To appear in: *Human Factors in Cybersecurity,* Vol. 53, 2022, 1–9 doi: 10.54941/xxxxxx

Chapter XX

Privacy Concerns about Smart Home Devices: A Comparative Analysis between Non-Users and Users

Chola Chhetri, Vivian Genaro Motti

[1] *George Mason University*

Fairfax, VA 22030, USA

ABSTRACT

Privacy concerns of smart home device (SHD) users have been largely explored but those of non-users are under-explored. Understanding of non-user concerns is essential to inform the design of user-centric privacy-preserving SHDs and facilitate acceptance. To address this gap, we conducted a survey of SHD non-users and analyzed their privacy concerns. We followed a mixed-methods approach to analyze and compare privacy concerns, explore non-use reasons, and provide design suggestions. We make recommendations based on our study findings. This paper contributes to improve privacy controls of SHD considering non-user privacy concerns.

Keywords: Internet of Things, Smart Home, Smart Home Devices, Privacy, Home

IoT, User, Non-User, User Studies, Survey

INTRODUCTION

The growth of Internet of Things (IoT) devices worldwide has led to an increase in home devices that are connected to the Internet, either directly or through a centralized device, such as a hub or controller (Bugeja, Davidsson and Jacobsson, 2018). In this paper, we refer to the home IoT devices as smart home devices (SHDs).

Privacy concerns of SHD users have been widely studied (Page *et al.*, 2018; Chhetri, 2019; Chhetri and Motti, 2019; Yao *et al.*, 2019) as well as privacy risks in SHDs (Chhetri and Motti, 2020). However, SHD non-user concerns have been understudied compared to those of users (Liao *et al.*, 2019; Yao *et al.*, 2019). Drawing from the Stakeholder theory (Freeman, 2010), non-users are as important stakeholders of SHDs as users because by exploring non-user concerns, academia, industry and policymakers are able to address open concerns. Improvements can make SHDs safer, facilitate SHD acceptance and adoption (Wyatt, 2003), and bridge digital divide (Baumer *et al.*, 2015) by turning non-users into users (Oostveen, 2014). An in-depth exploration of non-user privacy concerns is lacking, to the best of our knowledge.

To fill this gap, we conducted an online survey to analyze the privacy concerns of SHD non-users (n=41), and we explored non-use reasons and participants' suggestions to improve SHD. Lastly, we compared non-user privacy concerns with those of users (n=50). Our study makes two contributions: (a) We provide in-depth understanding of privacy concerns of SHD non-users, and (b) Our results provide novel insight into how SHD users' privacy concerns differ from those of non-users so that SHD designers can address such concerns in future designs.

RELATED WORK

In the SHD domain, non-user concerns are understudied so far. This occurs likely because they are difficult to locate and recruit for studies and experiments, and are not as coherently grouped as users (Wyatt, 2003). We highlight prior work that have included non-users among the participants. Liao et al. (2019) found that privacy impacted the decision to adopt intelligent personal assistants. Lau et al. (2018) interviewed 17 non-users of smart speakers and found that the deterrent factors to adoption included privacy. Yao et al. (2019) included three SHD non-users to explore perceived benefits and risks. Yao et al. (2019) included seven SHD non-users in another co-design study of 25 participants.

SURVEY METHOD

All survey-related documents were approved by our institution's Institutional Review Board. We announced the study through twitter, newsletters and email contacts. We received 95 responses, excluded 4 incomplete ones, and analyzed 91 responses.

Participants

Nearly 45% (n=41) reported not using any SHD and 55% (n=50) reported using at least one SHD. In this paper, we refer to the former group of respondents as 'non-users' and the latter as 'users'. About 43% were female and 52% were male. Nearly 43% were 18-29 years, 26% were 30-39, 14% were 40-49, 9% were 50-59, and 10% were 60 years or older. About 43% reported to be White, 26% Asian or Pacific Islander, 10% Black or African American, 9% Hispanic or Latino, and 12% other. About 36% had a Post-graduate (MS, PhD, etc.) degree, 24% Bachelor's, 15% Associate's, 19% High School, and 3% Other. Among the participants, 57% had an educational background related to technology (computer science or information technology) and nearly 43% had a non-technology background.

