Chuan Wu (Curriculum Vitae)

PRESENT APPOINTMENT	Professor Department of Computer Science *Office*: CB 427 The University of Hong Kong *Phone*: (852) 2857 8459 Pokfulam Road *Email*: cwu@cs.hku.hk Hong Kong *Web*: `http://www.cs.hku.hk/~cwu`

CURRENT RESEARCH INTERESTS

Modeling, optimization, design and implementation of cloud computing systems, big data analytics/distributed machine learning systems, smart elderly/health care technologies.

Google Scholar total number of citations: 8383, h-index: 51

ACADEMIC QUALIFICATION

◇ **B.Eng.**, Computer Science and Technology, **Tsinghua University**, Beijing, China, September 1996 — July 2000
◇ **M.Eng.**, Computer Science and Technology, **Tsinghua University**, Beijing, China, September 2000 — July 2002
◇ **Ph.D.**, Electrical and Computer Engineering, **University of Toronto**, Toronto, Ontario, Canada, September 2004 — July 2008

PROFESSIONAL MEMBERSHIP

◇ Senior Member, the Institute of Electrical and Electronic Engineers (IEEE) and IEEE Communications Society (ComSoc)
◇ Member, the Association for Computing Machinery (ACM)

HONORARY ACADEMIC APPOINTMENTS

◇ Visiting associate professor, Pengcheng Laboratory, Shenzhen China: 2018-2020
◇ Honorary chair professor, School of Computing, Wuhan University, China: 2018-2022
◇ Honorary Research Fellow, Sau Po Centre on Ageing (COA) of The University of Hong Kong: 2022 - present

HONORS AND AWARDS

◇ Brown Award for the invention "A Smart Elderly Walker" in the 48th Geneva International Exhibition of Inventions, 2023
◇ Among the World's Top 2% Scientists (by career-long impact), 2022.
◇ Amazon Research Award (ARA) on AWS AI, 2021.
◇ IEEE Transactions on Cloud Computing (TCC) Editorial Excellence and Eminence Award: 2019, 2020.
◇ MSRA Collaborative Research Award, 2018.
◇ Selected as one of the 10 female associate professors in HKU to participate in HKU-10 Mentorship Programme (inaugural cohort), 2017-2018.
◇ Best Paper Award in the Seventh ACM International Conference on Future Energy Systems (ACM e-Energy), 2016.
◇ Distinguished Member of TPC, IEEE INFOCOM 2015, 2016, 2019, 2020, 2023.
◇ Prolific Reviewer of IEEE Transactions on Parallel and Distributed Systems, 2012.
◇ Best Paper Award in the 4th International Workshop on Hot Topics in Peer-to-Peer Computing and Online Social Networking (HotPOST'12), in conjunction with IEEE ICDCS 2012.
◇ IEEE Outstanding Leadership Award, 2011.
◇ Amazon Web Service (AWS) in Education Research Grant Award, 2011.
◇ University of Toronto Fellowship, 2004 — 2008.
◇ Outstanding Master Thesis, Tsinghua University, 2002.
◇ Tsinghua University Excellent Graduate Scholarship, 2001.
◇ Outstanding Graduate, Tsinghua University, 2000.

◇ Tsinghua University Social Work Scholarship for Excellence in Social Work, 1999.
◇ Tsinghua-Legend Scholarship for Outstanding Undergraduate Students, 1999.
◇ Tsinghua University First-Class Scholarship (awarded to the top 5% undergraduate students), 1998.
◇ Tsinghua University Excellent Undergraduate Student Scholarship, 1997.
◇ Tsinghua-Haolaixi Scholarship for Outstanding Undergraduate Students, 1996.

PUBLICATIONS ◇ **Book Chapters:**

[B1] Linquan Zhang, **Chuan Wu**, Zongpeng Li, Chuanxiong Guo, Minghua Chen, Francis C.M. Lau. "Moving Big Data to the Cloud: Online Cost-Minimizing Algorithms", pp. 75-102 in *Networking for Big Data*, Chapman & Hall/CRC, August 3, 2015.

[B2] Yixin Bao, **Chuan Wu**. "Geo-Distributed Big Data Analytics Systems: An Online Learning Approach for Dynamic Deployment", pp. 161-190 in *Smart Data: State-of-the-Art Perspectives in Computing and Applications*, Chapman & Hall/ CRC Big Data Series, March 13, 2019.

◇ **Journal Publications:**

[J1] **Chuan Wu**, Yufei Ma, Yuwen He, Yuzhuo Zhong, Hongjiang Zhang. "Event Recognition by Semantic Inference in Sports Videos", in *Elsevier Journal of Tsinghua Science and Technology*, Vol. 43, No. 4, pp. 507-510, April 2003.

[J2] **Chuan Wu**, Baochun Li. "Diverse: Application-Layer Service Differentiation in Peer-to-Peer Communications," in *IEEE Journal on Selected Areas in Communications, Special Issue on Peer-to-Peer Communications and Applications*, Vol. 25, No. 1, pp. 222-234, January 2007.

[J3] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Characterizing Peer-to-Peer Streaming Flows," in *IEEE Journal on Selected Areas in Communications, Special Issue on Advances in Peer-to-Peer Streaming Systems*, Vol. 25, No. 9, pp. 1612-1626, December 2007.

[J4] **Chuan Wu**, Baochun Li. "rStream: Resilient and Optimal Peer-to-Peer Streaming with Rateless Codes," in *IEEE Transactions on Parallel and Distributed Systems*, Vol. 19, No. 1, pp. 77-92, January 2008.

[J5] **Chuan Wu**, Baochun Li, Zongpeng Li. "Dynamic Bandwidth Auctions in Multioverlay P2P Streaming with Network Coding," in *IEEE Transactions on Parallel and Distributed Systems*, Vol. 19, No. 6, pp. 806-820, June 2008.

[J6] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Exploring Large-Scale Peer-to-Peer Live Streaming Topologies," in *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol. 4, Issue 3, No. 19, August 2008.

[J7] Xuanjia Qiu, Wei Huang, **Chuan Wu**, Francis C.M. Lau, Xiaola Lin. "InstantLeap: an Architecture for Fast Neighbor Discovery in Large-scale P2P VoD Streaming," in *Springer Multimedia Systems Journal*, vol. 16, No. 3, pp. 183-198, Springer-Verlag, June 2010.

[J8] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "On Dynamic Server Provisioning in Multichannel P2P Live Streaming," in *IEEE/ACM Transactions on Networking*, vol. 19, No. 5, pp. 1317-1330, October 2011.

[J9] Zhi Wang, **Chuan Wu**, Lifeng Sun, Shiqiang Yang, "Strategies of Collaboration in Multi-Swarm Peer-to-Peer Content Distribution," in *Elsevier Journal of Tsinghua Science and Technology*, Vol. 17, No. 1, pp. 29-39, February 2012.

[J10] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Diagnosing Network-wide P2P Live Streaming Inefficiencies," in *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol. 8s, Issue 1, No. 13, February 2012.

[J11] **Chuan Wu**, Zongpeng Li, Xuanjia Qiu, Francis C.M. Lau, "Auction-based P2P VoD Streaming: Incentives and Optimal Scheduling," in *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol. 8s, Issue 1, No. 14, February 2012.

[J12] Jian Zhao, **Chuan Wu**. "Characterizing Locality-aware P2P Streaming," in *Journal of Communications*, Vol. 7, No. 3, pp. 222-231, March 2012.

[J13] Xingwei Wang, Jiajia Sun, Hongxing Li, **Chuan Wu**, Min Huang. "A Reverse Auction Based Allocation Mechanism in the Cloud Computing Environment", in *Applied Mathematics & Information Sciences*, vol. 7, pp. 75-84, February 2013.

[J14] Zhi Wang, **Chuan Wu**, Lifeng Sun, Shiqiang Yang. "Peer-Assisted Social Media Streaming With Social Reciprocity," in *IEEE Transactions on Network and Service Management*, vol. 10, No. 1, pp. 84-94, March 2013.

[J15] Zhizhong Zhang, **Chuan Wu**, David W.L. Cheung. "A Survey on Cloud Interoperability: Taxonomies, Standards, and Practice," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, No. 4, pp. 13-22, March 2013.

[J16] Yu Wu, Zhizhong Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Cloud-MoV: Cloud-based Mobile Social TV," in *IEEE Transactions on Multimedia, Special Issue on Cloud-Based Mobile Media – Infrastructure, Services and Applications*, vol. 15, No. 4, pp. 821-832, June 2013.

[J17] Ruiting Zhou, Zongpeng Li, **Chuan Wu**, Carey Williamson. "Signal Alignment: Enabling Physical Layer Network Coding for MIMO Networking," in *IEEE Transactions on Wireless Communications*, vol. 12, No. 6, pp. 3012-3023, June 2013.

[J18] Hongxing Li, **Chuan Wu**, Dongxiao Yu, Qiang-Sheng Hua, Francis C.M. Lau. "Aggregation Latency-Energy Tradeoff in Wireless Sensor Networks with Successive Interference Cancellation," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, No. 11, pp. 2160-2170, November 2013.

[J19] Linquan Zhang, **Chuan Wu**, Zongpeng Li, Chuanxiong Guo, Minghua Chen, Francis C.M. Lau. "Moving Big Data to the Cloud: An Online Cost-Minimizing Approach," in *IEEE Journal on Selected Areas in Communications, Special Issue on Networking Challenges in Cloud Computing Systems and Applications*, vol. 31, No. 12, pp. 2710-2721, December 2013.

[J20] Hongxing Li, Qiangsheng Hua, **Chuan Wu**, Francis C.M. Lau. "Latency-Minimizing Data Aggregation in Wireless Sensor Networks under Physical Interference Model," in *Elsevier Ad Hoc Networks Journal*, vol. 12, pp. 52-68, January 2014.

[J21] Tang Xiahou, Zongpeng Li, **Chuan Wu**, Jiaqing Huang. "A Geometric Per-

spective to Multiple-Unicast Network Coding," in *IEEE Transactions on Information Theory*, vol. 60, No. 5, pp. 2884 - 2895, May 2014.

[J22] Wei Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "The Performance and Locality Tradeoff in BitTorrent-like File Sharing Systems," in *Springer Peer-to-Peer Networking and Applications Journal*, vol. 7, no. 4, pp. 469 - 484, December 2014.

[J23] Jian Zhao, **Chuan Wu**, Xiaojun Lin. "Locality-aware Streaming in Hybrid P2P-Cloud CDN Systems", in *Springer Peer-to-Peer Networking and Applications Journal, Special Issue on P2P Cloud Systems*, vol. 8, no. 2, pp. 320-335, March 2015.

[J24] Zhi Wang, Lifeng Sun, **Chuan Wu**, Shiqiang Yang. "Enhancing Internet-scale Video Service Deployment Using Microblog-based Prediction", in *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 3, pp. 775 - 785, March 2015.

[J25] Yu Wu, **Chuan Wu**, Bo Li, Linquan Zhang, Zongpeng Li, Francis C.M. Lau. "Scaling Social Media Applications into Geo-Distributed Clouds," in *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 689 - 702, June 2015.

