

李则言

更新于 2022 年 8 月 7 日

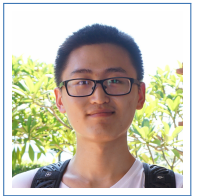
出生年月: 1997 年 5 月

☎ 184 0165 3327

✉ li_zeyan@icloud.com

📍 lizeyan

🌐 Zeyan Li



教育经历

- 博士, 清华大学计算机科学与技术系, 2018 – 至今
导师: 裴丹教授。预计 2023 年 6 月毕业。研究方向为智能运维 (AIOps), 故障诊断和根因定位
- 本科, 清华大学计算机科学与技术系, 2014 – 2018
GAP: 92/100, 排名前 10%。清华大学计算机系优秀毕业生

科研项目经历

故障定位

- 在线服务系统中基于图的可解释和无监督的根因定位方法, 2021.10 – 2022.08
该工作聚焦于通过寻找相似历史故障来进行根因定位。相比于有监督方法和无监督启发式方法, 查找相似历史故障既利用了海量的历史故障, 又不依赖大量的根因人工标注。我们提出了一个基于服务依赖图和监控指标的故障表现比较方法, 能够匹配发生在不同位置的同类根因。基于 287 个故障的实验结果表明, 该方法能够将 64.20% 的根因排在前两名, 超过现有的无监督方法 62.88%。
成果投稿在 INFOCOM 2023 (一作, CCF-A 类)
- 对在线服务系统中的重复类型故障的可操作和可解释的故障定位方法, 2021.06 – 2022.03
该工作研究从历史故障数据中有监督地学习故障表现和根因之间的关系, 针对此问题首次提出了一个可操作可解释的基于图注意力的根因定位模型。模型的输入是在线服务系统中的组件 (例如, 服务, 虚拟机, 服务器等) 之间的依赖关系, 以及每个组件上监控指标作为; 模型的输出是故障的根因位置 (例如, 故障服务器的名字) 和根因类型 (例如, 内存不足)。在四个系统上 600 余个故障的实验结果表明, 该方法能将正确的根因排序到第 1.66-5.03 位, 超过前人方法 54.52%-97.92%。
成果接收于 ESEC/FSE 2022 (一作, CCF-A 类), 代码开源在 <https://github.com/NetManAIOps/DejaVu>
- 基于调用轨迹分析的轻量化和实用的故障服务定位方法, 2019.08 – 2021.02
该工作基于微服务系统中的调用轨迹数据, 在故障发生时, 无监督地定位系统中的根因微服务。该工作首次将基于谱的故障定位 (SBFL) 方法引入微服务系统故障定位, 并提出了一套无监督的调用轨迹异常检测方法。在两个系统上 222 个故障的实验结果表明, 该方法能够将 83% 的故障的根因微服务排在第一, 超过前人方法 43%。该工作主要由我带领的包括一名实习本科生的团队完成。
成果发表于 IWQoS'21 (一作, CCF-B 类), 代码开源在 <https://github.com/NetManAIOps/TraceRCA>
- 通用和稳健的基于多维指标数据的故障根因线索定位方法, 2019.01 – 2019.06
该工作聚焦于故障服务内部, 通过在多维数据中 (无监督地) 定位根因属性组合, 为根因定位提供有用的线索。该工作扩展了在多维数据中描述根因属性组合性质的涟漪效应, 并据此首次提出了一个结合自底向上和自顶向下的方法, 在保证定位时效性的同时提高了准确度。实验结果表明, 我们的方法更加通用和稳健, 在异常不够显著等特殊条件下该方法的性能远超前人方法, 总的 F1-score 比前人方法提高了 0.40。该工作主要由我带领的包括三名实习本科生的团队完成。
成果发表于 ISSRE'19 (一作, CCF-B 类), 代码开源在 <https://github.com/NetManAIOps/Squeeze>

