“Tao” Uthaipon Tantipongpipat

Orlando, FL, USA +1 (804) 625-7798
www.uthaipon.com | Google Scholar uthaipon@gmail.com

Topics of Experience

- Fairness and privacy in machine learning – fair, explainable, and accountable machine learning; differential privacy, especially in deep learning and synthetic data generation; Human-Centered Interaction (HCI)
- Discrete and continuous optimization
- Optimal design in statistics, or design of experiments (DoE); diversity and representative sampling.

Other Interests: casual analysis, security

Work Experiences

Jun 2020 -Now

Machine Learning Researcher, Twitter, remote US
META (Machine learning Ethics, Transparency and Accountability), Cortex

- Led Twitter’s image cropping algorithmic bias audit resulting in a published academic paper and $1.5M press ad equivalency and 3B readership from 500 news articles in 49 countries. Led to another follow-up work by team members resulted in additional $1.4M, 2.7B reads, and 800 articles from 47 additional countries, and resulted in the production change to remove the algorithm
- Proposed a 13-18% precision-recall video classification model improvement with no additional cost to partnering team to fix offensive misclassifications on Tweet topic annotations, and discovered correlation bias with demographics despite a lack of private individual data
- Established a data-driven guideline for company-wide engineers to adopt a inequality metric in A/B statistical testing, as well as winning business approval with company leadership that finally led to shipping the metric
- Provided statistical analysis to partnering teams to evaluate and quantify bias in ML models; redesigned common ML statistical significance tests required for bias measurement
- Mentored a junior researcher
- Published two papers in social computing conference and one in data science journal
- Media coverage is positive on the team at the Twitter laid off event: https://www.wired.com/story/twitter-ethical-ai-team/

May - Jul 2019

Research Intern, Microsoft Research, Redmond, WA
Algorithms group. Supervisor: Janardhan Kulkarni and Sergey Yekhanin.

- Implemented privacy guarantee on large-scale natural language processing models (RNNs and LSTMs) to protect against personal deidentification due to model usage
- Developed novel correlation clustering algorithm with corresponding privacy analysis
- Researched private submodular optimization and surveyed literature for private stochastic gradient descent for improving training deep learning models
- Published one paper in a machine learning conference

Education

Aug 2016 -May 2020

Georgia Institute of Technology, Atlanta, GA, United States

	PhD in Algorithms, Combinatorics, and Optimization (ACO), School of Computer Science GPA 4.00/4.00 ● Minor in Computational Learning Theory ● Advisor: Dr. Mohit Singh ● Thesis: **Fair and Diverse Data Representation in Machine Learning**
Aug 2012 - May 2016	**University of Richmond**, Richmond, VA, United States BS in Mathematics (Honors). GPA: 3.97/4.00 ● Full-merit Robins Science scholarship covering tuition, fees, accommodations, and meals ● Thesis in algebraic combinatorics and discrete geometry ● Minor in Computer Science
Oct 2014 -Jun 2015	**University of Oxford**, Oxford, United Kingdom Study Abroad Program in Mathematics and Computer Science Grade: First-Class (equivalent to A/A+)

Publications

* papers whose authors are in alphabetical order or are with equal contributions

Conferences

1. * Zhiqi Bu, Sivakanth Gopi, Janardhan Kulkarni, Yin Tat Lee, Judy Hanwen Shen, and **Uthaipon Tantipongpipat**. *Fast and Memory Efficient Differentially Private-SGD via JL Projections.* Neural Information Processing Systems (NeurIPS), 2021
2. * Kyra Yee, **Uthaipon Tantipongpipat**, and Shubhanshu Mishra. *Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency.* Computer-Supported Cooperative Work and Social Computing (CSCW), 2021
3. **Uthaipon Tantipongpipat**, Chris Waites, Digvijay Boob, Amaresh (Ankit) Siva, and Rachel Cummings. *Differentially Private Mixed-Type Data Generation for Unsupervised Learning*. International Conference on Information, Intelligence, Systems and Applications (IISA), 2021
4. * Vivek Madan, Aleksandar Nikolov, Mohit Singh, and **Uthaipon Tantipongpipat**. *Maximizing Determinants under Matroid Constraints.* Symposium on Foundations of Computer Science (FOCS), 2020
5. **Uthaipon Tantipongpipat**, Samira Samadi, Mohit Singh, Jamie Morgenstern, and Santosh Vempala. *Multi-Criteria Dimensionality Reduction with Applications to Fairness.* Neural Information Processing Systems (NeurIPS), 2019, Spotlight (top 2.5% of submitted papers)
6. * Vivek Madan, Mohit Singh, **Uthaipon Tantipongpipat**, and Weijun Xie. *Combinatorial Algorithms for Optimal Design.* Conference on Learning Theory (COLT), pages 2210–2258, 2019
7. * Aleksandar Nikolov, Mohit Singh, and **Uthaipon Tantipongpipat**. *Proportional Volume Sampling and Approximation Algorithms for A-Optimal Design.* ACM-SIAM Symposium on Discrete Algorithms (SODA), 2019
8. Samira Samadi, **Uthaipon Tantipongpipat**, Jamie Morgenstern, Mohit Singh, and Santosh Vempala. *The Price of Fair PCA: One Extra Dimension.* Neural Information Processing Systems (NeurIPS), 2018

