Bachelor Project on Differential Privacy

Johan Ringmann Fagerberg (jofag17) - 1996-09-24

Motivation

Differential privacy is a recently proposed measure of the level of privacy provided by random functions. It provides a formal measure for how much information a potential attacker can achieve from a perturbed data set, when the data perturbation consists of random noise.

In global differential privacy, a global aggregator has access to the real data prior to perturbing it. This differs from local differential privacy, where users perturb their data before sending it to the aggregator.

The goal of this bachelor project is to gain an understanding of differential privacy in general, with a particular focus on local differential privacy, and use that understanding to implement and evaluate prior art in differentially private analytics.

Plan

The project will consist in 4 phases detailed below

- Survey of the state of the art in differential privacy (tentatively February-March). The survey will start by consulting [Des19], i.e., a reading list of key papers in differential privacy presented as a semi-technical blog series on differential privacy. Moreover, we will also read [Yan+20], i.e., a recent survey on local differential privacy methods for various use cases.
- Survey of existing frameworks for differential privacy within the field of analytics (tentatively March-April). Candidates for this include [EPK14], [Bit+17], [Wan+19] and [DKY17].
- Implementation and evaluation of a prototype based on one of the surveyed frameworks (tentatively April).
- Report writing (tentatively April-May). The final report will be written following the academic conventions.

Risks

TODO

Outcomes

At the end of the project, the outcomes of the project will be

- A report written in English and following the standard academic writing conventions. The report will contain the description of the state-of-the-art, the description of the prototype and its preliminary evaluation.
- The code of the prototype and the data for the experiments performed

References

[Bit+17] Andrea Bittau et al. "Prochlo: Strong Privacy for Analytics in the Crowd". In: (2017). DOI: `10.1145/3132747.3132769`. eprint: `arXiv:1710.00901`.

[Des19] Damien Desfontaines. *A reading list on differential privacy.* 2019. URL: `https://desfontain.es/privacy/differential-privacy-reading-list.html` (visited on 11/25/2020).

[DKY17] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. "Collecting Telemetry Data Privately". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems.* NIPS'17. Long Beach, California, USA: Curran Associates Inc., 2017, pp. 3574–3583. ISBN: 9781510860964.

[EPK14] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. "RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response". In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security.* CCS '14. Scottsdale, Arizona, USA: Association for Computing Machinery, 2014, pp. 1054–1067. ISBN: 9781450329576. DOI: `10.1145/2660267.2660348`. URL: `https://doi.org/10.1145/2660267.2660348`.

[Wan+19] N. Wang et al. "Collecting and Analyzing Multidimensional Data with Local Differential Privacy". In: *2019 IEEE 35th International Conference on Data Engineering (ICDE).* 2019, pp. 638–649. DOI: `10.1109/ICDE.2019.00063`.

[Yan+20] Mengmeng Yang et al. "Local Differential Privacy and Its Applications: A Comprehensive Survey". In: *arXiv:2008.03686 [cs]* (Aug. 2020). arXiv: 2008.03686. URL: `http://arxiv.org/abs/2008.03686` (visited on 11/25/2020).