Junyi Tao

Seattle, WA
3144050511 jaydent121208@gmail.com junyi-tao JaydenTaojy

EDUCATION

Stony Brook University **Stony Brook, NY**
MS in Computer Science (PhD Candidate) 08/2019–05/2022
University of Southern California **Los Angeles, CA**
MS in Computer Science 01/2018–05/2019
University of Electronic Science and Technology of China **Chengdu, China**
Bachelor of Engineering in Software Engineering 08/2013–06/2017

Work Experience

Amazon Web Services (AWS) **08/2022 – now**
Software Development Engineer *Seattle, WA*

- Work in the AWS networking team to build next generation software services for AWS Network: build services that form the foundation of **network configurations, validations and management**.
- Adopt the principles of **Software Defined Networking (SDN)** and **Intent Based Networking (IBN)** to change the way Networks are conceived, designed, built and operated.
- Put services and algorithms in place to control and manage all layers of the Amazon network, from top of the rack switches in Amazon data centers (TORs) to Amazon internet and private backbone facing routers.
- Derive Amazon's network topologies from higher level designs and intents, and a distributed software system controls the way to manage Amazon network.

Stony Brook University **08/2020 – 01/2022**
Research Assistant *Stony Brook, NY*

- Worked in the network and mobile computing lab in CS department, focusing on wireless network and mobile computing field.
- Worked on the review and study of data stream processing platform (i.e. Storm, Spark, Flink and some novel platforms from recent prestigious publications) in the fields of design architecture and code implementation.
- Implemented the **Apache Flink** with **Deep Learning** operation support on lightweight edge computing devices (i.e. Nvidia Jetson).
- Worked on several personal/leading projects, some highlighted projects are described under the **Selected Projects** section in detail.

SELECTED PROJECTS

Data stream processing on MIMIC-IV healthcare dataset **05/2021-12/2021**

- Extracted well organized data features from MIMIC dataset: with the original execel files, build local **PostgreSQL** database and use **Python** scripts for data extraction and organization.
- Designed and implemented several variants of **LSTM** deep learning models by using **Pytorch** to perform the in-hospital mortality predictions work on organized data.
- Based on the offline learning model, used **Apache Kafka** and **Apache Spark** to implement the real time online learning and inference frameworks for in-hospital-mortality work.

Octane Routers **01/2019-05/2019**

- Implement the software defined network system using **C++** under Linux.
- Use processes to work like network routers: one central router call fork() to initialize other routers; the central router makes route decisions and other routers just route traffic under central control.
- The Routers were interfaced with the outside world through RAW SOCKETS; implemented reliable transmission between routers by using timer to track each communication packet.
- Created sudo headers to send out TCP traffic through RAW SOCKETS to the outside world; performed load balancing in case of heavy traffic.

Group User Identification based on Acoustic Breathing Signals **08/2020-12/2020**

- Built a system for group user identifications. MFCC features were extracted from the acoustic breath signals, and a LSTM based RNN model is trained to perform user classification from the MFCC signals.
- Developed a group user authentication system, which enable the systems to identify a user from a group of legal users, to support more flexible application scenarios by providing multiple access control mechanisms.
- Developed a web app by using **html and javascript** that wraps up all the functions (MFCC feature extraction and person identity classification) as web APIs and enable human computer interaction in real time.

TECHNICAL SKILLS

Programming Languages: Python, Java, C/C++, SQL, Shell Scripts, Javascript
Tools: Linux, Docker, GDB, CMake, MySQL, PostgreSQL, Kafka, Flink, Spark