Xiaoying Wang

xiaoying_wang@sfu.ca • +1-(206)-816-5564 • wangxiaoying.github.io

Research Interest

My research interest is in bridging the gap between database and data science, including applying ML to database applications and accelerating query processing and data accessing for data science.

Education

Simon Fraser University, Canada *2021.01 - present*

- Doctor of Philosophy in Computing Science
- Supervisor: Prof. Jiannan Wang, GPA: 4.2/4.33

Simon Fraser University, Canada *2019.01 - 2020.12*

- Master of Science in Computing Science
- Supervisor: Prof. Jiannan Wang, GPA: 4.2/4.33

Tongji University, China *2012.09 - 2016.07*

- Bachelor of Engineering in Software Engineering
- Outstanding Graduates of Shanghai, GPA: 4.7/5.0

Experience

Simon Fraser University, Canada *2019.01 - present*

Research Assistant

- An Experimental Study of Learned Cardinality Estimation. [code]
  - Conducted comprehensive experiments for comparing 5 learned estimators with 3 databases and 6 traditional methods in both static and dynamic environment. Studied the bad cases of learned estimators and identified their potential problems in production. *Work published in VLDB 2021 and received best EA&B paper award.*
- ConnectorX, A Library for Accelerating Data Loading From Databases to Dataframes. [code] [blog]
  - Developed a library for loading data from databases to dataframes in both Python and Rust. Able to accelerate Pandas.read_sql by 10x with 3x less memory usage. Has been adopted by other popular open source projects like polars, dataprep, modin. *Work published in VLDB 2022 and received 2K+ stars on Github.*
- Accio, A Library for Enabling Efficient Query Federation for Relational Query Engines. [code]
  - Developed a bolt-on library that can enable and accelerate query federation for relational query engines through query rewrite. Designed a new algorithm for join pushdown and a mechanism for considering query partitioning within the cost-based rewrite process. *Work published in VLDB 2025.*

Microsoft Research, USA *2022.05 - 2022.08*

Research Intern - Data Systems Group *(Mentor: Wentao Wu)*

- Proposed and implemented a bound-based mechanism for improving what-if call allocation in index tuning that significantly increases the quality of the indexes recommended. *Work published in SIGMOD 2024.*
- Proposed a new mechanism for early stopping in index tuning that significantly improves the efficiency of index tuners without losing quality on the recommended indexes. *Work published in VLDB 2025.*

Qihoo 360 Technology Co. Ltd., China *2015.07 - 2018.04*

C++ Developer of 360 MAX *(Intern before 2016.07)*

- Developed, maintained, and optimized the performance of 360 MAX (360 AdExchange), a high-concurrency and low-latency real-time online advertising bidding platform.
- Owner of the API service. Communicated and collaborated with engineers from other teams and companies.
- Investigated and deployed new strategies for improving profits and saving costs.

Honors & Awards

VLDB Best Experiment, Analysis, & Benchmark Paper Award *2021*

Excellent Staff of Qihoo 360 *2017*

Outstanding Graduates of Shanghai *2016*

Undergraduate Scholarship *2013, 2014, 2015*

First Class in Tongji University Programming Competition *2014*

Other Activities

Student Volunteer, SIGMOD 2022, VLDB 2023

Team Leader, ConnectorX Project, SFU Data System Lab *2021.01 - 2022.08*

- Organized weekly team meetings and guided undergraduate research assistants to contribute to the project.

Vice Chairman, Google Camp, Tongji University *2014.09 - 2015.06*

- Organized club activities and applied for fundings from Google Beijing.

Publications

Accio: Bolt-on Query Federation *VLDB 2025*
***Xiaoying Wang**, Jiannan Wang, Tianzheng Wang, Yong Zhang.*

Esc: An Early-stopping Checker for Budget-aware Index Tuning *VLDB 2025*
***Xiaoying Wang**, Wentao Wu, Vivek Narasayya, Surajit Chaudhuri.*

Wii: Dynamic Budget Reallocation In Index Tuning *SIGMOD 2024*
***Xiaoying Wang**, Wentao Wu, Chi Wang, Vivek Narasayya, Surajit Chaudhuri.*

ConnectorX: Accelerating Data Loading From Databases to Dataframes *VLDB 2022*
Xiaoying Wang, Weiyuan Wu*, Jinze Wu, Yizhou Chen, Nick Zrymiak, Changbo Qu, Lampros Flokas, George Chow, Jiannan Wang, Tianzheng Wang, Eugene Wu, Qingqing Zhou.*

Are We Ready For Learned Cardinality Estimation? *VLDB 2021, Best EA&B Paper Award*
Xiaoying Wang, Changbo Qu*, Weiyuan Wu*, Jiannan Wang, Qingqing Zhou.*

Invited Talks

Learned Cardinality Estimation: Are We Ready For It? *LinkedIn, 2021.09*

Programming Skills

C/C++, Python, Rust, Java, AWK