

Yiming Qiao

Tsinghua University, Haidian District, Beijing, 100084, P. R. China
yimingqiao3163@gmail.com / qiaoym21@mails.tsinghua.edu.cn

I recently completed my Ph.D. in Computer Science at Tsinghua University and am an incoming postdoctoral researcher at Carnegie Mellon University. My research focuses on analytical database systems, particularly query execution engines and query optimization.

EDUCATION

Tsinghua University Ph.D. in Computer Science (Institute for Interdisciplinary Information Sciences, IIIS) GPA: 3.76/4.00, TOEFL iBT: 5.0/6.0 (CEFR C1), Advisor: Huanchen Zhang (huanchen@tsinghua.edu.cn)	Sept. 2021 - Jun. 2026 <i>Beijing, China</i>
Nanjing University of Posts and Telecommunications B.Eng. in Computer Science, with Honors Degree GPA: 92.9/100, CET-6: 542	Sept. 2017 - Jun. 2021 <i>Nanjing, China</i>

PUBLICATIONS

- [1] Ruipeng Wan, Xinyu Zeng, **Yiming Qiao**, Huanchen Zhang. “WaRP: Workload-Aware Request Planning for Vector Retrieval on Object Storage.” Under submission to VLDB’27, Revision, 2027.
- [2] Jiansheng Qiu, Fangzhou Yuan, Hengrui Wang, **Yiming Qiao**, Siyuan Bai, Lei Yu, Huanchen Zhang. “Entropy-based Compaction for LSM-trees.” Under submission to SIGMOD’27, Revision, 2027.
- [3] **Yiming Qiao**, Peter Boncz, Huanchen Zhang. “Robust Predicate Transfer with Dynamic Execution.” *Proceedings of the VLDB Endowment (VLDB’26)*, 19(6): 1278–1290.
- [4] **Yiming Qiao**, Huanchen Zhang. “Data Chunk Compaction in Vectorized Execution.” *Proceedings of the ACM on Management of Data (SIGMOD’25)*, 3(1): Article 26, 25 pages.
- [5] **Yiming Qiao**, Yihan Gao, Huanchen Zhang. “Blitzcrank: Fast Semantic Compression for In-memory Online Transaction Processing.” *Proceedings of the VLDB Endowment (VLDB’24)*, 17(10): 2528–2540.
- [6] **Yiming Qiao**, Chi Zhang, Yuxuan Zhu. “Reproducibility Report for ACM SIGMOD 2025 Paper: ‘PilotDB: Database-Agnostic Online Approximate Query Processing with A Priori Error Guarantees.’” *Reproducibility Reports of the 2025 International Conference on Management of Data (SIGMOD’25 Reproducibility)*, 47–48.
- [7] Hu Zhu*, **Yiming Qiao***, Guoxia Xu, Lizhen Deng, Yu-Feng Yu. “DSPNet: A Lightweight Dilated Convolutional Neural Network for Spectral Deconvolution with Self-paced Learning.” *IEEE Transactions on Industrial Informatics (TII)*, 16(12): 7392–7401, 2019. (*Equal contribution)
- [8] Huihui Wang, Shunmei Meng, **Yiming Qiao**, Jing Zhang. “Fast Classification Algorithms via Distributed Accelerated Alternating Direction Method of Multipliers.” *Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM’19)*, 1354–1359.

