Xuanhe Zhou

(+86) 188-1152-2118   wonder204zhou@gmail.com   GitHub   Google Scholar   Homepage

Profile

My research interest lies in autonomous database systems. The relevant research problems include but are not limited to: (1) Data Querying Optimization (e.g., query rewrite with relational algebra); (2) Anomaly Diagnosis (e.g., root cause analysis); and (3) AI Acceleration Techniques (e.g., large-scale 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: Autonomous Database Systems

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

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

Tsinghua Top Grade Scholarship, 2022 (清华大学特等奖学金)

MSRA Fellowship, 2022 (Top-12 in Asia)

ByteDance Fellowship, 2022 (Top-10 in China)

National Scholarship, 2023 | 2021 (9 PhD students from the CS department were selected)

Zhongshimo Fellowship, 2021 (Highest Student Honor in Tsinghua CST, 4 postgraduate students were selected)

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

Beijing University of Posts and Telecommunications, Bachelor   *09/2015 - 06/2019*
School of Computer Science, ranked 1/189

China Robotics and Artificial Intelligence Competition Runnerup, 2019

National Scholarship, 2017

Internship

Huawei GaussDB, AI-Native Database   *2018 - 2019*

Designing AI-native database system (published in IEEE Data Eng. Bull.) and implementing learned database configuration tuning methods (published in VLDB 2019).

4Paradigm OpenMLDB, Online Feature Platform   *2022 - 2023*

Implementing pre-aggregation techniques and designing feature computation benchmarking (published in VLDB 2023 and winned best industry paper runnerup).

Selected Publications

Top Conference (Journal) & First Author × 8, Others × 20+, Citation 1000+

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

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

VLDB Full Research Paper (2024). [paper] [code] Benchcouncil Top 100 Open Achievements

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

2. **DB-GPT: Large Language Model Meets Database.**

Xuanhe Zhou, Zhaoyan Sun, Guoliang Li.

Data Science and Engineering, 9(1): 102-111 (2024). [paper] [code]

3. **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]

4. **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

5. **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/

6. **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]

7. **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]

8. **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]

9. **Database Meets Artificial Intelligence: A Survey.**
Xuanhe Zhou, Chengliang Chai, Guoliang Li, Ji Sun.
TKDE Survey Paper, 1096-1116 (2022). [paper]

10. **DBMind: A Self-Driving Platform in openGauss.**
Xuanhe Zhou, Lianyuan Jin, Ji Sun, Xinyang Zhao, Xiang Yu, Jianhua Feng, Shifu Li, Tianqing Wang, Kun Li, Luyang Liu.
VLDB Demo, 14(12): 2743-2746 (2021). [paper]

11. **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] Last 5 Years' Most Cited VLDB Papers

12. **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]

13. **Automatic Database Knob Tuning: A Survey.**
Xinyang Zhao, **Xuanhe Zhou** (co-first author), Guoliang Li.
TKDE Survey Paper, 35(12): 12470-12490 (2023). [paper]

14. **Robustness of Updatable Learning-based Index Advisors against Poisoning Attack.**
Yihang Zheng, Chen Lin, Xian Lyu, **Xuanhe Zhou**, Guoliang Li.
SIGMOD Full Paper, 2(1): V2mod010:1-V2mod010:26 (2024). [paper] [code]

15. **Breaking It Down: An In-depth Study of Index Advisors.**
Wei Zhou, Chen Lin, **Xuanhe Zhou**, Guoliang Li.
VLDB Full Paper, (2024). [paper] [code]

16. **TRAP: Tailored Robustness Assessment for Index Advisors via Adversarial Perturbation.**
Wei Zhou, Chen Lin, **Xuanhe Zhou**, Guoliang Li.
ICDE Full Paper, (2024).