异常检测

- 基于图神经网络的调用轨迹异常检测方法, 2020.10 – 2021.04
微服务系统中的调用轨迹可能非常复杂, 每条调用轨迹都是包括几十上百个调用的复杂的图。我们发现这样的调用轨迹可能存在两种异常的模式: 调用轨迹出现异常的结构 (结构异常) 和其中一些调用的状态出现异常 (指标异常)。我们提出了一种基于图神经网络的方法, 以同时检测这两类异常。我们首先对每条调用链使用 FEATHER 算法进行 graph embedding, 其中节点的状态通过 one-hot, binary, hash 等不同方法进行编码。然后, 我们设计了一个深度贝叶斯网络, 它可以分别检测调用轨迹的结构异常和状态异常。
成果形成已授权专利 CN113516174B
- 基于变分自动编码器的稳健和无监督的业务指标异常检测方法, 2018.01 – 2018.09
该工作通过变分自动编码器模型对单条业务指标进行无监督的异常检测, 首次提出通过引入时间信息作为重建指标模式的条件变量, 相比于前人方法异常检测的 F1-score 提高了 0.08-0.43。
成果发表于 IPCCC'18 (一作, CCF-C 类), 代码开源在 <https://github.com/NetManAIOps/Bage1>

实习经历

- 算法工程师, 北京必示科技公司, 2019 年 1 月 – 2022 年 6 月
 - 基于服务间调用 KPI 数据的根因服务定位。我主要负责了算法研究, 产品化开发 (使用 Scala 语言开发算法核心模块, 同包括产品、前端等 5 名团队成员合作), 在客户现场调研实际运用效果和算法优化。产品已商用于国内多家银行, 取得良好效果。有授权专利 CN111597070B (第一发明人)
 - 基于调用轨迹数据的根因服务定位。我主要负责了前期数据情况调研, 算法研究, 和产品化开发

发表论文

以第一作者身份发表 CCF-A 类论文 1 篇, CCF-B 类论文 2 篇, CCF-C 类论文 1 篇; 在投 CCF-A 类一篇, CCF-B 类 1 篇。
以其他作者身份发表 CCF-A 类论文 5 篇, CCF-C 类会议 2 篇。

