Qinyi Luo

3740 McClintock Ave., EEB-201, Los Angeles, CA 90089, USA
+1 2132844080 • qinyiluo@usc.edu • Website

Education

Ph.D. in Computer Science
University of Southern California, Los Angeles, CA, USA — Aug. 2017 – May 2024
Advisor: Prof. Xuehai Qian
Thesis: High-performance heterogeneity-aware distributed training of machine learning models

B.Eng. in Electronic Engineering — June 2017
Tsinghua University, Beijing, China — Aug. 2013 – Jun. 2017
GPA: 92/100 Rank: 5/240

Professional Skills

Programming (C/C++, Python, CUDA, cuDNN, Pytorch, Tensorflow, Pthreads, OpenMP, MPI, Matlab, Verilog and Assembly), Parallel and Distributed Computation, Machine Learning, Data Analysis, Signal Processing

Publications (citations)

[1] Youwei Zhuo, Jingji Chen, Gengyu Rao, **Qinyi Luo**, Yanzhi Wang, Hailong Yang, Depei Qian, and Xuehai Qian. 2021. Distributed Graph Processing System and Processing-in-memory Architecture with Precise Loop-carried Dependency Guarantee. ACM Trans. Comput. Syst. 37, 1–4, Article 5 (June 2021).

[2] Youwei Zhuo, Jingji Chen, **Qinyi Luo**, Yanzhi Wang, Hailong Yang, Depei Qian, and Xuehai Qian. SympleGraph: Distributed graph processing with precise loop-carried dependency guarantee. *PLDI 2020.*

[3] **Qinyi Luo**, Jiaao He, Youwei Zhuo, and Xuehai Qian. Prague: High-performance heterogeneity-aware asynchronous decentralized training. *ASPLOS 2020*.

[4] **Qinyi Luo**, Jinkun Lin, Youwei Zhuo, and Xuehai Qian. HOP: Heterogeneity-aware decentralized training. *ASPLOS 2019*.

[5] Youwei Zhuo, Jinglei Cheng, **Qinyi Luo**, Jidong Zhai, Yanzhi Wang, Zhongzhi Luan, and Xuehai Qian. CSE: Parallel Finite State Machines with convergence set enumeration. *MICRO 2018.*

[6] Yirong Lv, Bin Sun, **Qinyi Luo**, Zhibin Yu, and Xuehai Qian. CounterMiner: Mining big performance data from hardware counters. *MICRO 2018*.

[7] **Qinyi Luo**, Rahul Gupta, and Shrikanth S. Narayanan. Transfer learning between concepts for human behavior modeling: An application to sincerity and deception prediction. *Interspeech 2017.*

Work Experience

Research Scientist Intern, Meta Platforms — May 2022 – Aug. 2022
Mentor: Wei Zhang. Manager: Yuxi Hu.

Research Intern, NVIDIA Corporation — May 2020 – Aug. 2020
Mentors: Eiman Ebrahimi, Isaac Gelado. Manager: David Nellans.

Awards and Honors

Meta PhD Research Fellowship *(top 1.6%)* — 2022
2021-2022 MHI Ph.D. Scholar (*6 students in ECE dept., USC*) — 2021

Facebook Fellowship Program Finalist *(top 3.5%)* 2021
WiSE Qualcomm Top-Off Fellowship 2020
8th Heidelberg Laureate Forum attendee *(224 young researchers worldwide)* 2020
Microsoft Research PhD Fellowship nominee *(3 students in CS dept., USC)* 2019
Microsoft Ada Lovelace Fellowship nominee *(3 students in CS dept., USC)* 2018
WeTech Qualcomm Global Scholars Award *(18 winners in China)* 2016
National Scholarship of China *(top 2% at Tsinghua University)* 2015
Provincial First Prize in China Mathematical Competition *(ranked 33rd in Jiangsu Province)* 2012

