

刘棋毓 LIU QIYU

手机：(+852) 51699044 · 邮箱：lqy@ust.hk keiyuk.liu@gmail.com

通讯地址：Academic Building, HKUST, Clear Water Bay, Kowloon, Hong Kong SAR

教育背景

香港科技大学，计算机科学与工程系，博士 2017.09 - 2022.02

- 导师：陈雷教授 (IEEE 会士, ACM 杰出科学家)
- 学位论文标题：On Effective and Efficient Learned Data Structure Design
- 研究兴趣：大规模数据管理及分析, 数据挖掘, AI4DB, DB4AI, 时空数据分析

电子科技大学，英才实验学院，本科 2013.09 - 2017.06

- 绩点：3.99/4 排名：1/89
- 主要获奖：四川省优秀毕业生、优秀毕业论文奖、唐立新奖学金

职业经历

香港应用科技研究院，高级研发工程师 2022.11 - 至今

- 从事文档 AI 的研发工作，包括：基于 AI 的文档语法分析、图表结构化数据提取、多源数据融合等
- 设计及研发基于应用场景的文档 AI 解决方案，客户包括：港交所、大湾区碳中和协会、德勤等

香港科技大学计算机科学与工程系，博士后研究员兼项目管理 2022.02 - 2022.11

- 从事数据库方向的前沿研究，主要研究内容为基于 AI 的数据库引擎优化及智能查询处理技术
- 作为 Co-PI 申请由香港政府 (RGC) 及工业界合作伙伴 (微软、微众银行等) 资助的科研项目
- 担任数据库领域顶级会议 (ICDE、VLDB、SIGMOD、DASFAA) 和期刊 (TKDE、VLDBJ) 的审稿人

Wework Inc., 应用科学家 (实习) 2019.01 - 2019.07

- 设计基于模型的价格敏感度预测模型，并基于此设计数据驱动的利润最大化算法
- 设计并部署从数据获取、数据清理、模型预测到模型更新的流程

香港科技大学大数据研究院，研究助理兼教学助理 2017.09 - 2022.02

- 进行多个数据库领域前沿问题的研究项目，并在顶级学术会议以第一作者身份发表学术论文 8 篇，包括 CCF A 类论文 4 篇，CCF B 类论文 2 篇，地理信息系统领域顶级会议论文 2 篇
- 协助进行香港科技大学本科及研究生课程的教学工作，包括：大数据挖掘 (MSBD5002)、数据库系统架构 (COMP3111)、Java 编程基础 (COMP1022)

主要项目经历

自学习数据库索引技术，合作单位：香港科技大学、上海交通大学 2020 - 2022

- 设计并实现基于机器学习的高维数据直方图 (Learned Multi-dimensional Histogram) 技术
- 提出一种新颖的数据驱动的自适应稳定布隆过滤器，可广泛应用于搜索引擎、web 缓存等应用
- 相关研究成果发表于数据库顶级会议 ICDE 2021、VLDB 2020

高维及时空数据查询技术，合作单位：香港科技大学、美国肯特大学、微众银行 2019 - 2020

- 设计了一种基于自适应数据划分和基于深度学习的损失估计的高性能汉明空间 (Hamming Space) 索引框架及相应的查询处理算法
- 提出并实现了一种高效的基于自定义评估函数的最优范围推荐算法
- 解决了空间数据库中的最大范围和查询 (Maximum Range-Sum Query) 在概率空间下的扩展问题
- 相关研究成果发表于数据库顶级会议 SIGMOD 2022、DASFAA 2021、SIGSPATIAL 2019

智慧交通及应用，合作单位：香港科技大学、滴滴出行 2020 - 2022

- 提出并实现了一种高效率、高有效性的基于时空众包的激励机制，可以应用于时空任务分发平台如网约车、外卖配送、共享单车等应用
- 提出并实现了一个数据驱动的城市充电桩位置优化系统
- 相关研究成果发表于数据库顶级会议 DASFAA 2020、ICDE 2020、SIGSPATIAL 2019

主要论文著作

- [1] HAP: An Efficient Hamming Space Index Based on Augmented Pigeonhole Principle, SIGMOD 2022. **Qiyu Liu**, Yanyan Shen, and Lei Chen. (CCF A 类)
- [2] LHist: Towards Learning Multi-dimensional Histogram for Massive Spatial Data, ICDE 2021. **Qiyu Liu**, Yanyan Shen, and Lei Chen. (CCF A 类)
- [3] Efficiently Discovering Regions of Interest with User-Defined Score Function, DASFAA 2021. **Qiyu Liu**, Libin Zheng, Xiang Lian, and Lei Chen. (CCF B 类)
- [4] Finish Them on the Fly: An Incentive Mechanism for Real-Time Spatial Crowdsourcing, DASFAA 2020. **Qiyu Liu**, Libin Zheng, Yanyan Shen, and Lei Chen. (CCF B 类、最佳论文奖 (第二名))
- [5] Stable Learned Bloom Filters for Data Streams, VLDB 2020. **Qiyu Liu**, Libin Zheng, Yanyan Shen, and Lei Chen. (CCF A 类)
- [6] RIDE: A System for Generalized Region of Interest Discovery and Exploration, ICDE 2020. **Qiyu Liu**, Libin Zheng, and Lei Chen. (CCF A 类)
- [7] Probabilistic maximum range-sum queries on spatial database, SIGSPATIAL 2019. **Qiyu Liu**, Xiang Lian, and Lei Chen. (地理信息系统领域顶级会议、最佳论文奖候选)
- [8] Social-Aware Optimal Electric Vehicle Charger Deployment on Road Network, SIGSPATIAL 2019. **Qiyu Liu**, Yuxiang Zeng, Lei Chen, and Xiuwen Zheng. (地理信息系统领域顶级会议)

主要获奖经历

- 最佳论文奖第二名 (**Best Paper Runner-up Award**), DASFAA 2020
- 最佳论文奖候选 (**Best Paper Candidate Nomination**), ACM SIGSPATIAL 2019
- 博士研究生奖学金, 香港科技大学, 2017 - 2022
- 优秀毕业论文奖, 电子科技大学, 2017
- 四川省优秀毕业生, 电子科技大学, 2017
- 唐立新奖学金, 电子科技大学, 2014 -2017

专业技能

- 语言: 普通话 (母语), 英语 (熟练掌握听说读写)
- 编程: Java (6 年), C++ (4 年), C (4 年), Python (5 年), Matlab (2 年)
- 数据分析: PostgreSQL (4 年), Spark (2 年), Tensorflow (2 年), Pytorch (2 年)
- 其他: Linux, Bash, Git, Microsoft Office, L^AT_EX