JIAYUN ZHANG

Email: jiayunzhang15@outlook.com ◇ Homepage: https://jiayunz.github.io/
Room 806, Building 1, No. 1433 Cailun Road, Shanghai, 201203, P.R.China

EDUCATION

Fudan University Shanghai, China
B.S. in Computer Science Graduation: Jul 2020 (Expected)

- Overall GPA: 3.52/4.00 GPA for the last two years: 3.67/4.00 (Rank 17/155)
- Core Courses: Programming (A)/ Design and Analysis of Algorithms (A)/ Probability Theory and Mathematical Statistics (A)/ Database (A-)/ Pattern Recognition (A-)/ Digital Image Processing (A-)/ Virtual Reality (A)/ Game development (A) etc.

Aalto University Espoo, Finland
Summer Intern supervised by Prof. Yu Xiao Jun 2019 - Sep 2019

- Research Assistant at Mobile Cloud Computing (mc^2) group

PUBLICATIONS

Detecting Malicious Accounts in Online Developer Communities Using Deep Learning. [paper]
*Qingyuan Gong, **Jiayun Zhang**, Yang Chen, Qi Li, Yu Xiao, Xin Wang, Pan Hui.*
Proc. of the 28th ACM International Conference on Information and Knowledge Management (**CIKM'19**).

Identifying Structural Hole Spanners in Online Social Networks Using Machine Learning. [paper]
*Qingyuan Gong, **Jiayun Zhang**, Xin Wang, Yang Chen.*
Proc. of the ACM SIGCOMM 2019 Conference Posters and Demos (**SIGCOMM Posters and Demos'19**).

DeepLoc: A Location Preference Prediction System for Online Lodging Platforms. [paper]
*Yihan Ma, Hua Sun, Yang Chen, **Jiayun Zhang**, Yang Xu, Xin Wang, Pan Hui.*
Proc. of the CCF Conference on Computer Supported Cooperative Work and Social Computing (**ChineseCSCW'19**).

Understanding the Working Time of IT Companies in China and the United States.
***Jiayun Zhang**, Yang Chen, Qingyuan Gong, Aaron Yi Ding, Yu Xiao, Xin Wang, Pan Hui.*
Under major revision at **IEEE Software Magazine**.

Understanding Work Rhythms in Software Development and Their Effects on Technical Performance.
***Jiayun Zhang**, Qingyuan Gong, Yang Chen, Yu Xiao, Xin Wang, Aaron Yi Ding.*
Submitted to International Conference on Mining Software Repositories (**MSR'20**).

A video dataset of a wooden box assembly process.
***Jiayun Zhang**, Petr Byvshev, Yu Xiao.*
Submitted to **Scientific Data Journal**.

RESEARCH EXPERIENCE

Identifying Structural Hole Spanners in Online Social Networks Mar 2019 – Present
Research Assistant supervised by Prof. Yang Chen, Fudan University

- Proposed a machine learning-based model for identifying structural hole spanners; leveraged the ego networks and the cross-site linking function to enhance the identification.
- Implemented the classifier by CatBoost. Achieved a test F1-Score of 0.857 and an AUC value of 0.856 on the Foursquare dataset.
- Contributed to a paper published in **SIGCOMM Posters and Demos'19**.

User Behavior Analysis in Online Developer Communities May 2018 – Present
Research Assistant supervised by Prof. Yang Chen, Fudan University

◦ **A Representative User-centric Dataset of GitHub Developers** [code] May 2018 – Sep 2018

- Crawled user data and dependencies on GitHub in an unbiased manner.
- Built a representative user-centric dataset including the information of over 10 million GitHub Developers.

◦ **Malicious User Identification on Version Control Systems** Jun 2018 – Jun 2019

- Did a comparative study between the behaviors of legitimate users and malicious users on GitHub.

