Wei Wang

PRESENT APPOINTMENT	Associate Professor

<table>
<tr><td>**PRESENT APPOINTMENT**</td><td>Associate Professor
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong</td><td>*Office*: Room 3524
Phone: +852 2358 6972
Email: weiwa@cse.ust.hk
Web: www.cse.ust.hk/~weiwa</td></tr>
<tr><td>**RESEARCH INTERESTS**</td><td colspan="2">Cloud computing, distributed systems, computer network</td></tr>
</table>

EDUCATION

University of Toronto, Toronto, Ontario, Canada
Department of Electrical and Computer Engineering

⬦ **Ph.D.**, Electrical and Computer Engineering, August 2015
  ▷ *Dissertation:* "Fair Scheduling in Cloud Datacenters with Multiple Resource Types"
  ▷ *Advisors:* Prof. Baochun Li and Prof. Ben Liang
  ▷ *Thesis Committee:* Jorg Liebeherr, Ding Yuan, Rong Zheng (McMaster University)

Shanghai Jiao Tong University, Shanghai, China
Department of Electrical and Computer Engineering

⬦ **M.Eng.**, Information Engineering, March 2010
  ▷ *Thesis:* "On Selfish Spectrum Sensing Policy and Congestion Games in Cognitive Radio Network"
  ▷ *Advisor:* Chen He

Shanghai Jiao Tong University, Shanghai, China
Department of Electrical and Computer Engineering

⬦ **B.Eng.** (*SJTU Distinguished Graduate Award*), Information Engineering, June 2007

HONORS AND AWARDS

⬦ WeBank Scholar (1 out 5 in China)	2021
⬦ Best Paper Runner Up, IEEE ICDCS 2021	2021
⬦ Distinguished TPC Member of IEEE INFOCOM	2018 – 2020, 2022
⬦ Chinese Government Award for Outstanding Students Abroad	2014 – 2015
⬦ Best Paper Runner Up, USENIX ICAC 2013	2013
⬦ University of Toronto Fellowship	2010 – 2014
⬦ University Distinguished Graduate (2%), Shanghai Jiao Tong University	2007

PROFESSIONAL EXPERIENCE

Hong Kong University of Science and Technology, Department of Computer Science and Engineering, Clear Water Bay, Kowloon, Hong Kong, China

Associate Professor (with tenure)	July 2021 – Present
Assistant Professor (tenure-track)	September 2015 – June 2021

University of Toronto, Department of Electrical and Computer Engineering, Toronto, Ontario, Canada

Research/Teaching Assistant	September 2010 – August 2015

Microsoft Server and Tools Business, Shanghai, China

Software Development Engineer (intern)
  ⬦ Placement and routing algorithms for Microsoft WorkFlow 4.0 — March – June, 2010
  ⬦ Deployment of the cloud identity service for BizTalk — June – September, 2008

Microsoft Research Asia, Beijing, China

Research Assistant, Internet Media Group   March – June 2007

<small>SCHOLARLY PUBLICATIONS</small>   In the areas of computer systems and networking, publications in top conferences are usually considered more significant than in journals. All publications are sorted in a reverse chronological order, where the underlined authors are HKUST students or research assistants under my supervision.

Refereed Journal Articles

[J13] Zhifeng Jiang, **Wei Wang**, Bo Li, Qiang Yang, "Towards Efficient Synchronous Federated Training: A Survey on System Optimization Strategies," accepted to appear in *IEEE Transactions on Big Data*, May 2022.

[J12] Jun Li, Wei Song, Yongbin Gao, Huixing Wang, Yier Yan, Bo Huang, Jun Zhang, **Wei Wang**, "Monocular 3D Object Detection Based on Depth Guided Local Convolution for Smart Payment in D2D systems," accepted to appear in *IEEE Internet of Things Journal*, November 2021.

[J11] Chen Chen, Qizhen Weng, **Wei Wang**, Baochun Li, Bo Li, "Accelerating Distributed Learning in Non-Dedicated Environments," accepted to appear in *IEEE Transactions on Cloud Computing*, July 2021.

[J10] Mingzhe Li, **Wei Wang**, Jin Zhang, Qian Zhang, "Incentivizing WiFi-based Multilateration Location Verification," accepted to appear in *IEEE Internet of Things Journal*, June 2021.

[J9] Mingzhe Li, Jingrou Wu, **Wei Wang**, Jin Zhang, "Towards Privacy-Preserving Task Assignment for Fully Distributed Spatial Crowdsourcing," accepted to appear in *IEEE Internet of Things Journal*, March 2021.

[J8] Shaohuai Shi, Zhenheng Tang, Xiaowen Chu, Chengjian Liu, **Wei Wang**, Bo Li, "A Quantitative Survey of Communication Optimizations in Distributed Deep Learning," accepted to appear in *IEEE Network Magazine*, October 2020.

