Shengze (William) Wang

Computer Science Doctoral Student, Graduate Student Researcher, UC Santa Cruz

Homepage: https://shengze.io
Email: shengze@ucsc.edu

Technical Skills

- **Academic**: Computer Networks & Distributed Systems and their applications in AI Infrastructure (Systems for AI), AI for Systems & Databases, LLM Inference, VectorDB, Datacenter Networks, Edge Computing, Connected Autonomous Vehicles
- **Languages & Databases**: C/C++, Rust, Python, Bash, Coq, Redis, LevelDB, Cassandra, DynamoDB, Oracle, MySQL
- **Platforms & Tools**: Linux, Network Operating Systems, ROS, Supercomputers (TACC), AWS, Azure, GCP, Docker, Git
- **Highlights**: Applied AI for Systems, Query processing, Key-value Pair Storage, Vector Database, Advanced Network Protocols, Network Security, Hashing Algorithms, Workload Characterization & Balancing, ML System Design, LLM Inference, Content Delivery Network, Quantum Networks, Programmable Switches, Edge-Cloud Systems, Vehicle Computing

Education

- **University of California, Santa Cruz (UCSC)** — San Francisco Bay Area
  Ph.D. in Computer Science and Engineering; Regents Fellowship; BE Dean's Fellowship — *2023 - present*
  Relevant Courses: *Adv Computer Networks, Network Security, Adv Computer Architecture, Adv Distributed Systems, Adv Stream Processing, Programming Languages, Adv Machin Learning, Adv Natural Language Processing, Adv Artificial Intelligence*

- **University of North Texas (UNT)** — Dallas - Fort Worth
  B.S. in Computer Science; GPA: 4.0; Outstanding Award (Top 1 of Class 2023); President's List — *2020 - 2023*
  Relevant Courses: *Algorithms, Machine Learning, Software Engineering, Systems Programming, Database Systems, Computer Networks, Computer Security, Operating Systems, Probability Models, Linear Algebra, IT Project Management, Technical Writing*

- **King's College London (KCL)** — London, United Kingdom
  Visiting Student in Computer Science; Scored: 95/100; JEISE Government Scholarship — *2019*

Working Experience

- **Graduate Student Researcher** — Baskin School of Engineering, UCSC — June. 2023 - present
- **NSF REU Research Mentor** — The VEC Lab, NSF eCAT Center — Jun. 2022 - Aug. 2022
- **Research Assistant** — Department of Computer Science and Engineering, UNT — Dec. 2021 - May. 2023
- **Full-stack Web Engineer** — DS Creative Office, UNT — Sept. 2021 - Jan. 2022

Selected Projects

- **Resource Storage and Discovery in Network and Database Systems — C/C++, Rust** — 2023 - present
  Qian Lab, https://users.soe.ucsc.edu/~qian/
  - Investigate fundamental problems in emerging networks, e.g., Data Center Networks, CDN, Quantum Networks
  - Design and optimize network protocols, routing algorithms, hash functions, and load balancers for Enterprise Infrastructures, e.g., Network Switches, HPC Clusters, IoT
  - Implement and evaluate Network and Database Systems in event-driven simulators, server clusters, and Clouds.
  - Proposed the first Distributed Learned Hash Table with LEAD, a novel system incorporating learned models within DHT structures to significantly optimize range query performance. LEAD achieves tremendous advantages in system efficiency compared to existing range query methods in large-scale distributed systems while maintaining high scalability and resilience to network churn. (https://github.com/ShengzeWang/LEAD)

- **Vehicular Edge Computing and Connected Autonomous Vehicles — Python, ROS** — 2021 - 2023
  NSF eCAT Center for Electric, Connected and Autonomous Technologies, https://www.ecat.center/
  - Profiled the hardware resource utilization of real-time detectors such as Yolo, Faster R-CNN, and SSD.
  - Assessed variations such as quantization, architectural reduction, and floating point precision reduction.
  - Characterized memory contention into three behaviors and quantified the impacts of memory contention for CAVs.
  - Engineered features and modeled the workload behavior for the Edge based on the device configurations.
  - Implemented and tested the Vehicle-to-edge (V2X) communication solutions on edge nodes with AWS Edge Services.

- **False Discovery Rates(FDR) Control in Metaproteomics Search — C++, Python** — 2021 - 2023
  Center for Computational Epidemiology and Response Analysis (CeCERA), https://cerl.unt.edu/
  - Applied computational methodologies and probability models to address FDR estimation bias in metaproteomics studies.
  - Developed the first fine-grained FDR assessment framework with an open-source tool FineFDR to filter the Comet and Percolator results in different taxonomic ranks. (https://github.com/Biocomputing-Research-Group/FDR)
  - Proposed Expectation-maximization General-mixture model for clustering proteomic samples based on abundances.
  - Evaluated six FDR control solutions on 10 datasets. Our novel fine-grained frameworks achieved higher precision, and more peptide and protein identifications compared to state-of-the-art methods such as Comet, Percolator, and Tailor.

- **Fatigue Detection for Medical Staffs**: Constructed a face-masked data set and trained the CNN model for face-masked facial landmarks detection and the LSTM network to measure PERCLOS. (Registered patent: 2020SR1233854)
- **DeepEmo.tech**: A real-time facial expression recognition WebApp using Tiny Face Detector and SSD Mobilenet.
- **Intelligent Traffic Management System**: Designed an effective traffic signal control solution for traffic congestion control in smart cities using Reinforcement Learning and Computer Vision. (Registered patent: 2020SR1235776)

Selected Publications

- **A Distributed Learned Hash Table** September. 2024
  Submitted to 2024 International Conference on Very Large Databases (VLDB) *First Author*
- **Poster: Distributed Learned Hash Table** September. 2024
  Accepted by 2024 IEEE International Conference on Network Protocols (ICNP) *First Author*
- **CALID: Collabrative Accelerate LLM Inference with Draft Model with Filter Decoding** August. 2024
  Accepted by BayLearn 2024 - Machine Learning Symposium *Co-Author*
- **Improving the Faithfulness of LLM-based Summarization with Unlikelihood Training** July. 2024
  Accepted by BayLearn 2024 - Machine Learning Symposium *Co-Author*
- **Perception Workload Characterization and Prediction on the Vehicular Edges** July. 2023
  2023 IEEE International Conference on Edge Computing (EDGE) *Co-First Author*
- **Fine-grained Taxonomy-specific False Discovery Rates Control in Metaproteomics** Nov. 2022
  2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) *First Author & Oral*
- **Applications of Computer Vision Techniques in Industrial Fields: A Review** Apr. 2021
  Journal of Network Security Technology & Application, 2021 (04), ISSN 1009-6833 *First Author*

PROFESSIONAL SERVICES

- **Reviewer** IEEE Transactions on Dependable and Secure Computing (TDSC)
- **Reviewer** IEEE International Conference on Computer Communications (INFOCOM)
- **Teaching Assistant** CSE 13S: Computer Systems and C Programming - 2023 Winter
- **Mentor** NSF Research Experiences for Undergraduates (REU) in Vehicular Edge Computing and Security