Journal of Privacy and Confidentiality
Vol. 9 (2) 2019
DP + TPDP 2017
Submitted Sept 2019
Published Sept 2019

EDITORIAL FOR VOLUME 9 ISSUE 2

JONATHAN ULLMAN AND LARS VILHUBER

Northeastern University
e-mail address: jullman@ccs.neu.edu

Cornell University and Managing Editor
e-mail address: managing-editor@journalprivacyconfidentiality.org

Differential privacy is a promising approach to privacy-preserving data analysis that provides strong worst-case guarantees about the harm that a user could suffer from contributing their data, but is also flexible enough to allow for a wide variety of data analyses to be performed with a high degree of utility. Researchers in differential privacy span many distinct research communities, including algorithms, computer security, cryptography, databases, data mining, machine learning, statistics, programming languages, social sciences, and law.

Two articles in this issue describe applications of differentially private, or nearly differentially private, algorithms to data from the U.S. Census Bureau:

- Raj Chetty and John Friedman (2019) apply noise to statistics of interest generated from samples with quite small number of observations. Because it uses a data-driven method to adapt the sensitivity of the algorithm, it is does not offer a formal privacy guarantee. However, it greatly outperforms traditional methods, while limiting the non-formal leakage of information to state-level statistics. The data released via this mechanisms, after review by the Census Bureau's Disclosure Review Board, can be found at `https://opportunityinsights.org/`.
- Andrew D. Foote, Ashwin Machanavajjhala and Kevin McKinney (2019) tackle a problem that is a poster child for the application of formal privacy mechanisms. In their case, state educational institutions already release earnings outcomes for their graduates based on in-state administrative records. The Census Bureau in turn publishes these outcomes, also per educational institutions, for nation-wide earnings outcomes, but is mandated to protect disclosure of individual records even from the state institutions that provide the original in-state data. They develop an algorithm that creates a differentially private estimate of the histogram of earnings, from which they release protected percentiles. They demonstrate that their mechanisms is more accurate most of the time when compared to smooth sensitivity (Nissim, Raskhodnikova and Smith, 2007). The data they release can be explored at `https://lehd.ces.census.gov/data/pseo_beta_viz.html` and downloaded at `https://lehd.ces.census.gov/data/pseo_beta.html`.

Received by the editors September 2019.
Key words and phrases: differential privacy; TPDP 2017.

JPrivacy Confidentiality
www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.731

The third article highlights a thorny issue that applies to all implementations of differential privacy: how to choose the key privacy parameter ϵ:

- Cynthia Dwork, Nitin Kohli and Deirdre Mulligan (2019) note that "there is little understanding of what is the optimal value of ϵ for a given system or classes of systems, purposes, data, etc., or how to go about figuring it out." They report on a survey of current implementations at various organizations, and propose the creation of the *Epsilon Registry* – "a publicly available communal body of knowledge about differential privacy implementations."

Theory and Practice of Differential Privacy 2017

This issue also includes selected contributions from the *3rd Workshop on Theory and Practice of Differential Privacy*, which was held in Dallas, TX on October 30, 2017 as part of the *ACM Conference on Computer Security (CCS).* The workshop brought researchers from these communities together to discuss recent developments in both the theory and practice of differential privacy. Three articles were retained:

- Victor Balcer and Salil Vadhan (2019) design differentially private algorithms for computing histograms using only discrete operations and finite running time. This work is motivated in particular by attacks on standard differentially private algorithms when implemented using floating-point operations (Mironov, 2012).
- Brendan Avent, Aleksandra Korolova, David Zeber, Torgeir Hovden and Benjamin Livshits (2019) propose a hybrid model of differential privacy for distributed data for companies whose user bases are too small to effectively use local differential privacy. Their work shows that if just a small fraction of users are willing to accept central differential privacy, then it is possible to achieve much greater utility than the local model.
- Steven Wu, Aaron Roth, Katrina Ligett, Bo Waggoner and Seth Neel (2019) propose designing differentially private algorithms that satisfy a fixed *accuracy constraint* at minimal cost in privacy, in contrast to satisfying a fixed privacy constraint at minimal cost in accuracy. They show an effective way to achieve this goal for linear and logistic regression problems.

The full program is available as part of this issue (Ullman, 2019) and online at `https://tpdp.cse.buffalo.edu/2017/`. TPDP 2019 will be held as part of CCS 2019 on 11 November 2019, the program is available at `https://tpdp.cse.buffalo.edu/2019/`.

References

Avent, Brendan, Aleksandra Korolova, David Zeber, Torgeir Hovden, and Benjamin Livshits. 2019. "BLENDER: Enabling Local Search with a Hybrid Differential Privacy Model." *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.680 .

Balcer, Victor, and Salil Vadhan. 2019. "Differential Privacy on Finite Computers." *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.679 .

Chetty, Raj, and John Friedman. 2019. "A Practical Method to Reduce Privacy Loss when Disclosing Statistics Based on Small Samples." *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.716 .

Dwork, Cynthia, Nitin Kohli, and Deirdre Mulligan. 2019. "Differential Privacy in Practice: Expose your Epsilons!" *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.689 .

Foote, Andrew D., Ashwin Machanavajjhala, and Kevin McKinney. 2019. "Releasing Earnings Distributions using Differential Privacy: Disclosure Avoidance System For Post-Secondary Employment Outcomes (PSEO)." *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.722 .

Mironov, Ilya. 2012. "On significance of the least significant bits for differential privacy." ACM Press. https://doi.org/10.1145/2382196.2382264 .

Nissim, Kobbi, Sofya Raskhodnikova, and Adam Smith. 2007. "Smooth Sensitivity and Sampling in Private Data Analysis." *STOC '07*, 75–84. New York, NY, USA:ACM. https://doi.org/10.1145/1250790.1250803 . http://doi.acm.org/10.1145/1250790.1250803.

Ullman, Jonathan. 2019. "Program for TPDP 2017." *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.698 .

Wu, Steven, Aaron Roth, Katrina Ligett, Bo Waggoner, and Seth Neel. 2019. "Accuracy First: Selecting a Differential Privacy Level for Accuracy-Constrained ERM." *Journal of Privacy and Confidentiality*, 9. https://doi.org/10.29012/jpc.682 .