CONTACT

✉ cheryllee@link.cuhk.edu.hk

○ cherylLbt

in baitongli

☎ +852 92866527

BAITONG LI

CUHK CSE Ph.D. Student

PROFILE

- Looking for an internship position related to AIOps & software reliability research.

- Proficient in deep learning model development and programming in Python.

- Operation & Information M.eng. at Cornell and CS B.s. at Peking University.

- Rich internship & research experiences in AI-driven problem-solving.

EDUCATION

Ph. D. - **Computer Science & Engineering**   `Aug 2022 - Present`
The Chinese University of Hong Kong - **Hong Kong (China)**

Developing deep learning-based / optimization models to facilitate cloud-scale software systems based on KPIs and logs (AIOps).

Master - **Operation Research & Information Engineering**   `Aug 2020 - May 2021`
Cornell - **New York City (U.S.)**

- GPA: 3.9/4.0 (ranking 5%)

- Major Courses: Applied Machine Learning, Deep Learning, Optimization Methods, Modeling Under Uncertainty, E-Logistics.

Bachelor - **Computer Science & Technology**   `Sep 2016 - Jun 2020`
Peking University - **Beijing (China)**

- GPA: 3.5/4.0 (ranking 30%)

- Major Courses: Probability Theory, Statistics, Game Theory, Applications of Big Data Techniques, Database Systems, Data Structure and Algorithm.

PUBLICATIONS

- 1st author, accepted by ICSE 23 – *EADRO: Integrating Anomaly Detection and Root Cause Localization on Multi-source Monitoring Data for Microservices*

- 1st author, accepted by ICSE 23 – *Heterogeneous Anomaly Detection for Software Systems via Semi-supervised Cross-modal Attention*

- 3rd author, accepted by ICSE 23 – *A Semantic-aware Parsing Approach for Log Analytics*

- 2nd author, submitted to FSE 23 – *AVERT: A Self-adaptive Resilience Testing Framework for Microservice Systems*

- 2nd author, submitted to FSE 23 – *EvLog: Log-based Root Cause Analyzer over Software Evolution*

- Corresponding author, published in International Journal of Electrical Power & Energy Systems (SCI Q1) – *A Novel Probabilistic Framework with Interpretability for Generator Coherency Identification*, ISSN 0142-0615

SKILLS

Deep Learning	Expert
PyTorch	Expert
Python	Expert
Data Mining	Advanced
Modeling	Advanced
C++	Basic

ACHIEVEMENTS

TOFEL

105/120
R30, L27, S22, W26

GRE

V152,Q169

Patents

Big data analysis platform for AC / DC power grid with a high proportion of alternative energy.

Awards

Merit Scholarship

Merit Scholarship in 2020 of Cornell

National 3rd Prize

The 17th Challenge Cup

1st Prize

The 4th Baidu star entrepreneurship competition

WORK EXPERIENCE

Research Assistant   `Jul 2021 - Jul 2022`
The Chinese University of Hong Kong, Hong Kong

- Conducted research in developing deep learning models for automated frameworks to guarantee the reliability of large-scale cloud software systems.

- Proposed a cross-modal attention-based approach to fuse heterogeneous data (i.e., text-based logs and multivariate KPIs) for anomaly detection (accepted by ICSE 23).

- Proposed a multi-modal GNN-based approach to troubleshoot microservices by integrating anomaly detection and root cause localization into an end-to-end manner (accepted by ICSE 23).

Machine Learning Engineer   `Jul 2020 - Dec 2020`
Apple Inc., Beijing (China)

- Proposed an unsupervised "Gradual Clustering" algorithm to extract templates from factory logs, saving half of the memory compared with baselines.

- Constructed a Transformer-based model to process over 400,000 log files in parallel, achieving over 98.29% top-1 accuracy.

- Designed an approach for identifying duplicated issues based on the longest common sequence for fault diagnosis.

Risk Analyst   `Jul 2019 - Sep 2019`
Deloitte Touche Tohmatsu CPA Ltd., Beijing (China)

- Recognized name entities and extracted relationships from the reports of listed companies in the electric power industry via a BiLSTM-CRF model.

- Aligned the entities via the BIRCH clustering algorithm to help automate the process of industrial analysis.

- Devised a grading system to evaluate the operation and potential of focused electric companies.

Data Analyst   `Jul 2018 - Sep 2018`
Baidu Inc., Beijing (China)

- Monitored the behavior logs of 3 billion users interacting with a voice-assisted smart device on Hadoop.

- Incubated a traffic routing function by mining user needs from function request logs, whose Page View achieved top 5.

- Investigated and reported the market performance of the low-price selling strategy of similar devices.

PROJECTS

Bitcoin Transaction Strategy Construction Based on Deep Reinforcement Learning   `Graduation Design`
Published in Applied Soft Computing (SCI Q1)

- Proposed a framework for automatic high-frequency bitcoin transactions based on a deep reinforcement learning algorithm, i.e., proximal policy optimization.

- Utilized LSTM as the policy function based on test-back results of static price predictions.

- Extensive empirical studies validated the superiority of the proposed strategy compared to baselines.