

谢哲 Zhe Xie

邮箱: xiez22@mails.tsinghua.edu.cn | 主页: xiez22.github.io | Google Scholar: Link | 手机: 18970816980

清华大学·计算机科学与技术系·博士研究生 研究方向: Multimodal LLM / AIOps / Time Series

教育背景

清华大学 计算机科学与技术 博士研究生

2022.09 – 预计 2027.06

导师: 裴丹教授 研究方向: 多模态大语言模型、时间序列分析、异常检测、根因定位

上海交通大学 计算机科学与技术 本科

2018.09 – 2022.06

课程均分: 91.94/100 GPA: 3.98/4.3 专业排名: 4/147 荣誉计划: 致远工科荣誉计划

代表论文

- **ChatTS: Aligning Time Series with LLMs via Synthetic Data for Enhanced Understanding and Reasoning** CCF-A 一作 VLDB 2025 GitHub
首个支持原生时序模态的多模态对话 LLM (TS-MLLM) 之一, 通过 SFT 将时间序列作为原生模态直接接入 LLM, 支持多变量、不定长时序与数值推理, 性能超越基于文本/图像表示的模型, 支持 vLLM 推理引擎。获 430+ GitHub Stars。
技术栈: 可控多模态数据合成与 SFT 语料合成, 支持对齐与推理两大类时序对话任务; 基于 Qwen2.5/3 进行时序模态的模型改造, 将编码后的时序 Patch 作为多模态 Token 输入 LLM; 实现了时序多模态的模型 SFT 工程代码, 实现基于 vLLM 的时序多模态推理代码, 支持 OpenAI API 推理部署。
贡献与成果: 完成数据合成、对齐训练、推理框架、评测、论文写作; GitHub Stars 439+, Huggingface Stars 143+, 引用 76+, 模型下载量 19k+。
- **FoundRoot: Towards Foundation Model for Root Cause Analysis via Structured Deep Thinking** CCF-A 一作 ICSE 2026
首个基于 LLM 的根因定位基础模型, 模仿人类的根因定位分析过程, 将基于时序多模态的语义特征、异常波动与传播关系的推理过程显式组织为结构化推理链路, 通过 RL 大幅提升模型在未见故障类型上的泛化能力与可解释性。
技术栈: 多领域 RCA 数据整合与清洗、交替式推理链路构造; GRPO 强化学习。
贡献与成果: 完成问题建模、推理格式设计、RL 训练方案与数据整理; 模型已在线上生产环境部署应用。
- **... Multimodal Chain-of-Thought ... Time Series ...** (部分标题隐藏) 一作 在投
探索支持多模态工具调用的时序 MLLM 强化学习训练的 Pipeline, 使模型能够在文本推理、时序理解与多模态 Python 工具 (代码可返回多模态结果, 类似 OpenAI o3 的 zoom-in 操作) 执行反馈之间进行交错式推理 (Thinking with Time Series)。
技术栈: 多模态工具调用语料构建 (迭代式优化思维链); 交错式多模态 RL 训练 (基于 TRL 改造, 实现时序多模态 + 工具调用支持)。
贡献与成果: 完成框架设计与实现, 首次实现时序 MLLM 动态工具调用, 为 TRL GRPOTrainer 增加时序模态支持, 支持 Agentic RL 训练。
- **Microservice Root Cause Analysis With Limited Observability Through Intervention Recognition in the Latent Space** CCF-A 一作 KDD 2024
面向可观测性受限的工业微服务系统, 首次提出基于隐空间干预识别根因定位方法, 缓解可观测性不足下的根因定位不准确的问题; 完成算法设计与实现、论文写作, 算法已被 eBay 工程化应用。
- **From Point-wise to Group-wise: A Fast and Accurate Microservice Trace Anomaly Detection Approach** CCF-A 一作 ESEC/FSE 2023
首创分组调用链异常检测概念, 以图算法驱动调用链树形结构分组, 替代传统的逐点检测, 在 eBay K8S 线上环境的检测速度较深度学习 SOTA 提升 20 余倍且保持同等准确率; 独立完成故障注入、算法设计与系统工程化实现。
- **Unsupervised Anomaly Detection on Microservice Traces through Graph VAE** CCF-A 一作 WWW 2023
将微服务调用链建模为图结构, 以 Graph VAE (GNN + 变分自编码器) 联合捕获拓扑关系与节点行为, 实现比序列化方法更精准的无监督端到端异常识别; 本人为第一作者, 引用 59+。
- **Adversarial and Contrastive Autoencoder for Sequential Recommendation** CCF-A 共同一作第一 WWW 2021
联合引入对抗训练与对比学习, 提升序列推荐的表示质量与推荐效果; 共同一作 (排名第一), 引用 145+。
- **AutoDA-Timeseries: Automated Data Augmentation for Time Series** CCF-A 三作 ICLR 2026
面向时间序列的自动化数据增强方法, 系统性探索增强策略搜索空间以提升下游任务性能。

实习经历

- **字节跳动** 2024.08 – 至今
时间序列多模态大语言模型研究, 一作 CCF-A 论文 2 篇, 其中 ChatTS 在工业/学术界具有一定影响力
- **eBay** 2021.11 – 2023.02 & 2023.04 – 2024.08
AIOps 研究实习, 一作 CCF-A 论文 3 篇, 成果已被工程化实现与应用

奖励与荣誉

- 2023、2024 年 清华大学校级一等奖学金
- 2018–2020 年 上海交通大学致远荣誉奖学金; 2019 年 校级 B 等奖学金