Ziwei Wang

(+86)18810917622 | wangziwei@stu.pku.edu.cn
No.5, Yiheyuan Road, Haidian District, Beijing

EDUCATION

Peking University   Sep 2021 - Jun 2024

Master of Data Science, Academy for Advanced Interdisciplinary Studies

- Advisor: Prof. Bin Cui (School of Computer Science)
- Research Interests: Database, Machine Learning
- Publications:
  - **Ziwei Wang**, Zheng Zhong, Jiarui Guo, Yuhan Wu, Haoyu Li, Tong Yang, Yaofeng Tu, Huanchen Zhang, Bin Cui. REncoder: A Space-Time Efficient Range Filter with Local Encoder. **IEEE ICDE 2023 (CCF A)**
  - Shiyue Huang, **Ziwei Wang**, Xinyi Zhang, Yaofeng Tu, Zhongliang Li, Bin Cui. DBPA: A Benchmark for Transactional Database Performance Anomalies. **ACM SIGMOD 2023 (CCF A)**
  - Chunhui Chen, **Ziwei Wang**, Ruixin Wang, Tong Yang, Yaofeng Tu. Fast Learned Multi-Dimensional Indexes in Limited Memory. (**in preparation**)
- Awards: Academic Merit Award

Beijing Institute of Technology   Aug 2017 - Jun 2021

Bachelor of Computer Science and Technology, School of Computer Science

- Cumulative GPA: 93.2/100 | Rank: 1/226
- English Proficiency: CET-4: 569 | CET-6: 559
- Awards: National Scholarship (3 times), First-Class Excellent Student Scholarship (4 times), Merit Student of BIT (2 times)

RESEARCH EXPERIENCE

Text-to-SQL using Large Language Model   Apr 2023 - Present

Intern, Tencent

We leverage large language models with LoRA and prompt engineering to achieve SQL generation, SQL translation, SQL correction, and SQL optimization for real-world business scenarios. The performance of our solution is comparable to ChatGPT.

SQL Rewrite using Machine Learning   Jul 2022 - Present

Graduate Student, PKU-DAIR

Base on the rewrite rules provided by Apache Calcite, we employ the Monte Carlo Tree Search (MCTS) to explore the optimal rewrite order, and use the random forest model to select the optimal rewrite rule set for higher efficiency.

Diagnosis of Performance Anomalies in Transactional Database   Oct 2021 - Oct 2022

Graduate Student, PKU-DAIR

We propose reproduction procedures of 9 common database performance anomalies, and construct a large dataset containing normal and anomalous data collected from various environments to evaluate algorithms for anomaly detection and diagnosis.

Query Performance Optimization of LSM-tree-based Database   Sep 2020 - Jan 2023

Intern, Peking University

We propose a novel range filter to improve range query performance of LSM-tree-based databases. By leveraging local encoding and SIMD acceleration, our range filter achieves a remarkable speedup of 4 to 5 times compared to existing methods.

Learned Index Structures of Databases   Aug 2020 - Present

Intern, Peking University

We propose a novel learned index which exhibits exceptional time and space efficiency while accommodating updates. Our proposed method demonstrates superior read and write throughput, surpassing the B-tree by 141% and the updateable learning index PGM by 87%. The algorithm has been deployed in ZTE.

COMPETITION AWARDS

Third Prize of the Computer System Development Capability Competition	2020.08
Champion in the Checkers Group of the 13th China Computer Games Tournament	2019.10
Runner-up in the Surakarta Group of the 13th China University Student Computer Games Competition	2019.10
First Prize in the Academic and Technological Works Competition of the 16th "Century Cup"	2019.05
Third Prize of the 14th "Lianshan Technology" Programming Contest	2019.04
Second Prize of the 10th "Blue Bridge Cup" Programming Contest	2019.03