Questionnaire Design

The participants were first presented with an informed consent form and an option to continue or exit the survey. If they continued, they were presented with a basic definition of the term 'smart home' to ensure a common understanding among participants. The survey was titled "Smart Home Survey" to avoid bias recruiting participants concerned about privacy. To prevent invoking privacy-related opinions, the word 'privacy' was not used in the questionnaire until the middle of the questionnaire where specific privacy concerns were asked.

We utilized skip logic and presented a subset of different questions depending on whether the participants answered 'yes' or 'no' to first question, "Do you use a smart home device?". If a participant responded 'yes' (user), we asked them about number and type of SHDs used, use reasons, concerns, suggestions and demographic information. If a participant responded 'no' (non-user), we asked them about non-use reasons, concerns, suggestions and demographic information.

We avoided binary questions to prevent introducing acquiescence bias (Baxter, Courage and Caine, 2015). To ensure validity, questions were inspired or adapted from related work (Acquisti and Grossklags, 2005; Williams, Nurse and Creese, 2017). The questionnaire design followed multiple iterations for evaluation and refinement. Questions were tested and revised multiple times for clarity. We pilot tested the survey with three participants and used their feedback to revise the questions. Participants did not receive any monetary compensation.

Data Analysis

For open-ended responses (non-use reasons, privacy concerns, and suggestions), two researchers conducted thematic analysis using open coding with selective (or axial) coding (Corbin and Strauss, 1990). We familiarized ourselves with the data by reading the responses and coded each sentence independently by observing the concept in it. We then generated categories by observing the similarity of concept in the codes. We achieved an inter-rater reliability of 0.89 using Cohen's kappa, which is considered 'almost perfect' (Landis and Koch, 1977). For codes that were different between coders, we discussed and agreed on revised codes. Finally, we agreed on all codes and ensured the codes were correctly assigned to categories, leading to a final codebook.

We then conducted descriptive and inferential statistical analyses to summarize non-use reasons, suggestions and privacy concerns. Lastly, we also performed quantitative analysis to compare the privacy concerns of non-users and users.

RESULTS

This section reports on the SHD distribution of the user group, reasons for non-use of SHDs, privacy concerns, and suggestions made by participants.

SHD Distribution in User Participants

User respondents reported use of 87 SHDs in total. The SHD data includes 11 types of devices (such as cameras and locks) and 20 brands of SHDs (such as Amazon Echo and Google Home). About 66% (n=33) of users reported using a single SHD and 34% of them (n=17) reported using multiple (2 to 8) devices. Table 1 lists the device type, brand or model, and the respective number of SHD users. The majority of participants used intelligent speakers (Amazon Echo or Google Home). This distribution is representative of current United States SHD market (Olmstead and Smith, 2017).

Table 1: SHDs used by participants (user group). UNK (Unknown) indicates no brand was reported. Door lock, television, and security system had a frequency of 1 and are not shown in table for brevity. Numbers in parentheses in the third column represent the frequency of that device. Frequency of devices is greater than number of user participants as some participants used more than one device.

Smart Device	Total (n=87)	Frequency Breakdown by Brand/Model
Speaker	38	Amazon Echo (25), Google Home (13)
Hub	15	Samsung SmartThings (10), Vera (4), Nexia (1)
Thermostat	11	Nest (7), Radio (1), Emerson (1), Carrier (1), UNK (1)
Camera	6	Arlo (2), Canary (2), Kuna (1), Blink (1)

Light	5	Philips Hue (3), Halo (1), UNK (1)
Door bell	4	Ring (4)
Plug	3	Kasa (1), Wemo (1), UNK (1)
Vacuum	2	Roomba 960 (1), Eufy (1)

Reasons for SHD Non-Use

Most non-users reported their non-use reason to be privacy concerns (68%). Other non-use reasons included lack of interest in SHDs (32%), cost (22%), lack of perceived usefulness (12%), insecurity or potential of hacking (10%), and perceived difficulty of usage (7%).

Privacy Concerns: Non-Users vs. Users

The thematic analysis resulted in 17 codes and three thematic areas of privacy concerns. Table 2 shows the themes and breaks down the percentage of non-users, users, and total participants for each code.