	9. * Rachel Cummings, Sara Krehbiel, Kevin A Lai, and **Uthaipon Tantipongpipat**. *Differential Privacy for Growing Databases.* Neural Information Processing Systems (NeurIPS), 2018
Manuscripts	1. **Uthaipon Tantipongpipat**. *λ-Regularized A-Optimal Design and its Approximation by λ-Regularized Proportional Volume Sampling.* 2020
Journals	1. Tomo Lazovich, Luca Belli, Aaron Gonzales, Amanda Bower, **Uthaipon Tantipongpipat,** Kristian Lum, Ferenc Huszar, Rumman Chowdhury. Measuring Disparate Outcomes of Content Recommendation Algorithms with Distributional Inequality Metrics. Patterns Journal. 2022. 2. * Aleksandar Nikolov, Mohit Singh, and **Uthaipon Tantipongpipat**. *Proportional Volume Sampling and Approximation Algorithms for A-Optimal Design.* Mathematics of Operation Research. 2022. 3. **Uthaipon Tantipongpipat**. *A Combinatorial Approach to Ebert's Hat Game with Many Colors*. The Electronic Journal of Combinatorics, 21(4):4–33, 2014
Workshops	1. * Digvijay Boob, Rachel Cummings, Dhamma Kimpara, **Uthaipon Tantipongpipat**, Chris Waites, and Kyle Zimmerman. *Differentially Private Synthetic Data Generation via GANs.* Theory and Practice of Differential Privacy (TPDP 2018) workshop, 2018
Theses	*1.* **Uthaipon Tantipongpipat.** *Fair and Diverse Data Representation in Machine Learning.* PhD Thesis, Georgia Institute of Technology, 2020 2. **Uthaipon Tantipongpipat**. *Cameron-Liebler Line Classes and Partial Difference Sets.* Undergraduate Thesis, University of Richmond, 2016

Talks and Presentations

1. *Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency*

- **a. Conference Presentation:** ACM Conference On Computer-Supported Cooperative Work And Social Computing, Virtual, October 2021

2. *Multi-Criteria Dimensionality Reduction with Applications to Fairness* (earlier version: Fair Dimensionality Reduction and Iterative Rounding for SDPs)

- a. **Invited talk**: Second Conference on Discrete Optimization and Machine Learning at RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan, July 2019
- b. **Invited talk**: Cornell Operations Research and Information Engineering (ORIE) workshop, Ithaca, NY, USA, October 2019
- c. **Invited talk**: INFORMS Annual Meeting, Seattle, WA, USA, October 2019
- d. **Spotlight and accepted for poster**: Conference on Neural Information Processing Systems (NeurIPS), Vancouver, Canada, December 2019

3. The Price of Fair PCA: One Extra Dimension

- a. **Accepted for poster**: Conference on Neural Information Processing Systems (NeurIPS), Montreal, Canada, December 2018

4. Proportional Volume Sampling and Approximation Algorithms for A-Optimal Design

- a. **Accepted paper presentation**: ACM-SIAM Symposium on Discrete Algorithms (SODA), San Diego, California, January 2019
- b. **Talk**: Machine learning theory group, Georgia Institute of Technology, November 2018

c. **Poster**: Machine Learning in Science and Engineering (MLSE) Conference, Carnegie Mellon University, June 2018
d. **Poster**: Algorithms and Randomness, Algorithms and Randomness Center (ARC) workshop, Georgia Institute of Technology, May 2018
e. **Talk**: Algorithms, Combinatorics, and Optimization (ACO) seminar, Georgia Institute of Technology, April 2018

5. *Differential Privacy for Growing Databases*
   a. **Accepted for poster**: Conference on Neural Information Processing Systems (NeurIPS), Montreal, Canada, December 2018
   b. **Talk**: Privacy reading group, Georgia Institute of Technology, February 2018
   c. **Accepted for poster**: Theory and Practice of Differential Privacy workshop (TPDP), Dallas, Texas, October 2017