[J7] Chengliang Zhang, Minchen Yu, **Wei Wang**, Feng Yan, "Enabling Cost-Effective, SLO-Aware Machine Learning Inference Serving on Public Cloud," accepted to appear in *IEEE Transactions on Cloud Computing*, June 2020.

[J6] Yinghao Yu, **Wei Wang**, Renfei Huang, Jun Zhang, Khaled Ben Letaief, "Achieving Load-Balanced, Redundancy-Free Cluster Caching with Selective Partition," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 2, pp. 439–454, February 2020.

[J5] Yinghao Yu, Chengliang Zhang, **Wei Wang**, Jun Zhang, Khaled Ben Letaief, "Towards Dependency-Aware Cache Management for Data Analytics Applications," accepted to appear in *IEEE Transactions on Cloud Computing*, September 2019.

[J4] **Wei Wang**, Ben Liang, Baochun Li, "Optimal Online Multi-Instance Acquisition in IaaS Clouds," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 12, pp. 3407–3419, December 2015.

[J3] **Wei Wang**, Ben Liang, Baochun Li, "Multi-Resource Fair Allocation in Heterogeneous Cloud Computing Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 10, pp. 2822–2835, October, 2015.

[J2] **Wei Wang**, Di Niu, Ben Liang, Baochun Li, "Dynamic Cloud Resource Reservation via IaaS Cloud Brokerage," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 6, pp. 1580–1593, June, 2015.

[J1] **Wei Wang**, Ben Liang, Baochun Li, "Designing Truthful Spectrum Double Auctions with Local Markets," *IEEE Transactions on Mobile Computing*, vol. 13, no. 1, pp. 75–88, January 2014.

Refereed Papers in Conference and Workshop Proceedings

[C57] Minchen Yu, Tingjia Cao, **Wei Wang**, Ruichuan Chen, "Following the Data, Not the Function: Rethinking Function Orchestration in Serverless Computing," in the *Proceedings of the 20th USENIX Symposium on Networked Systems Design and Implementation (NSDI '23)*, Boston, MA, USA, April 2023 (acceptance ratio: 18%).

[C56] Mingzhe Li, You Lin, Jin Zhang, **Wei Wang**, "Jenga: Orchestrating Smart Contracts in Sharding-Based Blockchain for Efficient Processing," in the *Proceedings of the 42nd IEEE International Conference on Distributed Computing Systems (ICDCS '22)*, Bologna, Italy, July 2022 (acceptance ratio: 20%).

[C55] Da Yan, **Wei Wang**, Xiaowen Chu, "POSTER: An LLVM-based Open-Source Compiler for NVIDIA GPUs," poster paper in the *Proceedings of ACM Symposium on Principles and Practice of Parallel Programming (PPoPP '22)*, Virtual Conference, April 2022.

[C54] Qizhen Weng, Wencong Xiao, Yinghao Yu, **Wei Wang**, Cheng Wang, Jian He, Yong Li, Liping Zhang, Wei Lin, Yu Ding, "MLaaS in the Wild: Workload Analysis and Scheduling in Large-Scale Heterogeneous GPU Clusters," in the *Proceedings of the 19th USENIX Symposium on Networked Systems Design and Implementation (NSDI '22)*, Renton, WA, USA, April 2022 (acceptance ratio: 20%).

[C53] Chengliang Zhang, Junzhe Xia, Baichen Yang, Huancheng Puyang, **Wei Wang**, Ruichuan Chen, Istemi Ekin Akkus, Paarijaat Aditya, Feng Yan, "Citadel: Protecting Data Privacy and Model Confidentiality for Collaborative Learning," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '21)*, Seattle, WA, USA, November 2021.

[C52] Luping Wang*, Lingyun Yang*, Yinghao Yu, **Wei Wang**, Bo Li, Xianchao Sun, Jian He, Liping Zhang, "Morphling: Fast, Near-Optimal Auto-Configuration for Cloud-Native Model Serving," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '21)*, Seattle, WA, November 2021 (*equal contribution).

[C51] Suyi Li*, Luping Wang*, **Wei Wang**, Yinghao Yu, Bo Li, "George: Learning to Place Long-Lived Containers in Large Clusters with Operation Constraints," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '21)*, Seattle, WA, November 2021 (*equal contribution).

[C50] Huangshi Tian, Minchen Yu, **Wei Wang**, "CrystalPerf: Learning to Characterize the Performance of Dataflow Computation through Code Analysis," in the *Proceedings of USENIX Annual Technical Conference (ATC '21)*, Virtual Conference, July 2021 (acceptance ratio: 19%).

[C49] Chen Chen, Hong Xu, **Wei Wang**, Baochun Li, Bo Li, Li Chen, Gong Zhang, "Communication-Efficient Federated Learning with Adaptive Parameter Freezing," in the *Proceedings of the 41st IEEE International Conference on Distributed Computing Systems (ICDCS '21)*, Virtual Conference, July 2021 (acceptance ratio: 20%).