- (在投) **Zeyan Li**, Mingjie Li, Qingyang Yu, Dan Pei. 2022. Interpretable Graph-Based Root Cause Analysis in Online Service Systems. INFOCOM 2023. (CCF A 类会议)
- (在投) **Zeyan Li**, Junjie Chen, Yihao, Chen, Chengyang Luo, Yiwei Zhao, Yongqian Sun, Kaixin Sui, Xiping Wang, Dapeng Liu, Xing Jin, Qi Wang, and Dan Pei. 2022. Generic and Robust Root Cause Localization for Multi-Dimensional Data in Online Service Systems. Journal of Systems and Software. (CCF B 类期刊)
- **Zeyan Li**, Nengwen Zhao, Mingjie Li, Xianglin Lu, Lixin Wang, Dongdong Chang, Li Cao, Wenchi Zhang, Kaixin Sui, Yanhua Wang, Xu Du, Guoqing Duan, and Dan Pei. 2022. Actionable and Interpretable Fault Localization for Recurring Failures in Online Service Systems. ESEC/FSE 2022. (CCF A 类会议)
- Mingjie Li, **Zeyan Li**, Kanglin Yin, Xiaohui Nie, Wenchi Zhang, Kaixin Sui, and Dan Pei. 2022. Causal Inference-Based Root Cause Analysis for Online Service Systems with Intervention Recognition. KDD'22. (CCF-A 类会议)
- Xianglin Lu, Zhe Xie, **Zeyan Li**, Mingjie Li, Xiaohui Nie, Nengwen Zhao, Qingyang Yu, Shenglin Zhang, Kaixin Sui, Lin Zhu, and Dan Pei. 2022. Generic and Robust Performance Diagnosis via Causal Inference for OLTP Database Systems. CCGRID'22. (CCF-C 类会议, 清华 B 类)
- Nengwen Zhao, Honglin Wang, **Zeyan Li**, Xiao Peng, Gang Wang, Zhu Pan, Yong Wu, Zhen Feng, Xidao Wen, Wenchi Zhang, Kaixin Sui, and Dan Pei. 2021. An empirical investigation of practical log anomaly detection for online service systems. ESEC/FSE 2021. (CCF A 类会议)
- **Zeyan Li**, Junjie Chen, Rui Jiao, Nengwen Zhao, Zhijun Wang, Shuwei Zhang, Yanjun Wu, Long Jiang, Leiqin Yan, Zikai Wang, Zhekang Chen, Wenchi Zhang, Xiaohui Nie, Kaixin Sui, and Dan Pei. 2021. Practical Root Cause Localization for Microservice Systems via Trace Analysis. IWQOS'21. (CCF B 类会议)
- Ruming Tang, Zheng Yang, **Zeyan Li**, Weibin Meng, Haixin Wang, Qi Li, Yongqian Sun, Dan Pei, Tao Wei, Yanfei Xu, and Yan Liu. 2020. ZeroWall: Detecting Zero-Day Web Attacks through Encoder-Decoder Recurrent Neural Networks. INFOCOM 2020. (CCF A 类会议)
- **Zeyan Li**, Chengyang Luo, Yiwei Zhao, Yongqian Sun, Kaixin Sui, Xiping Wang, Dapeng Liu, Xing Jin, Qi Wang, and Dan Pei. 2019. Generic and Robust Localization of Multi-dimensional Root Causes. ISSRE'19. (CCF B 类会议)
- Wenxiao Chen, Haowen Xu, **Zeyan Li**, Dan Pei, Jie Chen, Honglin Qiao, Yang Feng, and Zhaogang Wang. 2019. Unsupervised Anomaly Detection for Intricate KPIs via Adversarial Training of VAE. INFOCOM 2019. (CCF A 类会议)
- **Zeyan Li**, Wenxiao Chen, and Dan Pei. 2018. Robust and Unsupervised KPI Anomaly Detection Based on Conditional Variational Autoencoder. IPCCC 2018. (CCF C 类会议)
- Guoshun Nan, Xiuquan Qiao, Jiting Wang, **Zeyan Li**, Jiahao Bu, Changhua Pei, Mengyu Zhou, and Dan Pei. 2018. The Frame Latency of Personalized Livestreaming Can Be Significantly Slowed Down by WiFi. IPCCC 2018. (CCF C 类会议)
- Haowen Xu, Wenxiao Chen, Nengwen Zhao, **Zeyan Li**, Jiahao Bu, Zhihan Li, Ying Liu, Youjian Zhao, Dan Pei, Yang Feng, Jie Chen, Zhaogang Wang, and Honglin Qiao. 2018. Unsupervised Anomaly Detection via Variational Auto-Encoder for Seasonal KPIs in Web Applications. WWW '18. (CCF A 类会议)

■ 发明专利

- (已授权) 一种故障定位方法、装置、电子设备及存储介质 (CN111597070B)
发明人: 李则言; 张文池; 程博; 黄成; 陈哲康; 沈梦家; 隋楷心; 刘大鹏
- (已授权) 调用链异常检测方法、计算机设备以及可读存储介质 (CN113516174B)
发明人: 裴丹 (博士导师), 李则言
- (实质审查生效) 基于条件变分自动编码器的 KPI 异常检测方法和装置 (CN113114529A)
发明人: 裴丹 (博士导师), 李则言

■ 技能

- 熟练使用 Python, 包括常见的机器学习和数据科学库, 包括 NumPy, SciPy, Matplotlib, scikit-learn, PyTorch 等等
- 熟悉 Linux 系统上常见的工具, 管理过实验室的服务器集群 (30+ 服务器)
- 会使用 Kubernetes, Docker 等, 熟悉 Elasticsearch, Prometheus, Grafana, Jaeger 等云原生系统监控常会使用到的工具
- 有 C++, Java, Scala 等开发经验, 熟悉 Qt5 等框架。
- 有基于 Python (Django 框架) 和 Javascript (Vue.js, backbone 等框架) 的全栈 Web 应用开发经验