CV

JIAO ZIYANG

Mobile: 3145849450 Email: zjiao04@syr.edu
Address: 121 Lafayette RD, Syracuse, NY 13205

Washington University in St. Louis (WUSTL) **08/2019 – 08/2020**

Major: Computer Science and Engineering

GPA: 4.0 Ranking: NA

Degree: Master of science in computer science (expected in 2021)

Syracuse University(SU) **08/2020 – Now**

Major: Computer and Information Science and Engineering

GPA: 4.0 Ranking: NA

Degree: Ph.D. in computer science (expected in 2023)

Publications

Ziyang Jiao, Byran S. Kim, *Generating Realistic Wear Distributions for SSDs,* FAST 22 WiP

Skills & Tools

Languages: C, C++, Python, HTML5, Java
Tracing: BPF(BCC Tools, bpftrace), blktrace, blkparse, btrecord, btreplay
SSD development platforms: FTLSim, Amber, FEMU

WORK EXPERIENCE

Syracuse University, Research Assistant, Syracuse, NY Aug 2020 - Present

- Storage Systems Research. Currently working on NSF grant – CPR for Flash-Based Storage Systems under Prof. Bryan Kim.
- Capacity-variant storage systems: exploiting the tradeoffs among capacity, performance, and reliability (CPR) and demonstrating the effectiveness of a capacity variant SSD.
- Self-learning systems: imbuing intelligence to the devices so that they can self-learn, self-configure, and self-manage.
- Next-generation storage stack with key-value / ZNS devices: exploring the design of a storage stack using key-value / ZNS devices instead of block devices.

Washington University in St. Louis, Teaching Assistant, St. Louis, MO Jan 2020 – May 2020

- Course Link: CSE 417T - Introduction to Machine Learning
- Advisor: Prof. Chien-Ju Ho
- Topics: Generalization in finite and infinite hypothesis spaces; Model complexity, the VC bound, the bias-variance tradeoff; Linear models: the perceptron, regression, logistic regression; Nonlinear transformations of data; The problem of overfitting; Modern supervised learning techniques, including decision trees, neural networks, nearest neighbor methods, support vector machines, boosting, and random forests.

Chinese Academy of Sciences(CAS), Research Assistant, BeiJing Nov 2018 - Jan 2019

- Laboratory for Face Recognition Based on Matlab+PCA+SVM.
- Advisor: Prof. Chao Liu
- Designing and building data pre-processing and training system
- Dataset: ORL face database + Real face image data
- Feature Engineering: Correlation analysis + PCA
- Model: NN, SVM, GAN

ACADEMIC EXPERIENCES

A Principled Approach for Selecting Block I/O Traces Oct 2021 – Now

Summary: Analyzing and profiling block I/O traces. Presenting a trace recommendation tool for selecting a subset of traces for benchmarking.

- Extract important features from I/O traces and computes (dis)similarities among them.
- Apply principal component analysis (PCA) and K-means to characterize I/O traces.
- Provide a guideline for selecting traces when benchmarking storage systems.

Generating Realistic Wear Distributions for SSDs Aug 2021 – Now

Summary: Building Fast-Forwardable SSD, a machine learning-based SSD aging framework that generates representative future wear-out states.

- Study on SSD aging and fail-slow symptoms.
- Quantify the low-performance variation under various I/O traces.
- Develop FF-SSD, an ML-based framework for SSD aging, trace replay acceleration, and drive failure estimation

Capacity Performance Reliability(CPR) for Flash-Based Storage Systems Oct 2020 – Sep 2023

Summary: Exploiting tradeoffs among CPR and designing a capacity-variant interface that allows the SSD to maintain performance while gracefully reducing the capacity.

- link:nsf.gov/awardsearch/showAward?AWD_ID=2008453
- Quantify the error-induced performance degradation.
- Build a capacity-variant system and demonstrate the effectiveness of a capacity variant SSD.
- Develop new filesystems and RAID systems to study how capacity-variance can be extended to a heterogeneous set of SSDs

Creating Synergies between Memory, Disk and Log in Log Structured KV Stores Aug 2020 – Dec 2020

Summary: Improving the background I/O performance on LevelDB – an open sourced key-value store by Google

- Study on LevelDB and analyze performance under different configurations
- Implement techniques discussed in TRIAD on LevelDB
- Smarter categorizing and scheduling to amortize background IO costs in LevelDB
- Self-adaptive database adjusting to different workloads

Spam Email Detection Dec 2019 – Feb 2020

Summary: Spam Email classifier based on Naïve Bayes (NB)

- Implement a system that mitigates the spamming problem
- Python programming to classify emails received into three categories: safe, spam, and suspicious email
- Based on naive Bayes method

Creative Programming and Rapid Prototyping Aug 2019 – Dec 2019

Summary: Web development

- Module description
- Github main page: https://github.com/ZiyangJiao
- Techniques: HTML5, CSS, AWS, MySQL, PHP, NodeJS, Git, Python

Machine Learning Algorithm Optimization*(supervised learning)* Aug 2019 – Dec 2019

Summary: Implementing and optimizing ML models based on MATLAB

- Implementing modern supervised learning techniques, including decision trees, neural networks, nearest neighbor methods, support vector machines, boosting, and random forests.
- Feature engineering based on filter, wrapper, and embedded methods
- Improve model performance based on mathematical analysis, such as Hoeffding's inequality, VC-dimension and the bias-variance tradeoff.

CONFERENCES

FAST ’22: 20th USENIX Conference on File and Storage Technologies

FAST ’21: 19th USENIX Conference on File and Storage Technologies

OSDI ’20: 14th USENIX Symposium on Operating Systems Design and Implementation

SOCC ’20: 11th ACM Symposium on Cloud Computing 2020

HONORS AND AWARDS

University scholarship (school-level) 2015/2016/2017/2018

National scholarship 2016

Syracuse University Fellowship 2020 / 2021

EXTRA-CURRICULUM ACTIVITIES

WHO Program: COVID-19: methods for detection, prevention, response and control	2021
Member of College Students Association for Science and Technology	2015-2019
Volunteer for the Asia-Pacific Mathematical Contest in Modeling	2016