EXPERIENCE

Suanzhi Future Systems Research Intern	Dec. 2025 - Apr. 2026 <i>Beijing, China</i>
Tencent Summer Intern (Qingyun Program, 青云计划)	Jul. 2025 - Oct. 2025 <i>Shanghai, China</i>
Centrum Wiskunde & Informatica Visiting Student Advisor: Peter Boncz Funding: Partial financial support from CWI; remaining expenses covered by personal funds.	Feb. 2025 - June 2025 <i>Amsterdam, Netherlands</i>
Oracle Software Engineering Intern	Nov. 2019 - May 2020 <i>Nanjing, China</i>
Nanjing University Exchange Student	Sept. 2018 - Jun. 2019 <i>Nanjing, China</i>

RESEARCH PROJECTS

Robust Query Execution in DuckDB

Feb. 2025 - Jan. 2026

Advisor: Peter Boncz, Huanchen Zhang

We address performance regressions in Yannakakis-style joins caused by unnecessary materialization. We design a novel “early-cancel” mechanism that monitors query progress, terminates inefficient work at runtime, and reroutes execution to more efficient pipelines. Evaluated on the SQLStorm benchmark with over 10,000 queries, the approach prevents major slowdowns and delivers robust and predictable performance.

Data Chunk Compaction

Aug. 2023 - Nov. 2024

Advisor: Huanchen Zhang

We identify and systematically study the “small chunk” problem in vectorized query execution, and show that filter and hash join operators are the main sources of the issue. We propose two solutions: Learning Compaction for filters and Logical Compaction for hash joins. Logical Compaction has been incorporated into modern OLAP engines such as Velox and DuckDB (Pull Request #14956).

Compression for Relational Tables

Feb. 2021 - May 2024

Advisor: Yihan Gao, Huanchen Zhang

We propose Delayed Coding, a new entropy coding method in the family of arithmetic coding and asymmetric numeral systems. Based on this method, we design a compressor for row-store OLTP databases. It achieves a high compression ratio while supporting fast random access to individual tuples, around 1 μ s per tuple—slower than main-memory access (200 ns), but much faster than SSD access (40 μ s).

Neural Network-Based Spectrum Deblurring

Jul. 2018 - Dec. 2019

Advisor: Hu Zhu

Developed an end-to-end neural network framework for spectrum deblur in infrared spectrometers, addressing band overlap and noise in aging instruments. By combining dilated convolutions with self-paced learning, the method outperformed traditional partial differential equation (PDE)-based approaches and improved spectral reconstruction quality.

OTHERS

Awards

- Outstanding Graduate, Tsinghua University, 2026.
- Outstanding Graduate, Institute for Interdisciplinary Information Sciences, Tsinghua University, 2026.
- SIGMOD Programming Contest 2025, Third Place Team, 2025.
- Zhang Mingwei Comprehensive Scholarship, Tsinghua University, 2024.
- Excellent Comprehensive Scholarship of IIIS, Tsinghua University, 2023.
- Private Equity Comprehensive Competency Enhancement Program Certificate, Tsinghua GPEI, 2023.
- Bell Honors School Graduate Gold Medal, Nanjing University of Posts and Telecommunications, 2020.
- First-Class Scholarship, Nanjing University of Posts and Telecommunications, 2019.

Teaching & Service

- Artifact Availability and Reproducibility Committee (ARC), SIGMOD 2025.
- Teaching Assistant - Database Systems (Tsinghua 40470414) - Spring 2026
- Teaching Assistant - Quantitative Investment and Financial Optimization (Tsinghua 80470273) - Fall 2023
- Teaching Assistant - Data Mining (Tsinghua 40470333) - Fall 2021
- Invited Talk - Dynamic Execution in Robust Predicate Transfer, CWI Amsterdam, Sep. 12, 2025.
- Invited Talk - Robust Predicate Transfer with Dynamic Execution, Tencent Cloud, Sep. 2, 2025.

Patents

- Zhiqiang Zou, Linrui Li, Shuyu Chang, **Yiming Qiao**, “A Classification Method for Outlier Celestial Objects Based on Astronomical Spectral Data,” Patent CN114282585B, China, Filed Sept. 2020, Granted Nov. 2024.
- Xingguo Chen, **Yiming Qiao**, Wei Liu, Jie Zhu, “A User-Oriented Method for Enhancing Custom Sports Commentary,” Patent CN111477244B, China, Filed Apr. 2020, Granted Sept. 2023.