Erchi Wang

La Jolla, CA erw011@ucsd.edu 217-693-8227 erchiw.github.io erchi-wang erchiw

Summary

Ph.D. in Data Science at UC San Diego, specializing in privacy-preserving machine learning. Experienced in developing data-adaptive and practical differentially private algorithms with provable guarantees[1, 2]. Currently, I am also exploring the use of differential privacy techniques to address safety risks in generative models, such as protecting privacy in Retrieval-Augmented Generation (RAG) [3] and quantifying LLM memorization.

Education

University of California, San Diego, San Diego, US — *Jul. 2024 – Est. 2027*
Ph.D. in Data Science, GPA: 3.90/4.0 — Advised by *Prof. Yu-Xiang Wang*

University of California, Santa Barbara, Santa Barbara, US — *Aug. 2021 – Jul. 2024*
Ph.D. in Statistics (transferred to UCSD), GPA: 3.91/4.0 — Advised by *Prof. Yu-Xiang Wang*

University of Illinois at Urbana-Champaign, Urbana, US — *Aug. 2018 – Dec. 2020*
M.S. in Statistics, GPA: 3.82/4.0

Ocean University of China, Qingdao, China — *Aug. 2013 – Jul. 2018*
B.S. in Applied Math and Biological Science, GPA: 3.76/4.0

Publications & Manuscripts (* denotes equal contribution)

[1] **Erchi Wang**, Yuqing Zhu, Yu-Xiang Wang. Adapting to Linear Separable Subsets with Large Margin in Differentially Private Learning. *Accepted by ICML-2025. Oral presentation at TPDP 2025. Arxiv link*

[2] Yingyu Lin*, **Erchi Wang***, Yi-An Ma, Yu-Xiang Wang. Purifying Approximate Differential Privacy with Randomized Post-processing. *In Submission. Oral presentations at TPDP 2025 and Crypto-PPML 2025. Arxiv link*

[3] Ruihan Wu*, **Erchi Wang***, Yu-Xiang Wang. Beyond Per-Question Privacy: Multi-Query Differential Privacy for RAG Systems. *Manuscript*

[4]Erchi Wang, Arinbjörn Kolbeinsson, Luca Foschini, Yu-Xiang Wang. Revisiting Differentially Private XGboost: Are Random Decision Trees Really Better than Greedy Ones? *In Submission.*

Selected Projects

Multi-Query Retrieval-Augmented Generation with differential privacy guarantee

- Designed a novel DP-RAG framework enabling multiple-query retrieval-augmented generation with significantly reduced privacy budget and enhanced generation utility.
- Demonstrated practical performance on four QA benchmarks and three LLMs (OPT-1.3B, Pythia-1.4B, and Mistral-7B), achieving up to 100× privacy savings, while maintaining stronger utility on privacy-sensitive tasks compared to non-private LLM without RAG.

Differential Private Adaptive Margin Learning

- Designed a computationally efficient differentially private algorithm for classification problems. Implemented advanced private hyperparameter tuning methods and refined the analysis of DP-SGD, allowing the algorithm to adapt to large data margins without requiring prior knowledge of the margin value. Theoretically, the proposed method guarantees utility adaptation to both separable and non-separable cases.

Converting Approximate DP Mechanisms into Pure DP Mechanisms

- Developed a black-box converter from approximate to pure differential privacy and leveraged it to design efficient pure DP optimization and data-dependent algorithms that were previously difficult to construct.

(Ongoing) Quantifying Per-instance Memorization in Large Language Model

- Systematically reviewed various concepts of LLM memorization and proposed a per-instance memorization framework analyzed through the lens of a data reconstruction attack. Developing algorithmic tools for

auditing memorization.

Differential Private Greedy XGBoost on Tabular Data

- Designed and implemented an enhanced differentially private greedy XGBoost algorithm, leveraging modern privacy accounting techniques, including Rényi Differential Privacy-based composition and bounded range analysis for the exponential mechanism. (GitHub Repo)
- Conducted extensive empirical studies on 18 UCI tabular datasets, achieving state-of-the-art performance with DP-XGBoost by reducing the number of trees by 30% to 50%, thereby enhancing model explainability and accelerating inference speed.

Programming Skills

Languages: Python, R, Bash, Git

Libraries & Frameworks: Pytorch, Pandas, SciPy, Scikit-learn, Opacus, AutoDP,

Professional Service

Reviewers for NeurIPS (2024, 2025), ICLR (2025), AISTATS (2025), ICML (2025)