The first theme was 'data collection concerns' which included five codes: recording audio/video, tracking occupancy, listening to private conversations, monitoring usage/behavior, and identity theft. About 34% of the codes fell under this theme. This category included participants' concerns regarding the initial temporal data collection feature where the SHD (e.g. smart speaker) collects data by constantly listening to the user and recording audio, video or both. Most of these concerns were about recording audio or video. Participants were also concerned about SHDs listening to private conversations and purposefully or accidentally recording them. They were also concerned about the consequences of data collection where the SHD allows tracking (e.g. occupancy) and monitoring the usage behavior or patterns.

The second theme was 'data sharing concerns' which included 22% of the privacy concerns under four codes: selling data, third party data access, leakage without consent, and marketing data. Participants raised concerns about the selling of SHD data to business partners or data brokers, third party (e.g. government) access to SHD data, leakage of the data, and the potential of the data being used in marketing.

The third theme was 'data protection concerns' which included eight codes: hacking, data handling, protecting data, secondary use, aggregation, data abuse, data loss, and fraud. About 32% of the codes fell in this category. Participants were mostly concerned about SHD devices being hacked and data being improperly handled by the SHD company. For example, one participant (P1) noted: *"My concerns are [...] being hacked and used against me. Such as, someone could hack smart locks unlock my doors, or with smart thermometers if it was hacked one could tell if I was home."*

A chi-square test between non-users and users showed that the privacy concerns of non-users differed significantly from users (χ^2=8.46, $p<0.05$). Non-users reported a higher frequency of concerns in data collection and data protection themes than those of users (46% vs 24% and 34% vs 30% respectively). This is likely because

privacy was a major reason for non-use and collection of data is the first stage in the SHD data life cycle that raises privacy concerns. However, non-users reported fewer concerns in the data sharing theme than those of users (15% vs 28% respectively). This is likely because user participants are already having their data collected and used by SHDs, which naturally raises concerns about how those data will be shared for marketing and other purposes.

Table 2: Breakdown of privacy concerns categories (bold) and codes by non-users (NU, n=41) and users (U, n=50). Non-users were more concerned about data collection than users (NU>U). Users were more concerned about data sharing.

Privacy Concerns	% of Non-Users	% of Users	% of All
Data Collection Concerns			
Recording audio/video	24.39	20.00	21.98
Tracking occupancy	9.76	2.00	5.49
Listening to private conversations	4.88	2.00	3.30
Monitoring usage/behavior	4.88	0.00	2.20
Identity theft	2.44	0.00	1.10
Data Sharing Concerns			
Selling data	4.88	10.00	7.69
Third party data access	7.32	8.00	7.69
Leakage without consent	2.44	6.00	4.40
Marketing data	0.00	4.00	2.20
Data Protection Concerns			
Hacking potential	21.95	12.00	16.49
Data handling	7.32	4.00	5.49
Protecting data	2.44	4.00	3.30
Secondary use	0.00	2.00	2.20
Aggregation	0.00	2.00	1.10
Data abuse	0.00	2.00	1.10
Data loss	0.00	2.00	1.10
Fraud likelihood	0.00	2.00	1.10

Participant Suggestions to Improve SHDs

The thematic analysis of participants' suggestions for developers resulted in four main themes: (a) data anonymization and minimization, (b) data protection and security, (c) transparent data use policies, and (d) user-centric practices.

Data Anonymization and Minimization. This theme included twenty suggestions that SHD devices collect only anonymized data, not store any personal information, and not track SHD users. Five participants suggested that anonymization should be 'guaranteed' before data is sent to the cloud. Three participants suggested that vendors provide choice to opt out of unsupervised monitoring involving SHD data collection.

Data Protection and Security. Fifteen participants suggested that manufacturers build devices so that they are less likely to be hacked. Specifically, participants suggested developers to conduct vulnerability testing, research cybersecurity

incidents, and improve security on SHDs. Participants also suggested building SHDs with more computing power so that security technologies can be implemented. In the words of one non-user, "*Put security first*". Six participants suggested developers to use encryption to protect the data collected. Three participants suggested third-party verification of security features to raise users confidence and trust. Participants also suggested implementing authentication, such as passwords, in SHDs.

