

Hailiang Zhao

ZJU 100 Young Professor, School of Software Technology, Zhejiang University
hliangzhao.me Google Scholar ZJU Profile

hliangzhao@zju.edu.cn
Ningbo and Hangzhou, China

RESEARCH PROFILE

I work on algorithms and systems for intelligent services, with current interests in learning-augmented systems, agentic services, generative AI runtime, and cloud-edge intelligence. My research connects online optimization, machine learning, and systems design. Related methods have been evaluated in deployments involving domestic intelligent-computing chips and enterprise AI services.

ACADEMIC APPOINTMENTS

Zhejiang University, School of Software Technology Aug. 2024 - Present
ZJU 100 Young Professor; Ph.D. Supervisor

EDUCATION AND VISITING RESEARCH

Zhejiang University, College of Computer Science and Technology 2019 - 2024
Ph.D. in Computer Science and Technology; Advisor: Prof. Shuiguang Deng
Nanyang Technological University, PDCL Lab Sep. 2022 - Sep. 2023
Visiting Ph.D. Student; Host: Prof. Xueyan Tang
Wuhan University of Technology 2015 - 2019
B.Eng.

RESEARCH INTERESTS

Learning-augmented online algorithms; intelligent computing systems; agentic services; generative AI systems; cloud-edge intelligence; service computing and resource scheduling.

SELECTED HONORS

- Huawei Spark Award, 2026.
- Zhejiang Provincial Outstanding Doctoral Dissertation Nomination, 2026.
- SICI Best Paper Award, 2025.
- CCF Service Computing Doctoral Dissertation Incentive Program, 2025.
- ZJU Outstanding Young Scholar, 2024.
- Zhejiang University Outstanding Doctoral Dissertation, 2024.
- IEEE ICWS Best Student Paper Award, 2019.

SELECTED RESEARCH PROJECTS

- **PI**, National Key R&D Program Project: Adaptive multimodal data-governance technologies for fully connected systems, 2024-2027.
- **PI**, NSFC Youth Fund: Hardware-software co-optimization for service scheduling and resource management, 2026-2028.
- **Participant**, Zhejiang Provincial Natural Science Foundation Major Project: Resource optimization and computing acceleration for generative AI on heterogeneous hardware, 2025-2027.
- **PI**, Kuaishou collaboration on intelligent-service systems, 2024-2025.

Young Editorial Board Member, *Computer Engineering & Science*. Reviewer or program-committee member for NeurIPS, IJCAI-ECAI, AAAI, PAKDD, IEEE Transactions on Services Computing, IEEE Transactions on Mobile Computing, Future Generation Computer Systems, ACM TKDD, and IEEE Internet of Things Journal.

PUBLICATIONS

Full publication is available at Google Scholar.

