

Yinda Zhang

✉ yindaz@seas.upenn.edu 🏠 yindazhang.github.io

EDUCATION

University of Pennsylvania

Advisor: Prof. Vincent Liu

Ph.D., Computer and Information Science

Sep. 2022 - Present

University of Chicago

Advisor: Prof. Junchen Jiang

Pre-Doctoral M.S., Computer Science

Sep. 2020 - Dec. 2021

Peking University

Advisor: Prof. Tong Yang

B.S., Computer Science and Technology

Sep. 2016 – Jul. 2020

PUBLICATION

Conference

1. **Yinda Zhang**, Peiqing Chen, Zaoxing Liu, OctoSketch: Enabling Real-Time, Continuous Network Monitoring over Multiple Cores, USENIX Symposium on Networked Systems Design and Implementation (NSDI 2024)
2. Yikai Zhao*, Wenchen Han*, Zheng Zhong*, **Yinda Zhang**, Tong Yang, Bin Cui, Double-Anonymous Sketch: Achieving Fairness for Finding Global Top-K Frequent Items, ACM International Conference on Management of Data (SIGMOD 2023)
3. Yikai Zhao*, **Yinda Zhang***, Yuanpeng Li*, Yi Zhou, Chunhui Chen, Tong Yang, Bin Cui, MinMax Sampling: A Near-optimal Global Summary for Aggregation in the Wide Area, ACM International Conference on Management of Data (SIGMOD 2022)
4. Zhuochen Fan, **Yinda Zhang**, Tong Yang, Mingyi Yan, Gang Wen, Yuhan Wu, Hongze Li, Bin Cui, PeriodicSketch: Finding Periodic Items in Data Streams, IEEE International Conference on Data Engineering (ICDE 2022)
5. Tong Yang, Jizhou Li, Yikai Zhao, Kaicheng Yang, Hao Wang, Jie Jiang, **Yinda Zhang**, Nicholas Zhang, QCluster: Clustering Packets for Flow Scheduling, International World Wide Web Conference (WWW 2022)
6. **Yinda Zhang**, Zaoxing Liu, Ruixin Wang, Tong Yang, Jizhou Li, Ruijie Miao, Peng Liu, Ruwen Zhang, Junchen Jiang, CocoSketch: High-Performance Sketch-based Measurement over Arbitrary Partial Key Query, Annual conference of the ACM Special Interest Group on Data Communication (SIGCOMM 2021)
7. **Yinda Zhang**, Jinyang Li, Yutian Lei, Tong Yang, Zhetao Li, Gong Zhang, Bin Cui, On-Off Sketch: A Fast and Accurate Sketch on Persistence, International Conference on Very Large Data Bases (VLDB 2021)
8. Xiangyang Gou*, Long He*, **Yinda Zhang***, Ke Wang, Xilai Liu, Tong Yang, Yi Wang, Bin Cui, Sliding Sketches: A Framework using Time Zones for Data Stream Processing in Sliding Windows, ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2020)

Journal

1. Ruijie Miao*, **Yinda Zhang***, Zihao Zheng*, Ruixin Wang, Ruwen Zhang, Tong Yang, Zaoxing Liu, Junchen Jiang, High-Performance Sketch-based Measurement over Arbitrary Partial Key Query, IEEE/ACM Transactions on Networking (TON 2023)
2. Xiangyang Gou*, **Yinda Zhang***, Zhoujing Hu*, Long He, Ke Wang, Xilai Liu, Tong Yang, Yi Wang, Bin Cui, A Sketch Framework for Approximate Data Stream Processing in Sliding Windows, IEEE Transactions on Knowledge and Data Engineering (TKDE 2022)

SELECTED RESEARCH PROJECTS

Network Measurement System

Collecting Network Telemetry Without Affecting User Traffic Nov. 2022 – Present

- Developed a communication substrate for collecting network telemetry from switches silently
- Implemented the system on the Tofino switch and run large-scale simulations on ns-3
- Achieved both near-zero overhead on user traffic and high sustainable throughput for telemetry data

Real-Time, Continuous Network Monitoring over Multiple Cores Sep. 2021 – Sep. 2022

- Designed a software monitoring framework scaling sketching algorithms to many cores
- Applied the framework to nine sketches on three software platforms (CPU, DPDK, and XDP)
- Improved accuracy by $15.6\times$ and throughput by $4.5\times$ compared to the state-of-the-art

High-Performance Sketch-based Measurement over Arbitrary Partial Key Query Jul. 2020 – Jul. 2021

- Designed a sketch-based measurement system supporting queries on multiple keys simultaneously
- Implemented the system on four platforms (CPU, Open vSwitch, P4, and FPGA)
- Improved accuracy by $10.4\times$ and throughput by $27.2\times$ when measuring multiple flow keys

Approximate Streaming Algorithm

Enhance Query Accuracy of Sketches by Machine Learning Oct. 2023 – Present

- Applied ML algorithms to improve query accuracy of sketches without any modification on update
- Investigated the impact of ML models, feature sets, and loss functions

Near-optimal Global Summary for Aggregation in the Wide Area Jun. 2021 – Mar. 2022

- Designed a fast, adaptive, and accurate sampling scheme for global aggregation in WAN
- Derived the error bound of the algorithm and proved the optimality
- Applied the algorithm to federated learning, distributed state aggregation, and hierarchical aggregation

Fast and Accurate Sketch on Persistence Jan. 2020 – Sep. 2020

- Designed a sketch to address persistence estimation and finding persistent items
- Derived the error bound and showed smaller space complexity
- Conducted experiments on multiple datasets and showed $6.17\times$ better accuracy

INDUSTRY EXPERIENCE

Conviva

Jul. 2020 – Jan. 2021

Research Intern

Beijing, China

- Measured the viewership of 7.6M video sessions in four popular content providers
- Analyzed the interplay between video content and quality sensitivity
- Managed different data schema from various devices and platforms

MISC

Languages: C++/C, Python, Java, SQL

Fellowship: Jonathan M. Smith Fellowship (2022)

Teaching Assistant: CIS 553: Networked Systems (Spring 2023, Fall 2023)

Shadow TPC Member: IMC 2022, EuroSys 2023