[C48] Minchen Yu, Zhifeng Jiang, Hok Chun Ng, **Wei Wang**, Ruichuan Chen, Bo Li, "Gillis:

Serving Large Neural Networks in Serverless Functions with Automatic Model Partitioning," in the *Proceedings of the 41st IEEE International Conference on Distributed Computing Systems (ICDCS '21)*, Virtual Conference, July 2021 (acceptance ratio: 20%; **Best Paper Runner Up**: 3 out of 489 submissions).

[C47] Da Yan, **Wei Wang**, Xiaowen Chu, "POSTER: Simplifying Low-Level GPU Programming with GAS," poster paper in the *Proceedings of ACM Symposium on Principles and Practice of Parallel Programming (PPoPP '21)*, Virtual Conference, February 2021.

[C46] Luping Wang*, Qizhen Weng*, **Wei Wang**, Chen Chen, Bo Li, "Metis: Learning to Schedule Long-Running Applications in Shared Container Clusters at Scale," in the *Proceedings of IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC20)*, Virtual Conference, November 2020 (*equal contribution; acceptance ratio: 18%).

[C45] Chen Chen, Qizhen Weng, **Wei Wang**, Baochun Li, Bo Li, "Semi-Dynamic Load Balancing: Efficient Distributed Learning in Non-Dedicated Environments," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '20)*, Virtual Conference, October 2020 (acceptance ratio: 24%).

[C44] Chengliang Zhang, Suyi Li, Junzhe Xia, **Wei Wang**, Feng Yan, Yang Liu, "BatchCrypt: Efficient Homomorphic Encryption for Cross-Silo Federated Learning," in the *Proceedings of USENIX Annual Technical Conference (ATC '20)*, Virtual Conference, July 2020 (acceptance ratio: 19%).

[C43] Jun Yi, Chengliang Zhang, **Wei Wang**, Cheng Li, Feng Yan, "Not All Explorations are Equal: Harnessing Heterogeneous Profiling Cost for Efficient MLaaS Training," in the *Proceedings of IEEE International Parallel and Distributed Processing Symposium (IPDPS '20)*, Virtual Conference, May 2020 (acceptance ratio: 25%).

[C42] Da Yan, **Wei Wang**, Xiaowen Chu, "Demystifying Tensor Cores to Optimize Half-Precision Matrix Multiply," in the *Proceedings of IEEE International Parallel and Distributed Processing Symposium (IPDPS '20)*, Virtual Conference, May 2020 (acceptance ratio: 25%).

[C41] Minchen Yu, Yinghao Yu, Yunchuan Zheng, Baichen Yang, **Wei Wang**, "RepBun: Load-Balanced, Shuffle-Free Cluster Caching for Structured Data," in the *Proceedings of IEEE INFOCOM '20*, Virtual Conference, April 2020 (acceptance ratio: 20%).

[C40] Da Yan, **Wei Wang**, Xiaowen Chu, "Optimizing Batched Winograd Convolution on GPUs," in the *Proceedings of ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP '20)*, San Diego, CA, USA, February 2020 (acceptance ratio: 23%).

[C39] Suyi Li, Yong Cheng, Yang Liu, **Wei Wang**, "Abnormal Client Behavior Detection in Federated Learning," in the *Proceedings of the 2nd International Workshop on Federated Learning for Data Privacy and Confidentiality*, in Conjunction with *NeurIPS 2019 (FL-NeurIPS '19)*, Vancouver, BC, Canada, December 2019.

[C38] Huangshi Tian, Yunchuan Zheng, **Wei Wang**, "Characterizing and Synthesizing Task Dependencies of Data-Parallel Jobs in Alibaba Cloud," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '19)*, Santa Cruz, CA, USA, November 2019 (acceptance ratio: 23%).

[C37] Huangshi Tian, Qizhen Weng, **Wei Wang**, "Towards Framework-Independent, Non-Intrusive Performance Characterization for Dataflow Computation," in the *Proceedings of ACM*

SIGOPS Asia-Pacific Workshop on Systems (APSys '19), Hangzhou, China, August 2019.

[C36] Chengliang Zhang, Minchen Yu, **Wei Wang**, Feng Yan, "MArk: Exploiting Cloud Services for Cost-Effective, SLO-Aware Machine Learning Inference Serving," in the *Proceedings of USENIX Annual Technical Conference (ATC '19)*, Renton, WA, USA, July 2019 (acceptance ratio: 20%).

[C35] Luping Wang, **Wei Wang**, Bo Li, "CMFL: Mitigating Communication Overhead for Federated Learning," in the *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS '19)*, Dallas, TX, USA, July 2019 (acceptance ratio: 20%).

[C34] Yinghao Yu, **Wei Wang**, Jun Zhang, Khaled Ben Letaief, "LACS: Load-Aware Cache Sharing with Isolation Guarantee," in the *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS '19)*, Dallas, TX, USA, July 2019 (acceptance ratio: 20%).