Awards and Fellowships

2021	Impact Recognitions Award, CSCW (Conference On Computer-Supported Cooperative Work And Social Computing) ● "intended to recognize papers that contribute to a potentially significant impact in CSCW research, in practice, in design, in policy, or in the real world in substantive ways."
2020	ACO Outstanding Student Award (best PhD student in the year, awarded at the graduation) ● "[For] his overall research contributions, spanning all three components of ACO, including his work on algorithmic foundations of experimental design, fairness in machine learning algorithms, and differential privacy; his scientific leadership and collaborative attitude; and the quality and breadth of his research."
2019, 2020	Best Reviewers (top 10%) of NeurIPS (Conference on Neural Information Processing Systems). Awarded free registration
2018	First prize winner and People's Choice Awards ($20,000 total prize), Privacy Engineering Challenge, National Institute of Standards and Technology (NIST), Public Safety Communications Research Divisions (PSCR). https://www.herox.com/UnlinkableDataChallenge
2018	Algorithm and Randomness Center (ARC) and Transdisciplinary Research Institute for Advancing Data Science (TRIAD) Fellowship, Georgia Institute of Technology
2016	Finalist, ITA Tech Challenge programming competition, Illinois Technology Association, IL
2016	David C. Evans Awards for Outstanding Achievement in Scholarship, Annual Honors Convocation, University of Richmond, VA. ● Awarded to a few students each year for achievements in arts and sciences. In press: https://news.richmond.edu/features/article/-/13415/2016-david-c.-evans-awards-school-of-arts-and-sciences-recognizes-outstanding-achievement.html
2012-2016	Robins Science Scholar, University of Richmond (merit scholarship covering full tuition, fees, accommodations, and meals for four years)
2016	Phi Beta Kappa (most prestigious honor society for liberal arts and sciences)
2015	Honorable Mention (top 2.5%), William Lowell Putnam Mathematical Competition ● Widely considered to be the most prestigious undergraduate-level mathematics examination
2015	Second Place, Mid-Atlantic Regional ACM Programming Contest, Christopher Newport University site

Prior to Undergraduate Education:

2010-2012	3-Year Finalist, International Mathematical Olympiad (IMO) selection, Thailand

2012	Honorable Mention, Nern-Thong-Khong-Mee-Ka National Contest in Economics, Thailand
2011	Honorable Mention, Finance and Economics National Competition, National Bank of Thailand
2010, 2011	Bronze Medal and Honorable Mention, Asia-Pacific Mathematics Olympiad (APMO)
2008, 2009	Gold and Bronze Medals, IWYMIC International Mathematics Competition
2008, 2009	Two Gold Medals, Thailand Mathematical Olympiad

Codes

1. **Twitter Image Cropping.** In communication internally to push to open-source the code of Twitter's saliency-based cropping algorithm. Publicly available at https://github.com/twitter-research/image-crop-analysis
2. **Fair PCA project.** Semi-definite program and multiplicative weight heuristics for solving multi-criteria principle component analysis. In MATLAB and CVXOPT on Python. Publicly available at https://github.com/sdpforall (a website of the project is at https://sites.google.com/site/ssamadi/fair-pca-homepage).
3. **DPautoGAN.** Combining autoencoder and GAN to generate synthetic data with privacy protection guarantee. In Python and using Pytorch for neural networks. Publicly available at https://github.com/DPautoGAN/DPautoGAN.

Academic Service

2018-2021	Reviewer for NeurIPS (Conference on Neural Information Processing Systems), AAAI Conference on Artificial Intelligence, FOCS (Symposium on Foundations of Computer Science), MAPR (Mathematical Programming journal)
2018-2019	Co-organizer of ACO student seminar, Georgia Institute of Technology

Teaching Experience

2019	Teaching Assistant, CS7520/ISYE8813 Approximation Algorithms, Georgia Institute of Technology
2018	Teaching Assistant, CS6550 Graduate Algorithms, Georgia Institute of Technology
2015-2016	Language Partner, Self-Directed Language Acquisition Program, University of Richmond
2014	Grader, MATH245 Linear Algebra, University of Richmond
2011-2017	Teacher and tutor for middle- and high-school competitive mathematics, Bangkok Christian College, Bangkok, Thailand

Skills

Technical	Responsible AI, model audit / model governance, cross-functional communications, differential privacy, statistics. Python (Pandas, numpy, scipy), Jupyter and Colab notebooks, SQL (BigQuery), Git, CVXOPT, Java, C++, MATLAB, Mathematica, LaTeX, MS Word, MS Excel, MS PowerPoint
Communication	Public speaking – Toastmaster
Languages	Thai (native speaker); English (fluent)