

Yuduo Wu

✉ Email: yuduow@gmail.com

🏠 Homepage: <http://www.yuduowu.com>

🌐 LinkedIn: <https://www.linkedin.com/in/yuduowu>

🔍 Google Scholar: <https://scholar.google.com/citations?user=fihuycMAAAAJ>

EXPERIENCES

NovuMind Inc., HPC & AI Infrastructure Team **Santa Clara, California**
Staff Software Engineer **2018 -**
Senior Software Engineer **2017 - 2018**

Built high-performance machine learning platform to accelerate AI training/inference in production;
Scaled distributed deep learning trainer, matched world record (ImageNet in 1 hour) with 256 GPUs;
Optimized, containerized deep learning workloads; Enabled one-key deployment on GPU clusters;
Implemented softwares and AI models for medical and health-care, tumor detections 97% accuracy;
Platform deployed at Hewlett Packard Enterprise (HPE) AI Innovation Center, top-3 best hospitals
in China and other partners; Work recognized/presented at Spark+AI Summit '18, GTC Europe '17.

IBM, Watson & Cloud Platform **San Jose, California**
Software Engineer **2015 - 2017**

Prototyped, implemented and deployed multiple Cloud Applications/Web Services for IBM Cloud
and SoftLayer; Developed health and monitor mechanisms for service reliability and data analytics.

University of California, Davis, Dept. of Electrical and Computer Engineering **Davis, California**
Graduate Student Researcher **2013 - 2015**

Initial contributor of an open source project (Gunrock) for GPU graph analytics (350+ GitHub stars);
Implemented Boruvka's Minimum Spanning Tree (MST) and other graph algorithms on GPU and
achieved up to 120× speedups against a popular serial C++ implementation (Boost Graph Library);
Benchmarked and analyzed high-level programming models for parallel graph processing libraries.

DARPA, Defense Advanced Research Projects Agency **Arlington, Virginia**
Research Internship **Summer 2014, Summer 2015**

Collaborated with multiple research groups, universities and companies for big-data analytics chal-
lenges; Designed and created pure C and Python interfaces for the Gunrock graph analytics library.

EDUCATION

University of California, Davis **Davis, California**
Master of Science, Electrical and Computer Engineering **2013 - 2015**

Advisor: Professor John D. Owens

Research: Parallel Computing, GPU Computing, Graph Analytics. Thesis:

"Performance Characterization of High-Level Programming Models for GPU Graph Analytics"

Macau University of Science and Technology **Taipa, Macau**
Bachelor of Science, Electronic Information Technology **2009 - 2013**

TECHNOLOGY SKILLS

Programing: C/C++, Python, Shell/Bash, Java, GoLang, R, CUDA, JavaScript
 Toolkits: Unix/Linux, Git, L^AT_EX, MySQL, MPI, Caffe, TensorFlow, Keras, PyTorch

AWARDS & HONERS

Excellence Achievement Award , NovuMind Inc.	2017
IBM Manager's Choice Award , IBM	2015, 2016
Distinguished Paper Award , Principles and Practice of Parallel Programming (PPoPP'16)	2016
IBM Solutions Excellence Award (EA) , IBM	2015
Best Paper Finalist , IEEE International Symposium on Workload Characterization (IISWC'15)	2015
Dean's Honor List Scholarship , Macau University of Science and Technology	2011, 2012, 2013
Crystal Cup Award , Top 3 in GPA, Macau University of Science and Technology	2011, 2012, 2013
Nam Kwong Academic Year Scholarship , Nam Kwong (Group) Co., Ltd, Macau	2012

PROFESSIONAL SERVICES

Program Committee/Reviewer: Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS'18), held as part of ACM/IEEE Supercomputing (SC'18) **2018**

TALKS & PUBLICATIONS

Rui Liu, **Yuduo Wu**. Talk: "High Performance Deep Learning with Apache Spark". In *Spark + AI Summit '18 | Artificial Intelligence & Apache Spark Conference*, June 4-6 2018.

Yuechao Pan, Yangzihao Wang, **Yuduo Wu**, Carl Yang, and John D. Owens. "Multi-GPU Graph Analytics". In *Proceedings of the 31st IEEE International Parallel and Distributed Processing Symposium, IPDPS '17*, May/June 2017.

Yangzihao Wang, Yuechao Pan, Andrew Davidson, **Yuduo Wu**, Carl Yang, Leyuan Wang, Muhammad Osama, Chenshan Yuan, Weitang Liu, Andy T. Riffel, and John D. Owens. "Gunrock: GPU Graph Analytics". *invited paper to ACM Transactions on Parallel Computing (TOPC), an extended version of PPoPP'16 paper "Gunrock: A High-Performance Graph Processing Library on the GPU"*, January 2017.

Yangzihao Wang, Andrew Davidson, Yuechao Pan, **Yuduo Wu**, Andy Riffel, and John D. Owens. "Gunrock: A High-Performance Graph Processing Library on the GPU". In *Proceedings of the 21st ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPoPP'16*, pages 11:1-11:12. March 2016. **Distinguished Paper**.

Yuduo Wu, Yangzihao Wang, Yuechao Pan, Carl Yang, and John D. Owens. "Performance Characterization of High-Level Programming Models for GPU Graph Analytics". In *IEEE International Symposium on Workload Characterization, IISWC'15*, pages 66-75. October 2015. **Best Paper finalist**.

Yuduo Wu. "Performance Characterization of High-Level Programming Models for GPU Graph Analytics". *ProQuest (Master's thesis at University of California, Davis)*, September 2015.