

# Shuo Yang

3993 N Campbell Ave, Apt 1212, Tucson, AZ, 85719

☎ 952-297-6289 | ✉ imsure95@gmail.com | 🏠 imsure.github.io | 📱 imsure | 📧 shuo-yang-12ab7047

## Education

### University of Arizona

M.S. IN COMPUTER SCIENCE, GPA: 3.5

Tucson, Arizona USA

Aug. 2014 - May. 2017

### University of St. Thomas

M.S. IN SOFTWARE ENGINEERING, GPA: 3.8

St. Paul, Minnesota USA

Aug. 2011 - May. 2014

### Harbin Institute of Technology

B.E. IN ELECTRICAL ENGINEERING

Harbin, Heilongjiang China

Aug. 2004 - May. 2008

## Academic Experience

### University of Arizona

Tucson, Arizona USA

RESEARCH ASSISTANT WITH DR. RICHARD SNODGRASS ON ANTARES PROJECT

Aug. 2016 - Mar. 2017

- Chief programmer of ANTARES (The Arizona-NOAO Temporal Analysis and Response to Events System) project.
- Deployed ANTARES system to the dedicated CentOS cluster using MySQL Cluster as Database backend.
- Analyzed and improved the performance of ANTARES data processing pipeline.
- Designed and proposed the data provenance framework for ANTARES to answer provenance questions that astronomers would ask.
- Designed the packet format specification for alert data and implemented the simulator for generating alert streams.

### University of Arizona

Tucson, Arizona USA

RESEARCH ASSISTANT WITH DR. RICHARD SNODGRASS ON ANTARES PROJECT

Aug. 2014 - Aug. 2015

- Designed and implemented the database used by ANTARES, as well as the web front-end for querying data produced by ANTARES pipeline using Django and MySQL.
- Collaborated with astronomers to design, implement and document the Python API for astronomical alert data manipulation.
- Designed and built the bootstrapping system to run ANTARES pipeline in a pseudo-distributed mode using Vagrant and Puppet.

### University of Arizona

Tucson, Arizona USA

INDEPENDENT STUDY WITH DR. BEICHUAN ZHANG ON NAMED DATA NETWORKING (NDN) PROJECT

Spring 2016 & Fall 2016

- Researched consumer-driven congestion control mechanisms in NDN.
- Implemented TCP-like congestion control algorithms in NDN consumer (in C++), specifically, TCP RENO, CUBIC and VEGAS. Evaluated and studied the performance of each of them in the context of NDN.

### University of Arizona

Tucson, Arizona USA

TEACHING ASSISTANT WITH DR. LESTER McCANN ON CS460: DATABASE SYSTEMS

Fall 2015 & Spring 2016

- Held office hours; graded programming and written assignments; prepared solutions for written assignments.
- Designed the final project "Database-driven Web Application" using Oracle, Tomcat and JSP; evaluated students' design and implementation.

### University of St. Thomas

St. Paul, Minnesota USA

STUDENT RESEARCHER WITH DR. BRAD RUBIN AND DR. JADIN JACKSON

Aug. 2013 - May. 2014

- We initiated the project "Large-scale Neural Network Modeling with Hadoop" to explore the use of big data technologies in computational neuroscience and I was responsible for investigating graph processing in Hadoop.
- Implemented the neural network with MapReduce model in Hadoop.
- Implemented the neural network with vertex-centric model using Apache Giraph in Hadoop.

## Work Experience

### Danfoss Power Solutions

Plymouth, Minnesota USA

SOFTWARE ENGINEER INTERN

June. 2012 - May. 2013

- Migrated the legacy software to the new hardware platform with other team members and conducted unit testing.
- Applied static code analyzer FlexeLint to the legacy software written in C.
- Built a tool (in Python) for summarizing, indexing and querying large volume of warning messages produced by FlexeLint.
- Improved overall software quality, identified and corrected several vulnerabilities existed in the legacy code.

### Beijing Farsight Technology and Information

Beijing, China

EMBEDDED SOFTWARE ENGINEER & TRAINING ASSISTANT

July. 2009 - July. 2011

- Developed Linux device drivers and applications (in C) for various ARM platforms (ARM 9, 11 & Cortex-A8).
- Assisted training instructors in preparing training content and developed technical how-to documents for ARM and Embedded Linux training.
- Instructed trainees in their lab sessions and final design project.

- Being part of a team that built the initial prototype for Samsung's Remote Test Lab (RTL), I was responsible for the development on RTL device.
- Implemented the communication protocol between RTL device ( LiMo platform) and RTL proxy.
- Developed a daemon process (in C++) for RTL device to handles requests from proxy, in order for RTL client to remotely control the device.

## Publication

---

Yang S, Spielman ND, Jackson JC, Rubin BS. Large-scale neural modeling in MapReduce and Giraph. In IEEE International Conference on Electro Information Technology. IEEE Computer Society. 2014. p. 556-561. 6871824. Available from, DOI: 10.1109/EIT.2014.6871824

## Selected Coursework

---

### UNIVERSITY OF ARIZONA

<b>Database Systems Implementation</b>	Implemented Heap File, Buffer Manager and B <sup>+</sup> Tree components of Minibase in C++.
<b>Principles of Computer Networking</b>	Implemented a virtual router that runs PWOSPF routing protocol and routes real IP packets in C.
<b>Principles of Compilation</b>	Implemented C- - (a subset of C) compiler (both code generation and optimizations) in C.
<b>Computer Security</b>	Implemented a secure chat program using a hybrid protocol (DES + RSA) in Java.
<b>Parallel and Distributed Programming</b>	Implemented a MPI critical-path profiler and Redundant MPI protocols in C.
<b>Introduction to Machine Learning</b>	Implemented a sparse autoencoder in Python, trained with MNIST handwritten digit dataset.

### UNIVERSITY OF ST. THOMAS

<b>Big Data Architecture</b>	Developed a new design pattern for improving the performance of graph processing in MapReduce.
<b>Data Visualization</b>	Explored what are the key factors that make a championship team with NBA dataset using Python, SQLite and R.
<b>Information Retrieval</b>	Built a search engine (web crawler + indexer + query processor) for Shakespeare's whole collection.
<b>Software Engineering</b>	Built a genetic programming system for automatically solving linear regression problem in C.

## Skills

---

<b>Programming Languages</b>	C/C++, Python, Java, R, JavaScript
<b>Web</b>	Django, HTML, CSS
<b>Database Systems</b>	MySQL, SQLite
<b>Parallel and Distributed Programming</b>	Pthreads, MPI, MapReduce
<b>Environment/Frameworks/Tools</b>	Linux, Hadoop, Giraph, $\LaTeX$ , Git, Vagrant