

Shaoyuan Huang

Ph.D. 3nd year，TJU Edge Big Bang Group，College of Intelligence and Computing，Tianjin University，Tianjin

[dblp](https://dblp.org/pid/303/9138.html) [Google Scholar](https://scholar.google.com/citations?user=nNfHu_QAAAAJ&hl=zh-CN)

Building No. 55, Tianjin University, Haihe Education District, Jinnan, Tianjin

hsy\_23@tju.edu.cn +86-15022618263

|  |  |
| --- | --- |
| ResearchField | * Distributed System Workload and Performance Modeling * AI Inference Serving Systems * FL-based Joint LLM Inference and Fine-tuning |
| Education Experience | * Visting Ph.D.(2024-2025) Department of Engineering, King’s College London   ***(Supervisor: Prof. Yansha Deng)***   * Ph.D. (2022-Now), M.S.(2020-2022), B.S. (2016-2020)   All from College of Intelligence and Computing, Tianjin University, Tianjin, China  *(**Supervisor: Prof. Xiaofei Wang,* *Peiyang Young Scholar, National Thousand Youth Talents Plan)* |
| Internship | * 2021.09-2022.01 Algorithm Development Intern, in PPIO Cloud Computing (Shanghai) Co.   - Responsible for the development of cloud server workload analysis and prediction algorithms, server utilization impact feature mining.   * 2022.03-2022.06 Algorithm Development Intern, in Paiou Cloud Computing (Shanghai) Co.   - Responsible for system integration of prediction algorithms and task deployment recommender system. |
| Publication | * **Journal**  1. **Shaoyuan Huang**, Zheng Wang, Heng Zhang, Xiaofei Wang, Cheng Zhang, Wenyu Wang"DynEformer: A Unified Framework for Robust Workload Prediction Under Dynamic Environment," Under Review. 2. **Shaoyuan Huang**, Heng Zhang, Xiaofei Wang∗, Min Chen, Jianxin Li, Victor C.M. Leung "Fine-grained Spatio-Temporal Distribution Prediction of Mobile Content Delivery in 5G Ultra-Dense Networks,'' in ***IEEE Transactions on Mobile Computing (TMC)***, 2022. (JCR-1, IF:7.9) 3. **Shaoyuan Huang**, Yuxi Zhang, Guozheng Peng, Juan Zhao, Keping Zhu, Heng Zhang, Xiaofei Wang∗, "MF-GCN-LSTM: A Cloud-Edge Distributed Framework for Key Positions Prediction in Grid Projects," in ***Journal of Cloud Computing***, 2022. (JCR-2, IF:4.0) 4. Heng Zhang, **Shaoyuan Huang**, Xin Wang, Jianxin Li, Xiaofei Wang∗, Victor C. M. Leung, "A Measurement-driven Analysis and Prediction of Content Propagation in the Device-to-Device Social Networks,'' in ***IEEE Transactions on Knowledge and Data Engineering (TKDE)***, 2022. (JCR-1, IF:8.9) 5. Heng Zhang, **Shaoyuan Huang**, Mengwei Xu, Deke Guo, Xiaofei Wang, Xin Wang, Victor CM Leung, Wenyu Wang, "Large-scale Measurements and Optimizations on Latency in Edge Clouds," in ***IEEE Transactions on Cloud Computing (IEEE TCC),*** 2024. (JCR-2, IF:5.4) 6. Hui Sun, Yiru Chen, Kewei Sha, **Shaoyuan Huang**, Xiaofei Wang, Weisong Shi, "A Proactive On-Demand Content Placement Strategy in Edge Intelligent Gateways,'' in ***IEEE Transactions on Parallel and Distributed Systems***, 2023. (JCR-1, IF:5.3)  * **Conference**  1. **Shaoyuan Huang,** Tiancheng Zhang, Zhongtian Zhang, Xiaofei Wang, Lanjun Wang, Xin Wang, "MetaEformer: Unveiling and Leveraging Meta-Patterns for Complex and Dynamic Systems Load Forecasting", in ***31TH ACM SIGKDD Conference on Knowledge Discovery and Data Mining (ACM SIGKDD)***, 2025, (CCF-A). 2. **Shaoyuan Huang**, Zheng Wang, Zhongtian Zhang and Heng Zhang, Xiaofei Wang, Wenyu Wang, "Seer: Proactive Revenue-Aware Scheduling for Live Streaming Services in Crowdsourced Cloud-Edge Platforms,", in ***IEEE International Conference on Computer Communications (IEEE INFOCM)***, 2024, (CCF-A). 3. **Shaoyuan Huang**, Zheng Wang, Heng Zhang, Xiaofei Wang, Cheng Zhang and Wenyu Wang, "One for All: Unified Workload Prediction for Dynamic Multi-tenant Edge Cloud Platforms," in ***29TH ACM SIGKDD Conference on Knowledge Discovery and Data Mining (ACM SIGKDD)***, 2023, (CCF-A). 