[J26] Zhi Wang, Lifeng Sun, **Chuan Wu**, Wenwu Zhu, Qidong Zhuang, Shiqiang Yang. "A Joint Online Transcoding and Delivery Approach for Dynamic Adaptive Streaming," in *IEEE Transactions on Multimedia*, vol. 17, no. 6, pp. 867-879, June 2015.

[J27] Xuanjia Qiu, Hongxing Li, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Cost-Minimizing Dynamic Migration of Content Distribution Services into Hybrid Clouds", in *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 12, pp. 3330-3345, December 2015.

[J28] Ruiting Zhou, Zongpeng Li, **Chuan Wu**, Minghua Chen. "Demand Response in Smart Grids: A Randomized Auction Approach," in *IEEE Journal on Selected Areas in Communications - Series in Green Communications and Networking (JSAC-SGCN)*, vol. 33, No. 12, pp. 2540-2553, December 2015.

[J29] Linquan Zhang, Zongpeng Li, **Chuan Wu**, Shaolei Ren. " Online Electricity Cost Saving Algorithms for Co-Location Data Centers," in *IEEE Journal on Selected Areas in Communications - Series in Green Communications and Networking (JSAC-SGCN)*, vol. 33, No. 12, pp. 2906-2919, December 2015.

[J30] Can Zhao, Xiaojun Lin, **Chuan Wu**. "The Streaming Capacity of Sparsely-Connected P2P Systems with Distributed Control", in *IEEE/ACM Transactions on Networking*, vol. 24, no. 1, pp. 58-71, February 2016.

[J31] Hongxing Li, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Virtual Machine Trading in a Federation of Clouds: Individual Profit and Social Welfare Maximization," in *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1827 - 1840, June 2016.

[J32] Weijie Shi, Linquan Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "An Online Auction Framework for Dynamic Resource Provisioning in Cloud Computing," in *IEEE/ACM Transactions on Networking*, vol 24, no. 4, pp. 2060 - 2073, August 2016.

[J33] Can Zhao, Jian Zhao, Xiaojun Lin, **Chuan Wu**. "Capacity of P2P On-Demand

Streaming with Simple, Robust and Decentralized Control", in *IEEE/ACM Transactions on Networking*, vol. 24, no. 5, pp. 2607 - 2620, October 2016.

[J34] Linquan Zhang, Shaolei Ren, **Chuan Wu**, Zongpeng Li. "A Truthful Incentive Mechanism for Emergency Demand Response in Geo-distributed Colocation Data Centers," in *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, vol. 1, no. 4, September 2016.

[J35] Qihang Sun, **Chuan Wu**, Zongpeng Li, Shaolei Ren. "Colocation Demand Response: Joint Online Mechanisms for Individual Utility and Social Welfare Maximization," in *IEEE Journal on Selected Areas in Communications - Series on Green Communications and Networking*, vol. 34, no. 12, pp. 3978 - 3992, December 2016.

[J36] Yu Wu, Zhizhong Zhang, **Chuan Wu**, Chuanxiong Guo, Zongpeng Li, Francis C.M. Lau. "Orchestrating Bulk Data Transfers across Geo-Distributed Datacenters", in *IEEE Transactions on Cloud Computing*, vol. 5, no. 1, pp. 112 - 125, January-March 2017.

[J37] Xiaoxi Zhang, Zhiyi Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "An Online Stochastic Buy-Sell Mechanism for VNF Chains in the NFV Market," in *IEEE Journal on Selected Areas in Communications (JSAC) - Special Issue on Game Theory for Networks*, vol. 35, No. 2, pp. 1-15, February 2017.

[J38] Sijia Gu, Zongpeng Li, **Chuan Wu**, Huyin Zhang. "Virtualized Resource Sharing in Cloud Radio Access Networks Through Truthful Mechanisms," in *IEEE Transactions on Communications*, vol. 65, No. 3, pp. 1105 - 1118, March 2017.

[J39] Weijie Shi, **Chuan Wu**, Zongpeng Li. "An Online Auction Mechanism for Dynamic Virtual Cluster Provisioning in Geo-Distributed Clouds," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, No. 3, pp. 677-688, March 2017.

[J40] Xiaoxi Zhang, Zhiyi Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Online Auctions in IaaS Clouds: Welfare and Profit Maximization with Server Costs", in *IEEE/ACM Transactions on Networking*, vol. 25, No. 2, pp. 1034 - 1047, April 2017.

[J41] Ruiting Zhou, Zongpeng Li, **Chuan Wu**, Zhiyi Huang. "An Efficient Cloud Market Mechanism for Computing Jobs with Soft Deadlines", in *IEEE/ACM Transactions on Networking*, vol. 25, No. 2, pp. 793-805, April 2017.

[J42] Jingpu Duan, **Chuan Wu**, Franck Le, Alex X. Liu, Yanghua Peng. "Dynamic Scaling of Virtualized, Distributed Service Chains: A Case Study of IMS," in *IEEE Journal on Selected Areas in Communications (JSAC) Special issue on Emerging Technologies in Software-driven Communication*, vol. 35, no. 11, pp. 2501-2511, November 2017.

[J43] Mohammad H. Hajiesmaili, Lok To Mak, Zhi Wang, **Chuan Wu**, Minghua Chen. "Cost-Effective Low-Delay Cloud Video Conferencing," in *IEEE Transactions on Multimedia*, vol. 19, no. 12, pp. 2760-2774, December 2017.

[J44] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "Virtualized Resource Sharing in Cloud Radio Access Networks: An Auction Approach", in *Elsevier Computer Communications Journal*, vol. 114, no. C, pp. 22-35, December 2017.

[J45] Weijie Shi, **Chuan Wu**, Zongpeng Li. "A Shapley-value Mechanism for Bandwidth On Demand between Datacenters," in *IEEE Transactions on Cloud Computing*, vol. 6, no. 1, pp. 19-32, January-March 2018.

[J46] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "Scheduling Frameworks for Cloud Container Services," in *IEEE/ACM Transactions on Networking*, vol. 26, no. 1, pp. 436-450, February 2018.

[J47] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "An Online Emergency Demand Response Mechanism for Cloud Computing," in *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, vol. 3, no. 5, February 2018.

[J48] Yongzheng Jia, **Chuan Wu**, Zongpeng Li, Franck Le, Alex Liu. "Online Scaling of NFV Service Chains across Geo-distributed Datacenters," in *IEEE/ACM Transactions on Networking*, vol. 26, no. 2, pp. 699-710, April 2018.

[J49] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "A Truthful Online Mechanism for Location-Aware Tasks in Mobile Crowd Sensing," in *IEEE Transactions on Mobile Computing*, vol. 17, no. 8, pp. 1737-1749, August 2018.

[J50] Jingpu Duan, Xiaodong Yi, Shixiong Zhao, **Chuan Wu**, Heming Cui, Franck Le. "NFVactor: A Resilient NFV System using the Distributed Actor Model," in *IEEE Journal on Selected Areas in Communications (JSAC) Special issue on Network Softwarization & Enablers*, vol. 37, no. 3, pp. 586-599, March 2019.

[J51] Jingpu Duan, Xiaodong Yi, Junjie Wang, **Chuan Wu**, Franck Le. "NetStar: a Future/Promise Framework for Asynchronous Network Functions," in *IEEE Journal on Selected Areas in Communications (JSAC) Special issue on Network Softwarization & Enablers*, vol. 37, no. 3, pp. 600-612, March 2019.

[J52] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "An Efficient Online Placement Scheme for Cloud Container Clusters," in *IEEE Journal on Selected Areas in Communications (JSAC) Special issue on Network Softwarization & Enablers*, vol. 37, no. 5, pp. 1046 - 1058, May 2019.

[J53] Ziyue Luo, **Chuan Wu**, Zongpeng Li, Wei Zhou. "Scaling Geo-distributed Network Function Chains: A Prediction and Learning Framework," in *IEEE Journal on Selected Areas in Communications (JSAC) Special issue on Network Softwarization & Enablers*, vol. 37, no. 8, pp. 1838 - 1850, August 2019.

[J54] Jin Wang, Jianping Wang, **Chuan Wu**, Kejie Lu, Naijie Gu. "Enhancing the Anonymity in Information Diffusion Based on Obfuscated Coded Data," in *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 4, pp. 968-982, October-December 2019.

[J55] Qihang Sun, Shaolei Ren, **Chuan Wu**. "Fair Online Power Capping for Emergency Handling in Multi-Tenant Cloud Data Centers," in *IEEE Transactions on Cloud Computing*, vol. 8, No. 1, pp. 152-166, January-March 2020.

[J56] Xiaotong Li, Ruiting Zhou, Lei Jiao, **Chuan Wu**, Zongpeng Li. "Online Placement and Scaling of Geo-distributed Machine Learning Jobs via Volume-discounting Brokerage," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 41, No. 4, pp. 948-966, April 2020.

[J57] Ziyue Luo, **Chuan Wu**. "An Online Algorithm for VNF Service Chain Scaling

in Datacenters," in *IEEE/ACM Transactions on Networking*, vol. 28, No. 3, pp. 1061-1073, June 2020.

[J58] Xiaoxi Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "A Truthful $(1-\epsilon)$-Optimal Mechanism for On-demand Cloud Resource Provisioning," in *IEEE Transactions on Cloud Computing*, vol. 8, No. 3, pp. 735-748, July-September 2020.

[J59] Xiaoyang Zhao, Zhi Zhu, MingShan Liu, Chongyu Zhao, Yafei Zhao, Jia Pan, Zheng Wang, **Chuan Wu**. "A Smart Robotic Walker with Intelligent Close-proximity Interaction Capabilities for Elderly Mobility Safety," in *Frontiers in Neurorobotics, October 2020*.

[J60] Xiaoxi Zhang, Zhiyi Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Dynamic VM Scaling: Provisioning and Pricing through an Online Auction," in *IEEE Transactions on Cloud Computing*, vol. 9, No. 1, pp. 131-144, January-March 2021.

[J61] Le Fang, Yu Wu, **Chuan Wu**, Yizhou Yu. "A Non-Intrusive Elderly Home Monitoring System," in *IEEE Internet of Things Journal*, vol. 8, No. 4, pp. 2603-2614, February 2021.

[J62] Yanghua Peng, Yixin Bao, Yangrui Chen, **Chuan Wu**, Chen Meng, Wei Lin. "DL^2: A Deep Learning-driven Scheduler for Deep Learning Clusters," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, No. 8, pp. 1947-1960, August 2021.

[J63] Bingqian Du, Zhiyi Huang, **Chuan Wu**. "Adversarial Deep Learning for Online Resource Allocation," in *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, vol. 6, No. 4, pp. 1-25, December 2021.

[J64] Menglu Yu, Jia Liu, **Chuan Wu**, Bo Ji, Elizabeth S. Bentley. "Toward Efficient Online Scheduling for Distributed Machine Learning Systems," in *IEEE Transactions on Network Science and Engineering*, vol. 9, No. 4, pp. 1951-1969, July-August 2022.

[J65] Le Fang, Bingqian Du, **Chuan Wu**. "Differentially Private Recommender System with Variational Autoencoders," *Elsevier Knowledge-Based Systems Journal*, vol. 250, 17 August 2022.