[C33] Chen Chen, **Wei Wang**, Bo Li, "Round-Robin Synchronization: Mitigating Communication Bottlenecks in Parameter Servers," in the *Proceedings of IEEE INFOCOM '19*, Paris, France, April-May 2019 (acceptance ratio: 20%).

[C32] Yuechen Tao, Jingjie Jiang, Shiyao Ma, Luping Wang, **Wei Wang**, Bo Li, "Unraveling the RTT-fairness Problem for BBR: A Queueing Model," in the *Proceedings of IEEE GLOBECOM '18*, Abu Dhabi, UAE, December 2018.

[C31] Mingzhe Li, Jin Zhang, **Wei Wang**, "Task Selection and Scheduling for Food Delivery: A Game-theoretic Approach," in the *Proceedings of IEEE GLOBECOM '18*, Abu Dhabi, UAE, December 2018.

[C30] Yinghao Yu, Renfei Huang, **Wei Wang**, Jun Zhang, Khaled Ben Letaief, "SP-Cache: Load-balanced, Redundancy-free Cluster Caching with Selective Partition", in the *Proceedings of IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC18)*, Dallas, TX, USA, November 2018 (acceptance ratio: 19%).

[C29] Huangshi Tian, Minchen Yu, **Wei Wang**, "Continuum: A Platform for Cost-Aware, Low-Latency Continual Learning," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '18)*, Carlsbad, CA, USA, October 2018 (acceptance ratio: 24%).

[C28] Chen Chen, Qizhen Weng, **Wei Wang**, Baochun Li, Bo Li, "Poster: Fast Distributed Deep Learning via Worker-adaptive Batch Sizing," poster paper in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '18)*, Carlsbad, CA, USA, October 2018.

[C27] Chengliang Zhang, Huangshi Tian, **Wei Wang**, Feng Yan, "Stay Fresh: Speculative Synchronization for Fast Distributed Machine Learning," in the *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS '18)*, Vienna, Austria, July 2018 (acceptance ratio: 20%).

[C26] Yinghao Yu, **Wei Wang**, Jun Zhang, Qizhen Weng, Khaled Ben Letaief, "OpuS: Fair and Efficient Cache Sharing for In-Memory Data Analytics," in the *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS '18)*, Vienna, Austria, July 2018 (acceptance ratio: 20%).

[C25] Luping Wang, **Wei Wang**, "Fair Coflow Scheduling without Prior Knowledge," in the *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS '18)*, Vienna, Austria, July 2018 (acceptance ratio: 20%).

[C24] <u>Luping Wang</u>, **Wei Wang**, Bo Li, "Utopia: Near-optimal Coflow Scheduling with Isolation Guarantee," in the *Proceedings of IEEE INFOCOM '18*, Honolulu, HI, USA, April 2018 (acceptance ratio: 19%).

[C23] <u>Chen Chen</u>, **Wei Wang**, Bo Li, "Performance-Aware Fair Scheduling: Exploiting Demand Elasticity of Data Analytics Jobs," in the *Proceedings of IEEE INFOCOM '18*, Honolulu, HI, USA, April 2018 (acceptance ratio: 19%).

[C22] <u>Yinghao Yu</u>, **Wei Wang**, Jun Zhang, Khaled Ben Letaief, "LERC: Coordinated Cache Management for Data-Parallel Systems," in the *Proceedings of GLOBECOM '17*, Big Data Track, Singapore, December 2017.

[C21] <u>Luping Wang</u>, **Wei Wang**, Bo Li, "Towards Online Checkpointing Mechanism for Cloud Transient Servers," in the *Proceedings of GLOBECOM '17*, Big Data Track, Singapore, December 2017.

[C20] <u>Chen Chen</u>, **Wei Wang**, Bo Li, "Speculative Slot Reservation: Enforcing Service Isolation for Dependent Data-Parallel Computations," in the *Proceedings of the 37th IEEE International Conference on Distributed Computing Systems (ICDCS '17)*, Research Track, Atlanta, GA, USA, June 2017 (acceptance ratio: 17%).

[C19] <u>Yinghao Yu</u>, **Wei Wang**, Jun Zhang, Khaled Ben Letaief, "LRC: Dependency-Aware Cache Management for Data Analytics Clusters," in the *Proceedings of IEEE INFOCOM '17*, Atlanta, GA, USA, May 2017 (acceptance ratio: 21%).

[C18] <u>Chen Chen</u>, **Wei Wang**, <u>Shengkai Zhang</u>, Bo Li, "Cluster Fair Queueing: Speeding up Data-Parallel Jobs with Delay Guarantees," in the *Proceedings of IEEE INFOCOM '17*, Atlanta, GA, USA, May 2017 (acceptance ratio: 21%).

