ZHEJIAN JIN

github.com/Jack-Kin • zj2324@columbia.edu • www.linkedin.com/in/zhejian-jin • (917) 609-4100

EDUCATION

Columbia University New York, NY
M.S. in Electrical Engineering Expected Dec 2022
- Coursework: Cloud Computing, Distributed Storage Systems, Large-scale Stream Processing, Heterogeneous Computing, Database System Implementation, Big Data Analysis, Intro to Blockchain Technology

Shanghai Jiao Tong University Shanghai, CN
B.S. in Electrical and Computer Engineering Aug 2021
- Coursework: Data Structures and Algorithms, Operating System, Computer Networks, Artificial Intelligence, etc.

EXPERIENCE

Jupiter Research Capital New York, NY
Quantitative Developer Intern Jun 2022 - Sep 2022
- Built high-speed exchange data parsers and data searching engines with self defined index builder for large binary sequential data from China A-shares stock exchange, Nasdaq ITCH and NYSE PILLAR INTEGRATED FEED
- Finished low-latency trades&orders FIX protocol message parsing module and high-throughput database module supporting 30k+ inserts/second for live trading message and historical log message
- Wrote data feeder protocols using REST API, websocket API and FIX protocol for 9 crypto markets (FTX, KuCoin, etc.)

Columbia University New York, NY
CSEE4121 Computer Systems for Data Science Course Assistant Jan 2022 - May 2022
- Designed programming project using Apache Spark, Spark Streaming, HDFS in Google Cloud Dataproc
- Held office hour weekly and answer questions on Piazza daily to answer students' questions from homework and class content, took teaching-team meeting and graded homework with other 9 CAs for the class with 300+ students

Shanghai Jiao Tong University, UM-SJTU Joint Institute Shanghai, CN
Research Assistant Sep 2020 - Apr 2021
- Proposed the MAGIC2 model and Stochastic Curriculum Learning method. Optimized the RL model is by Policy Gradient with a dynamicfurthest-insertion+2-opt baseline during training (Accepted by IEEE SSCI 2021)
- Achieved state-of-the-art results as a Deep Reinforcement Learning solver for TSP, taking both generalizability and computational time into account. E.g.: 8.79% Gap from Concorde Method on TSP200 problem while saving 75% time

East Money Information Co., Ltd. Shanghai, CN
Back-End Software Developer Intern Dec 2019 - Mar 2020
- Implemented network framework with Python to develop websites for large-scale stock data in MongoDB visualization
- Utilized Kafka to maintain the distributed publish-subscribe messaging system
- Developed the Data Quality Detection System independently: Wrote TCP private protocols for specific market messages in Golang to handle 5 problems including packet loss, delay, corruption, duplication and reordering
- Completed the Data Quality Detection System with WeChat alarming function. Improved the network latency through multithreading TCP technique and optimized data-processing time 3 times faster than original design

PROJECTS

Evaluation of DisaggregatedPersistent Memory System Oct 2021 - Dec 2021
Columbia University
- Benchmarked latency and throughput for both local and remote devices including Hard Disk, Ramdisk, DRAM, PM, remote DRAM and remote PM, where remote devices are accessed through RDMA over infiniband
- Designed disaggregated integral test framework, generated workload by memaslap and measured performance, stated guidelines to better design disaggregated PM systems

GPU Acceleration of K-Means Clustering Oct 2021 - Dec 2021
Columbia University
- Proposed multiple GPU parallelization ways to speed up the naive K-means, including using shared memory to calculate distance for each data point and using shared memory and parallel scan to sum up each centroid
- Implemented the algorithms on NVIDIA Tesla T4 in GCP. Beaten the Sklearn K-means algorithm by speeding up 5 times when the number of data points comes to 1,000,000 with dimension of 8 and the cluster number is set to 5

Cloud Restaurant Scheduling Platform Using AWS Services Sep 2021 - Nov 2021

Team Leader, Columbia University

- Formulated RESTful APIs of a restaurant scheduling platform and deployed 3 atomic microservices handling user and address, restaurant and schedule information on AWS EC2, docker and AWS Elastic Beanstalk
- Implemented multiple Lambda Functions to provide different functionalities with different components with API Gateway
- Integrated third party Oauth and encapsulated external cloud service to verify an address

Development of Cross Camera Video Analysing System at Edge May 2021 - Aug 2021

Team Leader, Shanghai Jiao Tong University

- Designed and accomplished a video analysing system dynamically assigning workloads across cameras embedded with Neural Processing Units and a NVIDIA edge cluster
- Applied ZeroMQ for data transmissions among 5 cameras and between cameras and the edge cluster
- Transferred YOLO PyTorch model to RKNN type to be compatible with the NPU API. Pre-compiled the model and deployed the model to cameras by Docker

Development of Multi-thread Database Oct 2020 - Nov 2020

Team Member, Shanghai Jiao Tong University

- Classified queries with table, maintained a dependency table and built a thread pool to run queries in parallel. Protected threads with mutexes, semaphores and condition variables
- Divided large tables into small parts and executed operations such as searching and calculation simultaneously
- Utilized the random-access property of std::vector and kept a keymap to store indices for given keys in order to improve the efficiency of conditional queries

SKILLS & INTERESTS

Languages: C, C++, Golang, Python, CUDA, OpenCL, mySQL, MATLAB
Tools: AWS, GCP, Docker, Kafka, Hadoop, Spark, Airflow
Interests: All levels of knowledge supporting low latency and high throughput applications