[J66] Xiaodong Yi, Shiwei Zhang, Lansong Diao, **Chuan Wu**, Zhen Zheng, Shiqing Fan, Siyu Wang, Jun Yang, Wei Lin. "Optimizing DNN Compilation for Distributed Training with Joint OP and Tensor Fusion," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, No. 12, pp. 4694 - 4706, December 2022.

[J67] Xiaoyang Zhao, **Chuan Wu**. "Large-scale Machine Learning Cluster Scheduling via Multi-agent Graph Reinforcement Learning," in *IEEE Transactions on Network and Service Management*, vol. 19, No. 4, pp. 4962-4974, December 2022.

[J68] Yangrui Chen, Jiaxuan Youb, Jun He, Yuan Lin, Yanghua Peng, **Chuan Wu**, Yibo Zhu. "SP-GNN: Learning Structure and Position Information from Graphs Learning Systems," in *Elsevier Neural Networks Journal*, vol. 161, pp. 505-514, April 2023.

[J69] Shiji Zhou, Zhi Wang, Chenghao Hu, Yinan Mao, Haopeng Yan, **Chuan Wu**,

Shanghang Zhang, Wenwu Zhu. "Caching in Dynamic Environments: a Near-optimal Online Learning Approach," in *IEEE Transactions on Multimedia*, vol. 25, pp. 792-804, 2023.

[J70] Yixin Bao, Yanghua Peng, **Chuan Wu**. "Deep Learning-based Job Placement in Distributed Machine Learning Clusters with Heterogeneous Workloads," in *IEEE/ACM Transactions on Networking*, vol. 31, No. 2, pp. 634-647, April 2023.

[J71] Shiwei Zhang, Xiaodong Yi, Lansong Diao, **Chuan Wu**, Siyu Wang, Wei Lin. "Expediting Distributed DNN Training with Device Topology-Aware Graph Deployment," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, No. 4, pp. 1281-1293, April 2023.

[J72] Ruiting Zhou, Jinlong Pang, Qin Zhang, **Chuan Wu**, Lei Jiao, Yi Zhong, Zongpeng Li. "Online Scheduling Algorithm for Heterogeneous Distributed Machine Learning Jobs," in *IEEE Transactions on Cloud Computing*, vol. 11, No. 2, pp. 1514-1529, April-June 2023.

[J73] Yihong Li, Xiaoxi Zhang, Tianyu Zeng, Jingpu Duan, **Chuan Wu**, Di Wu, Xu Chen. "Task Placement and Resource Allocation for Edge Machine Learning: A GNN-based Multi-Agent Reinforcement Learning Paradigm," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 12, pp. 3073-3089, December 2023.

[J74] Cheng Pan, Hao Luo, Gary Cheung, Huiquan Zhou, Reynold Cheng, Sarah Cullum, **Chuan Wu**. "Identifying Frailty in Older Adults Receiving Home Care Assessment Using Machine Learning: Examining the Role of Classifier, Feature Selection, and Sample Size," in *JMIR AI*, vol. 3, January 2024.

[J75] Tianyu Zeng, Xiaoxi Zhang, Jingpu Duan, Chao Yu, **Chuan Wu**, Xu Chen. "An Offline-Transfer-Online Framework for Cloud-Edge Collaborative Distributed Reinforcement Learning," to appear in *IEEE Transactions on Parallel and Distributed Systems*.

◇ **International Conference Papers:**

[C1] **Chuan Wu**, Yuwen He, Li Zhao, Yuzhuo Zhong. "Motion Feature Extraction Scheme for Content-based Video Retrieval", in the Proceeding of *SPIE Storage and Retrieval for Media Databases 2002*, Vol. 4676, pp. 296-305, San Jose, California, January 23-25, 2002.

[C2] **Chuan Wu**, Yufei Ma, Hongjiang Zhang, Yuzhuo Zhong. "Events Recognition by Semantic Inference for Sports Video", in the Proceeding of the *IEEE International Conference on Multimedia and Expo 2002 (IEEE ICME 2002)*, pp. 805-808, Lausanne, Switzerland, August 26-29, 2002.

[C3] **Chuan Wu**, Baochun Li. "rStream: Resilient Peer-to-Peer Streaming with Rateless Codes," in the Proceedings of *ACM Multimedia 2005*, pp. 307-310, Singapore, November 6-12, 2005.

[C4] **Chuan Wu**, Baochun Li. "Optimal Peer Selection for Minimum-Delay Peer-to-Peer Streaming with Rateless Codes," in the Proceedings of *ACM Workshop on Advances in Peer-to-Peer Multimedia Streaming (P2PMMS 2005)*, in conjunction with *ACM Multimedia 2005*, pp. 69-78, Singapore, November 6-12, 2005.

[C5] **Chuan Wu**, Baochun Li. "Echelon: Peer-to-Peer Network Diagnosis with Net-

work Coding," in the Proceedings of the *Fourteenth IEEE International Workshop on Quality of Service (IEEE IWQoS 2006)*, pp. 20-29, Yale University, New Haven, Connecticut, June 19-21, 2006.

[C6] **Chuan Wu**, Baochun Li. "Strategies of Conflict in Coexisting Streaming Overlays," in the Proceedings of *IEEE INFOCOM 2007*, pp. 481-489, Anchorage, Alaska, May 6-12, 2007.

[C7] **Chuan Wu**, Baochun Li. "Optimal Rate Allocation in Overlay Content Distribution," in the Proceedings of *IFIP Networking 2007, also Lecture Notes in Computer Science, Springer*, vol. 4479, pp. 678-690, Atlanta, Georgia, May 14-18, 2007.

[C8] **Chuan Wu**, Baochun Li. "Outburst: Efficient Overlay Content Distribution with Rateless Codes," in the Proceedings of *IFIP Networking 2007, also Lecture Notes in Computer Science*, Springer, vol. 4479, pp. 1208-1216, Atlanta, Georgia, May 14-18, 2007.

[C9] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Magellan: Charting Large-Scale Peer-to-Peer Live Streaming Topologies," in the Proceedings of the *27th IEEE International Conference on Distributed Computing Systems (IEEE ICDCS 2007)*, pp. 62-69, Toronto, Ontario, Canada, June 25-29, 2007.

[C10] **Chuan Wu**, Baochun Li. "On Meeting P2P Streaming Bandwidth Demand with Limited Supplies," in the Proceedings of the *Fifteenth Annual SPIE/ACM International Conference on Multimedia Computing and Networking (MMCN 2008)*, Vol. 6818, pp. 68180C-1—68180C-12, San Jose, California, January 30-31, 2008.

[C11] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Multi-channel Live P2P Streaming: Refocusing on Servers," in the Proceedings of *IEEE INFOCOM 2008*, pp. 1355-1363, Phoenix, Arizona, April 14-17, 2008.

[C12] Zimu Liu, **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Distilling Superior Peers in Large-Scale P2P Streaming Systems," in the Proceedings of *IEEE INFOCOM 2009*, pp. 82-90, Rio de Janeiro, Brazil, April 19-25, 2009.

[C13] **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Diagnosing Network-wide P2P Live Streaming Inefficiencies," in the Proceedings of *IEEE INFOCOM 2009* Mini-Conference, pp. 2731-2735, Rio de Janeiro, Brazil, April 19-25, 2009.

[C14] Zimu Liu, **Chuan Wu**, Baochun Li, Shuqiao Zhao. "Why Are Peers Less Stable in Unpopular P2P Streaming Channels?" in the Proceedings of *IFIP Networking 2009*, also Lecture Notes in Computer Science, Springer, vol. 5550, pp. 274-286, Aachen, Germany, May 11-15, 2009.

[C15] Xuanjia Qiu, **Chuan Wu**, Xiaola Lin, Francis C.M. Lau, "InstantLeap: Fast Neighbor Discovery in P2P VoD Streaming," in the Proceedings of the *19th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2009)*, pp. 19-24, Williamsburg, Virginia, June 3-5, 2009.

[C16] Zimu Liu, **Chuan Wu**, Baochun Li, Shuqiao Zhao. "UUSee: Large-Scale Operational On-Demand Streaming with Random Network Coding," in the Proceedings of *IEEE INFOCOM 2010*, pp. 2070-2078, San Diego, California, March 15-19, 2010.

[C17] Wei Huang, **Chuan Wu**, Francis C.M. Lau. "The Performance and Local-

ity Tradeoff in BitTorrent-like P2P File-Sharing Systems," in the Proceedings of *IEEE International Conference on Communications (IEEE ICC 2010)*, Cape Town, South Africa, May 23-27, 2010.

[C18] Shun-Yun Hu, **Chuan Wu**, Eliya Buyukkaya, Chien-Hao Chien, Tzu-Hao Lin, Maha Abdallah, Jehn-Ruey Jiang, Kuan-Ta Chen. "A Spatial Publish Subscribe Overlay for Massively Multiuser Virtual Environments," in the Proceedings of the *2010 International Conference on Electronics and Information Engineering (ICEIE 2010)*, pp. V2-314 — V2-318, Kyoto, Japan, August 1-3, 2010.

[C19] Hongxing Li, Qiangsheng Hua, **Chuan Wu**, Francis C.M. Lau. "Minimum-Latency Aggregation Scheduling in Wireless Sensor Networks under Physical Interference Model," in the Proceedings of the *13th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (ACM MSWiM 2010)*, pp. 360-367, Bodrum, Turkey, October 17-21, 2010.

[C20] Zhi Wang, **Chuan Wu**, Lifeng Sun, Shiqiang Yang. "Strategies of Collaboration in Multi-channel P2P VoD Streaming," in the Proceedings of *IEEE Globecom 2010*, Miami, Florida, December 6-10, 2010.

[C21] Jin Wang, Jianping Wang, **Chuan Wu**, Kejie Lu, Naijie Gu. "Anonymous Communication with Network Coding against Traffic Analysis Attack," in the Proceedings of *IEEE INFOCOM 2011*, pp. 1008-1016, Shanghai, China, April 10-15, 2011.

[C22] Ajay Gopinathan, Zongpeng Li, **Chuan Wu**. "Strategyproof Auctions for Balancing Social Welfare and Fairness in Secondary Spectrum Markets," in the Proceedings of *IEEE INFOCOM 2011*, pp. 3020-3028, Shanghai, China, April 10-15, 2011.

[C23] Can Zhao, Xiaojun Lin, **Chuan Wu**. "The Streaming Capacity of Sparsely-Connected P2P Systems with Distributed Control," in the Proceedings of *IEEE INFOCOM 2011*, pp. 1449-1457, Shanghai, China, April 10-15, 2011.

[C24] Chenguang Kong, **Chuan Wu**, Victor O.K. Li. "SMS: Collaborative Streaming in Mobile Social Networks," in the Proceedings of *IFIP Networking 2011*, also Lecture Notes in Computer Science, Springer, vol. 6641, pp. 275-287, Valencia, Spain, May 9-13, 2011.

[C25] Zhi Wang, **Chuan Wu**, Lifeng Sun, Shiqiang Yang. "Peer-Assisted Online Games with Social Reciprocity," in the Proceedings of the *ACM/IEEE International Workshop on Quality of Service (ACM/IEEE IWQoS 2011)*, San Jose, California, USA, June 6-7, 2011.