[C17] **Wei Wang**, <u>Shiyao Ma</u>, Bo Li, Baochun Li, "Coflex: Navigating the Fairness-Efficiency Tradeoff for Coflow Scheduling," in the *Proceedings of IEEE INFOCOM '17*, Atlanta, GA, USA, May 2017 (acceptance ratio: 21%).

[C16] **Wei Wang**, <u>A-Long Jin</u>, "Friends or Foes: Revisiting Strategy-Proofness in Cloud Network Sharing," in the *Proceedings of the 24th IEEE International Conference on Network Protocols (ICNP '16)*, Singapore, November 2016 (acceptance ratio: 20%).

[C15] **Wei Wang**, Baochun Li, Ben Liang, Jun Li, "Multi-Resource Fair Sharing for Datacenter Jobs with Placement Constraints," in the *Proceedings of IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC16)*, Salt Lake City, UT, USA, November 2016 (acceptance ratio: 18%).

[C14] **Wei Wang**, Baochun Li, Ben Liang, Jun Li, "Towards Multi-Resource Fair Allocation with Placement Constraints," poster paper in the *Proceedings of the ACM SIGMETRICS '16*, Antibes Juan-les-Pins, France, June 2016 (acceptance ratio: 24%).

[C13] **Wei Wang**, Chen Feng, Baochun Li, Ben Liang, "On the Fairness-Efficiency Tradeoff for Packet Processing with Multiple Resources," in the *Proceedings of the 10th ACM International Conference on emerging Networking EXperiments and Technologies (CoNEXT '14)*, Sydney, Australia, December 2014 (acceptance ratio: 20%).

[C12] **Wei Wang**, Baochun Li, Ben Liang, "Dominant Resource Fairness in Cloud Computing Systems with Heterogeneous Servers," in the *Proceedings of IEEE INFOCOM '14*, Toronto, ON, Canada, April 2014 (acceptance ratio: 19%).

[C11] **Wei Wang**, Ben Liang, Baochun Li, "Low Complexity Multi-Resource Fair Queueing with Bounded Delay," in the *Proceedings of IEEE INFOCOM '14*, Toronto, ON, Canada, April 2014 (acceptance ratio: 19%).

[C10] **Wei Wang**, Baochun Li, Ben Liang, "Multi-Resource Round Robin: A Low Complexity Packet Scheduler with Dominant Resource Fairness," in the *Proceedings of the 21st IEEE International Conference on Network Protocols (ICNP '13)*, Göttingen, Germany, October 2013 (acceptance ratio: 18%).

[C9] **Wei Wang**, Baochun Li, Ben Liang, "To Reserve or Not to Reserve: Optimal Online Multi-Instance Acquisition in IaaS Clouds," in the *Proceedings of the 10th USENIX International Conference on Autonomic Computing (ICAC '13)*, San Jose, CA, USA, June 2013 (acceptance ratio: 22%; one of three **Best Paper Runners Up**).

[C8] **Wei Wang**, Ben Liang, Baochun Li, "Multi-Resource Generalized Processor Sharing for Packet Processing," in the *Proceedings of the 21st ACM/IEEE International Symposium on Quality of Service (IWQoS '13)*, Montreal, QC, Canada, June 2013 (acceptance ratio: 28%).

[C7] **Wei Wang**, Ben Liang, Baochun Li, "Revenue Maximization with Dynamic Auctions in IaaS Cloud Markets," in the *Proceedings of the 21st ACM/IEEE International Symposium on Quality of Service (IWQoS '13)* (short paper), Montreal, QC, Canada, June 2013.

[C6] **Wei Wang**, Ben Liang, Baochun Li, "On Fairness-Efficiency Tradeoffs for Multi-Resource Packet Processing," invited paper in the *Proceedings of the 3rd IEEE International Workshop on Data Center Performance (DCPerf)*, in conjunction with *IEEE ICDCS '13*, Philadelphia, PA, USA, July 2013.

[C5] **Wei Wang**, Di Niu, Baochun Li, Ben Liang, "Dynamic Cloud Resource Reservation via Cloud Brokerage," in the *Proceedings of the 33rd International Conference on Distributed Computing Systems (ICDCS '13)*, Philadelphia, PA, USA, July 2013 (acceptance ratio: 13%).

[C4] **Wei Wang**, Baochun Li, Ben Liang. "Towards Optimal Capacity Segmentation with Hybrid Cloud Pricing," in the *Proceedings of the 32nd International Conference on Distributed Computing Systems (ICDCS '12)*, Macau, China, June 2012 (acceptance ratio: 13%).

[C3] **Wei Wang**, Baochun Li, Ben Liang. "District: Embracing Local Markets in Truthful Spectrum Double Auctions," in the *Proceedings of the 8th IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '11)*, Salt Lake City, UT, USA, June 2011 (acceptance ratio: 22%).

