

# Cheryl (Baitong) LEE

 <https://cherylllee.vip> |  [cherylllee@link.cuhk.edu.hk](mailto:cherylllee@link.cuhk.edu.hk)  Google Scholar

## EDUCATION

- The Chinese University of Hong Kong** Hong Kong SAR  
*Mphil - Computer Science and Engineering; Advisor: Prof. Michael R. Lyu* Aug 2022 – June 2024
  - **Research Interest:** Human-centered software; Program testing and repair; AI for software reliability.
  - **Main Courses:** Foundations of Optimization, Natural Language Processing, Graph Mining.
- Cornell University (Cornell Tech)** NYC, U.S.  
*Master - Operation Research & Information Engineering* Aug 2020 – June 2021
  - **Main Courses:** Machine Learning, Deep Learning, Optimization Methods, Modeling Under Uncertainty, E-Logistics.
- Peking University** Beijing, China  
*Bachelor - Computer Science & Technology* Aug 2016 – June 2020
  - **Main Courses:** Probability Theory, Statistics, Game Theory, Applications of Big Data Techniques, Database Systems, Data Structure and Algorithm, C++ Programming.

## PUBLICATIONS

- ISSTA** 2024 **C. Lee**, S. C. Xia, J. Huang, S. Li, L. Zhang, M. R. Lyu, “RAMA: A Unified Programming Tutoring Framework via LLM-based Agent Synergy,” in *the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA) 2024*, Under Review
- ICDE** 2024 **C. Lee**, T. Yang, Z. Zhu, Y. Huo, Y. Su, P. He, M. R. Lyu, “SPES: Runtime Cold Start Optimization for Serverless Computing via Differentiated Function Provision,” the 40th IEEE International Conference on Data Engineering (ICDE), Under review
- ISSTA** 2024 T. Yang, **C. Lee**, J. Shen, Y. Su, Y. Yang, Y. Yang, M. R. Lyu, “An Adaptive Resilience Testing Framework for Microservice Systems,” in *the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA) 2024*, Under Review,  URL: [arxiv.org/abs/2212.12850](https://arxiv.org/abs/2212.12850)
- TOSEM** 2024 Z. Zhu, **C. Lee**, Y. Tang, P. He, “HeMiRCA: Fine-Grained Root Cause Analysis for Microservices with Heterogeneous Data Sources,” *Transactions on Software Engineering and Methodology (TOSEM)*, Major Revision
- ICSE** 2023 **C. Lee**, T. Yang, Z. Chen, Y. Su, Y. Yang, M. R. Lyu, “Heterogeneous anomaly detection for software systems via semi-supervised cross-modal attention,” in *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, 2023, pp. 1724–1736,  DOI: [10.1109/ICSE48619.2023.00148](https://doi.org/10.1109/ICSE48619.2023.00148)
- ICSE** 2023 **C. Lee**, T. Yang, Z. Chen, Y. Su, M. R. Lyu, “Eadro: An End-to-End Troubleshooting Framework for Microservices on Multi-source Data,” in *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, 2023, pp. 1750–1762,  DOI: [10.1109/ICSE48619.2023.00150](https://doi.org/10.1109/ICSE48619.2023.00150)
- ASE** 2023 **C. Lee**, T. Yang, Z. Chen, Y. Su, Y. Yang, M. R. Lyu, “Maat: Performance Metric Anomaly Anticipation for Cloud Services with Conditional Diffusion,” in *Proceedings of IEEE/ACM 38th International Conference on Automated Software Engineering (ASE)*, 2023,  URL: [arxiv.org/abs/2308.07676](https://arxiv.org/abs/2308.07676)
- ICSE** 2023 Y. Huo, Y. Su, **C. Lee**, M. R. Lyu, “SemParser: A Semantic Parser for Log Analytics,” in *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, 2023, pp. 881–893,  DOI: [10.1109/ICSE48619.2023.00082](https://doi.org/10.1109/ICSE48619.2023.00082)
- ISSRE** 2023 Y. Huo, **C. Lee**, Y. Su, S. Shan, J. Liu, M. R. Lyu, “EvLog: Identifying Anomalous Logs over Software Evolution,” in *Companion Proceedings of IEEE 34th International Symposium on Software Reliability Engineering (ISSRE)*, 2023, pp. 391–402,  DOI: [10.1109/ISSRE59848.2023.00018](https://doi.org/10.1109/ISSRE59848.2023.00018)
- ISSREW** 2022 T. Yang, **C. Lee**, J. Shen, Y. Su, Y. Yang, Y. Yang, M. R. Lyu, “Managing Service Dependency for Cloud Reliability: The Industrial Practice,” in *Companion Proceedings of IEEE 33rd International Symposium on Software Reliability Engineering Workshops (ISSREW)*, 2022, pp. 67–68,  DOI: [10.1109/ISSREW55968.2022.00041](https://doi.org/10.1109/ISSREW55968.2022.00041)