[C26] Yu Wu, **Chuan Wu**, Bo Li, Xuanjia Qiu, Francis C.M. Lau. "CloudMedia: When Cloud on Demand Meets Video on Demand," in the Proceedings of the *31st International Conference on Distributed Computing Systems (IEEE ICDCS 2011)*, pp. 268-277, Minneapolis, Minnesota, USA, June 20-24, 2011.

[C27] Hongxing Li, Wei Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Utility-Maximizing Data Dissemination in Socially Selfish Cognitive Radio Networks," in the Proceedings of the *8th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS 2011)*, pp. 212-221, Valencia, Spain, October 17-21, 2011.

[C28] Ruiting Zhou, Zongpeng Li, **Chuan Wu**, Carey Williamson. "Physical Layer

Network Coding with Signal Alignment for MIMO Wireless Networks," in the Proceedings of the *8th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS 2011)*, pp. 530-539, Valencia, Spain, October 17-21, 2011.

[C29] Yu Wu, **Chuan Wu**, Bo Li, Linquan Zhang, Zongpeng Li, Francis C.M. Lau. "Scaling Social Media Applications into Geo-Distributed Clouds," in the Proceedings of *IEEE INFOCOM 2012*, pp. 684-692, Orlando, Florida, USA, March 25-30, 2012.

[C30] Hongxing Li, **Chuan Wu**, Zongpeng Li, Wei Huang, Francis C.M. Lau. "Stochastic Optimal Multirate Multicast in Socially Selfish Wireless Networks," in the Proceedings of *IEEE INFOCOM 2012*, pp. 172-180, Orlando, Florida, USA, March 25-30, 2012.

[C31] Xuanjia Qiu, Hongxing Li, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Cost-Minimizing Dynamic Migration of Content Distribution Services into Hybrid Clouds," in the Proceedings of *IEEE INFOCOM 2012* Mini-Conference, pp. 2853-2857, Orlando, Florida, USA, March 25-30, 2012.

[C32] Zhi Wang, Lifeng Sun, **Chuan Wu**, Shiqiang Yang. "Guiding Internet-Scale Video Service Deployment Using Microblog-based Prediction," in the Proceedings of *IEEE INFOCOM 2012* Mini-Conference, pp. 3183-3187, Orlando, Florida, USA, March 25-30, 2012.

[C33] Xuanjia Qiu, Hongxing Li, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Dynamic Scaling of VoD Services into Hybrid Clouds with Cost Minimization and QoS Guarantee," in the Proceedings of *the 19th International Packet Video Workshop (PV 2012)*, pp. 137-142, Munich, Germany, May 10-11, 2012.

[C34] Xingwei Wang, Jiajia Sun, Min Huang, **Chuan Wu**, Xueyi Wang. "A Resource Auction based Allocation Mechanism in the Cloud Computing Environment," in the Proceedings of the *26th IEEE International Parallel & Distributed Processing Symposium (IPDPS) Workshops*, pp. 2111-2115, Shanghai, China, May 21, 2012.

[C35] Hongxian Sun, **Chuan Wu**. "Epidemic Forwarding in Mobile Social Networks," in the Proceedings of *IEEE International Conference on Communications (IEEE ICC 2012)*, pp. 1421-1425, Ottawa, Canada, June 10-15, 2012.

[C36] Xuanjia Qiu, Wei Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Exploring the Sustainability of Credit-incentivized Peer-to-Peer Content Distribution," in the Proceedings of *the Fourth International Workshop on Hot Topics in Peer-to-Peer Computing and Online Social Networking (HotPOST), in conjunction with IEEE ICDCS 2012*, pp. 118-126, Maucau, June 18, 2012 [**Best Paper Award**]

[C37] Tang Xiahou, **Chuan Wu**, Jiaqing Huang, Zongpeng Li. "A Geometric Framework for Investigating the Multiple Unicast Network," in the Proceedings of *the 2012 IEEE International Symposium on Network Coding (NetCod 2012)*, pp. 37-42, Cambridge, Massachusetts, USA, June 29-30, 2012.

[C38] Zongpeng Li, **Chuan Wu**. "Space Information Flow: Multiple Unicast," in the Proceedings of *the 2012 IEEE International Symposium on Information Theory (ISIT 2012)*, Cambridge, Massachusetts, USA, July 1-6, 2012.

[C39] Ruiting Zhou, Zongpeng Li, **Chuan Wu**, Carey Williamson. "Buddy Rout-

ing: A Routing Paradigm for NanoNets Based on Physical Layer Network Coding," in the Proceedings of *the IEEE International Conference on Computer Communication Networks (ICCCN 2012)*, Munich, Germany, July 30-August 2, 2012.

[C40] Qiang-Sheng Hua, Haisheng Tan, Yuexuan Wang, Hongxing Li, Dongxiao Yu, Francis C.M. Lau, **Chuan Wu**. "Computing Capacity and Connectivity in Cognitive Radio Ad-Hoc Networks," in the Proceedings of *2012 International Symposium on Pervasive Systems, Algorithms, and Networks (ISPAN)*, pp. 9-16, San Marcos, Texas, December 13-15, 2012.

[C41] Hongxing Li, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Socially-optimal Multi-hop Secondary Communication under Arbitrary Primary User Mechanisms," in the Proceedings of *IEEE INFOCOM 2013*, pp. 1765-1773, Turin, Italy, April 14-19, 2013.

[C42] Can Zhao, Jian Zhao, Xiaojun Lin, **Chuan Wu**. "Capacity of P2P On-Demand Streaming with Simple, Robust and Decentralized Control," in the Proceedings of *IEEE INFOCOM 2013*, pp. 2797-2805, Turin, Italy, April 14-19, 2013.

[C43] Linquan Zhang, **Chuan Wu**, Zongpeng Li, Chuanxiong Guo, Minghua Chen, Francis C.M. Lau. "Moving Big Data to the Cloud," in the Proceedings of *IEEE INFOCOM 2013* Mini-Conference, pp. 405-409, Turin, Italy, April 14-19, 2013.

[C44] Hongxing Li, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Profit-Maximizing Virtual Machine Trading in a Federation of Selfish Clouds," in the Proceedings of *IEEE INFOCOM 2013* Mini-Conference, pp. 25-29, Turin, Italy, April 14-19, 2013.

[C45] Wei Deng, Fangming Liu, Hai Jin, **Chuan Wu**, Xue Liu. "MultiGreen: Cost-Minimizing Multi-source Datacenter Power Supply with Online Control," in the Proceedings of *the Fourth ACM International Conference on Future Energy Systems (ACM e-Energy)*, pp. 149-160, Berkley, California, USA, May 22-24, 2013.

[C46] Xuanjia Qiu, Wai Leong Yeow, **Chuan Wu**, Francis C.M. Lau. "Cost-Minimizing Preemptive Scheduling of MapReduce Workloads on Hybrid Clouds," in the Proceedings of the *ACM/IEEE International Symposium on Quality of Service (ACM/IEEE IWQoS 2013)*, Montreal, Canada, June 3-4, 2013.

[C47] Tang Xiahou, Zongpeng Li, **Chuan Wu**. "Information Multicast in (Pseudo-)Planar Networks: Efficient Network Coding over Small Finite Fields," in the Proceedings of *the 2013 IEEE International Symposium on Network Coding (NetCod 2013)*, Calgary, Canada, June 7-9, 2013.

[C48] Wei Deng, Fangming Liu, Hai Jin, **Chuan Wu**. "SmartDPSS: Cost-Minimizing Multi-source Power Supply for Datacenters with Arbitrary Demand," in the Proceedings of *the 33rd International Conference on Distributed Computing Systems (ICDCS 2013)*, pp. 420-429, Philadelphia, USA, July 8-11, 2013.

[C49] Ming He, Zongpeng Li, **Chuan Wu**, Xunrui Yin. "On Uniform Matroidal Networks," in the Proceedings of *the 2013 IEEE International Symposium on Information Theory (ISIT 2013)*, Istanbul, Turkey, July 7-12, 2013.

[C50] Yu Wu, **Chuan Wu**, Bo Li, Francis C.M. Lau. "vSkyConf: Cloud-assisted Multi-party Mobile Video Conferencing," in the Proceedings of the *Second Mobile Cloud Computing (MCC) Workshop, in conjunction with ACM SIGCOMM 2013*, pp. 33-38, Hong Kong, August 12, 2013.

[C51] Jian Zhao, Hongxing Li, **Chuan Wu**, Zongpeng Li, Zhizhong Zhang, Francis C.M. Lau. "Dynamic Pricing and Profit Maximization for the Cloud with Geo-distributed Data Centers," in the Proceedings of *IEEE INFOCOM 2014*, pp. 118-126, Toronto, Canada, April 27-May 2, 2014.

[C52] Zhi Wang, Lifeng Sun, **Chuan Wu**, Wenwu Zhu, Shiqiang Yang. "Joint Online Transcoding and Geo-distributed Delivery for Dynamic Adaptive Streaming," in the Proceedings of *IEEE INFOCOM 2014*, pp. 91-99, Toronto, Canada, April 27-May 2, 2014.

[C53] Linquan Zhang, Zongpeng Li, **Chuan Wu**. "Dynamic Resource Provisioning in Cloud Computing: A Randomized Auction Approach," in the Proceedings of *IEEE INFOCOM 2014*, pp. 433-441, Toronto, Canada, April 27-May 2, 2014.

[C54] Linquan Zhang, Zongpeng Li, **Chuan Wu**, Minghua Chen. "Online Algorithms for Uploading Deferrable Big Data to The Cloud," in the Proceedings of *IEEE INFOCOM 2014*, pp. 2022-2030, Toronto, Canada, April 27-May 2, 2014.

[C55] Weijie Shi, **Chuan Wu**, Zongpeng Li. "RSMOA: A Revenue and Social Welfare Maximizing Online Auction for Dynamic Cloud Resource Provisioning," in the Proceedings of the *ACM/IEEE International Symposium on Quality of Service (ACM/IEEE IWQoS 2014)*, pp. 166-175, Hong Kong, May 26-27, 2014.

[C56] Shengkai Shi, Zhi Wang, **Chuan Wu**, Xiaojun Lin. "Characterizing Cascade Dynamics in A Microblogging System," in the Proceedings of *IEEE International Conference on Communications (IEEE ICC 2014)*, Sydney, Australia, June 10-14, 2014.

[C57] Weijie Shi, Linquan Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "An Online Auction Framework for Dynamic Resource Provisioning in Cloud Computing," in the Proceedings of *ACM SIGMETRICS 2014*, pp. 71-83, Austin, Texas, USA, June 16-20, 2014.

[C58] Linquan Zhang, Zongpeng Li, **Chuan Wu**. "Randomized Auction Design for Electricity Markets between Grids and Microgrids," in the Proceedings of *ACM SIGMETRICS 2014*, pp. 99-110, Austin, Texas, USA, June 16-20, 2014.

