

# Xuanhe Zhou

[zhouxuan19@mails.tsinghua.edu.cn](mailto:zhouxuan19@mails.tsinghua.edu.cn) [GitHub](#) [Google Scholar](#) [Homepage](#)

## Profile

---

My research interest lies in self-driving databases, i.e., automating and optimizing the logical/physical designs of databases so as to better serve AI/DB users. The relevant research problems include but are not limited to: (1) Database Optimization (e.g., query rewrite with relational algebra); (2) Database Diagnosis (e.g., root cause analysis); (3) AI-relevant Techniques (e.g., real-time feature computation). I have published papers in SIGMOD/VLDB/ICDE conferences and TKDE journal ([over 1000 citations in total](#)). I have led several open projects, e.g., DBMind, FEBench, DB-GPT, and gained over one thousand stars in GitHub.

## Education & Awards

---

**Tsinghua University, Ph.D. Student (Advised by Guoliang Li)**

*09/2019 - present*

**Department of Computer Science and Technology**

Research Direction: Self-Driving Database System

**VLDB Best Industry Paper Runnerup Award, 2023 (First Author)**

**BenchCouncil Top-100 Open Source Achievements, 2023 (Major Contributor)**

**Outstanding Scholarship of Tsinghua University, 2022 (清华特奖)**

**MSRA Fellowship, 2022 (Top-12 in Asia)**

**ByteDance Fellowship, 2022 (Top-10 in China)**

National Scholarship, 2023 | 2021 | 2017

Apple Scholars in AI/ML Nomination, 2021 (One Tsinghua student)

**Beijing University of Posts and Telecommunications, Bachelor**

*09/2015 - 06/2019*

## Selected Publications

---

1. **D-Bot: Database Diagnosis System using Large Language Models.**

**Xuanhe Zhou**, Guoliang Li, Zhaoyan Sun, Zhiyuan Liu, Weize Chen, et al.

VLDB Full Paper (2024).

<http://dbgpt.dbmind.cn/dashboard>, over 400 GitHub stars.

2. **Grep: A Graph Learning Based Database Partitioning System.**

**Xuanhe Zhou**, Guoliang Li, Wei Guo, Luyang Liu.

SIGMOD Full Paper, 1(1): 94:1-94:24 (2023). [[paper](#)] [[code](#)]

3. **FEBench: A Benchmark for Real-Time Relational Data Feature Extraction.**

**Xuanhe Zhou**, Cheng Chen, Kunyi Li, Bingsheng He, Mian Lu, Qiaosheng Liu, Wei Huang, et al.

VLDB Full Paper, 16(12): 3597-3609 (2023). [[paper](#)] [[code](#)]

**Best Industry Paper Runnerup Award**

4. **A Learned Query Rewrite System.**

**Xuanhe Zhou**, Guoliang Li, Jianming Wu, Jiesi Liu, Zhaoyan Sun, Xinning Zhang.

VLDB Demo, 16(12): 4110-4113 (2023). [[paper](#)] [[code](#)]

[http://rewrite\\_demo.dbmind.cn/](http://rewrite_demo.dbmind.cn/)

---

5. **AutoIndex: An Incremental Index Management System for Dynamic Workloads.**

**Xuanhe Zhou**, Luyang Liu, Wenbo Li, et al.

ICDE Full Research, 2196-2208 (2022). [[paper](#)] [[code](#)]

6. **A Learned Query Rewrite System using Monte Carlo Tree Search.**

**Xuanhe Zhou**, Guoliang Li, Chengliang Chai, Jianhua Feng.

VLDB Full Paper, 15(1): 46-58 (2021). [[paper](#)] [[code](#)]

7. **Query Performance Prediction for Concurrent Queries using Graph Embedding.**

**Xuanhe Zhou**, Ji Sun, Guoliang Li, Jianhua Feng.

VLDB Full Paper, 13(9): 1416-1428 (2020). [[paper](#)] [[code](#)]

8. **Database Meets Artificial Intelligence: A Survey.**

**Xuanhe Zhou**, Chengliang Chai, Guoliang Li, Ji Sun.

TKDE Survey Paper, 1096-1116 (2022). [[paper](#)]

9. **QTune: Query-Aware Database Tuning System with Deep Reinforcement Learning.**

Guoliang Li, **Xuanhe Zhou**, Shifu Li, Bo Gao.

VLDB Full Paper, 12(12): 2118-2130 (2019). [[paper](#)] [[code](#)]

10. **openGauss: An Autonomous Database System.**

Guoliang Li, **Xuanhe Zhou**, Ji Sun, Xiang Yu, et al.

VLDB Full Paper, 14(12): 3028-3041 (2021). [[paper](#)] [[code](#)]

11. **Automatic Database Knob Tuning: A Survey.**

Xinyang Zhao, **Xuanhe Zhou** (co-first author), Guoliang Li.

TKDE Survey Paper, 35(12): 12470-12490 (2023). [[paper](#)]

12. **Robustness of Updatable Learning-based Index Advisors against Poisoning Attack.**

Yihang Zheng, Chen Lin, Xian Lyu, **Xuanhe Zhou**, Guoliang Li.

