines/mobile-apps/end-user-agreements/

Embry, S., & Love, G. (2017). Smart home technology. National Association of Home Builders. https://www.nahb.org/advocacy/legal-issues/smart-home-technology

Ferraris, D., Bastos, D., Fernandez-Gago, C., & El-Moussa, F. (2021). A trust model for popular smart home devices. *International Journal of Information Security*, 20(4), 571–587. https://doi.org/10.1007/s10207-020-00519-2

Georgiev, D. (2023). *20 eye-opening smart home statistics to know in 2023*. Techjury; Techjury.net. https://techjury.net/blog/smart-home-statistics/

Gilliard, C. (2019). *How smart home tech could perpetuate discrimination and racial profiling | CBC Radio*. CBC. https://www.cbc.ca/radio/spark/how-smart-home-tech-could-perpetuate-discrimination-and-racial-profiling-1.5324608

Herold, R. (2020). *Five common privacy problems in an era of smart devices*. ISACA; ISACA. https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2020/five-common-privacy-problems-in-an-era-of-smart-devices

Kelly, G. (2019). *Learn to navigate the privacy challenges of popular smart speakers*. Common Sense Media. https://www.commonsense.org/education/articles/compare-the-privacy-practices-of-the-most-popular-smart-speakers-with-virtual-assistants

Keshavarz, M., & Anwar, M. (2018). Towards improving privacy control for smart homes: A privacy decision framework. 2018 16th Annual Conference on Privacy, Security and Trust (PST), 1–3. https://doi.org/10.1109/PST.2018.8514198

Legal—Privacy policy—Apple. (n.d.). Apple Legal. Retrieved March 5, 2023, from https://www.apple.com/legal/privacy/

Malkin, N., Egelman, S., & Wagner, D. (2019). Privacy controls for always-listening devices. *Proceedings of the New Security Paradigms Workshop*, 78–91. https://doi.org/10.1145/3368860.3368867

Moustafa, A. A., Bello, A., & Maurushat, A. (2021). The role of user behaviour in improving cyber security management. *Frontiers in Psychology*, *12*, 561011. https://doi.org/10.3389/fpsyg.2021.561011

Nelson, B. W., & Allen, N. B. (2018). Extending the passive-sensing toolbox: Using smart-home technology in psychological science. *Perspectives on Psychological Science*, *13*(6), 718–733. https://doi.org/10.1177/1745691618776008

Open Connectivity Foundation. (2022). *OCF Device to Cloud Services Specification*. Open Connectivity Foundation, inc. https://openconnectivity.org/specs/OCF_Device_To_Cloud_Services_Specification.pdf

Piasecki, S., & Chen, J. (2022). Complying with the GDPR when vulnerable people use smart devices. *International Data Privacy Law*, *12*(2), 113–131. https://doi.org/10.1093/idpl/ipac001

Phan, L.-A., & Kim, T. (2020). Breaking down the compatibility problem in smart homes: A dynamically updatable gateway platform. *Sensors*, *20*(10), 2783. https://doi.org/10.3390/s20102783

Privacy policy—Samsung. (n.d.). Samsung Electronics America. Retrieved March 5, 2023, from https://www.samsung.com/us/account/privacy-policy/

S.3195 – 117th Congress (2021-2022): Consumer Online Privacy Rights Act. (2021, November 4). https://www.congress.gov/bill/117th-congress/senate-bill/3195

Sadique, K. M., Rahmani, R., & Johannesson, P. (2018). Towards security on internet of things: Applications and challenges in technology. *Procedia Computer Science*, *141*, 199–206. https://doi.org/10.1016/j.procs.2018.10.168

Shelby, Z., Hartke, K., & Bormann, C. (2014). *The Constrained Application Protocol (CoAP)*. Internet Engineering Task Force. https://www.rfc-editor.org/rfc/rfc7252

Wilson, C., Hargreaves, T., & Hauxwell-Baldwin, R. (2017). Benefits and risks of smart home technologies. *Energy Policy*, *103*, 72–83. https://doi.org/10.1016/j.enpol.2016.12.047

Zainab, A., S. Refaat, S., & Bouhali, O. (2020). Ensemble-based spam detection in smart home iot devices time series data using machine learning techniques. *Information*, 11(7), 344. https://doi.org/10.3390/info11070344