[C59] Xuanjia Qiu, **Chuan Wu**, Hongxing Li, Zongpeng Li, Francis C.M. Lau. "Federated Private Clouds via Broker's Marketplace: A Stackelberg-Game Perspective," in the Proceedings of *the 7th IEEE International Conference on Cloud Computing (IEEE Cloud 2014)*, Alaska, USA, June 27-July 2, 2014

[C60] Haoming Fu, Zongpeng Li, **Chuan Wu**, Xiaowen Chu. "Core-Selecting Auctions for Dynamically Allocating Heterogeneous VMs in Cloud Computing," in the Proceedings of *the 7th IEEE International Conference on Cloud Computing (IEEE Cloud 2014)*, Alaska, USA, June 27-July 2, 2014

[C61] Jian Zhao, **Chuan Wu**. "Socially-optimal ISP-aware P2P Content Distribution via a Primal-Dual Approach," in the Proceedings of *the Sixth International Workshop on Hot Topics in Peer-to-Peer Computing and Online Social Networking (Hot-POST), in conjunction with IEEE ICDCS 2014*, Madrid, Spain, June 30, 2014

[C62] Xiaoxi Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "A Truthful $(1 - \epsilon)$-Optimal Mechanism for On-demand Cloud Resource Provisioning," in the Proceedings of *IEEE INFOCOM 2015*, pp. 1053-1061, Hong Kong, April 26-May 1,

2015.

[C63] Linquan Zhang, Shaolei Ren, **Chuan Wu**, Zongpeng Li. "A Truthful Incentive Mechanism for Emergency Demand Response in Colocation Data Centers," in the Proceedings of *IEEE INFOCOM 2015*, pp. 2632-2640, Hong Kong, April 26-May 1, 2015.

[C64] Hongxing Li, **Chuan Wu**, Zongpeng Li. "Socially-Optimal Online Spectrum Auctions for Secondary Wireless Communication," in the Proceedings of *IEEE INFOCOM 2015*, pp. 2047-2055, Hong Kong, April 26-May 1, 2015.

[C65] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "An Online Procurement Auction for Power Demand Response in Storage-Assisted Smart Grids," in the Proceedings of *IEEE INFOCOM 2015*, pp. 2641-2649, Hong Kong, April 26-May 1, 2015.

[C66] Jingpu Duan, Zhi Wang, **Chuan Wu**. "Responsive Multipath TCP in SDN-based Datacenters," in the Proceedings of *IEEE International Conference on Communications (IEEE ICC 2015)*, London, UK, June 8-12, 2015.

[C67] Xiaoxi Zhang, Zhiyi Huang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Online Auctions in IaaS Clouds: Welfare and Profit Maximization with Server Costs," in the Proceedings of *ACM SIGMETRICS 2015*, pp. 3-16, Portland, Oregon, USA, June 15-19, 2015.

[C68] Linquan Zhang, Zongpeng Li, **Chuan Wu**, Shaolei Ren. "Online Electricity Cost Saving Algorithms for Co-Location Data Centers," in the Proceedings of *ACM SIGMETRICS 2015* (poster paper), pp. 463-464, Portland, Oregon, USA, June 15-19, 2015.

[C69] Qihang Sun, **Chuan Wu**, Shaolei Ren, Zongpeng Li. "Fair Rewarding in Colocation Data Center: Truthful Mechanism for Emergency Demand Response," in the Proceedings of the *ACM/IEEE International Symposium on Quality of Service (ACM/IEEE IWQoS 2015)*, Portland, Oregon, USA, June 15-16, 2015.

[C70] Xiaoxi Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Online Cost Minimization for Operating Geo-distributed Cloud CDNs," in the Proceedings of the *ACM/IEEE International Symposium on Quality of Service (ACM/IEEE IWQoS 2015)*, Portland, Oregon, USA, June 15-16, 2015.

[C71] Shengkai Shi, **Chuan Wu**, Zongpeng Li. "Cost-Minimizing Online VM Purchasing for Application Service Providers with Arbitrary Demands," in the Proceedings of *the 8th IEEE International Conference on Cloud Computing (IEEE Cloud 2015)*, pp. 146-154, New York, USA, June 27-July 2, 2015.

[C72] Linquan Zhang, Xunrui Yin, Zongpeng Li, **Chuan Wu**. "Hierarchical Virtual Machine Placement in Modular Data Centers," in the Proceedings of *the 8th IEEE International Conference on Cloud Computing (IEEE Cloud 2015)*, pp. 171-178, New York, USA, June 27-July 2, 2015.

[C73] Mohammad H. Hajiesmaili, Lok To Mak, Zhi Wang, **Chuan Wu**, Minghua Chen, Ahmad Khonsari. "Cost-Effective Low-Delay Cloud Video Conferencing," in the Proceedings of the *35th International Conference on Distributed Computing Systems (IEEE ICDCS 2015)*, pp. 103-112, Columbus, Ohio, USA, June 29-July 2, 2015.

[C74] Shunyi Xu, **Chuan Wu**, Zongpeng Li. "Software Defined Mobile Multicast," in the Proceedings of *the 12th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS 2015)*, Dallas, Texas, USA, October 19-22, 2015.

[C75] Ajay Gopinathan†, Niklas Carlsson, Zongpeng Li, **Chuan Wu**. "Revenue-maximizing and Truthful Online Auctions for Dynamic Spectrum Access," in the Proceedings of *2016 IEEE/IFIP Wireless On-demand Network systems and Services Conference (WONS 2016)*, Cortina d'Ampezzo, Italy, January 20-22, 2016.

[C76] Weijie Shi, **Chuan Wu**, Zongpeng Li. "An Online Mechanism for Dynamic Virtual Cluster Provisioning in Geo-Distributed Clouds," in the Proceedings of *IEEE INFOCOM*, San Francisco, CA, USA, April 10-15, 2016.

[C77] Sijia Gu, Zongpeng Li, **Chuan Wu**, Chuanhe Huang. "An Efficient Auction Mechanism for Service Chains in The NFV Market," in the Proceedings of *IEEE INFOCOM*, San Francisco, CA, USA, April 10-15, 2016.

[C78] Yixin Bao, Xiaoke Wang, Zhi Wang, **Chuan Wu**, Francis C.M. Lau. "Online Influence Maximization in Non-Stationary Social Networks," in the Proceedings of the *ACM/IEEE International Symposium on Quality of Service (ACM/IEEE IWQoS 2016)*, Beijing, China, June 20-21, 2016.

[C79] Ruiting Zhou, Zongpeng Li, **Chuan Wu**. "An Emergency Demand Response Mechanism for Cloud Computing," in the Proceedings of *the GreenMetrics Workshop, in conjuction with ACM SIGMETRICS 2016*, Antibes Juan-les-Pins, France, June 14, 2016. [**Best Student Paper Runner-up**]

[C80] Qihang Sun, Shaolei Ren, **Chuan Wu**, Zongpeng Li. "An Online Incentive Mechanism for Emergency Demand Response in Geo-Distributed Colocation Data Centers," in the Proceedings of *the Seventh ACM International Conference on Future Energy Systems (ACM e-Energy)*, Waterloo, Canada, June 21-24, 2016. [**Best Paper Award**]

[C81] Xiaoke Wang, **Chuan Wu**, Franck Le, Alex Liu, Zongpeng Li, Francis C.M. Lau. "Online VNF Scaling in Datacenters," in the Proceedings of *the 9th IEEE International Conference on Cloud Computing (IEEE Cloud 2016)*, San Francisco, California, USA, June 27-July 2, 2016.

[C82] Zhixiong Niu, Zhi Wang, Hong Xu, **Chuan Wu**, Francis C.M. Lau. "More is Better? Measurement of MPTCP based Cellular Bandwidth Aggregation in the Wild," in the Proceedings of *the 13th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS 2016)*, Brasilia, Brazil, October 10-13, 2016.

[C83] Zhichao Zhao, Fei Chen, T-H. Hubert Chan, **Chuan Wu**. "Double Auction for Resource Allocation in Cloud Computing," in the Proceedings of the Seventh International Conference on Cloud Computing and Services Science (CLOSER), Porto, Portugal, April 24-26, 2017.

[C84] Xiaoxi Zhang, **Chuan Wu**, Zongpeng Li, Francis C.M. Lau. "Proactive VNF Provisioning with Multi-timescale Cloud Resources: Fusing Online Learning and Online Optimization," in the Proceedings of *IEEE INFOCOM*, Atlanta, GA, USA, May 1-4, 2017.

[C85] Linquan Zhang, Shangqi Lai, **Chuan Wu**, Zongpeng Li, Chuanxiong Guo.

"Virtualized Network Coding Functions on the Internet," in the Proceedings of *the 37th IEEE International Conference on Distributed Computing Systems (ICDCS)*, Atlanta, GA, USA, June 5-8, 2017.

[C86] Zijun Zhang, Zongpeng Li, **Chuan Wu**, Chuanhe Huang. "A Scalable and Distributed Approach for NFV Service Chain Cost Minimization," in the Proceedings of *the 37th IEEE International Conference on Distributed Computing Systems (ICDCS)* (short paper), Atlanta, GA, USA, June 5-8, 2017.

[C87] Zijun Zhang, Zongpeng Li, **Chuan Wu**. "Optimal Posted Prices for Online Cloud Resource Allocation," in *ACM SIGMETRICS 2017*, Urbana-Champaign, Illinois, USA, June 5-9, 2017.

[C88] Xiaoke Wang, **Chuan Wu**, Franck Le, Francis C.M. Lau. "Online Learning-Assisted VNF Service Chain Scaling with Network Uncertainties," in the Proceedings of *the 10th IEEE International Conference on Cloud Computing (IEEE Cloud 2017)*, Honolulu, Hawaii, USA, June 25-30, 2017.

[C89] Yanghua Peng, Ji Yang, **Chuan Wu**, Chuanxiong Guo, Chengchen Hu, Zongpeng Li. "deTector: a Topology-aware Monitoring System for Data Center Networks," in the Proceedings of *the 2017 USENIX Annual Technical Conference (USENIX ATC'17)*, Santa Clara, CA, USA, July 12-14, 2017.

[C90] Xiaodong Yi, Jingpu Duan, **Chuan Wu**. "GPUNFV: a GPU-Accelerated NFV System," in the Proceedings of *the 1st Asia-Pacific Workshop on Networking (APNet'17)*, Hong Kong, August 3-4, 2017.

[C91] Yixin Bao, Yanghua Peng, **Chuan Wu**, Zongpeng Li. "Online Job Scheduling in Distributed Machine Learning Clusters," in the Proceedings of *IEEE INFOCOM*, Honolulu, HI, USA, April 15-19, 2018.

[C92] Xiaoxi Zhang, **Chuan Wu**, Zhiyi Huang, Zongpeng Li. "Occupation-Oblivious Pricing of Cloud Jobs via Online Learning," in the Proceedings of *IEEE INFOCOM*, Honolulu, HI, USA, April 15-19, 2018.

[C93] Yanghua Peng, Yixin Bao, Yangrui Chen, **Chuan Wu**, Chuanxiong Guo. "Optimus: An Efficient Dynamic Resource Scheduler for Deep Learning Clusters," in the Proceedings of *EuroSys 2018*, Porto, Portugal, April 23-26, 2018.

[C94] Ziyue Luo, Zongpeng Li, **Chuan Wu**. "Online Cloud Resource Allocation and Pricing with Server Speed Scaling," in the Proceedings of *IEEE International Conference on Communications (IEEE ICC 2018)*, Kansas City, MO, USA, May 20-24, 2018.

[C95] Bingqian Du, **Chuan Wu**, Zhiyi Huang. "Learning Resource Allocation and Pricing for Cloud Profit Maximization," in the Proceedings of *the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, Honolulu, Hawaii, USA, January 27-February 1, 2019.