4. **Shaoyuan Huang**, Heng Zhang, Xiaofei Wang, Min Chen, Jianxin Li, Victor C.M. Leung, "Spatial-Temporal-Social Multi-Feature-based Fine Grained Hot Spots Prediction for Content Delivery Services in 5G Era," in ***30th ACM International Conference on Information and Knowledge Management (ACM CIKM)***, 2021, (CCF-B). 5. Yuting Li, **Shaoyuan Huang**, Tengwen Zhang Cheng Zhang Xiaofei Wang and Victor C.M. Leung, "Sentinel: Scheduling Live Streams with Proactive Anomaly Detection in Crowdsourced Cloud-Edge Platforms", in ***IEEE International Conference on Computer Communications (IEEE INFOCM)***, 2025, (CCF-A). 6. Heng Zhang, **Shaoyuan Huang**, Mengwei Xu, Deke Guo, Xiaofei Wang, Victor C. M. Leung and Wenyu Wang, "How Far Have Edge Clouds Gone? A Spatial-Temporal Analysis of Edge Network Latency In the Wild," in ***IEEE/ACM International Symposium on Quality of Service (IWQoS)***, 2023, (CCF-B). 7. Heng Zhang, Zixuan Cui, **Shaoyuan Huang**, Deke Guo, Xiaofei Wang, Wenyu Wang, "QM-RGNN: An Efficient Online QoS Measurement Framework with Sparse Matrix Imputation for Distributed Edge Clouds'', in ***IEEE International Conference on Computer Communications (IEEE INFOCM)***, 2024, (CCF-A). 8. Tiancheng Zhang, **Shaoyuan Huang**, Cheng Zhang, Xiaofei Wang, Wenyu Wang, "EasyTS: The Express Lane to Long Time Series Forecasting", ***in AAAI 2024 Demonstration Program***, 2024, (CCF-A). |
| Talk | * 43th IEEE International Conference on Computer Communications (IEEE INFOCOM) 2024, presentation and talk. * 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (ACM SIGKDD), 2023, Long Beach, CA, USA, online participation and presentation. * IEEE Global Communications Conference (Globecom), 2023, Kuala Lumpur, Malaysia, participation and presentation. * 30th ACM International Conference on Information and Knowledge Management (ACM CIKM), 2021, online participation and presentation. |
| Leadership | * Tianjin University, department president of the Youth Culture Club, 2016-2019   Participated in organizing more than 50 lectures with more than 13,000 audience. |
| Patent | * " ***Multi-feature based neural network for content delivery hotspots prediction***", Chinese Patent, CN112822045B (Patent Authorized) * " ***Edge cloud server utilization prediction method, prediction device and storage medium based on boosting algorithm***", Chinese Patent, CN114721898A (Patent Pending)   Including those not listed, totaling 12 patents. |
| Award | * 2025, China Scholarship Council * 2024, CCF DPCS Distinguished Doctorate * 2021, 2023, “Suzhou Talent Scholarship”, Suzhou Government Talent Group * 2023, “Merit Student” of Tianjin University * 2021, “Suzhou Talent Scholarship”, Tianjin University * 2020, “Outstanding Graduate” of Tianjin University * 2017-2019, “Merit Student” of Tianjin University |
| Technical Contributions | * **Open source system models and datasets** * [**KDD25\_MetaEformer: MetaEformer v1.0 - Initial Release**](https://zenodo.org/records/15502862) * [**Edge Cloud Server Latency Measurements**](https://github.com/henrycoding/IWQoS23EdgeMeasurements) * [**DynEformer: Edge Cloud Server Workload Prediction Framework**](https://github.com/hsy23/KDD23_DynEformer) * [**ECW:**  **Edge Cloud Server Workload Dataset**](https://github.com/hsy23/ECWDataset) * **Internship Achievements** * Successfully designed the workload and utilization prediction model based on Xgboosting and residual learning, with an accuracy of over 90% across thousands of servers, through several rounds of improvement and A/B testing. * Participated in the development of a prototype predictive modeling-based task deployment recommendation system, responsible for algorithm integration, data flow automation, and recommendation algorithms components. |