

QI LUO

 [Github@rocky-lq](https://github.com/rocky-lq)  qluo615@connect.hkust-gz.edu.cn  [Google Scholar](#)  +86-188-9859-1764

EDUCATION

The Hong Kong University of Science and Technology (Guangzhou) <i>Ph.D. in Microelectronics, GPA: N/A</i> <i>Courses: CUDA Programming, Efficient Computing in Heterogeneous Systems, etc.</i>	<i>Guangzhou, China</i> <i>Aug.2025 - Sep.2029 (exp.)</i>
Southern University of Science and Technology <i>M.Eng. in Computer Science and Technology, GPA: 3.5 / 4.0</i> <i>Courses: Advanced Algorithms, Advanced Artificial Intelligence, Numerical Computing, Reinforcement Learning, etc.</i>	<i>Shenzhen, China</i> <i>Sep.2022 - Jun.2025</i>
Xidian University <i>B.Eng. in Computer Science and Technology, GPA: 3.6 / 4.0</i> <i>Courses: Operating Systems, Computer Networks, Advanced Mathematics, Linear Algebra, Data Structures and Algorithms, etc.</i>	<i>Xi'an, China</i> <i>Sep.2015 - Jun.2019</i>

RESEARCH PAPERS

- [TSE'2025] **Condor: A Code Discriminator Integrating General Semantics with Code Details.** (CCF-A)
Qingyuan Liang, Zhao Zhang, Chen Liu, Zeyu Sun, Wenjie Zhang, Yizhou Chen, Zixiao Zhao, **Qi Luo**, Wentao Wang, Yanjie Jiang, Yingfei Xiong, Lu Zhang. [\[Paper\]](#)
- [NeurIPS'2025] **Decompile-Bench: Million-Scale Binary-Source Function Pairs for Real-World Binary De-compilation.** (CCF-A)
Hanzhuo Tan, Xiaolong Tian, Hanrui Qi, Jiaming Liu, Zuchen Gao, Siyi Wang, **Qi Luo**, Jing Li, Yuqun Zhang. [\[Paper\]](#) [\[Code\]](#) [\[Dataset\]](#)
- [ACL'2025 Findings] **Grammar-Based Code Representation: Is It a Worthy Pursuit for LLMs?** (CCF-A)
Qingyuan Liang, Zhao Zhang, Zeyu Sun, Zheng Lin, **Qi Luo**, Yueyi Xiao, Yizhou Chen, Yuqun Zhang, Haotian Zhang, Lu Zhang, Bin Chen, Yingfei Xiong. [\[Paper\]](#)
- [TOSEM'2025] **Prompt-based Code Completion via Multi-Retrieval Augmented Generation.** (CCF-A)
Hanzhuo Tan*, **Qi Luo***, Ling Jiang, Zizheng Zhan, Jing Li, Haotian Zhang, Yuqun Zhang [\[Paper\]](#)
- [EMNLP'2024] **LLM4Decompile: Decompiling Binary Code with Large Language Models.** (CCF-B)
Hanzhuo Tan, **Qi Luo**, Jing Li, Yuqun Zhang. [\[Paper\]](#) [\[Code\]](#) [\[Models\]](#)
- [FSE'2023] **Enhancing Coverage-Guided Fuzzing via Phantom Program.** (CCF-A)
Mingyuan Wu, Kunqiu Chen, **Qi Luo**, Jiahong Xiang, Ji Qi, Junjie Chen, Heming Cui, Yuqun Zhang. [\[Paper\]](#)

RESEARCH EXPERIENCE

HKUST(GZ) Secure Architecture Lab, Research Assistant <i>Supervised by Prof. Yun Chen</i>	<i>Guangzhou, China</i> <i>Jun.2025 - Aug.2025</i>
SUSTech ARiSE Lab, Research Assistant <i>Supervised by Prof. Yuqun Zhang</i>	<i>Shenzhen, China</i> <i>May.2022 - Aug.2022</i>

- Worked on dynamic voltage and frequency scaling (DVFS) for large language model training.

- Worked on the effectiveness of fuzz testing for both binary applications and deep learning models.