[C2] Qingjiang Shi, Chen He, Hongyang Chen, Lingge Jiang, **Wei Wang**, "Sequential Greedy Localization in Wireless Sensor Networks with Inaccurate Anchor Positions", in the *Proceedings of IEEE GLOBECOM '09*, Honolulu, HI, USA, December 2009.

[C1] **Wei Wang**, Chen He. "A Noncooperative Spectrum Sensing Game with Maximum Network Throughput," in the *Proceedings of IEEE GLOBECOM '09*, Honolulu, HI, USA, December 2009.

PATENT

[P1] **Wei Wang**, Chen He, Lingge Jiang, "An Optimal Spectrum Sensing Policy with Maximum Network Throughput in Spectrum Agile Networks," China Patent CN101730254A, June 9, 2010.

EXTERNAL RESEARCH GRANTS

[G12] Principle Investigator, "Building a Scalable, Low-Latency Serverless Platform: A Data-Centric Approach," RGC General Research Fund (Ref No. 16210822), 2023–25 (total amount:

$1,086,719 HKD).

[G11] Principle Investigator, "Autopiloting Serverless Computing at Scale," Alibaba Innovative Research Program, 2022–23 (total amount: $500,000 CNY).

[G10] Principle Investigator, "Learning to Characterize the Performance of Dataflow Computation: A Code Analysis Approach," RGC General Research Fund (Ref No. 16202121), 2022–24 (total amount: $838,393 HKD).

[G9] Principle Investigator, WeBank Faculty Fellowship Award, 2021–23 (total amount: $1,000,000 CNY).

[G8] Principle Investigator, "Enabling SLO-Aware Resource Overcommitment in Heterogeneous GPU Clusters at Scale," Alibaba Innovative Research Program, 2021–22 (total amount: $500,000 CNY).

[G7] Co-Principle Investigator, "Enabling Secure and Efficient Cross-Silo Federated Learning at Scale," RGC Research Impact Fund (Ref. No. R6021-20F), 2021–26 (total amount: $8,171,429 HKD; my share: $1,300,000 HKD).

[G6] Principle Investigator, "Intelligent Scheduling for Serverless Functions in Shared Platforms at Scale," Alibaba Innovative Research Program, 2020–21 (total amount: $500,000 CNY).

[G5] Principle Investigator, "Learning to Schedule Long-Running Applications in Shared Clusters at Scale," RGC General Research Fund (Ref No. 16213120), 2021–23 (total amount: $845,055 HKD).

[G4] Principle Investigator and Project Coordinator, "Federated Learning at Scale: Systems, Security and Applications," WeBank Collaborative Research Program, 2019–21 (total amount: $1,638,051 HKD).

[G3] Principle Investigator, "Toward Application-Aware Cache Management for In-Memory Data Analytics," RGC General Research Fund (Ref No. 26213818), 2019–21 (total amount: $750,000 HKD).

[G2] Co-Investigator, "People-Aware Smart City: A People-Centered Integration, Mining and Decision Framework," ITC Innovation and Technology Fund (Ref No. ITS/391/15FX), 2016–18 (total amount: $19,550,000 HKD; my share: $800,000 HKD).

[G1] Principle Investigator, "Performance-Centric Scheduling with Service Guarantees for Datacenter Jobs," Microsoft Research Asia Collaborative Research Program, 2016–17 (total amount: $155,000 HKD).

PROFESSIONAL SOCIETY MEMBERSHIPS	IEEE Member	2015 – Present
	IEEE Student Member	2010 – 2015
	ACM Member	2019 – Present

PROFESSIONAL SERVICES

Journal Editorship

◇ *Guest Editor*, IEEE Network: The Special Issues on Interplay between Machine Learning and Networking Systems   2020

◇ Guest Editor, Springer Mobile Networks and Applications: The Special Issues on Advances in Mobile, Edge and Cloud Computing   2020

◇ Editor, China Communications   2017 – 2019

Membership in Conference Organization Committees

◇ *Publicity Chair*, the 17th IEEE International Conference on Mobility, Sensing and Networking (MSN 2021)

◇ *Session Chair*, IEEE INFOCOM 2021

◇ *Publication Chair*, the 15th EAI International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (Qshine 2019)

◇ *Finance Chair*, the 1st ACM Asia-Pacific Workshop on Networking (APNet 2017)

◇ *Finance Chair*, the 25th IEEE International Conference on Network Protocols (ICNP 2017)

◇ *Student Travel Grant Chair*, the 7th ACM SIGOPS Asia-Pacific Workshop on Systems (APSys 2016)

◇ *Student organizer*, IEEE INFOCOM 2014

◇ *Student organizer*, IEEE INFOCOM 2014 TPC Meeting

◇ *Student organizer*, the 21st IEEE/ACM International Symposium on Quality of Service (IWQoS 2013)

◇ *Student organizer*, the 22nd ACM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2012)

