XIE, HAOTIAN

E-mail: haotian.xie@rutgers.edu Tel: +1 (732) 470-6920

Education

Rutgers University, New Brunswick 01/2023- Present

- Master of Science in Computer Science

The Chinese University of Hong Kong, Shenzhen 09/2018-06/2022

- Bachelor of Engineering in Computer Science and Engineering

Publication

Deep Learning or Classical Machine Learning? An Empirical Study on Log-Based Anomaly Detection

Boxi Yu, Jiayi Yao, Qiuai Fu, Zhiqing Zhong, **Haotian Xie**, Yaoliang Wu, Yuchi Ma, Pinjia He

ICSE 2024 (46th International Conference on Software Engineering)

Professional Experiences

Infrastructure Engineer, **Autra.Tech (autonomous trucking start-up)** 12/2022- Present

- Autonomous Driving Scene Mining System: Implemented a system based on RAM (an image tagging model) and Ray (a distributed framework) for distributed deep learning tasks to mine scene information from autonomous driving videos. Increased utilization of autonomous driving raw records by 10x.
- Hybrid Cloud Object Storage System: Formulated metadata-driven strategies for atomic operations on datasets (a collection of files), enabling tiered storage without compromising data integrity or atomicity. Allowed users to use remote object storage as a mounted directory. Realized an 80% cost reduction in storage.
- Hybrid Cloud Resource Management System: Implemented a GitOps approval-release process based on Terraform and GitHub webhooks, enabling real-time, minute-level adjustments to cloud resources in a hundred-machines cluster. Achieved a 10% cost reduction in cloud resources expenditures.
- Module Package Manager for Autonomous Driving System: Established Bazel-based binary packaging rules and developed a package tool for binary package version control, enabling effective decoupling of various modules within the autonomous driving system, resulting in faster development cycles and more efficient CI/CD pipelines.

Infrastructure Engineer, **Pony.ai (autonomous mobility start-up)** 06/2022-11/2022

- Automatic Virtual Machines Creation Pipeline: Designed and implemented an automated system for creating virtual machine, featuring resource and IP pre-allocation tailored to user needs. Successfully managed the large-scale deployment of several hundred machines. Reduced the deployment time to minutes while maintaining 99% reliability.
- Kubernetes Eviction Policies for Autonomous Driving Workloads: Developed a disk space-based job queuing system to prioritize pod retention. Addressed limitations in default Kubernetes eviction policies and better aligning with training system requirements.
- Automatic Certificates Management System: Utilized Argo workflow, Jira webhooks, Grafana, and Vault to automate monitoring, alerting, and renewal of all certificates. Enhanced management efficiency and minimizing security risks due to certificate expiration.

Intern of Infrastructure Engineer, **Pony.ai** 06/2021-12/2021

- Configuration Management Database (CMDB): Developed and deployed a centralized CMDB system to manage over a thousand servers, streamlining the procurement-to-delivery workflow and significantly enhancing operational efficiency and asset tracking.
- Automatic Hardware Data Collection System with Network Topology Generation: Implemented an automation framework using Ansible and Argo Workflow for hardware data collection and network topology mapping. Enriched existing CMDB data and improved asset management and network visibility.

Intern of Software Engineer, **Shenzhen AgriFuture Technology Co, Ltd.** 07/2020-09/2020

- Keyword Search System: Implemented a robust fuzzy search system based on Solr (an enterprise-search platform), enabling coarse-grained matching for accurate retrieval even with partial character mismatches. Achieved sub-second search times across tens of millions of Chinese characters.

Research Experiences

Research Assistant, Visitor **CUHK Software Engineer and Software Testing Lab** 03/2022-Present

Advisors: Professor He Pinjia

- The paper *Testing Gremlin-Based Graph Database Systems via Traversal Partitioning* (currently under review): Collaborated on the development of an innovative metamorphic testing method that simplifies complex Gremlin queries into sub-queries for enhanced testing. Discovered 28 new bugs across five major GDBMSs, underscoring the method's impact on improving GDBMS stability.
- The paper *Deep Learning or Classical Machine Learning? An Empirical Study on Log-Based Anomaly Detection* (accepted by ICSE 2024): Developed an automated pressure testing methodology that enabled comprehensive collection of diverse error logs. Performed a comparative analysis between traditional machine learning and deep learning in the context of log anomaly detection, finding that traditional machine learning yielded better results.

Team Leader, **CUHK Supercomputing Group** 09/2019-06/2022

Advisors: Professor Chung Yeh-Ching

- Optimized the PRESTO (Pulsar Search Toolkit) - Asia Student Supercomputer Challenge 2021: Employed techniques like vectorization, cache-blocking, and cache-reordering to improve runtime performance by 40x.
- Optimized QuEST (Quantum Exact Simulation Toolkit) - Asia Student Supercomputer Challenge 2020: Introduced a parallel logic gate merging method, boosting simulation speeds by 50x.
- Led and Managed Supercomputing Team Operations: Orchestrated regular meetings focused on topics such as compiler optimization and distributed systems. Handled group operations, server maintenance, and training sessions.

Awards & Honors

2nd Prize, Asia Student Supercomputer Challenge 2020-2021 02/2021
2nd Prize, China Undergraduate Mathematical Contest in Modeling 03/2020
2nd Prize, China Undergraduate Mathematical Contest in Modeling 03/2019

Extracurricular Activities

Founder & President, **CUHK (Shenzhen) Association of Lawn-Bowling & Floor-curling** 09/2019-06/2021

- Established the association and managed a community of over fifty lawn-bowling and floor-curling enthusiasts.
- Led a three-person team in the Wong Tai Sin District Floor-curling Championship, competing against over 30 teams from Asia to secure a notable second-place finish.

Coordinator & Volunteer, **Pacific-Asia Curling Championships** 10/2019-11/2019

- Led the recruitment and management of a volunteer team with fifty people for the Pacific-Asia Curling Championships, implementing a highly effective shift rotation system to ensure smooth event operations.
- Provided translation and guidance services to assist participants and spectators.

Skills & Interests

Technical Skills: Python, C++, Rust, Cloud computing, Distributed system, Automatic driving system
Languages: English (Fluent); Mandarin (Native Language)
Interests: Philosophy (relevant courses: history of philosophy, metaethics, epistemology, etc.), Tennis, Bowls, Curling