WORK EXPERIENCE

Kwai <i>LLM Algorithm Engineer Intern</i>	<i>Beijing, China</i> <i>Jun.2023 - Dec.2024</i>
<ul style="list-style-type: none">• Contributed to Kwaipilot, an LLM-powered code assistant plugin with code completion and Q&A capabilities.• Conducted data collection, cleaning, classification, filtering, and processing of over 100TB domain code data.• Conducted the evaluation of code LLMs by constructing diverse code completion benchmarks through extracting abstract syntax trees from code and human labeling, supporting over 10 programming languages.	

- Conducted Continue Pre-Training and Post-Training on various base models, including StarCoder, DeepSeek, Gemma, Llama, etc., ranging in size from 1B to 70B parameters.
- Explored and designed the Block-FIM to enhance the model's ability to automatically stop completions.
- Built a Retrieval-Augmented Generation system, boosting the code completion performance of LLMs.
- Conducted training of user behavior analysis models utilizing real auto-completion usage data, focusing on contextual filtering, to enhance the capabilities of LLMs on edge devices.
- Contributed to the data platform's AI development assistant, including SQL completion, SQL repair, and Text2SQL features, achieving over 2,000 DAU with an apply rate of 30%+.
- Provided intelligent code completion services to Xiaomi.
- During the internship, DAU increased from 60 to 4,000+, and the apply rate grew from 5% to 40%.

Streamax Technology AI Center
Algorithm Engineer

Shenzhen, China
Jul.2019 - Jun.2021

- Contributed to ADAS (Advanced Driver Assistance Systems), including FCW (Forward Collision Warning), LDW (Lane Departure Warning), and PCW (Pedestrian Collision Warning).
- Contributed to FCW. FCW utilizes object detection and regression algorithms to precisely locate vehicle position information in the driving area precisely. FCW can monitor collision risks in real-time during driving, integrated with ranging, modeling systems, and lane perception information. FCW can issue timely warning signals 2.7 seconds before danger occurs, effectively preventing traffic accidents such as rear-end collisions.
- Conducted an automated simulation testing tool based on Qt Creator. This tool can extract frame information from H.264 videos of real-world driving scenarios and use it as input for the algorithm pipeline, enabling the simulation of real-world testing and significantly improving algorithm testing efficiency.

Huawei
Software Development Engineer Intern

Xi'an, China
May.2018 - Aug.2018

- Developed back-end for a build pipeline client to streamline the microservices deployment process.

OPEN-SOURCE CONTRIBUTIONS

- LLM4Decompile:** Decompiling Binary Code with Large Language Models.
 - Accumulated over 6000 stars on GitHub.
 - Featured on Hacker News.
 - LLM4Decompile models have been downloaded over 100k times on Hugging Face.
 - Implemented at ByteDance as an IDA Pro/Ghidra plugin, serving reverse engineers.

SKILLS

- General:** Python, C/C++, Java, Shell, SQL, Git, Matlab, LaTeX, etc.
DevOps: ElasticSearch, gRPC, Docker, Kubernetes, Qt Creator, Ghidra, faiss, Spark, vLLM, etc.
MLOps: Pytorch, WanDB, TensorFlow, LangChain, DeepSpeed, Megatron etc.
Languages: Chinese (native), English (CET-6, TOEFL 83)
AI-assisted: Proficient with code agents for coding, debugging, and rapid prototyping.

SELECTED HONORS & AWARDS

- Excellence Award, GBA CS Institutes Joint Poster Competition. *SUSTech, 2025*
- SF Technology Scholarship (~1%). *XDU, 2018*
- Honorable Mention in the American Mathematical Contest in Modeling (MCM). *XDU, 2018*
- Second Prize in the National College Student Computer Design Competition. *XDU, 2018*
- Outstanding Student Award. *XDU, 2017*
- Third Prize in ACM Collegiate Programming Contest, Shaanxi Province. *XDU, 2017*
- Second Prize in the Provincial College Student Mathematics Competition. *XDU, 2016*
- First-Class & Third-Class Merit Student Scholarship. *XDU, 2015 - 2017*

TEACHING ASSISTANT

- **Object-oriented Analysis and Design:** Teaching Assistant for SUSTech Computer Scicence. *SUSTech, 2023*

STUDENT ACTIVITIES

President of the IBM Club; President of the Streamax Club.

- Organized enterprise-related activities and guided students in entrepreneurial competitions. *XDU, 2016 - 2018*