[C96] Yixin Bao, Yanghua Peng, **Chuan Wu**. "Deep Learning-based Job Placement in Distributed Machine Learning Clusters," in the Proceedings of *IEEE INFOCOM*, Paris, France, April 29 - May 2, 2019.

[C97] Qihang Sun, Shaolei Ren, **Chuan Wu**. "A Provably-Efficient Online Algo-

rithm for Re-Utilizing Unused VM Resources for Edge Providers," in the Proceedings of *IEEE International Conference on Communications (IEEE ICC 2019)*, Shanghai, China, May 20-24, 2019.

[C98] Jing Li, Walter Yu Hang Lam, Richard Tai Chiu Hsung, Edmond Ho Nang Pow, **Chuan Wu**, Zheng Wang. "Control and Motion Scaling of A Compact Cable-driven Dental Robotic Manipulator," in the Proceedings of *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Hong Kong, July 8-12, 2019.

[C99] Xiaodong Yi, Junjie Wang, Jingpu Duan, Wei Bai, **Chuan Wu**, Yongqiang Xiong, Dongsu Han. "FlowShader: a Generalized Framework for GPU-accelerated VNF Flow Processing," in the Proceedings of *The 27th IEEE International Conference on Network Protocols (IEEE ICNP 2019)*, Chicago, Illinois, USA, October 7-10, 2019.

[C100] Yanghua Peng, Yibo Zhu, Yangrui Chen, Yixin Bao, Bairen Yi, Chang Lan, **Chuan Wu**, Chuanxiong Guo. "A Generic Communication Scheduler for Distributed DNN Training Acceleration," in the Proceedings of *the 27th ACM Symposium on Operating Systems Principles (ACM SOSP 2019)*, Huntsville, Ontario, Canada, October 27-30, 2019.

[C101] Mengdi Wang, Chen Meng, Guoping Long, **Chuan Wu**, Jun Yang, Wei Lin. "Characterizing Deep Learning Training Workloads on a Large AI Cloud," in the Proceedings of *2019 IEEE International Symposium on Workload Characterization (IISWC-2019)*, Orlando, Florida, USA, November 3-5, 2019.

[C102] Hanpeng Hu, Dan Wang, **Chuan Wu**. "Distributed Machine Learning through Heterogeneous Edge Systems," in the Proceedings of *the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)*, New York, USA, February 7-12, 2020.

[C103] Yixin Bao, Yanghua Peng, Yangrui Chen, **Chuan Wu**. "Preemptive All-reduce Scheduling for Expediting Distributed DNN Training," in the Proceedings of *IEEE INFOCOM*, Toronto, Canada, July 6-9, 2020.

[C104] Xiaoyang Zhao, **Chuan Wu**, Franck Le. "Improving Inter-domain Routing through Multi-agent Reinforcement Learning," in the Proceedings of *the Second International Workshop on Data Driven Intelligence for Networks* (DDINS), colocated with IEEE INFOCOM 2020, Toronto, Canada, July 6, 2020.

[C105] Qin Zhang, Ruiting Zhou, **Chuan Wu**, Lei Jiao, Zongpeng Li. "Online Scheduling of Heterogeneous Distributed Machine Learning Jobs," in the Proceedings of *ACM International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (Mobihoc)*, online, October 11-14, 2020.

[C106] Yangrui Chen, Yanghua Peng, Yixin Bao, **Chuan Wu**, Yibo Zhu, Chuanxiong Guo. "Elastic Parameter Server Load Distribution in Deep Learning Clusters," in the Proceedings of *ACM Symposium on Cloud Computing 2020 (SOCC'20)*, Renton, WA, USA, October 19-21, 2020.

[C107] Xiaodong Yi, Shiwei Zhang, Ziyue Luo, Guoping Long, Lansong Diao, **Chuan Wu**, Zhen Zheng, Jun Yang, Wei Lin. "Optimizing Distributed Training Deployment in Heterogeneous GPU Clusters," in the Proceedings of *the 16th ACM International Conference on emerging Networking EXperiments and Technologies (ACM CoNEXT)*, Barcelona, Spain, December 1-4, 2020.

[C108] Xiaodong Yi, Ziyue Luo, Chen Meng, Mengdi Wang, Guoping Long, **Chuan Wu**, Jun Yang, Wei Lin. "Fast Training of Deep Learning Models over Multiple GPUs," in the Proceedings of *ACM/IFIP Middleware 2020*, Delft, The Netherlands, December 7-11, 2020.

[C109] Shiqing Fan, Yi Rong, Chen Meng, Zongyan Cao, Siyu Wang, Zheng Zhen, **Chuan Wu**, Guoping Long, Jun Yang, Lixue Xia, Lansong Diao, Xiaoyong Liu, Wei Lin. "DAPPLE: A Pipelined Data Parallel Approach for Large Models Training," in the Proceedings of *the 26th ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming (PPoPP'21)*, Seoul, South Korea, February 27-March 3, 2021.

[C110] Zhe Zhang, **Chuan Wu**, Zongpeng Li. "Near-Optimal Topology-adaptive Parameter Synchronization in Distributed DNN Training," in the Proceedings of *IEEE INFOCOM*, online, May 10-13, 2021.

[C111] Menglu Yu, **Chuan Wu**, Bo Ji, Jia Liu. "A Sum-of-Ratios Multi-Dimensional-Knapsack Decomposition for DNN Resource Scheduling," in the Proceedings of *IEEE INFOCOM*, online, May 10-13, 2021.

[C112] Jingyan Jiang, Ziyue Luo, Chenghao Hu, Zhaoliang He, Zhi Wang, Shutao Xia, **Chuan Wu**. "Joint Model and Data Adaptation for Cloud Inference Serving," in the Proceedings of *the 42th IEEE Real-time Systems Symposuim (RTSS)*, Dortmund, Germany, December 7-10, 2021.

[C113] Ziyue Luo, Xiaodong Yi, Guoping Long, Shiqing Fan, **Chuan Wu**, Wei Lin. "Efficient Pipeline Planning for Expedited Distributed DNN Training," in the Proceedings of *IEEE INFOCOM*, online, May 2-5, 2022.

[C114] Ziyue Luo, Yixin Bao, **Chuan Wu**. "Optimizing Task Placement and Online Scheduling for Distributed GNN Training Acceleration," in the Proceedings of *IEEE INFOCOM*, online, May 2-5, 2022.

[C115] Menglu Yu, Ye Tian, Bo Ji, **Chuan Wu**, Hridesh Rajan, Jia Liu. "GADGET: Online Resource Optimization for Scheduling Ring-All-Reduce Learning Jobs," in the Proceedings of *IEEE INFOCOM*, online, May 2-5, 2022.

[C116] Chongyu Zhao, Wenzhi Guo, Rongwei Wen, Zheng Wang, **Chuan Wu**. "Deep Learning-driven Front-Following within Close Proximity: a Hands-Free Control Model on a Smart Walker," in the Proceedings of *the 2022 IEEE International Conference on Robotics and Automation (ICRA)*, Philadelphia, USA, May 23-27, 2022.

[C117] Bingqian Du, **Chuan Wu**. "Federated Graph Learning with Periodic Neighbour Sampling," in the Proceedings of *IEEE/ACM International Symposium on Quality of Service (IWQoS)*, online, June 10-12, 2022.

[C118] Hanpeng Hu, Chenyu Jiang, Yuchen Zhong, Yanghua Peng, **Chuan Wu**, Yibo Zhu, Haibin Lin. "dPRO: A Generic Performance Diagnosis and Optimization Toolkit for Expediting Distributed DNN Training," in the Proceedings of *the Fifth Conference on Machine Learning and Systems (MLSys)*, August 29 - September 1, 2022.

[C119] Shiwei Zhang, Lansong Diao, **Chuan Wu**, Siyu Wang, Wei Lin. "Accelerating Large-Scale Distributed Neural Network Training with SPMD Parallelism,"

in the Proceedings of *the 13th ACM Symposium on Cloud Computing (SOCC'22)*, San Francisco, CA, Nobember 7-11, 2022.

[C120] Yangrui Chen, Cong Xie, Meng Ma, Juncheng Gu, Yanghua Peng, Haibin Lin, **Chuan Wu**, Yibo Zhu. "SAPipe: Staleness-Aware Pipeline for Data Parallel DNN Training," in the Proceedings of *the Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS)*, New Orleans, USA, November 29-December 1, 2022.

[C121] Yangrui Chen*, Tianfeng Liu*, Dan Li, **Chuan Wu**, Yibo Zhu, Jun He, Yanghua Peng, Hongzheng Chen, Hongzhi Chen, Chuanxiong Guo. "BGL: GPU-Efficient GNN Training by Optimizing Graph Data I/O and Preprocessing," in the Proceedings of *the 20th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Boston, MA, USA, April 17-19, 2023. (*: co-first authors)

[C122] Yuchen Zhong, Guangming Sheng, Juncheng Liu, Jinhui Yuan, **Chuan Wu**. "Swift: Expedited Failure Recovery for Large-scale DNN Training," in the Proceedings of *the 28th ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming (PPoPP'23)*, Montreal, Canada, Feb 25 - Mar 1, 2023 (poster paper).

[C123] Zhe Zhang, Ziyue Luo, **Chuan Wu**. "Two-level Graph Caching for Expediting Distributed GNN Training," in the Proceedings of *IEEE INFOCOM*, New York, USA, May 17-20, 2023.

[C124] Yihong Li, Tianyu Zeng, Xiaoxi Zhang, Jingpu Duan, **Chuan Wu**. "TapFinger: Task Placement and Fine-Grained Resource Allocation for Edge Machine Learning," in the Proceedings of *IEEE INFOCOM*, New York, USA, May 17-20, 2023.

[C125] Borui Wan, Juntao Zhao, **Chuan Wu**. "Adaptive Message Quantization and Parallelization for Distributed Full-graph GNN Training," in the Proceedings of *the Sixth Conference on Machine Learning and Systems (MLSys)*, Miami, USA, June 4 - 8, 2023.

[C126] Le Fang, **Chuan Wu**. "HES: Edge Sampling for Heterogeneous Graphs," in the Proceedings of *the International Joint Conference on Neural Networks (IJCNN)*, Queensland, Australia, June 18 - 23, 2023.

[C127] Jiyue Jiang, Sheng Wang, Qintong Li, Lingpeng Kong, **Chuan Wu**. "A Cognitive Stimulation Therapy Dialogue System with Multi-Source Knowledge Fusion for Elders with Cognitive Impairment," in the Proceedings of *The 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, Toronto, Canada, July 9-14, 2023.

[C128] Junwei Su, Difan Zou, Zijun Zhang, **Chuan Wu**. "Towards Robust Graph Incremental Learning on Evolving Graphs," in the Proceedings of *the Fortieth International Conference on Machine Learning (ICML)*, Honolulu, Hawaii, USA, July 23-29, 2023.

[C129] Yuran Sun, Alan Dougherty, Zhuoying Zhang, Yi King Choi, **Chuan Wu**. "MixSynthFormer: A Transformer Encoder-like Structure with Mixed Synthetic Self-attention for Efficient Human Pose Estimation," in the Proceedings of *the International Conference on Computer Vision (ICCV)*, Paris, France, October 2-6, 2023.

