Fixed-Parameter Tractability of Private Synthetic Data Generation

Badih Ghazi BADIHGHAZI@GMAIL.COM
Google DeepMind

Cristóbal Guzmán CRGUZMANP@UC.CL
Institute for Mathematical and Computational Engineering, Faculty of Mathematics and School of Engineering, Pontificia Universidad Católica de Chile

Pritish Kamath PRITISH@ALUM.MIT.EDU
Google Research

Alexander Knop ALEXANDERKNOP@GOOGLE.COM
Google Research

Ravi Kumar RAVI.K53@GMAIL.COM
Google Research

Pasin Manurangsi PASIN@GOOGLE.COM
Google Research

Editors: Steve Hanneke and Tor Lattimore

Abstract

We study the problem of generating synthetic data under differential privacy. We establish fixed-parameter tractability (FPT) for this problem where the parameter is the treewidth of the query family's incidence graph. Our algorithms attain optimal error rates across all regimes and are realized by two different approaches: the first is based on linear programming (LP) and the FPT of the separation problem for the LP dual; the second is based on a subsampled private multiplicative weights method, where we obtain FPT for sampling from Gibbs distributions. Both approaches are unified by a dynamic programming framework over a tree decomposition.

Keywords: Differential Privacy, Synthetic Data Generation, Fixed-Parameter Tractability

Acknowledgments

Part of C. Guzmán's work was done as a visiting researcher at Google. C. Guzmán's research was partially supported by ANID FONDECYT 1251029 grant, and National Center for Artificial Intelligence CENIA FB210017, Basal ANID. We would like to thank Victor Verdugo, José Verschae, and Jingcheng Liu for valuable discussions about this work, as well as the useful feedback from anonymous reviewers.

. Extended abstract. Full version appears as [arxiv reference,2606.11283]