Yibo Wang

wangyibo321.github.io LinkedIn wang7342@purdue.edu

Education

Purdue University, Department of Computer Science **Aug. 2025 – Present**

PhD *in Computer Science* *West Lafayette, IN*

- Advisor: Prof. Jianguo Wang

Sichuan University, College of Computer Science **Sept. 2021 – Jun. 2025**

B.Eng *in Computer Science* *Sichuan, China*

- GPA: 3.85/4
- Advisor: Prof. Jianguo Wang; Prof. Mingjie Tang

Publications

WAter: A Workload-Adaptive Knob Tuning System (Under Review)

- **Yibo Wang**, Jiale Lao, Chen Zhang, Cehua Yang, Yuanchun Zhou, Jianguo Wang, Mingjie Tang
- Under review at **SIGMOD** 2026, Paper, Project

QUITE: A Query Rewrite System Beyond Rules with LLM Agents (Under Revision)

- Yuyang Song, Hanxu Yan, Jiale Lao, **Yibo Wang**, Yufei Li, Yuanchun Zhou, Jianguo Wang, Mingjie Tang
- Under revision for **SIGMOD** 2026, Paper Project

GPTuner: An LLM-Based Database Tuning System

- Jiale Lao, **Yibo Wang**, Yufei Li, Jianping Wang, Yunjia Zhang, Zhiyuan Chen, Wanghu Chen, Mingjie Tang, Jianguo Wang
- **SIGMOD Record** 2025, Paper

GPTuner: A Manual-Reading Database Tuning System via GPT-Guided Bayesian Optimization

- Jiale Lao, **Yibo Wang**, Yufei Li, Jianping Wang, Yunjia Zhang, Zhiyuan Chen, Wanghu Chen, Mingjie Tang, Jianguo Wang
- **VLDB** 2024, Paper, Project
- **Selected as SIGMOD Research Highlight**

A Demonstration of GPTuner: A GPT-Based Manual-Reading Database Tuning System

- Jiale Lao, **Yibo Wang (Co-first)**, Yufei Li, Jianping Wang, Yunjia Zhang, Zhiyuan Chen, Wanghu Chen, Yuanchun Zhou, Mingjie Tang, Jianguo Wang
- **SIGMOD** 2024 Demo, Paper, Video

Research Experience

Automated DiskANN Index Tuning via Machine Learning **Aug. 2025 – Present**

Advisor: Prof. Jianguo Wang ***Project Leader***

- Developed a constrained multi-objective optimization pipeline to tune DiskANN, designed to jointly optimize recall, queries per second (QPS), and resource cost.
- Employed multi-fidelity Bayesian optimization on dataset subsamples to reduce high build-parameter tuning cost, a feasibility classifier to enforce constraints, and additive-kernel GP surrogates to model monotonic domain knowledge.
- Preliminary results: The method achieves up to **144%** higher QPS and **20%** lower monetary cost simultaneously over the **best-performing** baseline on SIFT100M, while satisfying the recall constraint.

A Query Rewrite System Beyond Rules with LLM Agents **Mar. 2025 – Present**

Advisors: Prof. Mingjie Tang; Prof. Jianguo Wang ***Research Assistant***

- Proposed QUITE, a training-free and feedback-aware query rewrite system that leverages LLM agents to support a broader range of query patterns and rewrite strategies than rule-based methods.
- Devised a multi-agent rewrite framework, a structured LLM-driven knowledge base, an agent context buffer, a hybrid SQL corrector, and a fine-grained hint injection technique.
- Evaluated QUITE on three OLAP workloads against state-of-the-art baselines, achieving up to **38.5%** lower latency and **24.1%** more successful rewrites over the **best-performing** baseline.
- Outcomes: a research paper **under revision** for **SIGMOD 2026** and an upcoming open-source project.

Runtime-Efficient Adaptive Knob Tuning System **Mar. 2024 – Oct. 2024**

Advisors: Prof. Jianguo Wang; Prof. Mingjie Tang ***Project Leader***

- Developed WAter, an adaptive knob tuning framework that uses runtime profile to significantly reduce benchmark evaluation costs by only selecting SQL subsets to evaluate at different time slices.
- Proposed a runtime-statistics-based representativity metric to continually refine subset, a history reuse method to achieve efficient subset tuning, and a hybrid scoring mechanism to choose the most promising configurations to evaluate.
- Evaluated WAter under four OLAP workloads, finding better configurations with up to **73.5%** less tuning time, achieving up to **16.2%** better performance than the **best-performing** alternative.
- Outcomes: a research paper under review at **SIGMOD 2026** and an upcoming open-source project.

Automatic Optimization of Database with Large Language Model **Sept. 2023 – Mar. 2025**

Advisors: Prof. Jianguo Wang; Prof. Mingjie Tang ***Research Assistant***

- Designed and implemented GPTuner, a novel manual-reading database tuning system that automatically exploits domain knowledge to enhance the knob tuning process.
- Developed an LLM-based data pipeline, a prompt ensemble algorithm, a workload-aware and training-free knob selection strategy, and a coarse-to-fine Bayesian optimization framework.
- Evaluated GPTuner under different benchmarks, metrics and DBMS. It identifies better configurations **16x** faster and achieves **30%** performance improvement over the **best-performing** alternative.
- Extended GPTuner for cost-efficient online query optimization using LLM-enhanced contextual Bayesian optimization and online learning that preserves a concise yet representative training dataset.
- Outcomes: a research paper accepted at **VLDB 2024** and awarded the **SIGMOD Research Highlight 2024**, a demo paper accepted at **SIGMOD 2024**, an open-source project with **110+ GitHub stars**, and a journal paper to be submitted to **VLDBJ**.

Services

- Reviewer - TKDE [2025]
- Reviewer - Information Sciences [2025]
- External Reviewer - VLDB [2024][2025]
- External Reviewer - ICDE [2024]
- Teaching Assistant, Purdue CS348: Information Systems [Fall 2025]

Awards

- SIGMOD Research Highlight, 2024