[C130] Chenyu Jiang, Zhen Jia, Yida Wang, Shuang Zheng, **Chuan Wu**. "DynaPipe: Optimizing Multi-task Training through Dynamic Pipelines," in the Proceedings of *EuroSys 2024*, Athens, Greece, April 22-25, 2024.

[C131] Hanpeng Hu, Junwei Su, Juntao Zhao, Yanghua Peng, Yibo Zhu, Haibin Lin, **Chuan Wu**. "CDMPP: A Device-Model Agnostic Framework for Latency Prediction of Tensor Programs," in the Proceedings of *EuroSys 2024*, Athens, Greece, April 22-25, 2024.

[C132] Shiwei Zhang, Lansong Diao, **Chuan Wu**, Zongyan Cao, Siyu Wang, Wei Lin. "HAP: SPMD DNN Training on Heterogeneous GPU Clusters with Automated Program Synthesis," in the Proceedings of *EuroSys 2024*, Athens, Greece, April 22-25, 2024.

[C133] Juntao Zhao, Borui Wan, **Chuan Wu**, Yanghua Peng, Haibin Lin. "LLM-PQ: Serving LLM on Heterogeneous Clusters with Phase-Aware Partition and Adaptive Quantization," in the Proceedings of *the 29th ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming (PPoPP'24)*, Edinburgh, United Kingdom, March 2-6, 2024 (poster paper).

[C134] Junwei Su, Difan Zou, **Chuan Wu**. "PRES: Toward Scalable Memory-Based Dynamic Graph Neural Networks," in the Proceedings of *The Twelfth International Conference on Learning Representations (ICLR)*, Vienna, Austria, May 7-11, 2024.

[C135] Xiaoxi Zhang, Haoran Xu, Longhao Zou, Jingpu Duan, **Chuan Wu**, Yali Xue, Zouzhou Chen, Xu Chen. "Rosevin: Employing Resource- and Rate-Adaptive Edge Super-Resolution for Video Streaming," in the Proceedings of *IEEE INFOCOM*, Vancouver, Canada, May 20-23, 2024.

[C136] Bingqian Du, Jun Liu, Ziyue Luo, **Chuan Wu**, Qiankun Zhang, Hai Jin. "Expediting Distributed GNN Training with Feature-only Partition and Optimized Communication Planning," in the Proceedings of *IEEE INFOCOM*, Vancouver, Canada, May 20-23, 2024.

[C137] Juntao Zhao, Borui Wan, Yanghua Peng, Haibin Lin, Yibo Zhu, **Chuan Wu**. "QSync: Quantization-Minimized Synchronous Distributed Training Across Hybrid Devices," in the Proceedings of *the 38th IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, San Francisco, USA, May 27-31, 2024.

[C138] Ye Tian, Zhen Jia, Ziyue Luo, Yida Wang, **Chuan Wu**. "DiffusionPipe: Training Large Diffusion Models with Efficient Pipelines," in the Proceedings of *the Seventh Conference on Machine Learning and Systems (MLSys)*, Santa Clara, USA, May 13 - 16, 2024.

[C139] Chenyu Jiang, Ye Tian, Zhen Jia, **Chuan Wu**, Yida Wang, Shuai Zheng. "Lancet: Accelerating Mixture-of-Experts Training by Overlapping Weight Gradient Computation and All-to-All Communication," in the Proceedings of *the Seventh Conference on Machine Learning and Systems (MLSys)*, Santa Clara, USA, May 13 - 16, 2024.

PATENTS

[P1] Invention: "Smart Walker"; Hong Kong Standard Patent No.: HK 30070283; Date of Grant: October 21, 2022; Duration: 8 years commencing on May 20, 2022

OPEN-SOURCE PROJECTS

[O1] "NFVActor: A Resilient NFV System using the Distributed Actor Model":

project source code available at https://github.com/duanjp8617/nfvactor

[O2] "deTector: a Topology-aware Monitoring System for Data Center Networks":
project source code available at
https://github.com/pengyanghua/deTector

[O3] "NetStar: a Future/Promise Framework for Asynchronous Network Functions": project source code available at
https://github.com/duanjp8617/netstar-project

[O4] "Optimus: An Efficient Dynamic Resource Scheduler for Deep Learning Clusters": project source code available at
https://github.com/pengyanghua/optimus

[O5] "A Generic Communication Scheduler for Distributed DNN Training Acceleration": project source code available at (Star: 3.5k)
https://github.com/bytedance/byteps/tree/bytescheduler/bytescheduler

[O6] "Optimizing Distributed Training Deployment in Heterogeneous GPU Clusters": project source code available at
https://github.com/eval-submissions/HeteroG

[O7] "Fast Training of Deep Learning Models over Multiple GPUs": project source
code available at
https://github.com/eval-submissions/FastT

[O8] "DL^2: A Deep Learning-driven Scheduler for Deep Learning Clusters": project
source code available at https://github.com/pengyanghua/DL2

[O9] "dPRO: A Generic Performance Diagnosis and Optimization Toolkit for Expediting Distributed DNN Training": project source code available at
https://github.com/joapolarbear/dpro

[O10] "Optimizing DNN Compilation for Distributed Training with Joint OP and
Tensor Fusion": https://github.com/TPDS-Submission/Disco

[O11] "Expediting Distributed DNN Training with Device Topology-Aware Graph
Deployment": https://github.com/ylxdzsw/tge

[O12] "Accelerating Large-Scale Distributed Neural Network Training with SPMD
Parallelism": https://github.com/ylxdzsw/hidup

[O13] "Adaptive Message Quantization and Parallelization for Distributed Full-
graph GNN Training": https://github.com/raywan-110/AdaQP

[O14] QSync, https://github.com/bytedance/QSync

[O15] GNNFlow, https://github.com/jasperzhong/GNNFlow

[O16] "CDMPP: A Device-Model Agnostic Framework for Latency Prediction of
Tensor Programs," https://github.com/joapolarbear/cdmpp

[O17] "DynaPipe: Optimizing Multi-task Training through Dynamic Pipelines,"
https://github.com/awslabs/optimizing-multitask-training-through-dynamic-pipelines

[O18] "HAP: SPMD DNN Training on Heterogeneous GPU Clusters with Automated Program Synthesis," https://github.com/alibaba/hap

EDITORIAL
BOARD
MEMBERSHIP

◇ Associate Editor, IEEE Transactions on Multimedia, July 2017 - July 2021

◇ Associate Editor, ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2017-present

◇ Associate Editor, IEEE Transactions on Cloud Computing, March 2018 - October 2022

◇ Associate Editor, IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), February 2019 - Dec 31, 2022

◇ Associate Editor, IEEE/ACM Transactions on Networking (ToN), January 2021 - present

KEYNOTE,
PLENARY AND
OTHER INVITED
TALKS/LECTURES

[K1] Invited talk, "Large-Scale Peer-to-Peer Streaming: Measurements and Optimizing Solutions," *The Hong Kong Applied Science and Technology Research Institute (ASTRI)*, June, 2009.

[K2] Invited lecture, "Introduction to Convex Optimization: Optimal Rate Control in Communication Networks", *Winter School, Institute of Theoretical Computer Science and Communications, The Chinese University of Hong Kong*, January, 2010.

[K3] Invited talk, "CloudMedia: When Cloud on Demand Meets Video on Demand," *Institute for Infocomm Research (I2R)*, Singapore, August 16, 2011.

[K4] Invited talk, "Dynamic, Optimal Migration of Content Distribution Applications into Geo-distributed Clouds," *Shenzhen Institutes of Advanced Technology*, Shenzhen, China, March 21, 2012.

[K5] Invited talk, "Online Auctions in IaaS Clouds," *Microsoft Research*, Redmond, Washington, U.S.A., May 19, 2016.

[K6] Invited lecture, "An Introduction to Convex Optimization," *Tsinghua-Berkeley Shenzhen Institute*, Shenzhen, China, May 12, 2017.

[K7] Keynote speech, "Mechanism Design for Emergency Demand Response in Colocation Data Centres," *the 6th International Workshop on Energy-Efficient Data Centres*, Colocated with *the 8th International Conference on Future Energy Systems (ACM e-Energy 2017)*, Hong Kong, May 16, 2017.

[K8] Invited talk, "Optimal Posted Prices for Online Cloud Resource Allocation," *Purdue University*, West Lafayette, IN, USA, June 9, 2017.

[K9] Invited talk, "Online Resource Scheduling and Pricing in IaaS Clouds," *Alibaba*, Hangzhou, China, Jan 11, 2018.

[K10] Invited talk, "Dynamic Resource Scheduling in Distributed Deep Learning Clusters," *Simon Fraser University*, Burnaby, BC, Canada, June 7, 2018.

[K11] Invited talk, "Dynamic Resource Scheduling in Distributed Deep Learning Clusters," *Wireless AI Research group, Huawei Technologies Co. Ltd*, Shanghai, China, August 24, 2018.

[K12] Plenary talk, "Efficient Resource Scheduling for Machine Learning Clusters," MSRA Computing in the 21st Century & Asia Faculty Summit, Beijing, China, November 6, 2018.

[K14] Keynote speech, "Efficient Resource Scheduling in Machine Learning Clusters," *the 9th International Workshop on Big Data and Cloud Performance (DCPerf'19)*, Colocated with *IEEE INFOCOM 2019*, Paris, France, April 29, 2019.

[K15] Plenary talk, "Resource Scheduling for Distributed Deep Training," *MSRA Academic Day System & Networking for AI Workshop*, Beijing, China, November 7, 2019.

[K16] Plenary talk, "Resource Scheduling for Distributed Deep Training in AI Clouds," *Huawei 2019 Network World - Data Center Network Technology Forum*, Nanjing, Jiangsu, China, November 18, 2019.

[K17] Plenary talk, "Optimizing Placement and Scheduling of Distributed DNN Training in AI clouds," *Huawei STW (Strategy and Technology Workshop) 2021*, online, October 15, 2021.

[K18] Invited talk, "Optimizing Placement and Scheduling of Distributed DNN Training in AI clouds," *Dalian University of Technology, China*, online, March 7, 2022.

[K19] Invited talk, "Performance Diagnosis and Optimization of Distributed DNN Training," *The 3rd Huawei Workshop on Intelligent Networks*, HKSTP, September 8, 2022.

[K20] Invited talk, "Performance Diagnosis and Optimization of Distributed DNN Training," *Hong Kong Science and Technology University (Guang Zhou)*, March 1, 2023.

[K21] Invited talk, "Optimizing Distributed GNN Training ," *Amazon Web Services AI Lab*, Shanghai, China, March 15, 2023.

[K22] Invited talk, "Optimizing Distributed GNN Training ," *Pengcheng Lab*, Shenzhen, China, May 8, 2023.

[K23] Invited talk, "Performance Diagnosis and Optimization of Distributed DNN Training," *Amazon Web Services (AWS)*, Santa Clara, California, USA, June 8, 2023.

[K24] Invited talk, "Performance Diagnosis and Optimization of Distributed DNN Training," *Meta*, Menlo Park, California, USA, June 9, 2023.

[K25] Invited talk, "Performance Diagnosis and Optimization of Distributed DNN Training," *Bytedance Ltd*, San Jose, California, USA, June 9, 2023.

