

Xiaozhuang Song

(+86) 18800172145 | xiaozhuangsong1@link.cuhk.edu.cn | [Google Scholar](#) |

EDUCATION

The Chinese University of Hong Kong, Shenzhen

Ph.D. in Computer Science

Shenzhen, China

Sep. 2023 – Mar. 2027 (expected)

Southern University of Science and Technology

M.Phil in Computer Science

Shenzhen, China

Sep. 2019 – Jul. 2022

Beijing University of Posts and Telecommunications

B.Eng. in Digital Media Technology

Beijing, China

Sep. 2015 – Jul. 2019

EXPERIENCE

Research Intern

Shanghai AI Laboratory

Shanghai, China

Jul. 2025 – Present

- Core member of an AI-assisted chemical synthesis project, responsible for platform development and molecular modeling research.
- Led research and system development for chemical synthesis planning.

Research Intern

Shanghai AI Laboratory

Shanghai, China

Dec. 2024 – Feb. 2025

- Conducted research on language model reasoning over structured data.

Research Intern

Microsoft Research Asia (MSRA)

Beijing, China

Jun. 2021 – Jun. 2023

- Conducted research on multi-task learning.
- Participated in the NeurIPS CityLearn Challenge (5th place) and contributed to tabular data research.

RESEARCH INTERESTS & OTHER EXPERIENCE

Research Interests: AI for Science; LLM reasoning and planning; Automated Scientific Discovery.

Teaching: Teaching assistant at CUHK-Shenzhen for Machine Learning (2023), Operating Systems (2024), Probability and Statistical Inference (2024), Introduction to Computer Science (2025, Lead TA, 700 students), and Data Structures (2026); at SUSTech for Digital Logic (CS201, 2021).

Competitions: NeurIPS 2022 CityLearn Challenge, 5th place; CCF-BDCI 2020 Road Spatiotemporal Prediction, top 10%.

PUBLICATIONS & SERVICE

Reviewer Service: Regularly served as a reviewer since 2022 for leading conferences and journals in the field, including NeurIPS, ICML, ICLR, WWW, AISTATS, CVPR, KDD, AAAI, ACL, COLM, IJCAI, TPAMI, and TKDE. Research contributions cover AI for Science, single-step and multi-step retrosynthesis, LLM reasoning, and structured graph generation.

Patents: Contributed to six publicly disclosed invention patents.

Under Review

- **Xiaozhuang Song**, Xuemin Chen, Xinjian Zhao, Yaoyao Xu, Tianshu Yu. *RetroGEF: Dynamic Graph Edit Flow for Single-Step Retrosynthesis*. **arXiv preprint**
- Xinjian Zhao, Yaoyao Xu, Xuemin Chen, **Xiaozhuang Song**, Tianshu Yu. *Latent JEPA: Abstract Future Prediction for Latent Reasoning in Chemistry*.
- Yaoyao Xu, Xinjian Zhao, **Xiaozhuang Song**, Xuemin Chen, Tianshu Yu. *ChemOPD: Multi-Teacher On-Policy Distillation for Multi-Task Chemical Reasoning*.
- Xuemin Chen, **Xiaozhuang Song**, Xinjian Zhao, Yaoyao Xu, Tianshu Yu. *Coverage Before Control: Route-Instruction Grounding and Steering for Controllable Retrosynthesis*.
- Yaoyao Xu, Xinjian Zhao, **Xiaozhuang Song**, Lei Bai, Tianshu Yu. *Closed-Loop Bayesian Molecular Inverse Design with Semantic LLM Surrogates*. **arXiv:2608.22967**, 2026.