SIGMOD Full Paper, (2024). [[paper](#)] [[code](#)]

13. **Breaking It Down: An In-depth Study of Index Advisors.**

Wei Zhou, Chen Lin, **Xuanhe Zhou**, Guoliang Li.

VLDB Full Paper, (2024). [[code](#)]

14. **TRAP: Tailored Robustness Assessment for Index Advisors via Adversarial Perturbation.**

Wei Zhou, Chen Lin, **Xuanhe Zhou**, Guoliang Li.

ICDE Full Paper, (2024). [[paper](#)] [[code](#)]

15. **Can LLM Already Serve as A Database Interface? A Big Bench for Large-Scale Database Grounded Text-to-SQLs.**

Jinyang Li, Binyuan Hui, Ge Qu, Jiayi Yang, Binhua Li, Bowen Li, Bailin Wang, Bowen Qin, Ruiying Geng, Nan Huo, Xuanhe Zhou, Ma Chenhao, Guoliang Li, Kevin Chang, Fei Huang, et al.

NeurIPS Full Paper, (2024). [[paper](#)] [[code](#)]

16. **AI Meets Database: AI4DB and DB4AI.**

Guoliang Li, **Xuanhe Zhou**, Lei Cao.

SIGMOD Tutorial, 2859-2866 (2021). (tutorial track)

## Selected Projects

---

### LearnedRewrite: Learning-Based Query Rewriting System

2021 - 2023

[http://rewrite\\_demo.dbmind.cn/](http://rewrite_demo.dbmind.cn/)

- Motivation: Query rewriting is a critical problem in query optimization (e.g., PostgreSQL, Calcite, and Soar). The goal of query rewriting is to transform a SQL query at the logical level (e.g., removing redundant operators, pulling up subqueries), such that the rewritten query is equivalent to the original one and has reduced execution time. However, existing methods rewrite SQL queries using pre-defined rule orders, such as attempting to pull up the subquery before pushing down predicates. This approach is limited because it applies rewrite rules in a default order, which may result in a local optimum.
- Solution: This project proposes a cost-aware query rewriting system, utilizing deep tree search algorithms to explore rewriting sequences with high benefits (performance can be improved by up to 600 times). Upon this work, we have been developing an online demonstration that inputs the slow query and schema only, and outputs (i) the rewritten query, (ii) the rewrite details like the differences in logical plans, and (iii) the rewrite analysis like detailed rewrite procedures together with the explanation by language models. *This work was published as a full research paper at VLDB 2022 and a demo paper in VLDB 2023.*

### FEBench: Real-Time Feature Computation for AI Services

2022 - 2023

<https://github.com/decis-bench/febench>

- Motivation: As the use of online AI inference services rapidly expands in various applications (e.g., fraud detection in banking, product recommendation in e-commerce), real-time feature extraction (RTFE) systems have been developed to compute the requested features from incoming data tuples in ultra-low latency. However, though RTFE is increasingly viewed as essential for deploying AI models in production, there is currently no research on the workload characteristics and benchmarks for RTFE, and especially the comparison with existing database workloads and benchmarks.
- Solution: This project studies the RTFE workload characteristics using over one hundred real datasets from open repositories and those from 4Paradigm and its customers. The study highlights the significant differences between RTFE workloads and existing database benchmarks in terms of application scenarios, operator distributions and query structures. Based on these findings, we propose to develop a real-time feature extraction benchmark named FEBench based on the four important criteria for a domain-specific benchmark proposed by Jim Gray. FEBench consists of selected representative datasets, query templates, and an online request simulator. We use FEBench to evaluate the effectiveness of feature extraction systems including OpenMLDB and Flink and find that each system exhibits distinct advantages and limitations in terms of overall latency, tail latency, and concurrency performance. *This work was published as a full industry paper at VLDB 2023 and won the Best Industry Paper Runnerup.*

### D-Bot: Database Diagnosis System using Large Language Models

2023 - Present

<https://github.com/TsinghuaDatabaseGroup/DB-GPT>

- Motivation: Currently, most companies still rely on DBAs for database maintenance (DM, e.g., tuning, configuring, diagnosing, optimizing) to ensure high performance, availability and reliability of the databases. However, there is a significant gap between DBAs and DM tasks. First, it takes a long time to train a DBA. There are numerous relevant documents (e.g., administrator guides), which can span over 10,000 pages for just one database product and consumes DBAs several years to partially grasp the skills by applying in real practice. Second, it is hard to obtain enough DBAs to manage a large number of database instances, e.g. millions of instance on cloud databases. Third, a DBA may not provide in-time response in emergent cases (especially for correlated issues across multiple database modules) and cause great financial losses.
- Solution: This project aims to build an LLM-enhanced database administrator, which can extract the knowledge chunks from numerous diagnosis documents, leverage these knowledge chunks and relevant tools for well-founded diagnoses, and solve complex anomalies with an asynchronous collaboration mechanism. *This project is open sourced at Github, gaining over 400 stars and over 200,000 views on Twitter.*