Guozhen She

GitHub: github.com/hazelnutsgz Email:gzshe15@fudan.edu.cn HomePage: https://sgzhazelnut.github.io/terminal

EDUCATION

Fudan University, Shanghai, China 2015.9 - Present
Bachelor of Computer Science (expected in 07.2020, **one year delay** because of a surgery)
GPA (overall): **3.55/4.0**; Ranking: **21/117(17.9%)**
Body:Computer Architecture(A)| Computer Network(A-)|Computer System(A)|Database Implementation(A-)
Brain: Distributed System(A)|Linear Algebra(A)|Mathematical Analysis(A)
Mouth: C Programming(A)|C++ Programming(A-)|Web Development(A)|Digital Logic Design Experiments(A)
Metaphysics: Neural Network and Deep Learning(A-)

SKILLS

Programming Language: Python, Java, Golang, C, C++, JavaScript, MATLAB(wanderer), Rust(dabbler), Haskell
Framework & Library: Tornado(Python), D3.js(JS), Node.js(V8), *nix System Call(C), Tensorflow, PyTorch,STL
Tools: LaTeX, Git, Docker**,** GNU toolchain
Soft Skills: Storytelling, Networking, Puns & Jokes, Information Retrieval, Open Source Engagement,

SELECTED PROJECTS

Multiple Pattern Text Matching Tool (https://github.com/hazelnutsgz/NaiveACAutomation)
Implement the **Aho–Corasick** automation(Tries with failed pointers), facilitating the text matching in multiple pattern.
Interactive Visualization of Coauthor Affiliation(https://github.com/hazelnutsgz/NaiveScholarMap**)**
Construct the co-author affiliation graph based on yearly data crawling from google scholar, then build an interactive web service by D3.js library, for comparison and analysis of co-author affiliations in different years.
Monitoring Service of WeChat Group(https://github.com/DaShiLar/Naive-WeChat-Monitor)
Build a **monitoring** backend service for all WeChat groups of the user, which captures the real-time chatting information(video, text, voice), and stores them to database and filesystem in backend. Also a front-end web UI is provided to users for authorization and inspection. The system is implemented in a **multi-processes** architecture for **the isolation** of different users and utilization of multi-cores on backend server.

INDUSTRY EXPERIENCES

Microsoft Research Asia, System and Network Group | Research Intern Jan.2019-

- Real-time Bot Detection System for Azure Cloud Service
  - Implemented a Golang web service to **build the preprocessing pipeline** of daily network log data(8,000,000) from Bing, which parse the raw log concurrently leveraging **goroutines** into heterogeneous files hosted on a distributed file system.
  - Based on that, provided an analytic service(React, D3.js) for visualizing and understanding data. For conservation of memory footprint on VM, the data was **fetched on demand** and cached as the data structures in memory, guaranteeing the data at **hotspot** would stay longer in memory to accelerate the analytic efficiency.
  - Implemented the algorithm to generate the behavior-based images for each request session, then develop a **CNN-based** model by PyTorch&Tensorflow to detect bot behavior by classifying the images generated, which reached **94.3%** accuracy on labeled Bing log data. Then optimized **IO performance** for the production environment.

Intel Asia-Pacific R&D, Open Source Technology Center | SDE Intern Aug.2018-Nov.2018

- **Contributed code** to StarlingX(OpenStack Foundation), assisted in deploying the StarlingX on bare-metal devices.
- Built a **rule-based** command line tool which migrate code from python2 to python3.
- Assisted the colleagues to setup the **compiling farm** based on K8S for building of StarlingX project.
- Implemented a static graph-based algorithm for **package dependency analysis** in the project.

Wish(ContextLogic) | SDE Intern Jan.2018-Apr.2018

- Developed an meta information adjustment service, using Tornado framework at backend, and backbone.js at front end.
- Built an adaptive notification service for accounts out of credits. Based on that, designed an algorithm to **spot zombie users**.
- Built **a channel search service** equipped with multiple filters to assist users to select the proper channel on their own conditions.

ACADEMIC EXPERIENCES

Fudan University, Mobile Systems And Networking Group, Advisor: Prof. **Yang Chen** April.2017-

- **LinkedIn** website:
  - Built a **cookie-based** crawling system to scraping profiles retaining, which **imitate** normal user behavior for anti-crawling.
  - Implemented automatic script to **expand the personal LinkedIn connection**s based on LinkedIn Recommendation.
  - Crawled the profiles of connections **concurrently** utilizing multiple mock accounts at the crawling system, then wrote an **error-tolerant** parser to generate JSON-like profile for each profile URL.
- **Google Scholar** website
  - Crawled the scholar's profile, integrated with **anti-captcha** service. Detected the untruthful profile by machine learning.
  - Crawled data from MS Academic, DBLP, Google Scholar, build the **heterogeneous graph** containing both authors and papers for given conferences in different year. Utilized **Graph Convolution Network(GCN)** to evaluate the conference.
- Qingyun Go
  - Built a **geo-based** social App. The communication between backend and frontend is hosted on HTTPS protocol based on RESTful API. The backend is FastCGI integrated with C++ code, while the frontend is a javascript runtime with asynchronous API. Then developed some streaming analytic tools to monitor the status of service and the behavior of users.

MISCELLANY

Interest: **Archeology** (on Computer Science), Soccer(DM), **Road Cycling**(Individual Time Trial, Mountain Climbing)
GRE:331