[K26] Invited talk, "Optimizing Distributed GNN Training," *Huawei Academic-Industrial Forum (Vision Forum)*, Kowloon Shangri-La, Hong Kong, July 14, 2023.

[K27] Plenary talk, "Optimizing Distributed DNN Training in AI Clouds," *Huawei STW (Strategy and Technology Workshop) 2023*, Shenzhen, September 26, 2023.

Teaching Activities

◇ Serving as Departmental Curriculum Development Committee Chair (2018—2021): handling various curriculum development and course revamp matters
◇ Serving as the programme coordinator of HKU-Tsinghua CS dual degree pro-

gram (to start in Fall 2021), and responsible for designing the dual-degree curriculum

◇ COMP9602 - Convex Optimization: taught in Fall 2009, Fall 2010, Fall 2011, Spring 2014, Spring 2015, Spring 2017, Fall 2018, Fall 2020, Fall 2023

◇ CSIS0234B/COMP3234 - Computer and Communication Networks: taught in Spring 2009, Spring 2010, Spring 2011, Spring 2012, Spring 2020, Spring 2021, Spring 2024

◇ CSIS0322/COMP3322 - Modern Technologies on World Wide Web: taught in Fall 2012, Fall 2013, Spring 2015, Fall 2015, Fall 2016, Fall 2017, Spring 2018, Spring 2019, Fall 2019, Spring 2022, Fall 2022

◇ CSIS7304 - The Wireless Internet and Mobile Computing: taught in Spring 2009

SUPERVISION OF MSC DISSERTATIONS/PROJECTS

◇ Yang Hu, Yunjing Li, Ningxin Zheng, Yufeng Chen, "Local feature aggregation in GNN training and analyze the impact on Model convergence", 2023
◇ Guohua Chen, Yifan Wang, "Hand Interaction in the Virtual Classroom", 2021-2022
◇ Mingshan Liu, "A Smart Elderly Walker", 2018-2019
◇ Owen Man Ho Yip, "A Scheduler Mobile App for Tutors and Parents", 2018-2019
◇ Chi Tin Wong, Pei Zhang, "Indoor Localization and Navigation for Mobile Robot", 2017-2018
◇ Ziqiao Wang, "Outdoor Short-Distance Tracker", 2017-2018
◇ Chao Tu, "A Scheduler Mobile App for Tutors and Parents", 2017-2018
◇ Ni Zhang, Haoning An, Xiaohan Jiang, "Planning, Pricing and Implementation of a Reservation-based Taxi Service", 2016 — 2017
◇ Shuran Li, Wangning Xu, Chen Wu, "Indoor Localization and Navigation for Mobile Robot", 2016 — 2017
◇ Duo Zhao, "A Personal Intelligent Fitness Assistant on iOS Platform", 2016 — 2017

SUPERVISION OF FINAL YEAR PROJECTS

◇ Jalik Yaw, "Automatic Music Transcription for Electric Bass Guitar", 2023-2024
◇ Youran Zhao, "AI-Based Speech Recognition and Dialog with the Elderly", 2021-2022
◇ Zhisheng Hua, "AI-Based Speech Recognition and Dialog with the Elderly", 2021-2022
◇ Arun Kaza, "My Assistant Pal - Smart Home Device Using 3D Printing", 2020-2021
◇ Ximing Wang, Yuyang Lin, Zhiying Ma, "An AI Piano Tutor", 2020-2021
◇ Anant Garg, Yujie Liu, "An AI Piano Tutor App", 2019-2020
◇ Anushka Vashishtha, "Robot Path Planning in Wireless Communication Using Reinforcement Learning", 2018-2019
◇ Fang Wan, "PDSecretary: A Modile App for Recommending Exercise to Parkinson's Disease Patients", 2018-2019
◇ Apoorva Katta, "Ping Based Digital Marketing Service", 2017 — 2018
◇ Ka Lok Chan, Hilary Ho, Siu Wai Lam, "A Cloud-based Mobile App for Tutors and Parents", 2017 — 2018
◇ Shun Cheung Lee, "An Easy-to-Use Mobile App for Personal Finance", 2016 — 2017
◇ Chun Man Cheung, Ka Hei Chow, "An Easy-to-Use Mobile App for Personal Buy and Sell", 2014 — 2015
◇ Brian Kwan Hon See, Chun Ka Hung, Kin Chung Wong, "A Crowdsourcing

Shopping Website with Incentives", 2013 — 2014

◇ Chun Wai Fu, Casey Wing Yan Tang, Cheuk Hang Wong, "An Easy-to-Use Mobile System for Calling Taxi", 2013 — 2014 (1st runner-up in the departmental FYP competition)

◇ Euky Chan, "A Drawing-Board Mobile Cloud App", 2012 — 2013

◇ Ka Shing Yip, "A Cloud-based Mobile Shopping Mashup", 2012 — 2013

◇ Kit Yin Lam, "A Crowdsourcing Online Shopping Network with Incentives", 2011 — 2012

◇ Zhidi Liang, Shining Sun, "A Crowdsourcing Online Shopping Network with Incentives", 2011 — 2012

◇ Tsz Wai Chow, Ping Lam Or, Chun Hin Yu, "A Group-on Mobile Cloud Application", 2011 — 2012

◇ Yan Kit Ng, Ho Tse, Chak Sum Lam, "An Online Social Network for Virtual Shopping", 2010 — 2011

◇ Shing Ki Wong, Tsz Fung Ho, "An Online Social Network for Virtual Shopping", 2010 — 2011

◇ Ka Yui Chan, Wai Yin Chung, Chi Long Wong, "An Emulation Framework for Peer-to-Peer Applications", 2009 — 2010

INTERNATIONAL PROFESSIONAL SERVICE

◇ Guest Editor: IEEE Network, Special Issue on "Interplay between Machine Learning and Networking Systems", 2023

◇ Steering committee member, IEEE/ACM International Symposium on Quality of Service (IWQoS), January 1, 2019 - January 1, 2022.

◇ Steering committee member, the 15th EAI International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine 2019), Hong Kong.

◇ Steering committee member, the International Workshop on Intelligent Cloud Computing and Networking (ICCN) in conjunction with IEEE INFOCOM, 2019-present.

◇ Area TPC Chair, IEEE INFOCOM, Vancouver, Canada, 2021.

◇ Workshop Co-Chair, IEEE INFOCOM, Toronto, Canada, 2020.

◇ Demo/Poster Co-Chair, IEEE INFOCOM, Paris, France, 2019.

◇ TPC Co-Chair, the 26th IEEE/ACM International Symposium on Quality of Service (IWQoS), Banff, Canada, 2018.

◇ TPC Co-Chair, the 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV), Amsterdam, The Netherlands, 2018.

◇ Area Chair ("Multimedia Transport and Delivery" Track), ACM Multimedia Conference, Seoul, South Korea, 2018.

◇ Co-Chair for Special Session "Cloud-based Multimedia", ACM Multimedia Systems Conference (ACM MMSys), Taipei, Taiwan, 2017.

◇ Demo/Poster Co-Chair, IEEE ICNP, Toronto, Canada, 2017.

◇ Poster Co-Chair, ACM e-Energy, Hong Kong, 2017.

◇ Demo Co-Chair, ITC 29, Genoa, 2017.

◇ Demo/Poster Co-Chair, IEEE INFOCOM, Hong Kong, 2015.

◇ Demo Chair, ACM Multimedia Systems Conference (ACM MMSys), Portland, Oregon, 2015.

◇ Chair of Interest Group on Multimedia services and applications over Emerging Networks (MEN), IEEE Multimedia Communication Technical Committee (MMTC), 2012-2014.

◇ Publication Chair, the 2011 International Conference on Cloud and Service Computing (CSC2011), Hong Kong, December 12-14, 2011.

◇ Program Track Co-Chair ("Pervasive, Grid, P2P and Cloud Computing" track), the 6th International Conference on Frontier of Computer Science and Technology (FCST-11), Changsha, China, November 16-18, 2011.

◇ Poster Session Co-Chair, the 17th IEEE International Workshop on Quality of Service (IEEE IWQoS 2010), Beijing, China, June 16-18, 2010.

◇ Regularly Refereeing for Academic Journals:

IEEE Transactions on Parallel and Distributed Systems (TPDS)
IEEE/ACM Transactions on Networking (TON)
IEEE Journal on Selected Areas in Communications (JSAC)
IEEE Transactions on Multimedia (TMM)
IEEE Transactions on Mobile Computing (TMC)
IEEE Transactions on Wireless Communications (TWC)
IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)
IEEE Transactions on Computers (TC)
IEEE Transactions on Communications (TCOM)
IEEE Transactions on Cloud Computing (TCC)
IEEE Transactions on Services Computing (TSC)
IEEE Transactions on Vehicular Technology (TVT)
IEEE Transactions on Network and Service Management (TNSM)
IEEE Transactions on Cognitive Communications and Networking (TCCN)
ACM Computing Surveys (CSUR)
ACM Transactions on Sensor Networks (TOSN)
ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)
ACM Transactions on Modeling and Performance Evaluation of Computing Systems (TOMPECS)
Springer Multimedia Systems Journal
Springer Peer-to-Peer Networking and Applications Journal (PPNA)
Elsevier Journal of Computer Networks (JCN)
Elsevier Journal of Parallel and Distributed Computing (JPDC)
Elsevier Computer Networks Journal (COMNET)
Elsevier Performance Evaluation Journal (PEVA)
IEEE Network
IEEE Multimedia
IEEE Communications Magazine
IEEE Communications Letters

◇ Membership in Technical Program Committees:

IEEE INFOCOM: 2011—2024
IEEE ICNP: 2023
ACM Mobihoc: 2020-2024
ACM International Conference on Information and Knowledge Management (CIKM): 2019
ACM SIGMETRICS: 2016

ACM International Conference on Multimedia (ACM Multimedia): 2023, 2018, 2016, 2014, 2012, 2011, 2009

ACM Multimedia Systems Conference (MMSys): 2013

IEEE International Conference on Data Mining (ICDM):2021, 2020, 2018

IEEE International Conference on Distributed Computing Systems (ICDCS): 2011, 2022, 2023

IEEE International Conference on Network Protocols (ICNP): 2022, 2021 2011

IEEE International Conference on Communications (ICC): 2018, 2010-2016

IEEE Global Communications Conference (GLOBECOM): 2017, 2012, 2011, 2010, 2009

IEEE International Conference on Parallel and Distributed Systems (ICPADS): 2016, 2012

IEEE International Conference on Computer Communications and Networks (ICCCN): 2011, 2010

IEEE International Conference on Multimedia and Expo (ICME): 2017

IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication System (MASCOTS): 2017

IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS): 2022, 2018

IEEE Wireless Communications and Networking Conference (WCNC): 2013

IEEE International Conference on Mobile Ad-hoc and Sensor Networks (MSN): 2012

IEEE Consumer Communications and Networking Conference (CCNC): 2011, 2010

IEEE International Conference on Cloud Engineering (IC2E): 2018

IFIP Networking: 2023, 2020, 2019, 2017, 2012

ACM/IEEE International Symposium on Quality of Service (ACM/IEEE IWQoS): 2014-2017, 2019 - 2023

ACM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV): 2017, 2014, 2013, 2012

ACM e-Energy: 2018

IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid): 2013

February 26, 2024