P. Kühtreiber and D. Reinhardt. Usable Differential Privacy for the Internet-of-Things. 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), pp. 426-427. 2021.

Usable Differential Privacy for the Internet-of-Things*

Patrick Kühtreiber
supervised by: Delphine Reinhardt
Institute of Computer Science
Georg-August-Universität Göttingen
Göttingen, Germany
kuehtreiber@cs.uni-goettingen.de

***Abstract*—Current implementations of Differential Privacy (DP) focus primarily on the privacy of the data release. The planned thesis will investigate steps towards a user-centric approach of DP in the scope of the Internet-of-Things (IoT) which focuses on data subjects, IoT developers, and data analysts. We will conduct user studies to find out more about the often conflicting interests of the involved parties and the encountered challenges. Furthermore, a technical solution will be developed to assist data subjects and analysts in making better informed decisions. As a result, we expect our contributions to be a step towards the development of usable DP for IoT sensor data.**

***Index Terms*—Pervasive Computing, Differential Privacy, Usable Privacy**

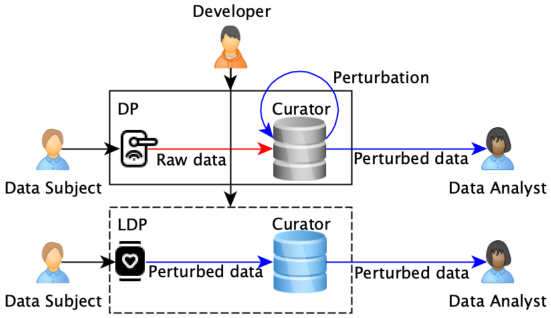

Fig. 1. IoT setup of DP and LDP.

I. Introduction

The IoT deployment and the associated data collection and processing has raised privacy concerns [1]. Even security systems designed for this context can threaten users' privacy [2]. To address these threats, different approaches like k-anonymity have been proposed. However, it has been shown that inference of personal data is still possible under k-anonymity and that anonymous data releases are not safe enough [3]. With the introduction of DP [4], a paradigm shift happened. Instead of receiving all the (anonymous) data, the data analyst queries the so-called curator that sends them the corresponding answers that will respect the data distribution but not reveal individual data. DP sets a statistical bound to the privacy risk of any individual contained in a data set. In short, DP guarantees that almost nothing can be learned from a data subject whether they are part of the data set or not. Since 2006, many DP-algorithms have been developed and published, including Local Differential Privacy (LDP). In contrast to DP, LDP users add noise locally to the collected data and send the already perturbed data to the curator (see Figure 1). While DP has been applied in the 2020 US Census and also by Apple, Google, and Microsoft, it has limitations: [5]–[7]. Therefore, it is still an open question whether DP is usable enough to provide both utility and privacy, especially in the context of IoT.

Moreover, very little research has been done on the party the most affected by the application of DP: the data subjects, i.e., the people about which data are collected. While a user study about their willingness to share data under DP has been conducted with a focus on health data [8], we are missing knowledge about IoT sensor data. Until now, it has also been shown that developers still struggle with the implementation of privacy-by-design [9] and generally do not feel responsible for privacy [10].

Our goal is therefore to consider three relevant actors: the data subjects, the application developers, and the data analysts in order to better understand what are the issues they encounter to be able to address them by proposing new technical solutions adapted to IoT. For example, this means providing transparency to both the data subjects and analysts by allowing them to see and understand the benefits, risks, or limitations associated to a particular DP implementation and parameterization. By doing so, we aim at assisting these actors in making better informed decisions and hence making DP usable in the context of IoT for all involved actors.

II. Methodology

We will investigate the problem of usable DP from three different angles, successively considering the data subjects, developers, and data analysts. All three groups have different needs and priorities and must be taken into account, so that we can build a generalized concept of a usable DP in the scope of IoT at the end.

A. Data Subjects

We will first explore the privacy attitudes and preferences of data subjects of IoT devices. Our study will investigate

*The final publication is available at IEEE Xplore via http://dx.doi.org/10.1109/PerComWorkshops51409.2021.9431047 whether the implementation of DP schemes (such as conventional DP, LDP or a hybrid model [11]) has an impact on their preferences. Like in [12]–[14], we will propose new methods to help data subjects visualizing, understanding, and controlling what is the trade-off between utility and privacy when considering the collected data. We will next measure the impact of these new methods on the data subjects' decisions to disclose data.

B. Developers

Our approach next focuses on DP and how developers can be supported in implementing DP and the visualization thereof in their applications. While privacy was historically not the first priority of developers [15], it is now becoming more important than ever due to legal regulations, such as the GDPR. We will hence perform user studies with IoT developers to understand their mental models of privacy protection in general. We will investigate whether they are aware of the concept of DP and what could be the limitations of not being able to implement them in practice. If necessary, we will explain DP and LDP to the participants and create a scenario in which a policy would be in place that requires a certain privacy budget which would define the upper bound of the DP-algorithm to be able to observe issues encountered by the application developers.