Teaching and Mentoring

Mentor for intern student(s) at my research lab Summer 2018, 2019, 2021, 2023; Fall 2019
Mentor for a new Master's student in the USC WiE MENTOR Program Fall 2022 & Spring 2023
TA for *Parallel and Distributed Computation* (EE-451) at USC Spring 2019, Spring 2020 & Spring 2021
Mentor for a new Ph.D. student in the USC WiSE Mentorship Program Fall 2020
TA for *Discrete Methods in Computer Science* (CSCI-170) at USC Fall 2020
TA for *Introduction to Computer Systems* (CSCI-356) at USC Fall 2018
Volunteer teaching at Qinghua Primary School, Dali, Yunnan, China Summer 2015

Research Projects

Automatic generation of hybrid strategies for distributed training of machine learning models
Advisor: Prof. Xuehai Qian, USC Apr. 2021 – present

- Devised a unified representation of data parallelism, model parallelism, pipeline parallelism and their hybrids
- Explored the benefits of fine-grained DNN partitioning and novel operation scheduling methods
- Derived the ideal stage partitioning ratio for synchronous pipelines using heterogeneous machines
- Applied a multi-level search algorithm to select parallelization strategies with superior performance

Distributed graph processing with precise loop-carried dependency guarantee
Advisor: Prof. Xuehai Qian, USC Nov. 2018 – Mar. 2021
* Collaborative project with senior Ph.D. students

- Removed redundant computation in existing frameworks by enforcing loop-carried dependency
- Proposed circulant scheduling and double buffering to boost performance
- Outperformed Gemini / D-Galois by 1.42x/3.30x on average and up to 2.30x/7.76x

Straggler mitigation in decentralized training via system-algorithm co-design
Advisor: Prof. Xuehai Qian, USC Sept. 2018 – Dec. 2019

- Proposed Partial All-Reduce that enables fast and randomized group-based synchronization
- Designed smart group generation strategies to avoid synchronization conflict
- Achieved 4.4x/1.2x speedup over Ring All-Reduce when stragglers are present / absent

Heterogeneity-aware decentralized machine learning training system
Advisor: Prof. Xuehai Qian, USC Aug. 2017 – Aug. 2018

- Identified a unique property of decentralized training, iteration gap, and analyzed its upper-bound
- Introduced queue-based synchronization to enable fast execution with controlled iteration gap
- Developed decentralized protocols for backup workers and bounded staleness to deal with transient stragglers; proposed skipping iterations to mitigate persistent stragglers
- Implemented the system upon TensorFlow; experiments showed 1.4-2.8x speedup using the proposed straggler-mitigation methods

Parallel finite-state machines with convergence set enumeration

Advisor: Prof. Xuehai Qian, USC Aug. 2017 – Jun. 2018

* Collaborative project with senior Ph.D. students

- Proposed state set → state set mapping to enable efficient enumerative FSM
- Developed a convergence set prediction algorithm and re-execution mechanisms to deal with inaccurate predictions
- Obtained on average 2.0x/2.4x speedup and maximum 8.6x/2.7x speedup across 13 benchmarks over Lookback Enumeration / Parallel Automate Processor

Mining big performance data from hardware counters

Advisor: Prof. Xuehai Qian, USC Aug. 2017 – Jun. 2018

* Collaborative project with senior Ph.D. students

- Proposed rigorous data mining methodology to make full use of performance data using data cleaning, event ranking, pruning and interaction analysis
- Performed experiments using CloudSuite, Spark version of HiBench and a real-world Spark application
- Showed significant error reduction and capability of identifying important configuration parameters

Data mining for indoor air quality monitoring

Advisor: Prof. Pingyi Fan, Tsinghua University Nov. 2016 – Jun. 2017

- Applied FEM-based clustering to indoor air quality data analysis
- Found several clusters with interpretable physical meanings
- Improved the spatial and temporal performance with dynamic programming

Transfer learning between sincerity and deception prediction

Advisor: Prof. Shrikanth Narayanan, USC Jun. 2016 – Aug. 2016

- Applied transfer learning to human behavioral trait modeling
- Proposed a stacked generalization framework for predicting sincerity and deception, using GFK transformation to ameliorate differences between in-domain and out-of-domain data characteristics
- Achieved significantly better results than baseline models trained using in-domain data only