Bokun Wang

📞 (+1) 319-512-3088   ✉ bokun.wang@utexas.edu

EXPERIENCE

Postdoctoral Researcher, hosted by Prof. Diana Marculescu   September 2025 -

Chandra Family Department of Electrical and Computer Engineering, The University of Texas at Austin

RESEARCH INTERESTS

Efficient Generative AI (Memory | Compute | Data)

EDUCATION

Texas A&M University, College Station, TX

Ph.D. in Computer Science   August 2025

Advisor: Prof. Tianbao Yang

University of California, Davis, Davis, CA

M.S. in Computer Science   June 2020

University of Electronic Science and Technology of China, Chengdu, Sichuan, China

B.S. in Computer Science (Yingcai Honors College Program)   June 2018

INTERNSHIP

Machine Learning Intern   May 2023 - August 2023

Central Technology / Machine Learning Group, ARM

- Proposed and implemented a new framework for efficient federated learning with quantized on-device training and communication, based on the new FP8 data type.
- 📑 *Supporting Publication:*
  ▷ **Bokun Wang**, et al. "Towards Federated Learning with on-device Training and Communication in 8-bit Floating Point." In *International Joint Workshop on Federated Learning for Data Mining and Graph Analytics (FedKDD), co-located with the ACM SIGKDD Conference*, 2024.

Research Intern advised by Prof. Peter Richtárik   Sept. 2020 - August 2021

OptML Lab, King Abdullah University of Science and Technology (KAUST)

- Proposed and implemented efficient algorithms for distributed ML with adaptive integer quantization.
- 📑 *Supporting Publication:*
  ▷ Konstantin Mishchenko, **Bokun Wang**, et al. "IntSGD: Adaptive Floatless Compression of Stochastic Gradients." In *International Conference on Learning Representations (ICLR)*, 2022 (**Spotlight**).

PROJECTS

Theory-driven Efficient Algorithms for (Multi-modal) Self-Supervised Representation Learning

- Developed efficient algorithms SOX and ALEXR with provable convergence guarantees, adopted by open-source projects including LibAUC, FastCLIP, and DRRho-CLIP.

- Proposed an algorithm based on discriminative probabilistic modeling and Monte Carlo integration for bi-modal self-supervised representation learning, offering improved generalization guarantees. It achieves superior overall performance across 8 downstream retrieval and zero-shot classification tasks compared to 5 representative baselines. Notably, it outperforms the widely used CLIP algorithm by over 15%.

- Implemented the project using Python and PyTorch, and open-sourced the code on GitHub.

- 📄 *Supporting Publication:*
  ▷ **Bokun Wang**, et al. "Finite-Sum Coupled Compositional Stochastic Optimization Theory and Applications." In *International Conference on Machine Learning (ICML)*, 2022;
  ▷ **Bokun Wang**, et al. "On Discriminative Probabilistic Modeling for Self-Supervised Representation Learning." In *International Conference on Learning Representations (ICLR)*, 2025;
  ▷ **Bokun Wang**, et al. "A Near-Optimal Single-Loop Algorithm for Convex Finite-Sum Coupled Compositional Stochastic Optimization." In *International Conference on Machine Learning (ICML)*, 2025.
  ▷ Vicente Balmaseda, **Bokun Wang**, et al. "Discovering Global False Negatives On the Fly for Self-supervised Contrastive Learning." In *International Conference on Machine Learning (ICML)*, 2025.

Efficient Algorithms for Imbalanced and Multi-Instance Medical Image Analysis

- Designed efficient, provably convergent algorithms for (partial) AUC maximization, tailored to the challenges of imbalanced medical datasets and high-resolution imaging data such as CT scans.

- Our proposed algorithms are GPU memory-efficient, unlike prior methods whose performance is limited by memory constraints (For example, they cannot handle large histopathological images).

- Demonstrated the effectiveness and superiority of our algorithms on benchmark datasets and real-world high-resolution medical image datasets.

- 📄 *Supporting Publication:*
  ▷ Dixian Zhu, **Bokun Wang**, et al. "Provable Multi-instance Deep AUC Maximization with Stochastic Pooling." In *International Conference on Machine Learning (ICML)*, 2023.
  ▷ Dixian Zhu, Gang Li, **Bokun Wang**, et al. "When AUC meets DRO: Optimizing Partial AUC for Deep Learning with Non-Convex Convergence Guarantee." In *International Conference on Machine Learning (ICML)*, 2022.

Graph Learning via Optimization-Based Approaches

- Proposed a technique called Feature Momentum (FM) to mitigate the neighborhood explosion and staleness problems in Graph Neural Network (GNN) training. Experiments on the Flickr, ogbn-arxiv, and Yelp datasets demonstrate the strong performance of the proposed method.

- Proposed a novel graph attack method, DGA, based on the continuous relaxation and parameterization of the graph structure. Compared to the state-of-the-art, DGA achieves nearly equivalent attack performance with 6 times less training time and 11 times smaller GPU memory footprint on different benchmark datasets.

- 📄 *Supporting Publication:*
  ▷ Haoran Liu, **Bokun Wang**, et al. "Everything Perturbed All at Once: Enabling Differentiable Graph Attacks." In *Companion Proceedings of the ACM Web (WWW) Conference*, 2024.
  ▷ Haiyang Yu, Limei Wang, **Bokun Wang**, et al. "GraphFM: Improving large-scale GNN training via feature momentum." In *International Conference on Machine Learning (ICML)*, 2022.

SKILLS

Languages and Libraries:	Python, PyTorch, Scikit-learn, NumPy, Pandas, Matplotlib
Tools:	Git, Docker, Slurm, LaTeX

HONORS & AWARDS

Notable Reviewer, International Conference on Learning Representations (ICLR), 2025

Oral Presentation, Conference on the Mathematical Theory of Deep Neural Networks (DeepMath), 2024

Scholar Award, 36th Conference on Neural Information Processing Systems (NeurIPS), 2022

Travel Award, 39th International Conference on Machine Learning (ICML), 2022

Spotlight Presentation, 10th International Conference on Learning Representations (ICLR), 2022

Outstanding Graduate, University of Electronic Science and Technology of China (UESTC), 2018

National Scholarship, Ministry of Education of the People's Republic of China, 2016

People's Scholarship (Special Class), University of Electronic Science and Technology of China (UESTC), 2015