

# Yifei Liu

1037 Luoyu Road, Wuhan, Hubei, China, 430074  
yfliucs@gmail.com • +86 182-0271-3791 • <https://yifeiliu.me/>

## RESEARCH INTERESTS

I am interested in distributed systems, file/storage systems, and machine learning for systems. I am currently seeking Computer Science Ph.D. openings for Fall 2019.

## EDUCATION

### Huazhong University of Science and Technology

Wuhan, Hubei, China

- Master of Science (M.Sc.) in Computer System Architecture 09/2016 – 06/2019 (Expected)
  - *Thesis (Ongoing)*: Graph-based Metadata Management in Storage Systems (Advised by Prof. Ke Zhou)
  - *Relevant Courses*: Parallel Processing; Theory and Technology of Information Storage; Advanced Computer Architecture; Theory of matrices; Data Mining and Retrieval.
  - Cumulative GPA: 85.02 / 100

### Huazhong Agricultural University

Wuhan, Hubei, China

- Bachelor of Engineering (B.Eng.) in Computer Science and Technology 09/2012 – 06/2016
  - *Thesis*: Predicting Disk Failures based on Machine Learning Methods
  - *Relevant Courses*: Data Structure; Operating System; Principles of Computer Composition; Computer Networks; Compilers Principles; Programming and Application of Linux.
  - Cumulative GPA: 3.43 / 4.00 Major GPA: 3.79 / 4.00 Rank: 9/118

## EXPERIENCE

### Wuhan National Laboratory for Optoelectronics

Wuhan, Hubei, China

*Master's Student/Research Assistant*

09/2016 – Present

- Use deep learning hash to design and implement a metadata system for integration of high-precision and low-latency semantic queries in storage systems
- Proposes a framework for assessing image “dark data” based on a novel semantic hash ranking (SHR) algorithm
- Perform theoretical analysis on hash-based graphs to facilitate rank algorithms and graph database operations

### Tencent (#1 social media, Internet corporation in Asia)

Shenzhen, Guangdong, China

*Backend Developer Intern*

12/2015 – 08/2016

- Predicted disk failures with disk data collected via machine learning algorithms to achieve high precision and recall
- Built infrastructure for collecting long-term disk S.M.R.A.T. data from over 10,000 servers in Tencent data centers
- Tested cache replacement policies on Tencent application traces

### Institute of Computing Technology, Chinese Academy of Sciences

Beijing, China

*Intern/Visiting Student*

11/2015 – 12/2015

- Explored the correlation between Apache Spark and Java Garbage Collection (GC) parameters

### Chinasoft International Ltd

Guangzhou, Guangdong, China

*Front-end Student Intern*

07/2015 – 08/2015

- Collaborated with other interns to develop a catering website using jQuery, AJAX, Java Servlet

### Huazhong Agricultural University

Wuhan, Hubei, China

*Undergraduate Research Assistant*

11/2014 – 06/2015

- Used Bayesian Network Reasoning to propose a web service recommendation approach for organizing and recommending a set of correlated services
- Designed an image line detection algorithm for automatically measuring character parameters of the rapeseed plant
- Developed a GUI for counting the number of cotton cells in microscopic images (for the Huazhong Agricultural University's College of Plant Science & Technology)

PUBLICATIONS	<div><div>[5] <a href="#">Yifei Liu</a>, Yu Liu, Yangtao Wang, Ke Zhou, and Yujuan Yang. “SwiftGraph: A Metadata Organization Approach for Enabling Semantic Queries in Storage Systems.” Under Review at <i>International Parallel &amp; Distributed Processing Symposium (IPDPS)</i>, 2019.</div><div>[4] Yangtao Wang, Yu Liu, <a href="#">Yifei Liu</a>, Ke Zhou, and Yujuan Yang. “Analysis and Management to Hash-Based Graph and Rank.” Under Review at <i>International Conference on Database Theory (ICDT)</i>, 2019.</div><div>[3] Yu Liu, Yangtao Wang, Ke Zhou, Guoliang Li, Yujuan Yang, and <a href="#">Yifei Liu</a>. “A Framework for Image Dark Data Assessment.” Under Review at <i>International Conference on Data Engineering (ICDE)</i>, 2019.</div><div>[2] Jianxiao Liu, Zonglin Tian, <a href="#">Yifei Liu</a>, and Liang Zhao. “Research of Web Service Recommendation Using Bayesian Network Reasoning.” In <i>Proceedings of the 15th International Conference on Services Computing (SCC)</i>, Seattle, WA, 2018.</div><div>[1] Pujuan Shi, <a href="#">Yifei Liu</a>, Yihang Fang, Chengda Lin, and Ruifang Zhai. “A new line detection algorithm - Automatic measurement of character parameter of rapeseed plant by LSD.” In <i>Proceedings of the 4th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)</i>, Istanbul, Turkey, 2015.</div></div>	
COMPETITION AWARDS	<div><div><div>▪ Second Prize, National Postgraduate Mathematic Contest in Modeling, China</div><div>12/2016</div></div><div><div>▪ Finalist, Interdisciplinary Contest in Modeling (MCM/ICM), USA</div><div>04/2015</div></div><div><div>One of 52 awardees in the world. First winner of my university. (52/9773 ~0.53%)</div></div><div><div>▪ First Prize, National Postgraduate Mathematic Contest in Modeling, China</div><div>12/2014</div></div><div><div>Won the highest award of this contest, (120/4900 ~2.4%)</div></div><div><div>▪ First Prize, Contemporary Undergraduate Mathematical Contest in Modeling, China</div><div>11/2014</div></div><div><div>Won the highest award of this contest, (293/22233 ~1.3%)</div></div></div>	
ACADEMIC AWARDS	<div><div><div>Huazhong University of Science and Technology</div><div><div>▪ Merit Graduate Student</div><div>2017 – 2019</div></div><div><div>▪ First-class Academic Scholarship</div><div>2016 – 2019</div></div></div><div><div><div>Huazhong Agricultural University</div><div><div>▪ Outstanding Graduate</div><div>Spring 2016</div></div><div><div>▪ First Scholarship for Study</div><div>Spring 2015</div></div><div><div>▪ Merit Undergraduate Student</div><div>2014 – 2016</div></div></div></div></div>	
TECHNICAL SKILLS	<div><div><div>▪ Daily-working on Storage Systems and Mathematical Modeling</div><div>▪ C/C++/Java</div><div>▪ Python/MATLAB/BASH/SQL/JavaScript</div><div>▪ Parallel Computing: Hadoop, Spark</div><div>▪ File and Storage: OpenStack Swift, HDFS, Redis, Extended File System (ext)</div><div>▪ Databases: HBase, SQL Server, MySQL</div><div>▪ Deep learning platforms: Caffe, TensorFlow</div><div>▪ Tools: Linux, git, gcc, gdb, pdb, Makefile, Scikit-learn, Sqoop, L<sup>A</sup>T<sub>E</sub>X</div></div></div>	
PROFESSIONAL TRAINING	<div><div>Machine Learning - Stanford Univeristy (2016), Coursera</div></div>	

[CV compiled on 2018-11-21 ]