

HAO SHI

Phone: (086) 186-215-27893
E-Mail Address: shisoftgenius@hotmail.com

401 Room, 23 Block, 1030 Alley, Zhenhua RD.
Shanghai, China

EDUCATION

BS Donghua University, Software Engineering 2010 - 2014

HONORS AND AWARDS

**13th “Challenge Cup” National Undergraduate
Curricular Academic Science and Technology Work Competition
Individual Group, National Region
- Third Prize** 2013

Project name: Web Service Integrated Personalized Recommendation System
Together with third prize in Shanghai Region

**12th “Challenge Cup” National Undergraduate
Curricular Academic Science and Technology Work Competition
Team Group, National Region
- First Prize (As Project Leader)** 2011

Project name: Web Service Integration Service Based On Cloud Computing Technology
Together with second prize in Shanghai Region

Campus Invention and Innovate - Excellence Prize 2011
Issuer: China Patent Conference Organizing Committee

Science and Technology Startup Cup 2011
Issuer: Shanghai Invention Association

Geek Award for Celebrate Ubuntu Hackathon 2016
Issuer: Canonical

Top 10 Team (No.2) for TechCrunch Hackathon Shanghai 2016
Issuer: TechCrunch, TechNode

RESEARCH EXPERIENCE

Graduate Dissertation, Donghua University 2013 - 2014
Topic: Community Detection and Automatic Tagging for Social Networks
Advisor: Li Son

- Predecessor research won the third prize of 13th national “Challenge Cup” competition in 2013 (same as the first entry in “Honors and Awards” section)
- Achieved “Excellent Graduate Dissertation” award in school

- Gathering training data from various social network, web services and Wikipedia.
- Building data center for project, 6 servers connected by 1000M Ethernet. Setting up Hadoop cluster, databases, caches and monitoring software.
- Primary data processing. Data cleaning, structuralize social network contents, convert and store data for model training processes.
- Model training. Including parallelized keyword extraction from social networks, parallelized keyword features calculation for term frequency and Latent Dirichlet allocation, real-time content feature extraction and index from live content sources for real-time update. User community detection based on author published contents categories, tagging by category names. The time of training process was long in the beginning (1 day per iteration). It was reduced dramatically (15 minutes per iteration) by parallelizing and various optimization, including cache mechanism and improved data structure.
- User interfaces. User can select one or more categories by their names and see what contents and authors are related with the categories recently published on various social networks instantly. User can also cluster authors by categories.

Donghua University

2012

Advisor: Qiao Pan

- Inspired the graduate research team on research and implement roadmap
- Raw data processing. Parallelized search query extraction from universal web access log data of telecom company.
- Chinese words segment
- Calculate top n searched words grouped by telecom account.

Patents, Development

2010 - 2015

Inventor, Architect, Developer

- Applied 3 patents, all granted by the authorities in China and Hong Kong “ZL201110186031.7”, “ZL201010271289.2”, “HK1150286” (see the “Patent” section for details)
- Winner of the first prize of 12th national “Challenge Cup” competition in 2011 (same as the second entry in “Honors and Awards” section)
- Business plan based on this project wined third prize on campus startup business competition of local university town
- Heterogeneous services integration approach. Analyzing design pattern and user behavior for each social networks and web services on communication functionality, abstract and implement those functions into one integrated messaging and feed service.
- Distributed architecture. Each parts of the system including web server, gateway server, data analytics server are all running in their own server process and able to expand by running new process to work as a cluster.
- Web services gateway. Get access to user service account, providers including twitter, facebook, facebook IM, MSN, XMPP, weibo, renren, tumblr, wordpress, flickr, email servers and may others (20+ total).

- User interfaces. Including one web page based client (javascript) and one native iOS application (Objective-C).

Patents, Development

2008 - 2009

Inventor, Developer

- Applied 4 patent and 2 of them granted by the authority in China “ZL201010005212.0”, “ZL201010005216.9” (see the “Patent” section for details)
- High school extracurricular research
- Achieved awards issued by Chinese government, Shanghai government, Intel corporation and other academic organizations
- Achieved award that exempted the “National Higher Education Entrance Examination” for colleges
- Design and implementation of chart generator on web pages. Including a chart image generator at server site based on parameters received from web clients and a chart data input user interface based on web page (javascript).
- Design and implementation of sketch pad on web pages. Including a sketch image generator at server site based on parameter received from web clients and a visualized drawing user interface based on web page (javascript).
- Other works that did not approved by the patent office. Including LaTeX editor, chemical equation editor, function graph generator based on web pages.

WORK EXPERIENCE

Factual Inc.

2017/3 - Present

Software Engineer

- Rebuilding on solr server cluster manager and transparent proxy. My application reduced response time to half of the original and solved server unresponding issue under heavy load by implementing asynchronous non-blocking I/O server.
- Data processing workflow tasks development.

