Yifu Wang

PHONE: (626) 226 6282 EMAIL: yifuw@andrew.cmu.edu
GITHUB: github.com/wangyifu LINKEDIN: www.linkedin.com/in/yifuwang

EDUCATION

M.S. Computer Science, Carnegie Mellon University, Pittsburgh **May / 16**

Courses: Advanced Data Structure and Algorithms, Machine Learning, Cloud Computing, Big Data System in Practice, Search Engine

Teaching: Data Structures for Application Programmers, JAVA for Application Programmers

B.S. Software Engineering, Dalian University of Technology, China **June / 14**

Courses: Operating System, Compiler Techniques, Database Systems, Computer Network, Computer Architecture, Data Structure and Algorithms

WORK EXPERIENCE

Software Engineer Intern at *Quixey, Inc.* May / 15 - Aug / 15

- Implemented Chinese query tokenizer based on Early parser.
- Designed a **dynamic programming** algorithm to reduce processing time to linear.
- **40% faster** and **1.6% gain of DCG score** compared with previous tokenizer.
- Created **RESTful** Web Services interface using Jetty.
- Released to **production stage** in Quixey app search and auto suggestion.
- Applied **U.S. provisional patent**(No. 62/216329) for core algorithm.

SDET Intern at *VMware, Inc.* Aug / 13 - Feb / 14

- Contributed to the automation testing framework using **C#**.
- Developed several web applications based on **LAMP** stack.
- Visualized product development progress, employee KPI, and bug report.
- Generated report in pdf and excel format using historical data.

PROJECTS

DOTA2 COUNTER-PICK:
- Hero recommendation engine for Dota 2 game. Written in **python**.
- Generated feature matrix from 30000 matches using **urllib, numpy, pandas**.
- Applied **K-nearest neighbors**, **SVM with RBF kernel** to predict result using **sci-kit**.
- Achieved **70% accuracy** in predicting match outcomes.

BIO-ASK:
- Question answering system in medical field based on **UIMA** architecture.
- Implemented document retrieval algorithm based on **Stanford NLP** and **Lingpipe**.
- Developed result evaluation system using **BM25** and **GMAP**.

SEAT FINDER:
- **Android** app for students to find nearest self study room.
- Designed **ranking algorithm** based on elapsed time, user rank, and votes.
- Developed HTTP server using **JAVA**, applied **thread pool** handle simultaneous requests.

DRAW SOMETHING:
- Online graphic multiplayer game developed on **Linux** using **C++** and **Qt**.
- Used **bezier curve** to represent and serialize canvas.
- Implemented **multi-thread TCP server** using **Qt** to support concurrent connections.

SKILLS

Languages: JAVA, Python, C++, shell script, Javascript, PHP, Scala, R

Other Technologies: Web Development(LAMP and MEAN), MySQL, AWS, Hadoop, Elastic MapReduce, ElasticSearch, MongoDB, Mahout, Lucene, Git, Scrum