C. Data Analysts

The main focus of this step lies in the utility of DP-perturbed data. DP guarantees the privacy of an individual's data to a certain degree. However, data analysts, while being aware of privacy risks [16], prefer unperturbed data. We will again conduct user studies with data analysts to better understand the common ground where privacy and utility meet for this target group. Also, there have been voices arguing that DP is useless [5] for certain kinds of data and applications. With the results from this step, we will also gain a better understanding of DP's usefulness for certain data. Visual representations of the perturbation factor could also improve the data utility.

III. Contributions

In addition to the results obtained in the aforementioned user studies, we will develop a new technical solution to help these actors in better understanding the implications of DP and LDP in IoT. By using it, data subjects will be able to see potential privacy risks associated with a certain implementation of DP. This will however require a careful design as data subjects are not experts and the associated overhead should be reduced to the minimum. Therefore, different designs and visualization options will be considered and evaluated with potential data subjects. Moreover, we will support data analysts to also understand the consequences of the application of a particular implementation of DP on different dimensions including privacy and utility. By doing so, we hope to increase the utility of privacy-preserving IoT data and also evaluate different visual designs in order to help data analysts to make informed decisions. We will finally address IoT developers to reduce the limitations and obstacles they might face when the implementation of DP is explicitly required.

IV. Conclusions

The goal of this thesis is to make DP more useful and understandable to the people affected by it: Data subjects, developers, and data analysts. We will conduct user studies including questionnaires and lab studies to understand the concerns, issues, and needs of these three actors. We will design and develop new technical solutions to address them and will test how they can help the actors with the understanding and handling of DP. A goal is to visualize DP to data subjects and analysts. Their behaviors and responses towards this clearer representation of DP can guide developers to implement usable DP in order to solve the conflicting interests.

References

[1] C. Bettini and D. Riboni, "Privacy Protection in Pervasive Systems: State of the Art and Technical Challenges," Pervasive and Mobile Computing, vol. 17, 2015.

[2] E. Toch, C. Bettini, E. Shmueli, L. Radaelli, A. Lanzi, D. Riboni, and B. Lepri, "The Privacy Implications of Cyber Security Systems: A Technological Survey," ACM Computing Surveys (CSUR), vol. 51, no. 2, 2018.

[3] N. Li, W. H. Qardaji, and D. Su, "Provably Private Data Anonymization: Or, k-Anonymity Meets Differential Privacy," CoRR, abs/1101.2604, vol. 49, 2011.

[4] C. Dwork, "Differential Privacy," in Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)(2), 2006.

[5] S. Ruggles, C. Fitch, D. Magnuson, and J. Schroeder, "Differential Privacy and Census Data: Implications for Social and Economic Research," in AEA papers and proceedings, vol. 109, 2019.

[6] J. Bambauer, K. Muralidhar, and R. Sarathy, "Fool's Gold: An Illustrated Critique of Differential Privacy," Vanderbilt Journal of Entertainment and Technology Law, vol. 16, 2013.

[7] J. Domingo-Ferrer, D. Sánchez, and A. Blanco-Justicia, "The Limits of Differential Privacy (and its Misuse in Data Release and Machine Learning)," arXiv preprint arXiv:2011.02352, 2020.

[8] A. Xiong, T. Wang, N. Li, and S. Jha, "Towards Effective Differential Privacy Communication for Users' Data Sharing Decision and Comprehension," arXiv preprint arXiv:2003.13922, 2020.

[9] S. Spiekermann, J. Korunovska, and M. Langheinrich, "Inside the Organization: Why Privacy and Security Engineering Is a Challenge for Engineers," Proceedings of the IEEE, vol. 107, no. 3, 2018.

[10] K. Bednar, S. Spiekermann, and M. Langheinrich, "Engineering Privacy by Design: Are Engineers Ready to Live Up to the Challenge?," The Information Society, vol. 35, no. 3, 2019.

[11] B. Avent, A. Korolova, D. Zeber, T. Hovden, and B. Livshits, "BLENDER: Enabling Local Search with a Hybrid Differential Privacy Model," in 26th USENIX Security Symposium (USENIX Security 17), 2017.

[12] P. Menard and G. J. Bott, "Analyzing IoT Users' Mobile Device Privacy Concerns: Extracting Privacy Permissions Using a Disclosure Experiment," Computers & Security, 2020.

[13] D. Christin, M. Michalak, and M. Hollick, "Raising User Awareness about Privacy Threats in Participatory Sensing Applications through Graphical Warnings," in Proceedings of the 11th International Conference on Advances in Mobile Computing and Multimedia (MoMM), 2013.

[14] D. Christin, F. Engelmann, and M. Hollick, "Usable Privacy for Mobile Sensing Applications," in Proceedings of the 8th IFIP Workshop on Information Security Theory and Practice (WISTP), 2014.

[15] R. Balebako and L. Cranor, "Improving App Privacy: Nudging App Developers to Protect User Privacy," IEEE Security & Privacy, vol. 12, no. 4, 2014.

[16] C. Bettini, S. Kanhere, M. Langheinrich, A. Misra, and D. Reinhardt, “Is Privacy Regulation Slowing Down Research on Pervasive Computing?,” Computer, vol. 53, no. 6, 2020.