MINGJU LIU

(908) 720-5062 ⋄ mliu9867@umd.edu

EDUCATION

University of Maryland, College Park *2023-present*
Ph.D. student (transferred), Computer Engineering (Advisor: Prof. Cunxi Yu)
Research interests: Hardware-software Co-design, ML for EDA, CAD for AI

University of Utah *2022-2023*
Ph.D. student, Computer Engineering (Advisor: Prof. Cunxi Yu)

Rutgers, The State University of New Jersey *2019-2022*
M.S., Electrical and Computer Engineering (GPA: 3.91/4.0)

University of Electronic Science and Technology of China *2016-2020*
B.S., Information and Communication Engineering

ACADEMIC EXPERIENCE

NB-IoT Lab *Summer 2018*
James Cook University, Cairns, Australia

AWARDS

DAC Young Fellow, 2023

Honor Student of School of Information and Communication Engineering, 2019

UESTC University Scholarship, 2017, 2018, 2019

PUBLICATIONS

2024

- **Mingju Liu**, Daniel Robinson, Yingjie Li, Cunxi Yu. *MapTune: Advancing ASIC Technology Mapping via Reinforcement Learning Guided Library Tuning* ACM/IEEE International Conference on Computer-Aided Design (*ICCAD'24*)
- **Mingju Liu**, Yingjie Li, Jiaqi Yin, Zhiru Zhang, Cunxi Yu. *Differentiable Combinatorial Scheduling at Scale.* International Conference on Machine Learning (*ICML'24*)
- Yingjie Li, **Mingju Liu**, Alan Mischenko, Mark Ren, Cunxi Yu. *DAG-aware Synthesis Orchestration.* IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (*TCAD'24*)

2023

- Yingjie Li, **Mingju Liu**, Alan Mishchenko, Cunxi Yu. *Verilog-to-PyG – A Framework for Graph Learning and Augmentation on RTL Designs.* ACM/IEEE International Conference on Computer-Aided Design (*ICCAD'23*)

2021

- Mingshuo Liu, Miao Yin, Kevin Han, **Mingju Liu**, Bo Yuan, Yu Bai. *Algorithm and Hardware Co-Design Co-Optimization Framework for LSTM Accelerator using Fully Decomposed Tensor Train.* Design Automation Conference (*DAC'21 WIP*)

PROJECTS

- **High-performance Algorithm Based on Nonconvex Optimization for Large-scale Tasks Scheduling** The ongoing research project focuses on investigating the potential of Gumbel-Softmax (GS) in developing a neural network-based framework to solve large-scale scheduling problems. The project aims to utilize GPU resources to achieve optimal Pareto Front of the runtime and the generation of the near-optimal solutions that can surpass the performance of state-of-the-art (SOTA) methods such as Heuristic based List-Scheduling (LS), Force-Directed-Scheduling (FDS) and Integer Linear Programming (ILP) solved by commercial tools CPLEX and Gurobi. This endeavor seeks to contribute to the advancement of efficient and effective solutions for complex scheduling tasks by integrating machine learning techniques.
  Publications: *ICML'24*
- **Algorithmic Exploration and Optimization of Fundamental Logic/Boolean Optimization Problems for Logic Synthesis** This research proposes a novel approach DAG-aware synthesis orchestration to enhancing both the quality-of-results and runtime in logic synthesis. This approach introduces new algorithmic concepts that strive to optimize the synthesis process by orchestrating multiple optimization methods in a single traversal. Our experimental findings demonstrate that current state-of-the-art (SOTA) DAG-aware logic synthesis algorithms tend to follow a common domain-specific design approach, where transformations are executed individually in a single DAG-aware traversal. In contrast, our proposed logic synthesis orchestration technique takes advantage of the interplay between optimization methods and integrates them on-the-fly in a single traversal, leading to improved efficiency and effectiveness in logic synthesis.
  Publications: *DAC'23 (WIP)*, *IWLS'23*, *TCAD'24*
- **Algorithm and Hardware Co-Design Optimization Framework for LSTM Accelerator Using Fully Decomposed Tensor Train** The Tensor Train (TT) decomposition is an emerging technology that holds great potential for compressing large deep neural network models, making them more suitable for deployment on resource-constrained edge devices. In this study, we introduce an effective hardware optimization methodology that utilizes a novel software solution to enable the application of TT-decomposition on the complete Long Short-Term Memory (LSTM) model. Additionally, we propose an efficient hardware accelerator that employs a hardware and algorithm co-design approach. The integration of our software and hardware solutions provides a practical approach to compressing intricate LSTM models and deploying them on resource-limited edge devices. This approach opens up the possibility for the development of more energy-efficient and high-performance edge computing systems.
  Publications: *DAC'21 (WIP)*
- **L2S - Large Scale MIMO Radar Antenna Selection Problem** The aim of this project is to implement Learn to Select (L2S), a machine learning-based method, to select antennas from a dense Uniform Linear Array (ULA) by utilizing a combination of multiple softmax layers subject to an orthogonalization criterion. The objective is to minimize energy costs while maintaining consistent communication rates and performance. This approach enables efficient scaling to handle larger problems by avoiding combinatorial explosion in the selection process. The project demonstrates potential in presenting a framework that addresses the selection problem, as it can be readily formulated for any metric.

TEACHING EXPERIENCE

Virtual Reality & Technology, Teaching Assistant, Spring'22, Spring'21, Spring'20

Computer Architecture, Teaching Assistant, Spring'18

SKILLS

Programming skills: Python, BASH, LaTeX, C/C++, Matlab

Platforms: Linux (Ubuntu)