ALLAN PENG

Email: megatron@berkeley.edu Website: www.allan.expert

WORK EXPERIENCE

Mod9 Technologies

Software Engineer *Spring 2018-Present*

- Lead development on multi-threaded Speech Recognition server, using C++, Boost, Asio, and Kaldi. The scalable server can process hundreds of concurrent audio streams at state of the art accuracy.
- Manage product releases, interface with clients, and write public-facing documentation.
- Built parallelized speech recognizer that transcribes audio up to 400x faster than real time.
- Implemented new speaker diarization (segmentation) system, using clustering and HMM-GMM models. The new system improved diarization error rate from 15% to 4%.
- Trained HMM-DNN models to transcribe English and Spanish using Kaldi.
- Designed finite state transducer that recognizes all possible pronunciations of American phone numbers.

UC Berkeley RISELab / AMPLab

Research Engineer *Spring 2017 – Winter 2017*

- Research on distributed computing on serverless cloud infrastructure (AWS Lambda).
- Migrated and benchmarked Pywren project on different cloud infrastructures: Google Cloud & Azure.

Facebook

Engineering Intern, Internet.Org Scalability Team *Fall 2015*

- Worked on i.org traffic monitor, a debugging tool for Android engineers on iorg team.
- Developed functionality for engineers to test Facebook products (FB-Lite, Free Basics, Native) on various carrier infrastructures from different countries.

TECHNICAL SKILLS

Languages: Python, C, C++ (Boost), Go, Bash, Java

Technologies: Linux, Kaldi, LaTeX, Docker, AWS, Google Cloud, OpenFST

PROJECTS

Paxos-based Distributed Datastore (Golang)

- General purpose Paxos library, and sequentially consistent partition-tolerant distributed datastore

Voice-Controlled Autonomous Guitar (Raspberry Pi, Python) *(www.allan.expert/guitar)*

- Embedded system built with an array of solenoids held over the frets and strings of an acoustic guitar
- Built speech recognition interface to start, stop and choose songs, using Google Cloud APIs.

Linux Firewall (Python)

- Stateful packet filter for Linux. Filters based on IP subnet, DNS, or geo-location
- Implemented DNS redirects, TCP resets, and logging HTTP transactions

EDUCATION

University of California, Berkeley

B.A. Double Major: Applied Mathematics & Computer Science, Regents & Chancellor's Scholarship

London School of Economics

International Relations & Law