Mahmoud Aghvamipanah
Email: m.aghvami@sharif.edu

Research Statement

Introduction

Smart devices collect data in a smart home environment, including presence state, temperature, humidity, electricity, and water consumption. This data is forwarded to the IoT platform's cloud to process and execute automation rules making life easier. With the expansion of IoT platform usage in smart homes, the collection and transmission of users' data through smart devices have also been expanded. While this data transfer helps the user to execute automation rules, it can also violate user privacy as it can be used for recognizing activities of daily living(ADLs) [1] and targeted advertisements [2] without the user's awareness or consent.

Privacy is a primary concern for IoT users, and studies[3] show that smart home users have different concerns about sharing data with advertisers and government entities. However, some commercial IoT platforms employ aggressive privacy policies. As an example, in Samsung SmartThings term of use 2017 [4], the below sentences shows that the platform clearly announces your data can be distributed through any media without any limitation*: "Subject to our privacy policy, you hereby do and shall grant SmartThings a worldwide, non-exclusive, perpetual, irrevocable, royalty-free, fully paid, sublicensable, and transferable license to use, modify, reproduce, distribute, share, prepare derivative works of, display, perform, and otherwise fully exploit the user submissions and device data in connection with the Services and SmartThings's business, including without limitation for promoting and redistributing part or all of the services in any media formats and through any media channels (including, without limitation, third-party websites and services)."*

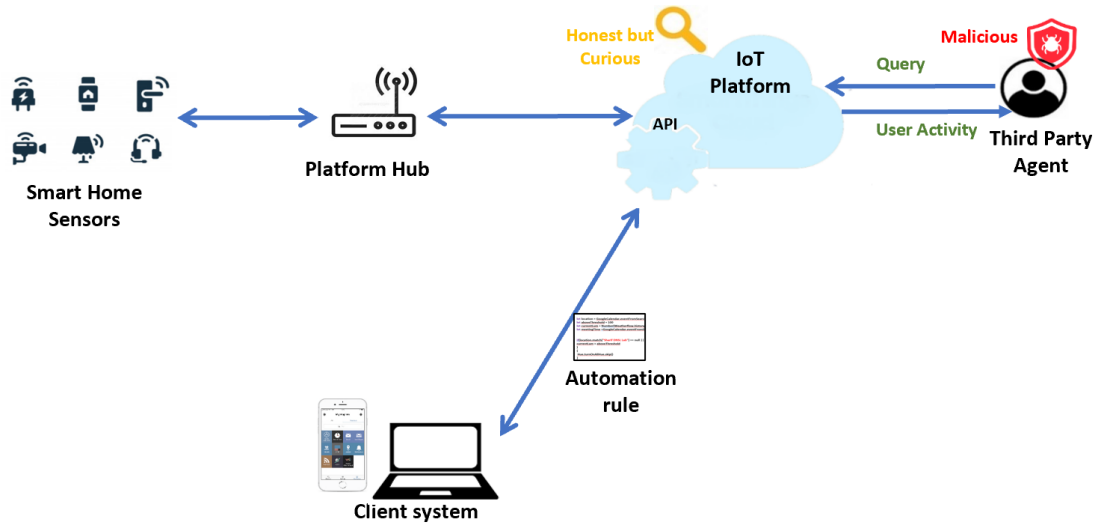

Figure 1-Our research Threat model

The section was toned down in recent versions of SmartThings term of use. Although in SmartThings Privacy Notice 2021 [5], the platform still states: "we may collect personal information about your online activities on connected devices over time and across third-party devices, apps, and other online features and services." The platform also informed using personal information for "data analysis" to provide "customized recommendations, promotions, marketing information and advertisement."

Human activity recognition (HAR) techniques are used in various contexts, such as smart healthcare systems to monitor patients and Ambient Assisted Living systems (AAL) to help older adults have an independent life [6]. Also, in typical commercial smart home platforms, user activity can be inferred using aggregated sensor data without the user's awareness or consent, which can be used for targeted advertisement. For example, suppose an Amazon agent as a third party has access to execute a query against user smart home data (without accessing the original data itself). In this case, He can infer the time profile of user activities, including eating, sleeping, and leaving home for a targeted advertisement scenario, violating the user's privacy.

As Figure1 shows, we consider the IoT platform an honest but curious party in our threat model. We also consider a malicious third-party agent who wants to infer the user's activity for targeted advertisement. Finally, we assume that the user trusts his client's devices, sensor devices, and platform hub.

Related work

Recent research have explored various aspects of the malicious/semi-honest IoT platform threat model. Some studies have investigated the integrity of automation rules in the threat model. DTAP [7] shows malicious rule makers can exploit overprivileged access to trigger/action APIs for violating the rule integrity and propose a decentralized trigger-

Figure 2- Malicious/semi-honest IoT platform threat model researches

action platform as a solution. Also, eTAP [8] solution grantee rule integrity using garbled circuit.

Data privacy is another important goal in the mentioned threat model. We found three research lines for preserving data privacy: data minimization, data encryption, and sensitive inference sanitization.