Transparent Data Use Policies. Transparency was the main theme of suggestions from twelve participants. Eight participants sought "more transparency and control over their data", and clarity on "how providers use collected data, how they store, and share the data." Twelve participants suggested that manufacturers inform the users clearly and succinctly (using not only text, but also images and videos) on "data usage, data collection, handling of breaches and technical issues, and research data." Several participants also suggested that manufacturers educate consumers about the potential privacy risks in their devices and the ways to mitigate risks.

User-centric Practices. Six responses contained users as the primary theme. Among these, four participants suggested that developers put users' interests and concerns first, not their own. For example, one non-user wrote: "*If they [vendors] can do anything to do improve the risks on the devices, they should do it—not put their own interests first and use it to their own advantage."* Another non-user emphasized the need to put consumers first: *"Develop with consumer in mind."* Participants sought features to control privacy and desired visible indicators of recording or data collection. For example, one participant wrote:

"*Give users options. Add features providing users the ability to control their privacy. Make it visible when recording, if the device is recording. Make their information accessible."* (P56).

DISCUSSION AND LIMITATIONS

We found that privacy was a major reason for non-use of SHDs and that SHD non-users had privacy concerns, which is confirmatory to Liao et al (2019) and Lau et al (2018) and extension to SHD domain. Based on the unique empirical analysis of user and non-user privacy concerns, it is evident that non-user privacy concerns are as important as those of users. We argue that addressing non-user privacy concerns will help reduce tension between users and non-users, especially in shared spaces, such as apartments. Another novel finding was that data collection and protection concerned non-users the most and data sharing concerned users the most. Our explanation is that users trade off privacy concerns with usage motivations. We suggest that SHD manufacturers address the data collection, protection and sharing concerns, so that users can continue to reap the benefits of SHDs with increased confidence, and non-users can consider using SHDs to reap the benefits of SHDs Designers and developers should take the concerns into consideration in the design, development and enhancement of SHDs to gain consumer trust and increase product adoption.

Our findings are skewed towards more educated users, but reflect the current SHD user population that largely includes educated and tech-savvy early adopters (Lau, Zimmerman and Schaub, 2018). Secondly, survey instruments have a tendency to measure attitudes rather than actual behavior. So, the findings reflect reported attitudes and perceptions. Despite these limitations, we believe our study provides valuable insights into SHD privacy concerns, and breaks ground into understanding reasons for SHD non-use and privacy concerns of non-users.

CONCLUSIONS

In this paper, we reported our work aimed at understanding non-user privacy concerns regarding SHD devices along with those of users. We analyzed and compared the privacy concerns and identified non-use reasons for SHDs. Our findings indicate that both users and non-users of SHDs are concerned about privacy violations caused by SHDs. Non-users of SHDs are concerned about data collection and its protection, and users are concerned about how their data might be shared and (ab)used by companies. While users reap the benefits of SHDs through usage, their usage is not concern-free. Privacy-concerned users are trading off privacy with other benefits. Enhanced data practices, data protection, and transparency from SHD manufacturers and application providers can lead to more confident users and attract non-users towards usage.

ACKNOWLEDGMENTS

This research was funded in part by 4-VA, a collaborative partnership for advancing the Commonwealth of Virginia, and Commonwealth Cyber Initiative (CCI). We thank our participants for their responses. We are grateful to anonymous referees for their comments in this article.

REFERENCES

Acquisti, A. and Grossklags, J. (2005) 'Privacy and rationality in individual decision making', *IEEE Security & Privacy*, 3(1), pp. 26–33. doi: 10.1109/MSP.2005.22.

Baumer, E. P. S., Burrell, J., Ames, M. G., Brubaker, J. R. and Dourish, P. (2015) 'On the Importance and Implications of Studying Technology Non-Use', *Interactions*. New York, NY, USA: Association for Computing Machinery, 22(2), pp. 52–56. doi: 10.1145/2723667.

Baxter, K., Courage, C. and Caine, K. (2015) *Understanding Your Users: A Practical Guide to User Research Methods*. 2nd edn. Morgan Kaufmann.

Bugeja, J., Davidsson, P. and Jacobsson, A. (2018) 'Functional Classification and Quantitative Analysis of Smart Connected Home Devices', in *2018 Global Internet of Things Summit (GIoTS)*, pp. 1–6. doi: 10.1109/GIOTS.2018.8534563.