Membership in Technical Program Committees

◇ IEEE INFOCOM 2023

◇ The 3rd International Workshop on Parallel AI and Systems for the Edge (PAISE 2022), in conjunction with IEEE IPDPS 2022

◇ IEEE INFOCOM 2022 (**Distinguished TPC Member**)

◇ The 12th ACM SIGOPS Asia-Pacific Workshop on Systems (APSys 2021)

◇ IEEE/ACM International Symposium on Quality of Service 2021 (IWQoS 2021)

◇ The 2nd International Workshop on Parallel AI and Systems for the Edge (PAISE 2021), in conjunction with IEEE IPDPS 2021

◇ The 22nd IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2021)

◇ IEEE INFOCOM 2021

◇ The 16th EAI International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (Qshine 2020)

◇ The 28th IEEE International Conference on Network Protocols (ICNP 2020)

◇ IEEE/ACM International Symposium on Quality of Service 2020 (IWQoS 2020)

◇ IEEE INFOCOM 2020 (**Distinguished TPC Member**)

◇ The 15th EAI International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (Qshine 2019)

◇ IEEE GLOBECOM 2019: Next-Generation Networking and Internet

◇ The 39th IEEE International Conference on Distributed Computing Systems (ICDCS 2019)

◇ IEEE INFOCOM 2019 (**Distinguished TPC Member**)

◇ The 26th IEEE International Conference on Network Protocols (ICNP 2018)

◇ The 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2018)

◇ IEEE GLOBECOM 2018: Big Data Track

◇ IEEE INFOCOM 2018 (**Distinguished TPC Member**)

◇ IEEE ICNP 2017 Workshop on Machine Learning and Artificial Intelligence in Computer Networks (ML&AI @ Network 2017)

◇ The 25th IEEE International Conference on Network Protocols (ICNP 2017)

- ◇ IEEE GLOBECOM 2017: Next-Generation Networking and Internet Symposium
- ◇ The 27th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2017)
- ◇ The 26th International Conference on Computer Communication and Networks (ICCCN 2017)
- ◇ IEEE International Conference on Cloud Engineering (IC2E) 2017
- ◇ IEEE ICNP Workshop on Machine Learning in Computer Networks (NetworkML) 2016
- ◇ The 25th International Conference on Computer Communication and Networks (ICCCN 2016), Hot Topics in Networking Track (HOT)
- ◇ IEEE GLOBECOM 2015: Communications Software Services and Multimedia Applications (CSSMA) Symposium

Reviewer for Journal Manuscript Submissions

- ◇ ACM Transactions on Computer Systems
- ◇ IEEE/ACM Transactions on Networking
- ◇ IEEE Transactions on Parallel and Distributed Systems
- ◇ IEEE Journal on Selected Areas in Communication
- ◇ IEEE Transactions on Cloud Computing
- ◇ IEEE Transactions on Mobile Computing
- ◇ IEEE Transactions on Knowledge and Data Engineering
- ◇ IEEE Transactions on Information Forensics & Security
- ◇ IEEE Transactions on Big Data
- ◇ ACM Transactions on Modeling and Performance Evaluation of Computing Systems
- ◇ IEEE Transactions on Services Computing
- ◇ IEEE Communications Letters
- ◇ Springer Multimedia Systems
- ◇ Springer Journal of Cloud Computing: Advances, Systems and Applications
- ◇ Springer Peer-to-Peer Networking and Applications

External Reviewer for Conference Manuscript Submissions

ACM CoNEXT, ACM SIGMETRICS, ACM MobiCom, ACM Multimedia, ACM MMSys, IEEE Cloud, IEEE SECON, IEEE MASS, IEEE ICC, IEEE ICME, IEEE NetCod, IFIP Networking, IEEE ChinaSIP.

Industry Education Program

- ◇ AWS Educate Cloud Ambassador   2020 – Present

UNIVERSITY
SERVICE
ACTIVITIES

Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

- ◇ *Associate Director of CPEG Program*, Department of Computer Science and Engineering, 2022 – 2025
- ◇ *MScIT Program Committee*, Department of Computer Science and Engineering, 2022 – 2025
- ◇ *CPEG Program Committee*, Department of Computer Science and Engineering, 2021 – 2022
- ◇ *Seminar Coordinator*, Department of Computer Science and Engineering, 2019 – 2021
- ◇ *Postgraduate Committee*, Department of Computer Science and Engineering, 2018 – present
- ◇ *Interview Panel for Joint Entrance Examination Candidates in Mainland China*, School of Engineering, 2016 – present
- ◇ *Local Arrangement Chair*, HKUST-SJTU Joint Workshop 2018, Department of Computer

Science and Engineering, June 2018

<table>
<tr><td rowspan="1"></td></tr>
</table>

GRADUATE **S**UPERVISION

Current Students in the Doctor of Philosophy Program

Name	Duration of Study
Huangshi Tian	2017 – Present
Qizhen Weng	2017 – Present
Minchen Yu	2018 – Present
Suyi Li	2019 – Present
Zhifeng Jiang	2019 – Present
Lingyun Yang	2020 – Present
Yongkang Zhang	2021 – Present
Peng Ye	2021 – Present