The data minimization approaches are based on *least privilege* and *need-to-know* principles. F&F[9] minimizes data transfer on a threat model with a SmartThings honest platform connecting to a malicious IFTTT platform. In a more general attempt, minTAP [10] leverages language-based data minimization to release only necessary user data attributes to the platform. Pfirewall [11] filters user data based on two types of policies, automation-dependent data-minimization policies(APs) and user-specified policies (UPs).

The data encryption approaches are based on the assumption that the confidentiality of each data record is important for the users. OTAP and ATAP [12] leverage encryption of user data. However, their solution restricts users to only have automation rules without filter code or computation, which is a limiting assumption. Also, eTAP executes automation rules without accessing users' private data in plaintext using a garbled circuit protocol. In another research line, Walnut [15] and My house, My rules [14] use hardware-based trusted execution environments (TEEs) or hardware security modules (HSMs) for preserving data confidentiality on a malicious platform.

Research Question

Although proposed solutions in data minimization, data encryption and hardware-based security approaches are promising, using them practically needs significant changes on the IoT platform or the connecting trigger/action services, which are not easy to apply in the real world. Therefore as an assumption, we can restrict user privacy concerns to the sensitive inference from aggregated data in the IoT platform, In contrast with the record-based confidentiality assumption.

OTAP, ATAP, and F&F attempt to generate fake events blindly to prevent all possible inferences from user data. In another research line, Replacement AutoEncoder (RAe) [13] transforms time-series sensor data in such a way that sensitive inference can not be inferred while desired and non-sensitive inference can be inferred from released data. In RAe, the privacy-utility trade-off is considered between sensitive inference and desired inference of user data. In contrast, in typical smart home platforms, executing automation rules correctly means preserving utility, and preventing sensitive inference means privacy enforcement. Therefore, while enforcing privacy in our research, we must guarantee no change for the user utility, as any unwanted, unknown utility change may violate the real-world IoT environment safety. Furthermore, in our threat model, we do not know the algorithm used by the malicious platform for inferring sensitive inference.

Our research tries to sanitize sensitive inference by considering the user privacy policy and the user automation rules in a real-world commercial IoT platform. The user's privacy policy in our solution is based on user activities. The privacy policies consist of explicit user "deny" policies and implicit auto-generated "allow" policies extracted from the user's automation rules.

References

[1] Bouchabou, D., Nguyen, S. M., Lohr, C., LeDuc, B., & Kanellos, I. (2021). A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning. Sensors, 21(18), 6037.

[2] Aksu, H., Babun, L., Conti, M., Tolomei, G., & Uluagac, A. S. (2018). Advertising in the IoT era: Vision and challenges. IEEE Communications Magazine, 56(11), 138-144.

[3] Zheng, S., Apthorpe, N., Chetty, M., & Feamster, N. (2018). User perceptions of smart home IoT privacy. Proceedings of the ACM on human-computer interaction, 2(CSCW), 1-20.

[4]"SmartThing Term of use 2017", https://edit.tosdr.org/services/3428/annotate#doc_5703. [Online Access: May 2022]

[5]"SmartThings U.S. Privacy Notice", https://eula.samsungiotcloud.com/legal/us/en/pps.html. [Online Access: May 2022]

[6] Ranasinghe, S., Al Machot, F., & Mayr, H. C. (2016). A review on applications of activity recognition systems with regard to performance and evaluation. International Journal of Distributed Sensor Networks, 12(8), 1550147716665520.

[7] Fernandes, E., Rahmati, A., Jung, J., & Prakash, A. (2018, February). Decentralized action integrity for trigger-action IoT platforms. In Proceedings 2018 Network and Distributed System Security Symposium.

[8] Chen, Y., Chowdhury, A. R., Wang, R., Sabelfeld, A., Chatterjee, R., & Fernandes, E. (2020). Data Privacy in Trigger-Action IoT Systems. arXiv e-prints, arXiv-2012.

[9] Xu, R., Zeng, Q., Zhu, L., Chi, H., Du, X., & Guizani, M. (2019). Privacy leakage in smart homes and its mitigation: IFTTT as a case study. IEEE Access, 7, 63457-63471..

[10] Chen, Y., Alhanahnah, M., Sabelfeld, A., Chatterjee, R., & Fernandes, E. Practical Data Access Minimization in Trigger-Action Platforms.

[11] Chi, H., Zeng, Q., Du, X., & Luo, L. (2021). PFirewall: Semantics-Aware Customizable Data Flow Control for Smart Home Privacy Protection. arXiv preprint arXiv:2101.10522.

[12] Chiang, Y. H., Hsiao, H. C., Yu, C. M., & Kim, T. H. J. (2020, September). On the privacy risks of compromised trigger-action platforms. In European Symposium on Research in Computer Security (pp. 251-271). Springer, Cham.

[13] Malekzadeh, M., Clegg, R. G., & Haddadi, H. (2018, April). Replacement autoencoder: A privacy-preserving algorithm for sensory data analysis. In 2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI) (pp. 165-176). IEEE.

[14] Zavalyshyn, I., Santos, N., Sadre, R., & Legay, A. (2020, December). My House, My Rules: A Private-by-Design Smart Home Platform. In MobiQuitous 2020-17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (pp. 273-282).

[15] Schoettler, S., Thompson, A., Gopalakrishna, R., & Gupta, T. (2020). Walnut: A low-trust trigger-action platform. arXiv preprint arXiv:2009.12447.