WANRONG ZHANG

wanrongzhang@fas.harvard.edu
https://wanrongz.github.io

ACADEMIC POSITION

Postdoctoral Research Fellow 07/2021- present

CRA/CCC's NSF-funded Computing Innovation Fellows Program
Theory of Computation group and OpenDP project
Harvard University, Cambridge, MA
Host: Salil Vadhan

EDUCATION

Georgia Institute of Technology, Atlanta, GA 08/2016 - 05/2021

Ph.D. in Industrial Engineering and Operation Research, Minor in Machine Learning
Dissertation: Privacy-preserving Statistical Tools: Differential Privacy and Beyonds
Advisors: Rachel Cummings and Yajun Mei

Peking University, Beijing, China 09/2012 - 06/2016

B.S. in Statistics and Probability

RESEARCH INTERNSHIPS AND LONG-TERM VISITS

Microsoft Research, Redmond WA 05/2020- 08/2020

Research Intern with Nalin Singal, Robert Sim, Jana Kulkarni and Priyanka Kulkarni

Differentially private sentence embedding models

Microsoft Research, Cambridge UK 05/2019- 08/2019

Research Intern with Olya Ohrimenko and Shruti Tople

Dataset-level attribute leakage in multi-party machine learning

University of California, Berkeley 01/2019- 05/2019

Visiting Graduate Student in Simons Institute for the Theory of Computing

RESEARCH INTERESTS

My research interests lie primarily in data privacy, in particular, in differential privacy, including (1) developing the theoretical foundations, (2) designing privacy-preserving algorithms for machine learning models and statistical analysis tools, and (3) adapting existing tools to solve domain-specific questions. In addition, I am interested in (4) broader privacy concerns, including understanding privacy vulnerabilities and proposing solutions.

PUBLICATIONS & PREPRINTS

Note: The convention in TCS is to list authors in alphabetical order. (* indicates primary author)

Conference Papers

DP-Fast MH: Private, Fast, and Accurate Metropolis-Hastings for Large-Scale Bayesian Inference, Wanrong Zhang, Ruqi Zhang. *International Conference on Machine Learning (ICML) 2023.*

Concurrent Composition Theorems for Differential Privacy, Salil Vadhan, Wanrong Zhang* (Alphabetical order),
The ACM Symposium on Theory of Computing (STOC) 2023

Private Sequential Hypothesis Testing for Statisticians: Privacy, Error Rates, and Sample Size, Wanrong Zhang, Yajun Mei, Rachel Cummings.
International Conference on Artificial Intelligence and Statistics (AISTATS) 2022

Attribute Privacy: Framework and Mechanisms, Wanrong Zhang, Olga Ohrimenko, Rachel Cummings.
ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT) 2022.
Symposium on Foundations of Responsible Computing (FORC) 2021 (non-archival track).

PAPRIKA: Private Online False Discovery Rate Control, Wanrong Zhang, Gautam Kamath, Rachel Cummings.
International Conference on Machine Learning (ICML) 2021).
Symposium on Foundations of Responsible Computing (FORC) 2021 (non-archival track).

Leakage of Dataset Properties in Multi-Party Machine Learning, Wanrong Zhang, Shruti Tople, Olga Ohrimenko.
30th USENIX Security Symposium (USENIX Security) 2021.

Privately Detecting Changes in Unknown Distributions, Rachel Cummings, Sara Krehbiel, Yuliia Lut, Wanrong Zhang* (Alphabetical order).
International Conference on Machine Learning (ICML) 2020.

Differentially Private Change-Point Detection, Rachel Cummings, Sara Krehbiel, Yajun Mei, Rui Tuo, Wanrong Zhang* (Alphabetical order).
Advances in Neural Information Processing Systems, (NeurIPS) 2018.

Journal Papers

Single and Multiple Change-Point Detection with Differential Privacy,
Wanrong Zhang, Sara Krehbiel, Rui Tuo, Yajun Mei, Rachel Cummings. *Journal of Machine Learn-*

ing Research (JMLR) 2021.

