nature COMMUNICATIONS

MATTERS ARISING

https://doi.org/10.1038/s41467-021-27567-z OPEN

Reply to: On the difficulty of achieving differential privacy in practice: user-level guarantees in aggregate location data

Aleix Bassolas[1], Hugo Barbosa-Filho[2], Brian Dickinson[3], Xerxes Dotiwalla[4], Paul Eastham[4], Riccardo Gallotti[5], Gourab Ghoshal[2,6]✉, Bryant Gipson[4], Surendra A. Hazarie[2], Henry Kautz[3,6], Onur Kucuktunc[4], Allison Lieber[4], Adam Sadilek[4] & Jose J. Ramasco[7]✉

REPLYING TO F. Houssiau et al. *Nature Communications* https://doi.org/10.1038/s41467-021-27566-0 (2021)

In the work developed in Bassolas et al.[1], we studied the structure of cities and their impact in city livability using a highly aggregated mobility dataset. In order to protect privacy, random noise was added using an automated Laplace mechanism (ε, δ)-differential privacy, with $\varepsilon = 0.66$ and $\delta = 2.1 \times 10^{-29}$. Where ε sets the noise intensity and δ stands for the deviation from pure ε-privacy.

To illustrate the protection provided by a layer of (ε, δ)-differential privacy, with $\varepsilon = 0.66$ and $\delta = 2.1 \times 10^{-29}$, we note that an attacker can improve their certainty about an individual's presence or absence in the dataset by at most 16%. This observation holds even if the attacker knows every individual's data, including that of the target, via some side channel. An attack model like this is known as *membership inference with perfect knowledge*.

In their analysis, Houssiau et al. assume that the dataset referred to in the statistic is the entry dataset of trips. However, we specify the layer of (ε, δ)-differential privacy as per metric, i.e., *the number of trips from location A to location B per week W*. In other words, the *unit of privacy* that is protected with the promised differential privacy guarantees is not an individual's contribution to the entire dataset, but rather whether the individual made a trip from A to B during week W. We agree with Houssiau et al. that it is important to communicate privacy protection precisely and we should have been more specific to avoid confusion.

It is worth pointing out that although Houssiau et al. correctly hypothesize that the 16% statistic does not hold when applied to the entire dataset, there are some discrepancies between their analysis and the privacy mechanisms we apply, resulting in stronger privacy protection in practice. In particular, we bound an individual's contribution to a particular aggregation partition, i.e., trips from A to B within a week W, to 1. Moreover, the geographical areas we consider are grid cells of size ~1.3 km^2 rather than exact locations, as Houssiau et al. assume. Thus, Houssiau et al.'s analysis of a single user (one of the authors), who reported 39 trips in total, likely translates to fewer contributions to the entire dataset and consequently also results in less privacy loss when evaluated over the entire dataset. Finally, we want to emphasize that membership inference with perfect knowledge of the entire dataset is a very strong attack model that is unrealistic in practice. So we stand by our claim that the dataset is highly aggregated and anonymous for all practical purposes.

Below we provide a clarified description of our data aggregation:

The automated Laplace mechanism adds random noise drawn from a zero mean Laplace distribution and yields (ε, δ)-differential privacy guarantee of $\varepsilon = 0.66$ and $\delta = 2.1 \times 10^{-29}$ per metric. Specifically, for each week W and each location pair (A, B), we compute the number of unique users who took a trip from location A to location B during week W. To each of these metrics, we add Laplace noise from a zero-mean distribution of scale 1/0.66. We then remove all metrics for which the noisy number of users is lower than 100, following the process described in ref. [2] and publish those remaining. Each metric published therefore satisfies (ε, δ)-differential privacy with values defined above.

The parameter ε controls the noise intensity in terms of its variance, while δ represents the deviation from pure ε-privacy.

[1] School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK. [2] Department of Physics & Astronomy, University of Rochester, Rochester, NY, USA. [3] Department of Computer Science, University of Rochester, Rochester, NY, USA. [4] Google Inc, 1600 Amphitheatre Parkway, Mountain View, CA, USA. [5] Bruno Kessler Foundation (FBK), Trento, Italy. [6] Goergen Institute for Data Science, University of Rochester, Rochester, NY, USA. [7] Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Campus UIB, 07122 Palma de Mallorca, Spain. ✉email: gghoshal@pas.rochester.edu; jramasco@ifisc.uib-csic.es

The closer they are to zero, the stronger the privacy guarantees. For example, with these values of the parameters, an attacker with perfect knowledge on all users except user U would increase the level of certainty as to whether U went from geographical area A to area B during a given week no more than 16%. Each user contributes at most one increment to each partition. If they go from a region A to another region B multiple times in the same week, they only contribute once to the aggregation count. No individual user data was ever manually inspected, only heavily aggregated flows of large populations were handled.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Code availability

Code sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Received: 16 May 2020; Accepted: 26 November 2021;
Published online: 10 January 2022

References

1. Bassolas, A. et al. Hierarchical organization of urban mobility and its connection with city livability. *Nat. Commun.* **10**, 4817 (2019).
2. Wilson, R. J. et al. Differentially private sql with bounded user contribution. *Proc. Priv. Enhancing Technol.* **2020**, 230–250 (2020).

Acknowledgements

A.B. is funded by the Conselleria d'Educacio, Cultura i Universitats of the Government of the Balearic Islands and the European Social Fund. A.B. and J.J.R. also acknowledge partial funding from the Spanish Ministry of Science and Innovation, the National Agency for Research Funding AEI MCIN/AEI/10.13039/501100011033/ and FEDER (EU) under the grant PACSS (RTI2018-093732-B-C22) and the Maria de Maeztu program for Units of Excellence in R&D (MDM-2017-0711). G.G. and S.H. acknowledge funding from the Department of Economic Development (DED), New York through the NYS Center of Excellence in Data Science at the University of Rochester (C160189). G.G. and H.B. also acknowledge support in part by the U. S. Army Research Office (ARO) under grant number W911NF-18-1-0421. Any opinions, findings, conclusions or recommendations expressed are those of the author(s) and do not necessarily reflect the views of the DED or the ARO.

Author contributions

A.B., H.B., B.D., R.G., G.G., S.A.H., A.S., and J.J.R. contributed to the work methodology. A.B., R.G., G.G., H.K., A.S., and J.J.R. wrote the paper. G.G., H.K., A.S., and J.J.R. coordinated the study. All authors read, edited, and approved the final version of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Gourab Ghoshal or Jose J. Ramasco.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.