

# Cai, Yifan

Email: [fyc1007261@live.com](mailto:fyc1007261@live.com) | Homepage: <http://www.yifancai.tech>

Address: 3-2-702, 9 Qianchao Road, Oriental Royal, Jianggan District, Hangzhou, Zhejiang, China

## EDUCATION

### Shanghai Jiao Tong University

Shanghai, China

*B.S. Software Engineering, System Software*

Expected in June 2020

- Overall GPA: 3.77 / 4.30 (88.19/100); Major GPA: 3.88 / 4.30 (89.95/100)
- Honors/Awards: SJTU Scholarships (2016-2017 & 2017-2018), Huawei Scholarships (2017-2018)
- Relevant Courses: Operating Systems (4.0), Distributed Systems (4.0), Computer Architecture (4.0)

## PUBLICATIONS

### Rethinking Data Management Systems for Disaggregated Data Centers.

Qizhen Zhang, **Yifan Cai**, Sebastian Angel, Ang Chen, Vincent Liu and Boon Thau Loo.

Submitted for *Conference on Innovative Data Systems Research (CIDR)*, 2020

### Consensus-based Data Statistics in Distributed Network Systems. [\[Paper\]](#)

**Yifan Cai**, Jianping He, Wenbin Yu, and Xinping Guan.

In the 57<sup>th</sup> *IEEE Conference on Decision and Control, Miami, USA*, December 2018

## RESEARCH EXPERIENCE

### University of Pennsylvania

PA, US

*Research Intern, Supervisor: Dr. Vincent Liu*

Jul 2019 – Dec 2019 (Expected)

- Conducted research on the impacts of disaggregated data centers on the design of relational databases, analyzed the potential performance drawbacks of the existed disaggregated operating system in the context of data processing, and outlined research challenges of remedying them
- Modified the kernel of LegoOS (A disaggregated operating system published in OSDI'18) to port it on a different type of hardware and deployed it on cloud machines; wrote code of nested loop join, hash join as well as grace hash join.
- Tested the performance of nested loop join and hash join and analyzed the relationship between local memory size, the number of remote memory access and the performance degradation of each operation. Designed possible solutions that might reduce the number of remote memory access and improve performance

### Lab of System Control and Information Processing

Shanghai, China

*Research Assistant, Supervisor: Dr. Jianping He*

Sep 2017 – Jun 2019

- Designed a protocol in distributed network systems to enable multiple compute nodes do statistics cooperatively, which was the first protocol proposed that could compute the probability density functions in a fully distributed way
- Designed and optimized two probability distribution function (PDF) Calculating algorithms which run under public network conditions and anonymous conditions respectively; responsible for numerical simulations and paper writing

### University of Michigan

MI, US

*Research Intern, Supervisor: Dr. Xianglei Huang*

Aug 2018 – Sep 2018

- Built an LSTM-based solar forecasting system on Keras deep learning library
- Designed the predicting model with an error of only approximately 10%, which was less than all of the predicting models in published papers based on the same dataset

## NOTABLE PROJECTS

---

### **Distributed File System**

Shanghai, China

*Computer System Engineering Project, SJTU*

Sep 2018 – Jan 2019

- Developed a distributed file system with Fuse. The file system supported basic Linux interfaces (touch, move, write, unlink, etc.), and worked under single server and multiple clients
- Implemented distributed locks with local cache in the system to ensure the atomicity and the lock performance
- Designed and implemented replication protocols using heartbeats and notifications in case of random failures to improve availability

### **Tiger Language Compiler**

Shanghai, China

*Project of the Compilers Course, SJTU*

Sep 2018 – Jan 2019

- Developed a compiler written in C for *Tiger* language followed x86-64 calling conventions, which could generate assembly code
- Realized optimizations such as liveness analysis and register allocations

### **Smart Electric Appliance Monitoring and Scheduling System**

Shanghai, China

*Project of the Project Management and Software Development Course, SJTU*

Jun 2018 – Jul 2018

- Collaborated with team members to build a smart electric appliance monitoring and scheduling system
- Focused on the back-end Java development of the web application, as well as the integrated hardware controlling system written in Python, which collected information from all the sensors and operates the appliances
- Managed to integrate a series of cross-process communications, including smart scheduling for appliances such as renewable energy simulating, deep learning based solar energy forecasting, and convenient human computer interactions
- Exhibited in the SJTU Software Exhibition, 2018

### **B+ Tree Based Database**

Shanghai, China

*Project of the Data Structure Course, SJTU*

Jul 2017 – Aug 2017

- Developed a B-plus-tree-based database system which supported CRUD operations at a speed of a million operations per second when running with over a million sets of data
- Designed and implemented a buffer system between disk and memory to improve the performance of reading and writing data

## SKILLS

---

**Languages:** Mandarin (Native); English (TOFEL 111).

**Programming Languages:** C/C++, Python, Java, JavaScript, SQL