- SCI Q1** 2022 F. Liu, Y. Li, **B. Li\*** (\*corresponding author), "A novel probabilistic framework with interpretability for generator coherency identification," *International Journal of Electrical Power & Energy Systems*, vol. 143, p.108474, 2022, [DOI: 10.1016/j.ijepes.2022.108474](#)
- SCI Q1** 2021 F. Liu, Y. Li, **B. Li\***, J. Li, H. Xie (\*corresponding author), "Bitcoin transaction strategy construction based on deep reinforcement learning," *Applied Soft Computing*, vol. 113, p. 107952, [DOI: 10.1016/j.asoc.2021.107952](#)

## EXPERIENCE

- The Chinese University of Hong Kong** Hong Kong SAR  
*Research Assistant* *Jul 2021 - Jul 2022*
  - **Cross-modal Anomaly Detection:** Proposed a cross-modal attention-based approach to fuse text-based logs and multivariate metric time series for heterogeneous anomaly detection.
  - **Microservice Root Cause Localization:** Proposed a multi-modal GNN-based approach for troubleshooting microservices, integrating anomaly detection and root cause localization into an end-to-end framework.
- Northeast Electric Power University** Jilin, China  
*Research Assistant Professor (Internship)* *Mar 2021 - Jun 2021*
  - **Coherency Identification:** Advised students to present an improved auto-encoder for probabilistic identification of coherent generators in power systems, using neural networks and the Gaussian mixture model to cluster phasor measurement units.
  - **Analysis Platform:** Designed a data analysis platform for ac/dc power grids with a high proportion of alternative energy; Patented as the first author, CN2020SR1652151.
- Apple Inc.** Beijing, China  
*Machine Learning Engineer (Internship)* *Jul 2020 - Dec 2020*
  - **Log Anomaly Detection:** Proposed an unsupervised "Gradual Clustering" log parser to analyze watch production logs; Designed a Transformer-based detector and achieved over 98.29% top-1 accuracy.
  - **Duplicated Issue Identification:** Proposed a longest common substring-based algorithm to identify duplicated issues.
- Deloitte Touche Tohmatsu CPA Ltd.** Beijing, China  
*Risk Analyst (Internship)* *Jul 2019 - Sep 2019*
  - **Report Analysis Automatization:** Leveraged BiLSTM-CRF to recognize name entities and extract relationships from the financial reports; Aligned the entities via BIRCH clustering.
  - **Company Appraisal:** Devised an appraisal system to quantify the operation and potential of electric companies.
- Baidu Inc.** Beijing, China  
*Data Analyst (Internship)* *Jul 2018 - Sep 2018*
  - **User Demand Incubation:** Incubated a traffic routing function by monitoring and mining behavior logs of users interacting with a voice-assisted smart device. The function's Page View achieved the top 5.
  - **Market Investigation:** Investigated the market performance of the low-price selling strategy of similar devices.

## HONORS AND AWARDS

- **Postgraduate Studentship** - The Chinese University of Hong Kong 2022
- **Merit Scholarship** - Cornell University 2020
- **National 3rd Prize** - The 17th Challenge Cup 2021
- **Bronze Prize** - Internet Plus College Students' Innovation and Entrepreneurship Competition 2021
- **Meritorious Winner** - Mathematical Contest in Modeling 2018

## SKILLS

- **Coding:** Python (proficient), HTML/CSS, JavaScript, C++ **Frameworks:** PyTorch, Pandas, Spark
- **Languages:** English, Mandarin **Tools:** Git, Docker, Linux/Unix

## ACTIVITIES

- **Teaching Assistant:** (CSCI3100) Software Engineering, (CSCI2720) Building Web Applications 2022-2024
- **Rural Women in STEM Initiative:** Founder 2022-Present
- **Ethnic Minority (Manchu) Culture Communication Club:** Founder 2018-2020
- **Event Hosting:** PKU Campus Top Ten Signers; PKU International Cultural Exchange Festival 2018-2019