# 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** — *Tucson, Arizona USA*

M.S. in Computer Science, GPA: 3.18 — *Aug. 2014 - May. 2017*

**University of St. Thomas** — *St. Paul, Minnesota USA*

M.S. in Software Engineering, GPA: 3.8 — *Aug. 2011 - May. 2014*

**Harbin Institute of Technology** — *Harbin, Heilongjiang China*

B.E. in Electrical Engineering — *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*

- Main developer 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 in Python.

**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 autoconfiguration and bootstrap system to run ANTARES in a pseudo-distributed mode using Puppet and Vagrant.

**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 analyer 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.

**Platomix Technologies. (Start-up company)** *Beijing, China*

Software Engineer *Feb. 2009 - July. 2009*

- 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 (in C++) 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

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

University of St. Thomas

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

## Skills

| | |
|---|---|
| **Programming Languages** | C/C++, Python, Java, R, JavaScript, SQL |
| **Web** | Django, HTML, CSS |
| **Database Systems** | MySQL, SQLite |
| **Parallel and Distributed Programming** | Pthreads, MPI, MapReduce |
| **Environment/Frameworks/Tools** | Linux, Hadoop, Giraph, LaTeX, Git, Vagrant |