Chhetri, C. (2019) 'Towards a Smart Home Usable Privacy Framework', in *Conference Companion*

Publication of the 2019 on Computer Supported Cooperative Work and Social Computing. New York, NY, USA: Association for Computing Machinery (CSCW '19), pp. 43–46. doi: 10.1145/3311957.3361849.

Chhetri, C. and Motti, V. (2020) 'Identifying Vulnerabilities in Security and Privacy of Smart Home Devices', in Choo, K.-K. R., Morris, T., Peterson, G. L., and Imsand, E. (eds) *National Cyber Summit {(NCS)} Research Track 2020, Huntsville, AL, USA, June 2-4, 2020*. Springer (Advances in Intelligent Systems and Computing), pp. 211–231. doi: 10.1007/978-3-030-58703-1_13.

Chhetri, C. and Motti, V. G. (2019) 'Eliciting Privacy Concerns for Smart Home Devices from a User Centered Perspective', in Taylor, N. G., Christian-Lamb, C., Martin, M. H., and Nardi, B. (eds) *Information in Contemporary Society*. Cham: Springer International Publishing, pp. 91–101. doi: https://doi.org/10.1007/978-3-030-15742-5_8.

Corbin, J. M. and Strauss, A. (1990) 'Grounded theory research: Procedures, canons, and evaluative criteria', *Qualitative Sociology*, 13(1), pp. 3–21. doi: 10.1007/BF00988593.

Freeman, R. E. (2010) *Strategic management: A stakeholder approach*. Cambridge university press.

Landis, J. R. and Koch, G. G. (1977) 'The Measurement of Observer Agreement for Categorical Data', *Biometrics*. [Wiley, International Biometric Society], 33(1), pp. 159–174. Available at: http://www.jstor.org/stable/2529310.

Lau, J., Zimmerman, B. and Schaub, F. (2018) 'Alexa, Are You Listening? Privacy Perceptions , Concerns and Privacy-seeking Behaviors with Smart Speakers', in *Proceedings of ACM Human-Computer Interaction*, pp. 102:1--31. doi: 10.1144/GSL.JGS.1939.065.01-04.12.

Liao, Y., Vitak, J., Kumar, P., Zimmer, M. and Kritikos, K. (2019) 'Understanding the Role of Privacy and Trust in Intelligent Personal Assistant Adoption', in *Proceedings of the 13th Annual iConference, Lecture Notes in Computer Science (iConference)*, pp. 102–113. doi: 10.1007/978-3-030-15742-5_9.

Olmstead, K. and Smith, A. (2017) *Americans and Cybersecurity*, *Pew Research Center*. Washington, DC. Available at: http://www.pewinternet.org/2017/01/26/americans-and-cybersecurity/.

Oostveen, A.-M. (2014) 'Non-Use of Automated Border Control Systems: Identifying Reasons and Solutions', in *Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014 - Sand, Sea and Sky - Holiday HCI*. Swindon, GBR: BCS (BCS-HCI '14), pp. 228–233. doi: 10.14236/ewic/hci2014.28.

Page, X., Bahirat, P., Safi, M. I., Knijnenburg, B. P. and Wisniewski, P. (2018) 'The Internet of What?: Understanding Differences in Perceptions and Adoption for the Internet of Things', *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* New York, NY, USA: ACM, 2(4), pp. 183:1--183:22. doi: 10.1145/3287061.

Williams, M., Nurse, J. R. C. and Creese, S. (2017) '"Privacy is the Boring Bit": User Perceptions and Behaviour in the Internet-of-Things', in *15th International Conference on Privacy, Security and Trust (PST)*. IEEE. Available at: http://www.cs.ox.ac.uk/files/9213/2017-pst-wnc-preprint.pdf.

Wyatt, S. (2003) 'Non-users also matter: The construction of users and non-users of the Internet', in *Now Users Matter: The Co-construction of Users and Technology*, pp. 67–79.

Yao, Y., Basdeo, J. R., Kaushik, S. and Wang, Y. (2019) 'Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes', in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '19), pp. 198:1---198:12. doi: 10.1145/3290605.3300428.