CloudZen Shanghai

2016/5 - 2016/8

Freelancer Software Engineer

- Go programming. Micro-service server architecture development for next generation cloud gaming and live video streaming platform
- Mass users chat server development

Shanghai Geyi Network Co. Ltd (上海格奕网络科技有限公司)

2014 - 2015

Software Architect, Developer

- Development of mission critical modules on game server. All of the server site work are written in a function programming called Clojure, it is an alternative implementation of lisp. Those modules include payment systems, messaging system and some game logical modules.
- Design and implementation of distributed server architecture middleware. Including sharding algorithm based on distributed hashing table, remote function invocation and server status synchronize based on zookeeper and message queue. This middleware

- solved problems on handling different level of server workload without interrupting services and reboot, server capability able to expand freely by starting new server processes.
- Design and implementation of server management portal. Web based control panel to remote start or stop servers, manage servers by groups and automatic server clusters deployment. Graphical server status chart report on resource usage. This portal made it easy to monitor and resolve issues on server programs.
 - Design and implementation of data analytics portal. Collecting business data including online users, user payment behaviors, user abnormal behaviors and a lot more from servers, aggregate by time span to produce reports. The analytics system is able handle at least 30,000 user activity data on one machine and provide reports in real-time. This is the entry of operation team get to know what's happening right now on the product and facilitate their decision making.
 - Performance analytics and implemented optimization solutions on server programs
 - Server deployment, install and maintain infrastructure software on Linux servers
 - Assist on client development, including payment system on iOS and network optimization on compression

Shanghai Open University,
National Laboratory of Information Security and Network Management 2013 - 2014
Intern, Technology Advisor, Developer

- Social network analytics. Find top influencer in one community according to their relationship in graph.
- Assist on building data center. Making decision on server hardware specifications. Supervising on equipment purchase and assemble.
- Negotiate with data vendor on social networks for data mining purpose
- General data visualize. Including using Gephi to depict social network relations in community, produce reports and display on screen array for demonstrate.
- Data mining workflow and data warehouse architecture planning
- Provide technical support and explanations

PATENTS

“Network service integration system and integration method thereof based on cloud computing architecture”

（基于云计算架构的网络服务聚合系统及其聚合方法）

Chinese Invention Patent, ZL201110186031.7

Applied at 5/7/2011

Granted at 13/8/2014

“Drawable web editor and web information editing system”

（可绘图的网络编辑器及网络信息输入编辑系统）

Chinese Invention Patent, ZL201010005212.0

First Applied as Innovate Patent at: 21/1/2009

Transferred to Invention Patent at: 15/1/2010

Granted at 19/2/2014

“Network information transmission/reception system and method”

(網絡信息收發系統及方法)

Hong Kong Patent, HK1150286

Applied at 29/4/2011

Granted at 30/8/2013

“Web editor capable of inputting diagrams and web information edition system”

(可輸入圖表的網絡編輯器及網絡信息輸入編輯系統)

Chinese Invention Patent, ZL201010005216.9

First Applied as Innovate Patent at: 21/1/2009

Transferred to Invention Patent at: 15/1/2010

Granted at 3/7/2013

“Network information transmission/reception system and method”

(網絡信息收發系統及方法)

Chinese Invention Patent, ZL201010271289.2

Applied at 31/8/2010

Granted at 8/5/2013

SELF-ASSIGNED PROJECTS

bifrost

2016/10 - Present

Distributed system framework for Rust applications. Built-in consensus framework with reliable data store, client membership and consistent hashing sharding based on raft algorithm. It also shipped with an ease of use RPC framework.

Morpheus

2016/3 - Present

General purpose distributed graph engine. With high-performance in-memory data store, graph query and parallel graph traversal capabilities (BFS). Inspired by Trinity graph engine from Microsoft.

Nebuchadnezzar

2016/2 - Present

RAMCloud implementation. Distributed full in-memory key-value store, with schema, disk write-back and rich data type support.

Cluster Connector

2015 - Present

Distributed system middleware for Clojure applications. Cluster sharding, message queue, cache, RPC and locks.

Proxy Entanglement

2015

Automatic proxy server (socks and http) discovery. Through search engines and port scanner. Designed to facilitate web clawing.

WebFusion 2010 - 2015
On-cloud web service integration

Web Page Based Rich Text editor 2008 – 2009
First patented work

COMPUTER SKILLS

Programming: Rust, Java, Clojure, Go, JavaScript, Objective-C, C#, Visual Basic

Applications:

Mac OS, Linux (Ubuntu, Arch Linux), Windows OS

PostgreSQL, MongoDB, Cassandra, Redis, ElasticSearch, Solr, Zookeeper, Kafka, Nginx

Platforms: Hadoop, Spark, Docker

LANGUAGES

Chinese: Native Language

English: Advanced Reading, Writing, Speaking and Listening

OTHER

Chinese Citizen

EXTERNAL REFERENCES

Code Repositories (See develop branch for last updates)

GitLab: <http://hq.shisoft.net:7770/>

GitHub: <https://github.com/shisoft>

Blogs

<http://shisoft.org/>

Work Demo

<http://demo.shisoft.net/>

Scanned Copy of Certifications and Patents

<http://goo.gl/loDhfr>

Translations

<http://goo.gl/mdNGmd>