17. **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, Jiaxi 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]

18. **Learned index: A comprehensive experimental evaluation.**
Zhaoyan Sun, **Xuanhe Zhou**, Guoliang Li.
VLDB Full Paper, 16(8): 1992-2004 (2023). [paper] [code]

19. **DBAugur: An Adversarial-based Trend Forecasting System for Diversified Workloads.**
Yuanning Gao, Xiuqi Huang, **Xuanhe Zhou**, Xiaofeng Gao, Guoliang Li, Guihai Chen.
ICDE Full Paper, 27-39 (2023). [paper] [code]

20. **Adaptive code learning for spark configuration tuning.**
Chen Lin, Junqing Zhuang, Jiadong Feng, Hui Li, **Xuanhe Zhou**, Guoliang Li.
ICDE Full Paper, 1995-2007 (2022). [paper] [code]

21. **LearnedSQLGen: Constraint-aware SQL Generation using Reinforcement Learning.**
Lixi Zhang, Chengliang Chai, **Xuanhe Zhou**, Guoliang Li.
SIGMOD Full Paper, 945-958 (2022). [paper] [code]

22. **Machine learning for data management: A system view.**
Guoliang Li, **Xuanhe Zhou**.
ICDE Tutorial, 3198-3201 (2022). (tutorial track)

23. **AI Meets Database: AI4DB and DB4AI.**
Guoliang Li, **Xuanhe Zhou**, Lei Cao.
SIGMOD Tutorial, 2859-2866 (2021). (tutorial track)

24. **基于机器学习的数据库技术综述.**
李国良, **周煊赫**, 孙佶, 余翔, 袁海涛, 刘佳斌, 韩越.
计算机学报 43 (11), 2019-2049 (2020). [paper]

25. **面向 AI 的数据管理技术综述.**
李国良, **周煊赫**.
软件学报 32 (1), 21-40 (2020). [paper]

Selected Projects

openGauss: An Autonomous Database System
2019 - Present

https://gitee.com/opengauss

• **Motivation:** Learning-based database optimization techniques have been widely studied. For example, Oracle utilizes learning-based methods to recommend the materialized views; DB2 is integrating learned cost estimation into optimizers. However, these techniques have not been widely deployed in commercial database systems and there lack an end-to-end autonomous framework.

• **Solution:** This project builds an autonomous database framework with five main components, including (i) learned query optimizer, (ii) learned advisors with self-monitoring, self-diagnosis, self-configuration, and self-optimization techniques, (iii) Model Validation that validate whether a model is effective for a workload, (iv) Model Management that achieves a unified resource scheduling and model management, (v) Training Data Management that collects the runtime database metrics, historical SQL queries and system logs as training data. *This work was published as a full paper in VLDB 2021 and was awarded the President's Commendation Order in Huawei.*

LearnedRewrite: Learning-Based Query Rewriting System
2021 - 2023

http://rewrite_demo.dbmind.cn

• **Motivation:** Query rewriting is a critical problem in query optimization (e.g., PostgreSQL, Calcite, and Soar). However, existing methods rewrite SQL queries using pre-defned 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 and (ii) the rewrite details like the differences in logical plans. *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, 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 characteristics using 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 and provides a real-time feature extraction benchmark based on the four important criteria proposed by Jim Gray. *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 diagnosis to ensure high performance, availability and reliability of the databases. However, it takes a long time to train a DBA. And there are numerous relevant documents (e.g., administrator guides) to learn from. Besides, it is hard to obtain enough DBAs to manage a large number of database instances. Moreover, a DBA may not provide in-time response in emergent cases.

• **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 was open sourced at Github, selected as one of the Benchcouncil Top 100 Open Achievements and gaining over 200,000 views on Twitter.*

Teaching

• **Database Systems, Tsinghua 30240262**
  Head Teaching Assistant [Fall 2020 | 2021 | 2022]

• **Database Technology and Applications, Tsinghua 20740063-92**
  Head Teaching Assistant [Spring 2020 | 2021 | 2022]

• **Database Systems, Tsinghua 30240262**
  Teaching Assistant [Fall 2019]