Bandit Change-Point Detection for Real-Time Monitoring High-Dimensional Data Under Sampling Control,
Wanrong Zhang, Yajun Mei. *Technometrics 2022.*

Preprints

A standardized differential privacy framework for epidemiological modeling with mobile phone data, M.K. Savi, A. Yavad, W. Zhang, N. Vembar, A. Schroeder, S. Balsari, C. Buckee, S. Vadhan, N. Kishore. Under Submission.

Continual Release of Differentially Private Synthetic Data, Wanrong Zhang, Marcel Neuhoeffer, Mark Bun, Marco Gaboardi. Under Submission

ONGOING PROJECTS

Currently, I'm particularly interested in privacy and security in AI tools. I have a few ongoing projects including (1) differentially private representation through contrastive language-image pre-training, (2) membership inference attacks and privacy in topic modeling, (3) machine unlearning from AI tools.

HONORS & AWARDS

Computing Innovation Fellowship (CI Fellowship), CCC/CRA/NSF	2021-2023
Rising Stars in EECS	2022
CDAC Rising Stars in Data Science	2021
ARC-TRIAD Fellowship, Georgia Tech	2019
The President Fellowship, Peking University	2015

TEACHING

Harvard CS208: Applied Privacy for Data Science (co-teaching with Salil Vadhan and James Honaker, Spring 2022)

Teaching Assistantships

ISyE 6412: Theoretical Statistics (Fall 2019)
ISyE 6669: Deterministic Optimization (Fall 2018)
ISyE 4031: Regression and Forecasting (Spring 2018)
ISyE 3039: Methods Quality Improvement (Summer 2017)
ISyE 2028: Basic Statistical Methods (Spring 2017)
ISyE 3770: Statistics and Applications (Fall 2016)

TALKS

Composition Theorems for Interactive Differential Privacy

· CATT 2022 Global Analytics Conference, UT Austin, November 2022.

Concurrent Composition Theorems for Differential Privacy

· Societal Considerations and Applications Workshop, Simons Institute for the Theory of Computing, November 2022.
· Privacy Tools DP meeting, September 2022.
· Google Privacy Seminar, August 2022.

Ensuring privacy in COVID-19 epidemiological mobility data sets

· Trust in Science Workshop, Harvard, September 2022 (with Koissi Savi and Nishant Kishore).

Differentially Private Approaches for Streaming Data Analysis

· ICSA Applied Statistics Symposium, June 2022.

Private Sequential Hypothesis Testing for Statisticians: Privacy, Error Rates, and Sample Size

· INFORMS ICS, Tampa, January 2022.

Leakage of Dataset Properties in Multi-Party Machine Learning

· USENIX Security Symposium, August 2021.

Privacy-Preserving Statistical Tools: Differential Privacy and Beyond

· Microsoft Research, February 2021.

Attribute Privacy: Framework and Mechanisms

· ACM FAccT, June 2022.
· FORC, June 2021.
· INFORMS annual meeting, November 2020.

PAPRIKA: Private Online False Discovery Rate Control

· ICML, July 2021.
· FORC, June 2021.
· CDAC Rising Stars in Data Science, January 2021.

Privately Detecting Changes in Unknown Distributions

· ICML, July, 2020.

Differentially Private Change-point Detection

· Boston-area DP seminar, December 2020.
· Cybersecurity Lecture Series, Georgia Tech, March 2020.
· INFORMS annual meeting, Seattle, October 2019.

Bandit Change-Point Detection for Real-Time Monitoring High-Dimensional Data Under Sampling Control

· INFORMS annual meeting, November 2020.

SERVICE

Workshop Organization 09/2021 - present

· The Boston-Area Data Privacy Seminar Series, co-organize with Maryam Aliakbarpour, Boston, MA

Program Committee

I have been (or will be) on the program committee (i.e., a reviewer) for
· Conferences: NeurIPS20, AAAI21, ICLR21, AISTATS21, ICML21, NeurIPS21, ICLR22, ICML22, FAccT22, COLT22, ISIT22, COLT23, FAccT23;
· Journals: Journal of Applied Statistics, Statistica Sinica, Journal of Machine Learning Research, Transactions on Machine Learning Research, Computers&Security.
· Workshops: TPDP20, TPDP21, TPDP22;