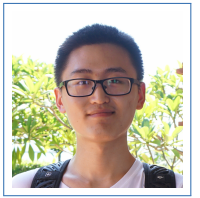


# Zeyan Li

Updated on August 7, 2022

Birth: May 1997  
☎ 184 0165 3327  
✉ [li\\_zeyan@icloud.com](mailto:li_zeyan@icloud.com)  
🌐 [lizeyan](#)  
📄 Zeyan Li



## Education

- **Ph.D. in Computer Science and Technology**, *Tsinghua University*, 2018 – June 2023 (Expected)  
Advisor: Dan Pei. Research interests: AIOps (artificial intelligent operations), failure diagnosis, and root cause localization.
- **B.S. in Computer Science and Technology**, *Tsinghua University*, 2014 – 2018  
GAP: 92/100, top 10%. Outstanding graduates of Department of Computer Science and Technology, Tsinghua University

## Research Projects

### Failure Diagnosis

- **Interpretable Graph-Based Root Cause Analysis in Online Service Systems**, 2021.10 – 2022.08  
We focus on root cause analysis by finding similar historical failures. Compared with supervised and unsupervised heuristic methods, it can leverage the numerous historical failures while not relying on manual labels. We propose a failure matching method based on service dependency graphs and monitoring metrics, which is able to match similar failures caused by root causes at different locations. Our experimental study based on 287 failures shows that our method can rank the root causes at top-2 for 64.20% failures, outperforming unsupervised baselines by 62.88%.  
Submitted to INFOCOM 2023 (first author, CCF-A)
- **Actionable and Interpretable Fault Localization in Online Service Systems**, 2021.06 – 2022.03  
We focus on learning the relationship between failure symptoms and root causes from historical failures. We design an actionable and interpretable graph attention network-based fault localization model. It takes the monitoring metrics and dependencies of the components (e.g., servers, virtual machines, containers, services) of an online service system as inputs and outputs fault locations and types (e.g., CPU exhaustion). Our experiments based on four systems and over 600 failures show that our method can rank the root causes at 1.66-5.03th, outperforming the baselines by 54.52%-97.92%.  
Accepted by ESEC/FSE 2022 (first author, CCF-A). Open source: <https://github.com/NetManAIOps/DejaVu>
- **Practical Root Cause Service Localization in Microservice Systems via Trace Analysis**, 2019.08 – 2021.02  
We localize the root-cause services based on the traces in microservice systems in an unsupervised manner. We adopt SBFL (spectrum-based fault location) and propose an unsupervised trace anomaly detection method. Our experiments based on two systems and 222 failures show that our method can rank the root causes at top-1 for 83% failures, outperforming the unsupervised baselines by 43%. The work was done by a team led by me, including an undergraduate student.  
Published in IWQoS'21 (first author, CCF-B). Open source: <https://github.com/NetManAIOps/TraceRCA>
- **Generic and Robust Root Cause Clues Localization based on Multidimensional Data**, 2019.01 – 2019.06  
We focus on localizing root cause clues based on the multi-dimensional data in the faulty services in an unsupervised manner. We propose *generalized ripple effect*, which describes the relationships between root-cause attribute combinations and other attribute combinations. Based on it, we design a bottom-up and top-up search method, achieving high accuracy while keeping efficient. Our experimental study shows that our approach is more generic and robust. It can significantly outperform the baselines in many situations (e.g., insignificant anomalies). The overall F1 score of our approach outperforms the baselines by 0.40. The work was done by a team led by me, including three undergraduate students.  
Published in ISSRE'19 (first author, CCF-B). Open source: <https://github.com/NetManAIOps/Squeeze>

### Anomaly Detection

- **Grauh Neural Network-Based Trace Anomaly Detection**, 2020.10 – 2021.04  
A trace in microservice systems can be complex, comprising tens to hundreds of spans. We categorize the anomalies of traces as metric anomalies and structural anomalies. We propose a neural network-based approach to detect these two categories of anomalies at the same time. We first perform graph embedding for each trace with FEATHER. The attributes of spans are encoded by one-hot, binary, or hashing. Then, we design a deep Bayesian network to detect the anomalies.  
Our method is patented: CN113516174B
- **Robust Unsupervised KPI Anomaly Detection based on Variational Autoencoder**, 2018.01 – 2018.09  
We focus on unsupervised anomaly detection for univariate business KPIs with variational autoencoders. For the first time, we introduce time information as a conditional variable for better finding posterior. The F1 score of our methods outperforms the baselines by 0.08-0.43.  
Published in IPCCC'18 (first author, CCF-C). Open source: <https://github.com/NetManAIOps/Bage1>

## Work Experience

- **R&D Intern**, *BizSeer*, Jan 2019 – June 2022
  - **Service call KPI-based Root Cause Service Localization**. I am mainly responsible for algorithm research, product development (using Scala language to develop the core module; working together in a 5-member team, including product manager and UI designer), and real-world usage study and improvement. The result is patented (CN111597070B, first inventor).
  - **Trace-based Root Cause Service Localization**. I am mainly responsible for preliminary data study, algorithm research, and product development.

## ■ Publications

First author: 2 CCF-A (1 under review), 3 CCF-B (1 under review), 1 CCF-C. Others: 5 CCF-A, 2 CCF-C

- (Under review) **Zeyan Li**, Mingjie Li, Qingyang Yu, Dan Pei. 2022. Interpretable Graph-Based Root Cause Analysis in Online Service Systems. INFOCOM 2023. (CCF-A)
- (Under review) **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)
- **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)
- 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)
- 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, TH-CPL B)
- 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)
- **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)
- 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)
- **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)
- 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)
- **Zeyan Li**, Wenxiao Chen, and Dan Pei. 2018. Robust and Unsupervised KPI Anomaly Detection Based on Conditional Variational Autoencoder. IPCCC 2018. (CCF-C)
- 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)
- 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)

## ■ Patents

- (Patented) A Fault Localization Approach, Device, Electronic Equipment and Storage Media (CN111597070B) **Zeyan Li**, Wenchi Zhang, Bo Cheng, Cheng Huang, Zhekang Chen, Mengjia Shen, Kaixin Sui, Dapeng, Liu
- (Patented) Trace Anomaly Detection Approach, Computer Devices and Readable Storage Media (CN113516174B) Dan Pei (Ph.D. advisor), **Zeyan Li**
- (Patent pending) A Conditional Autoencoder-based KPI Anomaly Detection Approach and Device (CN113114529A) Dan Pei (Ph.D. advisor), **Zeyan Li**

## ■ Skills

- Skilled in Python and common machine learn and data science libraries, such as NumPy, SciPy, Matplotlib, sklearn, and PyTorch.
- Skilled in using Linux. I have experience in managing a cluster with over 30 bare metals.
- Experienced in using Kubernetes Docker, and the system monitoring stacks (e.g., Elasticsearch, Prometheus, Grafana, Jaeger).
- Experienced in using C++, Java, and Scala. Familiar with Qt5 framework.
- Experienced in full-stack web application development based on Python (Django) and Javascript (Vue.js, backbone).