LIU Qiyu 劉棋毓

PhD · Computer Science

Academic Building, HKUST, Clear Water Bay, Kowloon, Hong Kong

+852 51699044 | lqy@ust.hk, qyliu.cs@gmail.com | https://qyliu-hkust.github.io

Education

The Hong Kong University of Science and Technology *Hong Kong SAR*

PhD Computer Science *2017–2022*

- **Advisor**: Prof. Lei CHEN (IEEE Fellow, ACM Distinguished Scientist)
- **Dissertation**: On Effective and Efficient Learned Data Structure Design
- **Research Interest**: Big Data Analytics, Data Mining, AI4DB, DB4AI, Spatial-Temporal Data

University of Electronic Science and Technology of China *Chengdu, China*

BS Computer Science *2013–2017*

- Yingcai Honors College (College for top-5% undergraduates)
- GPA: **3.99**/4; Academic Ranking: **1st**/89

Major Experience

Hong Kong Applied Science and Technology Research Institute *Hong Kong SAR*

Position: Senior Research Engineer *Nov. 2022 - Present*

- Perform research and prototype development on AI-aided document analysis.
- Develop and deploy a pipeline of document parsing, chart and table extraction, and joint information fusion.
- Design application-driven document analytics solutions for different customers including HKEX, GBCNA, Deloitte, etc.

Hong Kong University of Science and Technology *Hong Kong SAR*

Position: Post-doctorate Research Fellow & Project Manager *Mar. 2022 - Nov. 2022*

- Perform independent research on leveraging ML techniques to enhance query processing efficiency of database engines.
- Help with mentoring postgraduate students on their thesis studies.
- Serve as Co-PI of projects funded by the government (RGC) and industry partners (Microsoft and Webank).
- Serve as reviewers for international conferences (e.g., DASFAA, ICDE, VLDB, SIGMOD) and journals (e.g., TKDE, VLDBJ).

The Revenue Optimization Team of WeWork Inc. *Palo Alto, CA, USA*

Position: Applied Scientist Intern *Jan. 2019–July 2019*

- Designed a predictive model to depict the relationship between price and occupancy rate for revenue optimization.
- Designed the pipeline of data acquisition, data cleaning, model training, and model development.
- Developed a interactive dashboard for data visualization.

Big Data Institute of HKUST *Hong Kong SAR*

Position: Research Assistant & Teaching Assistant *Sep. 2017–Feb. 2022*

- Worked on 3 research projects related to query processing and big data analytics under the supervision of Prof. Lei CHEN.
- Published 8 first-authored academic papers on top-tier venues (3 papers on ML-enhanced DBMS, 3 papers on spatial-temporal query processing, and 2 papers on smart city applications).
- Help with proposal drafting and management for funded projects.
- Help with teaching undergraduate and graduate level courses in HKUST: MSBD5002, COMP3111, and COMP1022.

Project Experience

Self-tuned Database System *HKUST and SJTU*

Role: Principal Researcher *2020 - 2022*

- Designed and implemented a self-tuned ML-based multi-dimensional histogram for efficient query cardinality estimation.
- Proposed an adaptive and configuration-free Bloom filter for several DB and networking applications, i.e., distributed join processing, web caching, duplicate detection, etc.
- Published the research outputs on top-tier database conferences (VLDB 2020 and ICDE 2021).

Analytical Query Processing Engine *HKUST*

Role: Principal Researcher *2019 - 2020*

- Improved the Hamming space similarity search through data-aware partition and ML-based indexing.
- Proposed a framework to recommend region-of-interest with user-defined score function.
- Solved the fundamental maximum range-sum problem in a probabilistic setting (PMaxRS).
- Published the research outputs on top-tier database and GIS conferences (SIGMOD 2022, DASFAA 2021, SIGSPATIAL 2019).

Smart Transportation Application *HKUST and HK Government*

Role: Principal Researcher *2019 - 2020*

- Designed an efficient and effective monetary incentive mechanism for spatial spatial crowdsourcing APPs like shared mobility.
- Proposed a data-driven solution to the charger deployment problem for electric vehicles.
- Published the research outputs on top-tier database and GIS conferences (DASFAA 2020, ICDE 2020, SIGSPATIAL 2019).

Blockchain-based Crowdsourcing Platform *HKUST*

Role: Co-Researcher *2019*

- Solved non-transparent incentive mechanism and worker profiles problems in crowdsourcing platforms through Blockchain.
- Implemented a Blockchain-based crowdsourcing framework by leveraging the transparent data model and smart contract provided by Ethereum.

Publication List

HAP: An Efficient Hamming Space Index Based on Augmented Pigeonhole Principle, SIGMOD 2022 (**CCF A**).
Qiyu Liu, Yanyan Shen, and Lei Chen

LHist: Towards Learning Multi-dimensional Histogram for Massive Spatial Data, ICDE 2021 (**CCF A**).
Qiyu Liu, Yanyan Shen, and Lei Chen

Efficiently Discovering Regions of Interest with User-Defined Score Function, DASFAA 2021 (**CCF B**).
Qiyu Liu, Libin Zheng, Xiang Lian, and Lei Chen

Finish Them on the Fly: An Incentive Mechanism for Real-Time Spatial Crowdsourcing, DASFAA 2020 (**CCF B**).
Qiyu Liu, Libin Zheng, Yanyan Shen, and Lei Chen (**Best Paper Runnerup Award**)

Stable Learned Bloom Filters for Data Streams, VLDB 2020 (**CCF A**).
Qiyu Liu, Libin Zheng, Yanyan Shen, and Lei Chen

RIDE: A System for Generalized Region of Interest Discovery and Exploration, ICDE 2020 (**CCF A**).
Qiyu Liu, Libin Zheng, and Lei Chen

Probabilistic maximum range-sum queries on spatial database, SIGSPATIAL 2019 (**Top conference in GIS**).
Qiyu Liu, Xiang Lian, and Lei Chen (**Best Paper Candidate**)

Social-Aware Optimal Electric Vehicle Charger Deployment on Road Network, SIGSPATIAL 2019 (**Top conference in GIS**).
Qiyu Liu, Yuxiang Zeng, Lei Chen, and Xiuwen Zheng

Awards

2020	**Best Paper Runnerup Award**, DASFAA 2020
2019	**Nomination as Best Paper Candidate**, ACM SIGSPATIAL 2019
2019	**Student Travel Grant**, ACM SIGSPATIAL 2019
2017–2022	**Postgraduate Scholarship**, HKUST
2017	**Outstanding Graduate Student Award**, UESTC
2017	**Best Bachelor Thesis Award**, UESTC
2014–2017	**Lixin Tang Scholarship**, UESTC

Skill List

Language: Mandarin (native), English (proficient)
Programming: Java (6 years), C++ (4 years), C (4 years), Python (5 years), Matlab (2 years)
AI & ML: Scikit-learn (4 years), Matplotlib (4 years), Tensorflow (1 year), Pytorch (1 year)