- Proposed GitSec, a deep learning-based system with Phased LSTM and attention mechanism to detect malicious accounts on VCS. Achieved a test F1-Score of 0.920 and an AUC value of 0.938 on the GitHub dataset.
- Contributed to a paper published in **CIKM'19**.

○ **Discovering Work Patterns of Developers** Jan 2019 – Sep 2019

- Designed a data-driven approach with clustering algorithms to identify developers' work patterns with commit behaviors. Four developer-centric work rhythms and three organization-centric work rhythms were detected.
- Analyzed the relationship between work rhythms and demographics, collaboration role and productivity.
- Conducted a user survey to understand the situation of working overtime from developers' perspectives.
- Contributed to first-authored papers submitted to **MSR'20** and **IEEE Software Magazine**.

A Video and Sensor Dataset of a Wooden Box Assembly Process Jun 2019 – Sep 2019
Research Assistant supervised by Prof. Yu Xiao, Aalto University

- Recruited 17 subjects and acquired video and sensor data of a 9-step wooden box assembly process with multiple cameras and a sensor glove; performed data labeling, processing and analysis.
- Contributed to a first-authored paper submitted to **Scientific Data Journal**.

Data Mining on Health-Seeking Behavior May 2017 – Apr 2018
Research Assistant supervised by Prof. Yun Xiong, Fudan University

- Devised a model with SVM for pneumonia detection based on medication records. A test accuracy of 0.915 was obtained on a real-world dataset collected from hospitals in Shanghai.
- Devised a prediction model with Time-Aware LSTM to predict one's stage of diabetes based on previous diagnoses. Achieved a test F1-Score of 0.787 on the real-world dataset.
- Developed a web-based interactive system for diabetes prediction; the system could receive historical diagnoses from users, predict the stages of diabetes using the trained model and output the results on the webpage.

INDUSTRIAL EXPERIENCE

VMware Information Technology (China) Co., Ltd. Shanghai, China
MTS (Member of Technical Staff) Intern Apr 2018 – Oct 2018

- Developed a log analysis system for automatically detecting the causes of program failures. 67 types of error causes was detected with an accuracy of 0.936 on real-time data from an internal bug reporting platform.
- Developed web APIs for an internal cloud resource platform to support the use of virtual machine templates.
- Participated in the implementation of Template Validation Service, a system for security verification of virtual machine templates uploaded to database.

SELECTED PROJECTS

Raindrop Removal From a Single Image, advised by Prof. Junping Zhang [code] Summer 2019

- Devised a deep-learning-based model for raindrop removal. The model could identify the location and intensity of raindrops with ResNet and eliminate the raindrops with Dilated CNN and ConvLSTM.
- Incorporated Gaussian filtering in the model to remove the background interference; improved the network capability by focusing on high frequency detail of the images.
- Achieved raindrop removal results on real-world images with PSNR as 27.70 and SSIM as 0.8801.

3D Parkour Game [code] Winter 2017

- Developed a full-featured parkour game; built 3D game scenes in Unity, designed animation effects and user interactions; implemented the game logic with Unity Game scripts written in C#.

SELECTED AWARDS

2019 The First Prize of Shanghai Open Data Innovation Research Competition (Top 1 among 65 teams)
2019 Best Student Award, Mobile Systems and Networking Group at Fudan University (1 out of 32)
2019 Second Class Scholarship for Outstanding Students in Fudan University (Top 10%)
2019 Chun-Tsung Program (Research Endowment funded by Nobel Laureate Dr. Tsung-Dao Lee)
2018 Xiyuan Scholar (Undergraduate Research Program at Fudan University)

SKILLS

Programming: Python, C/C++, Ruby, C#, HTML/CSS, JavaScript, SQL.
Packages and Tools: Pytorch, Tensorflow, Scikit-learn, Matlab, Django, Bootstrap, Unity, Blender etc.
Standard Language Tests: TOEFL 104 (Reading 28, Listening 24, Speaking 24, Writing 28)