2026

1. Wenzhuo Qian, **Hailiang Zhao***, Ziqi Wang, Zhipeng Gao, Jiayi Chen, Zhiwei Ling, and Shuiguang Deng*. ARMOR: Missing-Aware Multimodal Fusion for Unified Microservice Incident Management *ASE, accepted*.
2. Peng Chen, **Hailiang Zhao***, Xueyan Tang, Yixuan Wang, and Shuiguang Deng*. Towards Optimal Robustness in Learning-Augmented Paging *ICML, Spotlight*.
3. Jiaji Zhang, Ruichao Sun, **Hailiang Zhao***, Jiaju Wu, Peng Chen, Hao Li, Yuying Liu, Kingsum Chow, Gang Xiong, and Shuiguang Deng*. SegQuant: A Semantics-Aware and Generalizable Quantization Framework for Diffusion Models *CVPR*.
4. Jiaji Zhang, **Hailiang Zhao***, Guoxuan Zhu, Ruichao Sun, Jiaju Wu, Xinkui Zhao, Hanlin Tang, Weiyi Lu, Kan Liu, Tao Lan, Lin Qu, and Shuiguang Deng*. Shiva-DiT: Residual-Based Differentiable Top-k Selection for Efficient Diffusion Transformers *arXiv preprint*.
5. Zhiwei Ling, **Hailiang Zhao***, Chao Zhang, Xiang Ao, Ziqi Wang, Cheng Zhang, Zhen Qin, Xinkui Zhao, Kingsum Chow*, Yuanqing Wu, and MengChu Zhou. Adaptive Dual-Weighting Framework for Federated Learning via Out-of-Distribution Detection *arXiv preprint*.
6. Yu Tang[†], **Hailiang Zhao***[†], Chuansheng Lu, Yifei Zhang, Kingsum Chow*, Shuiguang Deng, and Rui Shi. Morphis: SLO-Aware Resource Scheduling for Microservices with Time-Varying Call Graphs *arXiv preprint*.
7. **Hailiang Zhao**, Ziqi Wang, Daojiang Hu, Zhiwei Ling, Wenzhuo Qian, Jiahui Zhai, Yuhao Yang, Zhipeng Gao, Mingyi Liu, Kai Di, Xinkui Zhao, Zhongjie Wang, Jianwei Yin, MengChu Zhou, and Shuiguang Deng. Industrial Data-Service-Knowledge Governance: Toward Integrated and Trusted Intelligence for Industry 5.0 *arXiv preprint*.

2025

8. Jiahong Yu[†], Ziqi Wang[†], **Hailiang Zhao***, Wei Zhai, Xueqiang Yan, and Shuiguang Deng. TrackTeller: Temporal Multimodal 3D Grounding for Behavior-Dependent Object References *arXiv preprint*.
9. Ziqi Wang, **Hailiang Zhao***, Yuhao Yang, Daojiang Hu, Cheng Bao, Mingyi Liu, Kai Di, Schahram Dustdar, Zhongjie Wang, and Shuiguang Deng. OmniFuser: Adaptive Multimodal Fusion for Service-Oriented Predictive Maintenance *arXiv preprint*.
10. Shuiguang Deng*, **Hailiang Zhao***, Ziqi Wang, Guanjie Cheng, Peng Chen, Wenzhuo Qian, Zhiwei Ling, Jianwei Yin, Albert Y. Zomaya, and Schahram Dustdar. Agentic Services Computing *arXiv preprint*.
11. Peng Chen, Jiaji Zhang, **Hailiang Zhao***, Yirong Zhang, Jiahong Yu, Xueyan Tang, Yixuan Wang, Hao Li, Jianping Zou, Gang Xiong, Kingsum Chow, Shuibing He, and Shuiguang Deng*. Toward Robust and Efficient ML-Based GPU Caching for Modern Inference *arXiv preprint*.
12. Ziqi Wang, **Hailiang Zhao***, Cheng Bao, Wenzhuo Qian, Yuhao Yang, Xueqiang Sun, and Shuiguang Deng*. XFMNet: Decoding Cross-Site and Nonstationary Water Patterns via Stepwise Multimodal Fusion for Long-Term Water Quality Forecasting *arXiv preprint*.
13. Wenzhuo Qian, **Hailiang Zhao***, Tianlv Chen, Jiayi Chen, Ziqi Wang, Kingsum Chow, and Shuiguang Deng*. Learning Unified System Representations for Microservice Tail Latency Prediction *arXiv preprint*.
14. Yinuo Deng[†], **Hailiang Zhao***[†], Dongjing Wang, Peng Chen, Wenzhuo Qian, Jianwei Yin, Schahram Dustdar, and Shuiguang Deng*. PeerSync: Accelerating Containerized Model Inference at the Network Edge *IEEE Transactions on Services Computing*.
15. Peng Chen, **Hailiang Zhao***, Jiaji Zhang, Xueyan Tang, Yixuan Wang, and Shuiguang Deng*. Robustifying Learning-Augmented Caching Efficiently without Compromising 1-Consistency *NeurIPS*.
16. **Hailiang Zhao**, Xueyan Tang*, Peng Chen, Jianwei Yin, and Shuiguang Deng. Data-Locality-Aware Task Assignment and Scheduling for Distributed Job Executions *IEEE Transactions on Services Computing*.
17. **Hailiang Zhao**, Ziqi Wang, Guanjie Cheng*, Wenzhuo Qian, Peng Chen, Jianwei Yin, Schahram Dustdar, and Shuiguang Deng*. Online Workload Scheduling for Social Welfare Maximization in the Computing Continuum *IEEE Transactions on Services Computing*.
18. Cheng Zhang, Yinuo Deng, **Hailiang Zhao***, Tianlv Chen, and Shuiguang Deng*. Tail-Learning: Adaptive Learning Method for Mitigating Tail Latency in Autonomous Edge Systems *ACM Transactions on Autonomous and Adaptive Systems*.
19. Zhiwei Ling, Yachen Chang, **Hailiang Zhao***, Xinkui Zhao, Kingsum Chow*, and Shuiguang Deng. CADRef: Robust Out-of-Distribution Detection via Class-Aware Decoupled Relative Feature Leveraging *CVPR*.

