

# HAITIAN (HILTON) JIANG

☎ +86 136-8166-0596 ✉ [hiltonjiang@gmail.com](mailto:hiltonjiang@gmail.com) 🌐 [github.com/hiltonjiang](https://github.com/hiltonjiang)

## EDUCATION

### Fudan University

Sep. 2019 – Jul. 2023 (expected)

*B.S. in Data Science*

*Shanghai, China*

- **GPA:** 3.85/4.00 **Ranking:** 2/80
- Teaching Assistant for Algorithm & Data Structure(Honor), Advanced Big Data Analytics
- Core Courses: Algorithm & Data Structure, Computer System, Numerical Methods, Optimization, Database, Stochastic Process, Computer Vision, Natural Language Processing, Parallel Computing

### University of California, San Diego (UCSD)

Sep. 2021 – Dec. 2021

*Visiting Student, Computer Science and Engineering*

*La Jolla, CA*

- **GPA:** 3.94/4.00
- Courses: Theory of Computation, Recommender Systems, Probabilistic Reason & Learning

## RESEARCH EXPERIENCES

### Graph Neural Network (GNN) System with Historical Cache

May. 2022 – Dec. 2022

*Deep Graph Library (DGL) group, Amazon Web Services (AWS)*

*Shanghai*

- Accelerated and scaled up GNN training on large graphs with a selective historical cache to reduce computation and data movement. On average, **64.5%** of data movement is reduced.
- Gained a **23.6×** and **5.3×** speedup (with same test accuracy) compared with the popular GNN systems DGL and PyG, respectively, (GraphSAGE model on **ogbn-papers100M** dataset).
- Designed managing policies for nodes in cache and conducted experiments to compare performances.
- Provided theoretical analysis for the convergence of training GNNs with history cache.
- Submitted to OSDI 2023. Link to the paper: <https://hiltonjiang.github.io/pdf/FreshGNN.pdf>

### Scalable Graph Neural Networks using Subgraph Summarization

Jan. 2022 – Oct. 2022

*Advisor: Prof. Zengfeng Huang*

*Fudan University, Shanghai*

- Proposed a sub-graph sampling method that can cooperate with graph coarsening algorithms to reduce the size of sub-graphs by 30%.
- Gained accuracy increase by about 1% in small datasets with common models like GCN.
- Scaled up GNN training: implemented both pre-processing and training on multi-GPU; reduced the training time on papers100M from over 10 minutes to 27s for one epoch.

### Recommendation System with Faithful Textual Explanation

Oct. 2021 – Aug. 2022

*Advisor: Prof. Julian McAuley*

*University of California, San Diego*

- Evaluated current text-based recommenders through aspect-based sentiment analysis and found their text cannot faithfully explain their ratings because popular and positive aspects dominate the text.
- Designed a metric of faithfulness for recommendation explanation by learning and comparing the mapping from explanations to predicted ratings in order to measure the issue.
- Developed and trained a RoBERTa model to learn the underlying mapping between reviews and ratings on the Amazon Clothing and Google Local Restaurant dataset; got a mean squared error (MSE) of 0.33 against 0.9 from typical models using user and item embeddings.
- Designed and implemented a new pipeline that first generate text explanations of recommendations and then use the text to predict faithful ratings; achieved similar MSE compared with traditional methods that cannot generate explanations.

## PROJECT EXPERIENCE

---

### Parallel Computing | *C++, OpenMP*

Sep. 2022 – Present

- Accelerated matrix multiplication in C++ using optimized loop order, compiler optimizations, multi-core programming, tiling, and vectorization; got a cumulative 6941x speedup.
- Implemented parallel longest common sequence (LCS) algorithm with reduced space complexity and got a 5x speedup on a 8-thread machine.

### Image Registration System | *Python, numpy*

May. 2022 – Jun. 2022

- Built an image registration system using PyQt for GUI and numpy for implementing the stochastic gradient descent algorithm from scratch with user-friendly GUI and command line interface.
- Used optical-flow method for non-linear image transformation and different loss function for linear transformation: squared error, correlation coefficient, mutual information and K-L divergence, etc.
- Found the optical-flow method will break the global structure due to the coarse-to-fine procedure, so it should be used for local transformations; for the loss in linear transformation, K-L divergence is sensitive to parameters in this problem, while other loss functions all perform similarly good.

### Few-shot Oracle Character Classification | *Self-supervised learning*

May. 2022 – Jun. 2022

- Used sequence information extracted from images to generate augmented samples by adapting Sketch-BERT model.
- Added self-supervised learning on the direction prediction for the random rotation performed to the input image to enable better learning of representation; added ensembles on the feature vectors to make the network more robust.
- Gained about 15% higher accuracy than the SOTA method on all 1-shot, 3-shot and 5-shot tasks.

### Network Service | *Go*

Sep. 2021 – Nov. 2021

- Implemented a concurrent http server based on TCP sockets; can be visited through browsers or command line tools.
- Implemented a distributed file synchronizing client/server system based on consistent hashing for the load balance of servers and remote procedure call for the client.

### CPU Emulator with 5-stage Pipeline | *C*

Nov. 2020 – Dec. 2020

- Implemented a RISC CPU with the classic five-stage pipeline, guard for hazard, pipeline with stall.
- Added support for data forwarding and branch instruction.
- Tested on a binary compiled from a MIPS assembly code containing loops and arithmetic calculations and got the correct results.

## PROFESSIONAL EXPERIENCE

---

### Tencent Technology

Feb. 2021 – Jul. 2021

#### *IoT Security Intern*

*Shanghai, China*

- Collected, reproduced and exploited vulnerabilities(CVE) of IoT devices; conducted reverse-engineering on the firmware of devices.
- Implemented a crawler system using Scrapy and PostgreSQL to automatically gather information of the products and firmwares and store them in the database.
- Built an asynchronous firmware downloader that fetches the source URL from the database and uses message queue (RMQ) for communication; used docker as the container for portability.

## HONORS & AWARDS

---

National Scholarship (TOP 1%)

Nov. 2021

National Second Prize (TOP 3%), CUMCM

Dec. 2020

(Contemporary Undergraduate Mathematical Contest in Modeling)

Second Prize Scholarship of Fudan University (TOP 10 %)

Nov. 2020

## SKILLS

---

**Languages:** Mandarin (Native speaker), English (Fluent)

**Programming:** Python, C/C++, Go, CUDA, MATLAB, R, SQL, PHP, Bash

**Frameworks:** Linux, Git, PyTorch, GPU programming, OpenMP, Spark, Docker, Web Crawler(Scrapy)