

# SI CHEN

4703344900 ◊ si.chen2@emory.edu ◊ 400 Dowman Drive, Atlanta, Georgia 30322

## SUMMARY

---

Ph.D. student in Computer Science, under supervision of Dr. Avani Wildani. Currently, I am focusing on applying machine learning techniques to automatically characterize the block I/O workloads and optimize the storage system provisioning.

## EDUCATION

---

**Ph.D.** in Computer Science 2017 - 2023 expected  
Emory University, GPA : 3.98/4.00  
Concentration: Machine Learning in Storage System

**Master** in Electronic and Information Engineering 2004 - 2006  
Huazhong University of Science and Technology, China, GPA: 3.6/4.00

**Bachelor** in Electronic and Information Engineering, 2000 - 2004  
Huazhong University of Science and Technology, China, GPA: 3.84/4.00

## RESEARCH EXPERIENCE

---

**Research Assistant in SimBioSys Lab** 8/2017 - Present  
HPC application run-time performance prediction using Meta Learning  
Explored similarity measurement between HPC proxy/parent applications using ML-based methods  
Investigated time-series based feature extraction to identify the number of concurrent workloads, using gradient boosting and clustering.  
Detected and predicted workloads phase shift, combining FIO replay tool and feature extraction.  
Explored the workloads separation using blind source separation techniques such as Independent Component Analysis (ICA).  
Developed a cache prefetching framework for memory address with deep learning method LSTM.

## TEACHING EXPERIENCE

---

CS534 Machine Learning (graduate course) 1/2019 - 5/2019  
CS224 Discrete Structures 8/2018 - 12/2018  
CS170 Introduction to Computer Science (Java Programming) 1/2018 - 5/2018

## WORK EXPERIENCE

---

**Research Intern**, Bytedance 5/2021 - 8/2021  
Explored machine learning methods for AIOps. Improved the log diagnose and troubleshooting system for Deep Learning platform.

**Research Intern**, Netapp 5/2020 - 8/2020  
Enhanced the product function of performance headroom prediction and What-if Planning.

**Senior Engineer**, China Academy of Information and Communications Technology 07/2006-12/2016  
Developed the long-term strategy and plans for emergency communication and government communication network.

**Engineer**, Ministry of Industry and Information Technology, China 12/2011 - 09/2013  
Managed and coordinated Emergency Communication system

**Research assistant**, Internet Technology and Engineering R&D Center, Wuhan, China 09/2003 - 06/2006  
Designed and implemented a private communication system with high-quality voice using protocols of Voice over IP.

## PUBLICATIONS & PRESENTATION

---

**Si Chen**, Omar Aaziz, Jeanine Cook, Avani Wildani, *Similarity Measurement for Proxy Application Fidelity*, The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC21), November 2021

**Si Chen**, Jianqiao Liu, Avani Wildani, *CENSUS: Counting Interleaved Functional Tenants on Shared Storage*, 36th International Conference on Massive Storage Systems and Technology (MSST 2020), October 2020

**Si Chen**, Jianqiao Liu, Avani Wildani, *CENSUS: Counting Interleaved Functional Tenants on Shared Storage*, FAST'20 Work-in-Progress (Talk + Poster), February 2020

**Si Chen**, Avani Wildani, *Chasing the Signal: Statistically Separating Interleaved I/O Workloads*, FAST'19 Work-in-Progress (Talk + Poster), February 2019

**Si Chen**, Avani Wildani, *Chasing the Signal: Statistically Separating Multi-Tenant I/O Workloads*, workshop on ML for Systems (co-located with NeurIPS 2018), December 2018

## HONORS AND AWARDS

---

GHC Student Scholarship	2020
FAST '19 Student Grant	2019
Emory CS Grace Hopper Scholarship	2019
FAST '18 Diversity Grant	2018
Excellent Employee Award for China Academy of Information and Communications Technology	2012
Advanced Class of Elite Range (ACER), Huazhong University of Science and Technology	2000 - 2004
The Top Academic Outstanding Student (top 1% among 800 students), Huazhong University of Science and Technology	2000 - 2004

## TECHNICAL SKILL

---

<b>Program languages</b>	Python, C, Java
<b>Program libraries</b>	Keras, Tensorflow