2024

20. **Hailiang Zhao**, Xueyan Tang*, Peng Chen, and Shuiguang Deng. Learning-Augmented Algorithms for the Bahncard Problem *NeurIPS*.
21. Shuiguang Deng*, **Hailiang Zhao***, Bingbing Huang, Cheng Zhang, Feiyi Chen, Yinyao Deng, Jianwei Yin, Shahram Dustdar, and Albert Y. Zomaya. Cloud-Native Computing: A Survey from the Perspective of Services *Proceedings of the IEEE, Cover Paper*.
22. **Hailiang Zhao**, Shuiguang Deng*, Zhengzhe Xiang, Xueqiang Yan, Jianwei Yin, Shahram Dustdar, and Albert Y. Zomaya. Scheduling Multi-Server Jobs with Sublinear Regrets via Online Learning *IEEE Transactions on Services Computing*.

2023

23. **Hailiang Zhao**, Shuiguang Deng*, Feiyi Chen, Jianwei Yin, Shahram Dustdar, and Albert Y. Zomaya. Learning to Schedule Multi-Server Jobs with Fluctuated Processing Speeds *IEEE Transactions on Parallel and Distributed Systems*.

2022

24. Shuiguang Deng, **Hailiang Zhao**, Zhengzhe Xiang, Cheng Zhang, Rong Jiang, Ying Li*, Jianwei Yin, Shahram Dustdar, and Albert Y. Zomaya. Dependent Function Embedding for Distributed Serverless Edge Computing *IEEE Transactions on Parallel and Distributed Systems*.
25. **Hailiang Zhao**, Shuiguang Deng*, Zijie Liu, Zhengzhe Xiang, Jianwei Yin, Shahram Dustdar, and Albert Y. Zomaya. DPoS: Decentralized, Privacy-Preserving, and Low-Complexity Online Slicing for Multi-Tenant Networks *IEEE Transactions on Mobile Computing*.
26. **Hailiang Zhao**, Shuiguang Deng*, Zijie Liu, Jianwei Yin, and Shahram Dustdar. Distributed Redundant Placement for Microservice-based Applications at the Edge *IEEE Transactions on Services Computing*.

2020

27. Shuiguang Deng, **Hailiang Zhao**, Weijia Fang*, Jianwei Yin, Shahram Dustdar, and Albert Y. Zomaya. Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence *IEEE Internet of Things Journal*.

2019

28. **Hailiang Zhao**, Shuiguang Deng*, Cheng Zhang, Wei Du, Qiang He, and Jianwei Yin. A Mobility-aware Cross-edge Computation Offloading Framework for Partitionable Applications *IEEE ICWS, Best Student Paper Award*.