Current Students in the Master of Philosophy Program

Name	Duration of Study
Yunchuan Zheng	2019 – Present

Students Graduated with the Doctor of Philosophy Degree

Name	Duration & First Employment	Dissertation Title
Da Yan	2016 – 22, Systems Researcher at Anthropic, San Francisco, CA, USA	Principles and Automation of Low-Level Optimizations on GPUs
Mingzhe Li	2016 – 22, Research Scientist at A*STAR, Singapore	Towards Efficient and Secure Sharding-Based Blockchain
Luping Wang	2016–21 (informal advisee), Senior Software Engineer (P7, Alibaba Star) in Alibaba Group, Hangzhou, China	Optimizing Resource Scheduling for Cloud Workloads Running in Data Centers
Chengliang Zhang	2016–21, Data Scientist in Seleya Technologies, Hong Kong	Towards Efficient and Secure Large-Scale Systems for Distributed Machine Learning Training
Yinghao Yu	2015–19, Senior Software Engineer (P7, Alibaba Star) in Alibaba Group, Hangzhou, China	Towards Efficient, Fair and Load-Balanced Cluster Caching for Big Data Analytics
Chen Chen	2014–18, Research Scientist in Huawei Research, Hong Kong	Job Scheduling in the Cloud: A Tale on Fairness and Efficiency

Students Graduated with the Master of Philosophy Program

Name	Duration & First Employment	Dissertation Title
Xiandong Qi	2016–19, Software Engineer in AQUMON, Hong Kong	Multi-Resource Fair Sharing with Constraints in Heterogeneous Clusters

Ph.D. Thesis Committees

Chenyu Huang, Huangxun Chen, Zhuohang Lai, 2020
Shuihai Hu, Ming Wen, 2019
Li Chen, 2018
Jingjie Jiang, Wei Bai, Yi Zhang, 2017
Zhehui Wang, Zhice Yang, 2016

GRADUATE **S**UPERVISION **A**CHIEVEMENTS

Hong Kong PhD Fellowship Scheme (3-year fellowship for top PhD students in Hong Kong)

Name	Duration of Fellowship
Yongkang Zhang	2021 – 2024
Huangshi Tian	2017 – 2020
Qizhen Weng	2017 – 2020
Chengliang Zhang	2016 – 2019
Yinghao Yu	2015 – 2018

Huawei PhD Fellowship Scheme

Name	Duration of Fellowship
Yunchuan Zheng	2019 – 2020
Minchen Yu	2018 – 2021

Alibaba Star Program (Alibaba's top-tier recruitment program, aiming at the most brilliant graduates from all over the world)

⬦ Qizhen Weng, December 2021
⬦ Luping Wang, August 2021
⬦ Yinghao Yu, August 2019 (hired at the highest grade of the program in 2019)

IEEE INFOCOM '17 Best-In-Session Presentation Award   May 2017

⬦ Chen Chen, for the talk entitled "Cluster Fair Queueing: Speeding up Data-Parallel Jobs with Delay Guarantees," in Conference Session *Cloud Computing I.*

⬦ Yinghao Yu, for the talk entitled "LRC: Dependency-Aware Cache Management for Data Analytics Clusters," in Conference Session *Big Data Processing.*

TEACHING ACTIVITIES

Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
Department of Computer Science and Engineering

Academic Term	Course
Spring 2022	CSIT6000O: Advanced Cloud Computing
Spring 2022	COMP4651: Cloud Computing and Big Data Systems
Fall 2021	COMP4651: Cloud Computing and Big Data Systems
Spring 2021	COMP4651: Cloud Computing and Big Data Systems
Fall 2020	COMP3511: Operating System
Fall 2019	COMP3511: Operating System
	COMP4651: Cloud Computing and Big Data Systems
	CSIC5001: Introduction to Advanced Computing Systems
Spring 2019	COMP3511: Operating System
Fall 2018	COMP4651: Cloud Computing and Big Data Systems
Spring 2018	COMP3511: Operating System
	CSIC5001: Introduction to Advanced Computing Systems
Fall 2017	COMP4651: Cloud Computing and Big Data Systems
Spring 2017	COMP4651: Cloud Computing and Big Data Systems
Fall 2016	COMP6611B: Topics on Cloud Computing and Data Analytics Systems
Summer 2016	COMP4621: Computer Communication Networks I
Fall 2015	COMP6611B: Topics on Cloud Computing and Data Analytics Systems

University of Toronto, Toronto, Ontario, Canada
Department of Electrical and Computer Engineering

⬦ APS105: Computer Fundamentals (*Teaching Assistant*)   Fall 2010 – 2014

⬦ ECE344: Operating Systems (*Teaching Assistant*)   Spring 2011, 2012