

ZHIZHEN (CATHY) CAI

cathyc5@illinois.edu

<https://cathy-cai.page>

EDUCATION

University of Illinois Urbana-Champaign

Sep 2024 - Jun 2029

- PhD student in Computer Science

University of Science and Technology of China

Sep 2019 - Jun 2023

- Degree of Bachelor of Engineering in Computer Science and Technology

PUBLICATION

Q. Zhang, S. Xu, Qi Chen, Guxin Sui, Jiadong Xie, **Zhizhen Cai**, et al. VBase: A High-Performance Vector Database for Approximate Online Similarity Queries. *OSDI'23*

- Our team identified relaxed monotonicity in existing high-dimensional indices, incorporated vector indices into relational databases with noticeable performance gain.
- I proposed cost abstraction for vector search and implemented query-planning subsystem based on monotonicity we discovered, which minimized the latency for mixed query (query with both scalar and vector filter). I summarized the work in §4.3 and §5.5.
- I helped to optimize our implementation with profiling and coding to solve hot-spots.

RESEARCH & PROFESSIONAL EXPERIENCE

Researcher **XLab**

Sep 2024 - present

Advisor: Tianyin Xu, Xudong Sun

- Compositional verification of Kubernetes Cluster Controllers (ongoing project)

Researcher **Microsoft Research Asia Internship**

Aug 2022 - May 2023

Advisor: Qianxi Zhang & Qi Chen (System Research Group)

- Read through references about decades of progress in graph-based ANNS (approximate nearest neighbor search), and shared insights in during sharing activities.
- Inspected PostgreSQL's source code, communicated with experienced developers, and helped the whole team in extension development.
- Researched query planning & optimization in database, and gave an excellent performance in index selectivity and cost estimation. This work is accepted to **OSDI 2023**.

Technical Support **USTC Work-Study Program in Library**

Sep 2022 - Jul 2023

- Took part-time job as technical support in USTC library (in spare time), as a representative of USTC LUG.
- Helped the library maintain query machine (we ran netbooted linux on them).
- Used web crawler to collect metadata from core databases.

Teaching Assistant **Web Information Processing and Application**

Aug 2022 - Jan 2023

Teacher: Tong Xu

- Designed & Implemented experiment framework on information retrieval and recommendation system, scored homework
- Maintained & updated teacher's lecture, adding SOTA techniques and examples

Teaching Assistant Operating System (Honored)

Jan 2022 - Jun 2022

Teacher: Kai Xing

- Designed & Implemented OS experiment framework for spring 2022 Courses, introduced state of the art tools to class (introduced io_uring and fine-grained lock concept).
- Participated in daily QA & experiment scoring, assisted teacher in schedule making.

Independent Researcher MSRA-USTC Innovation Practice Project Apr 2021 - May 2021

- Investigated and presented on NER (Named Entity Recognition) & NEL (Named Entity linking) methods and current search directions, with Python as back end and Vue as front end
- Participated and got grade B in this project.

LEADERSHIP AND ACTIVITIES

Vice President Linux User Group of USTC

Jun 2021 - May 2022

- Planned and held activities including Software Freedom Day 2021, Hackergame 2021, Club Opening Day, and maintained daily operation of the club

Alternative & Outreach Leader Hackergame 2021

Nov 2021 - Dec 2021

- Drafted and completed outreach proposal, invited other universities' teams, set participants information sharing criteria and form with back-end team and other universities' representatives
- Team leader before and during the contest, delivered work and responded to contingencies (participant cheating, platform being attacked, etc.)

Organizer & Lecturer Software Freedom Day 2021

Sep 2021

- Organized and prepared for the whole event; designed the awards
- Delivered a speech on Wireguard's principle and potential

HONORS AND AWARDS

- Summer 2021 of Open Source Promotion Plan (scholarship of medium-class project)
I implemented an intranet mesh-network based on Wireguard and Bird to replace Tinc project. However, this project was not used in USTC LUG thereafter.

SKILLS

Technical Skills

- Quite familiar with (Arch) Linux
- Proficient in programming language of C++ & Python

Soft Skills

- Teamwork, Communication, Leadership

Language

- Fluent in English
- Native in Chinese (Mandarin)