Updated Sep 2024

ZIYU "SYLVIA" ZHANG

GitHub : SylviaZiyuZhang • sylziyuz@csail.mit.edu

EDUCATION

MASSACHUSETTS INSTITUTE OF TECHNOLOGY Cambridge, MA
Ph.D. student, CSAIL, EECS Sep 2023 - present
Co-advised by Professor Julian Shun and Professor Michael Cafarella

CARNEGIE MELLON UNIVERSITY Pittsburgh, PA
B.S., Computer Science Aug 2018 – Dec 2021
Concentration in Algorithms and Complexity
University Honors

AWARDS AND HONORS

Jacobs Presidential Fellowship, MIT, AY 2023-2024

Senior Leadership Recognition, Carnegie Mellon University, 2022

Phi Beta Kappa, Inducted 2022

RESEARCH EXPERIENCE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY Cambridge, MA

- Graduate Student, Parallel Algorithms Group Sep 2023 - present
  PI: Professor Julian Shun
  Working on approximate nearest neighbor searches on fully dynamic data.
- Graduate Student, Data Systems Group Sep 2023 - present
  PI: Professor Michael Cafarella
  Previously, I contributed to projects on applying causal inference to data systems diagnosis. I'm currently working on multimodal data integration for semantic search and retrieval augmented generation and contributing to efficiency optimizations in declarative AI-compatible data retrieval, including cost-based optimizations and scheduling algorithms for query processing.

CARNEGIE MELLON UNIVERSITY Pittsburgh, PA

- Undergraduate Independent Study, School of Computer Science Jan 2021 – Dec 2021
  Advisor: Professor David Woodruff
  worked on research on sketch-based tensor approximation, including algorithms for general tensor networks and a few other common decompositions, as well as hardness results.
- Undergraduate Research Assistant, Mohimani Lab Aug 2019 – March 2020
  Advisor: Professor Hosein Mohimani
  Worked on improving and benchmarking high performance algorithms on mass spectrometry matching.

RESEARCH INTERESTS

I'm interested in efficient data algorithms and data systems design that practically support reliability of augmented data retrieval and discovery in aspects like freshness, truthfulness, and semantic awareness.

PUBLICATIONS

- **Near-Linear Time and Fixed-Parameter Tractable Algorithms for Tensor Decompositions.** Arvind V. Mahankali, David P. Woodruff, **Ziyu Zhang**. (In alphabetical order) *ITCS 2024.*
- **Sawmill: From Logs to Causal Diagnosis of Large Systems.** Markos Markakis, Anbo Chen, Brit Youngmann, Trinity Gao, **Ziyu Zhang,** Rana Shahout, Peter Baile Chen, Chunwei Liu, Ibrahim Sabek, Michael Cafarella**.** S*IGMOD 2024 Demo*.
- **Press ECCS to Doubt (Your Causal Graph).** Markos Markakis, **Ziyu Zhang**, Rana Shahout, Trinity Gao, Chunwei Liu**,** Ibrahim Sabek, Michael Cafarella**.** *SIGMOD 2024 GUIDE-AI*. **(Best Paper)**
- **From Logs to Causal Inference: Diagnosing Large Systems.** Markos Markakis, Brit Youngmann, Trinity Gao, **Ziyu Zhang,** Rana Shahout, Peter Baile Chen, Chunwei Liu, Ibrahim Sabek, Michael Cafarella**.** *To appear in VLDB 2025.*

MANUSCRIPTS

- **Using Deep Learning Sequence Models to Identify SARS-CoV-2 Divergence. [link]**
  Yanyi Ding, Zhiyi Kuang, Yuxin Pei, Jeff Tan, Ziyu Zhang, Joseph Konan. (Students in alphabetical order)
  Graduate Level course project: CMU 11-785 Introduction to Deep Learning
- **Locality Sensitive Hashing on Causal Graphs.** Joshua Engels, **Ziyu Zhang**. (Equal contribution) (In submission)
  Graduate Level course project: MIT 6.838 Geometric Computing
- **CleANN: Efficient Dynamism of Graph-based Approximate Nearest Neighbor Search. Ziyu Zhang**, Yuanhao Wei, Joshua Engels, Julian Shun. (In submission)

PRESENTATIONS

Near-linear Time and Fixed Parameter Tractable Algorithms for Tensor Decompositions, ITCS, University of California Berkeley, Berkeley, CA. January 2024.

TEACHING

CARNEGIE MELLON UNIVERSITY Pittsburgh, PA

Undergraduate Teaching Assistant, School of Computer Science Jan 2019 – Dec 2021

15-451 Design and Analysis of Algorithms, Head Teaching Assistant	Fall 2021
15-451 Design and Analysis of Algorithms	2020-2021
15-151 Mathematical Foundations for Computer Science	Fall 2019
15-122 Principles of Imperative Programming	Spring 2019

MASSACHUSETTS INSTITUTE OF TECHNOLOGY Cambridge, MA

Graduate Teaching Assistant, EECS Sep 2024 - Dec 2024

Database Systems

Updated Sep 2024

LEADERSHIP AND SERVICE

- Webmaster and Media Officer, MIT EECS, Graduate Women in Course 6, 2024
- Coaching assistant, Egyptian national team for European Girls' Olympiad in Informatics, 2023
- Student Advisory Committee, Carnegie Mellon University, School of Computer Science, 2019-2020

PROFESSIONAL EXPERIENCE

OTTERTUNE Pittsburgh, PA
Software Engineer Summer 2021, Feb 2022 – Aug 2023

Conducted research on database health checks for PostgresSQL and MySQL. Developed a prototype end-to-end pipeline from database metrics collection to database health check result generation and storage. Designed and implemented several critical user management features. Developed some database related features. Helped with improving backend performance, architecture design, and technical maintenance. Mentored a summer intern on data engineering and analytics.

GOOGLE Mountainview, CA
STEP Intern Summer 2020

Developed a mentor-mentee matching web-app in Java and Javascript. Designed the matching algorithm and NoSQL storage layer.

BYTEDANCE Shanghai, China
Algorithms Engineering Intern, Audio Algorithms Team Summer 2019

Experimented with voice texture transformation framework research results, including GAN, variational encoder, and normalizing flow. Implemented chorus partn detection for pop songs in Python.

SKILLS

PROGRAMMING

Proficient: C++, golang, Python.
Familiar: React/Javascript, Rust, SML/OCaml, Java, PASCAL.

TECHNICAL

Adobe Photoshop, Illustrator, Premier, Audition. AWS, Google Cloud, Kubernetes.

LANGUAGES

Mandarin Chinese (native), English (bilingual proficiency), French (advanced), Russian (intermediate)