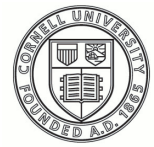


8-2018

The U.S. Census Bureau Adopts Differential Privacy

John M. Abowd
U.S. Census Bureau, john.maron.abowd@census.gov

The U.S. Census Bureau Adopts Differential Privacy

Abstract

The U.S. Census Bureau announced, via its Scientific Advisory Committee, that it would protect the publications of the 2018 End-to-End Census Test (E2E) using differential privacy. The E2E test is a dress rehearsal for the 2020 Census, the constitutionally mandated enumeration of the population used to reapportion the House of Representatives and redraw every legislative district in the country. Systems that perform successfully in the E2E test are then used in the production of the 2020 Census.
Motivation: The Census Bureau conducted internal research that confirmed that the statistical disclosure limitation systems used for the 2000 and 2010 Censuses had serious vulnerabilities that were exposed by the Dinur and Nissim (2003) database reconstruction theorem. We designed a differentially private publication system that directly addressed these vulnerabilities while preserving the fitness for use of the core statistical products.
Problem statement: Designing and engineering production differential privacy systems requires two primary components: (1) inventing and constructing algorithms that deliver maximum accuracy for a given privacy-loss budget and (2) insuring that the privacy-loss budget can be directly controlled by the policy-makers who must choose an appropriate point on the accuracy-privacy-loss tradeoff. The first problem lies in the domain of computer science. The second lies in the domain of economics.
Approach: The algorithms under development for the 2020 Census focus on the data used to draw legislative districts and to enforce the 1965 Voting Rights Act (VRA). These algorithms efficiently distribute the noise injected by differential privacy. The Data Stewardship Executive Policy Committee selects the privacy-loss parameter after reviewing accuracy-privacy-loss graphs.

Keywords

differential privacy, Economics of privacy

Disciplines

Labor Economics | Social Statistics | Statistical Methodology

Comments

Citation: Abowd, John M. "The U.S. Census Bureau Adopts Differential Privacy," KDD '18 Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK (August 2018): 2867, DOI: 10.1145/3219819.3226070.

The presentation files are available at https://hdl.handle.net/1813/60392.

The U.S. Census Bureau Adopts Differential Privacy

John M. Abowd
United States Census Bureau
Washington, DC, USA
john.maron.abowd@census.gov

ABSTRACT

The U.S. Census Bureau announced, via its Scientific Advisory Committee, that it would protect the publications of the 2018 End-to-End Census Test (E2E) using differential privacy. The E2E test is a dress rehearsal for the 2020 Census, the constitutionally mandated enumeration of the population used to reapportion the House of Representatives and redraw every legislative district in the country. Systems that perform successfully in the E2E test are then used in the production of the 2020 Census.

Motivation: The Census Bureau conducted internal research that confirmed that the statistical disclosure limitation systems used for the 2000 and 2010 Censuses had serious vulnerabilities that were exposed by the Dinur and Nissim (2003) database reconstruction theorem. We designed a differentially private publication system that directly addressed these vulnerabilities while preserving the fitness for use of the core statistical products.

Problem statement: Designing and engineering production differential privacy systems requires two primary components: (1) inventing and constructing algorithms that deliver maximum accuracy for a given privacy-loss budget and (2) insuring that the privacy-loss budget can be directly controlled by the policy-makers who must choose an appropriate point on the accuracy-privacy-loss tradeoff. The first problem lies in the domain of computer science. The second lies in the domain of economics.

Approach: The algorithms under development for the 2020 Census focus on the data used to draw legislative districts and to enforce the 1965 Voting Rights Act (VRA). These algorithms efficiently distribute the noise injected by differential privacy. The Data Stewardship Executive Policy Committee selects the privacy-loss parameter after reviewing accuracy-privacy-loss graphs.

CCS Concepts/ACM Classifiers

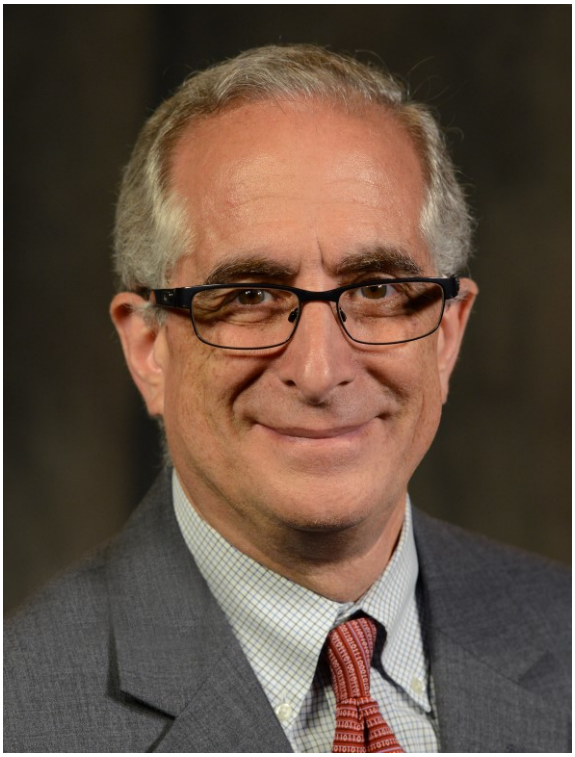
Privacy-preserving protocols; Privacy protections

Author Keywords

Differential privacy; Economics of privacy

BIOGRAPHY

John M. Abowd is Chief Scientist and Associate Director for Research and Methodology at the U.S. Census Bureau and the Edmund Ezra Day Professor of Economics, Professor of Statistics and Information Science at Cornell University. He is also Research Associate at the National Bureau of Economic Research, Research Affiliate at the Centre de Recherche en Economie et Statistique (CREST, Paris, France), Research Fellow at the Institute for Labor Economics (IZA, Bonn, Germany), and Research Fellow at IAB (Institut für Arbeitsmarkt-und Berufsforschung, Nürnberg, Germany). He is the past President (2014-2015) and Fellow of the Society of Labor Economists; Fellow of the American Statistical Association; elected member of the International Statistical Institute; and a fellow of the Econometric Society. He served as Distinguished Senior Research Fellow at the United States Census Bureau from 1998 to 2016, and on the National Academies' Committee on National Statistics (CNSTAT 2010-2016). He currently serves on the American Economic Association's Committee on Economic Statistics (2013-2018). He was the scientific lead on the team that implemented the first formally private production disclosure limitation system worldwide: OnTheMap (see Machanavajjhala et al. 2008). https://johnabowd.com



KDD 2018, August 19-23, 2018, London, United Kingdom.

ACM ISBN 978-1-4503-5552-0/18/08.
DOI: https://doi.org/10.1145/3219819.3226070

REFERENCES

1. Dinur, I. and K. Nissim (2003) "Revealing Information while Preserving Privacy" PODS 2003 3 ACM 1-58113-670-6/03/06.
2. Machanavajjhala, A., D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber (2008) "Privacy: Theory Meets Practice on the Map," ICDE, doi:10.1109/ICDE.2008.4497436