Accepted / Published

- **Xiaozhuang Song**, Xuanhao Pan, Xinjian Zhao, Hangting Ye, Shufei Zhang, Jian Tang, Tianshu Yu. *AOT*: Efficient Synthesis Planning via LLM-Empowered AND-OR Tree Search*. **ACL 2026 Findings**
- **Xiaozhuang Song**, Shufei Zhang, Tianshu Yu. *ReKG-MCTS: Reinforcing LLM Reasoning on KGs via Training-Free MCTS*. **ACL 2025 Findings**
- **Xiaozhuang Song**, Tianshu Yu. *Physically-Informed Flow Matching with Graph Neural Networks for Complex Fluid Dynamics*. **AAAI 2026 [CCF-A]**
- **Xiaozhuang Song**, Yuzhao Tu, Hangting Ye, Wei Fan, Qingquan Zhang, Xiaoxue Wang, Tianshu Yu. *Enhancing Generalizability in Molecular Conformation Generation with Metrization-Informed Geometric Diffusion Pretraining*. **AAAI 2025 [CCF-A]**
- **Xiaozhuang Song**, Shun Zheng, Wei Cao, James J.Q. Yu, Jiang Bian. *Efficient and Effective Multi-task Grouping via Meta Learning on Task Combinations*. **NeurIPS 2022 [CCF-A]**
- **Xiaozhuang Song**, Yuzhao Tu, Tianshu Yu. *Enhancing Molecular Conformer Generation via Fragment-Augmented Diffusion Pretraining*. **TMLR 2025**
- **Xiaozhuang Song**, Xu Liao, Hangting Ye, Yaoyao Xu, Wei Fan, Jin Liu, Tianshu Yu. *Single Cell Gene Expression Prediction via Prototype-based Proximal Neural Factorization*. **BIBM 2024 [CCF-B]**
- **Xiaozhuang Song**, Ying Wu, Chenhan Zhang. *TSTNet: A Seq2Seq Transformer for Spatial-temporal Traffic Prediction*. **ICANN 2021 [CCF-C]**
- Hangting Ye, Peng Wang, Wei Fan, **Xiaozhuang Song**, He Zhao, Dandan Guo, Yi Chang. *Deep Tabular Representation Corrector*. **TPAMI 2026 [CCF-A]**
- Xinjian Zhao, Wei Pang, Zhixuan Yu, Xiangru Jian, **Xiaozhuang Song**, Yaoyao Xu, Zhongkai Xue, Dingshuo Chen, Shu Wu, Philip Torr, Tianshu Yu. *When Vision Meets Graphs: A Survey on Graph Reasoning and Learning*. **IJCAI 2026 (Survey Track)**
- Xinjian Zhao, Xiangru Jian, Yaoyao Xu, **Xiaozhuang Song**, Wei Pang, Lei Bai, Tianshu Yu. *MolEmb: Multimodal Large Language Models Can Be Strong Molecular Embedding Models*. **ICML 2026 Workshop (FM4LS)**
- Xinjian Zhao, Wei Pang, Zhongkai Xue, Xiangru Jian, Lei Zhang, Yaoyao Xu, **Xiaozhuang Song**, Shu Wu, Tianshu Yu. *The Underappreciated Power of Vision Models for Graph Structural Understanding*. **NeurIPS 2025 [CCF-A]**
- Wei Fan, Jingru Fei, Dingyu Guo, Kun Yi, **Xiaozhuang Song**, Haolong Xiang, Hangting Ye, Min Li. *MedGNN: Multi-resolution Spatiotemporal Graph Learning for Medical Time Series*. **WWW 2025 [CCF-A]**
- Hangting Ye, Wei Fan, **Xiaozhuang Song**, Shun Zheng, He Zhao, Dandan Guo, Yi Chang. *PTaRL: Prototype-based Tabular Representation Learning*. **ICLR 2024 Spotlight [CCF-A]**
- Wei Fan, Jingru Fei, Junwen Han, Jinglan Lian, Hangting Ye, **Xiaozhuang Song**, Xiaoling Lv, Kun Yi, Min Li. *Medvia: Medical Time Series Classification with Vision Augmentation*. **Information Fusion 2025 [JCR Q1, IF 14.7]**
- Yunjie Huang, **Xiaozhuang Song**, Shiyao Zhang, Lei Li, James J.Q. Yu. *GT-TTE: Modeling Trajectories as Graphs for Travel Time Estimation*. **IEEE IoT Journal 2024 [JCR Q1]**
- Yunjie Huang, **Xiaozhuang Song**, Yuanshao Zhu, Shiyao Zhang, James J.Q. Yu. *Traffic Prediction with Transfer Learning*. **IEEE T-ITS 2023 [CCF-B / JCR Q1]**
- Ao Wang, Yongchao Ye, **Xiaozhuang Song**, Shiyao Zhang, James J.Q. Yu. *Traffic Prediction With Missing Data*. **IEEE T-ITS 2023 [CCF-B / JCR Q1]**
- Yaoyao Xu, Xinjian Zhao, **Xiaozhuang Song**, Benyou Wang, Tianshu Yu. *Boosting Protein Language Models with Negative Sample Mining*. **ECML-PKDD 2024 [CCF-B]**
- Kingsley Nweye, ..., **Xiaozhuang Song**, ..., Shun Zheng, Jiang Bian. *The CityLearn Challenge 2022*. **NeurIPS 2022 Competition Track**
- Yunjie Huang, **Xiaozhuang Song**, Shiyao Zhang, James J.Q. Yu. *Transfer Learning in Traffic Prediction with GNNs*. **IEEE ITSC 2021**
- **Xiaozhuang Song**, Christos Markos, James J.Q. Yu. *Multimix: Multi-task Deep Learning for Travel Mode Identification*. **IEEE ITSC 2020**