

## Education and Qualifications

Present	PhD	Computer Science and Engineering	University of Michigan
2017	MS	Computer Science and Engineering	University of Michigan
2015	B.Eng.	Software Engineering	Xi'an Jiaotong University

## Publications

### Research papers

1. Yu, P. and M. Chowdhury (2020). Salus: Fine-Grained GPU Sharing Primitives for Deep Learning Applications. In: *Proceedings of the 3rd Conference on Machine Learning and Systems*.

### Workshop papers

1. Nguyen, L., P. Yu, and M. Chowdhury (2017). No! Not Another Deep Learning Framework. In: *Proceedings of the 16th Workshop on Hot Topics in Operating Systems*, pp.88–93.

### Under Submission

1. Yu, P., J. Liu, and M. Chowdhury (2021). FluidExec: A Generic Resource-aware Hyperparameter Tuning Execution Engine. In: *MLSys'21*.

## Awards and Scholarships

- 2014 Google Excellent Scholarship (\$1500, only one in Xi'an Jiaotong University)
- 2014 Meritorious Winner of Interdisciplinary Contest in Modeling
- 2013 Silver Medal of the ACM-ICPC Asia China  
Shaanxi Provincial Programming Contest
- 2013 Fuji Xerox (China) Scholarship (RMB5000, 4 of top 20% student in the major)
- 2013 Merit Student in Xi'an Jiaotong University
- 2012 First prize for MCM/ICM of Xi'an Jiaotong University
- 2012 Third prize for ACM Programming Contest of Xi'an Jiaotong University
- 2012 Third prize for "Tengfei Cup" Undergraduate  
Curricular Academic Science and Technology Competition
- 2011 Excellent Student Cadre in Xi'an Jiaotong University
- 2011 "Lu Shidi" Scholarship (RMB6000, 2 of top 10% students in the major)

## Work Experience

- Internship at Facebook From May., 2019 to Aug., 2019
  - Build fleet-wide GPU utilization regression detection and attribution dashboard
  - Discover and fix data consistency issue in GPU performance data.
  - Identify optimization opportunity and give improvement suggestion via automated emails.

## Research Experience

- Generic Resource-aware Hyperparameter Tuning Execution Engine From May, 2020 to Present

- Generalized Hyperparameter tuning execution engine using efficient heuristics with theoretical guarantees to solve the scheduling problem.
- Improve resource utilization by both inter- and intra-GPU training mechanisms
- Salus: Fine-Grained GPU Sharing for Deep Learning Applications *From Sep., 2016 to Apr., 2020*
  - Fine-grained GPU sharing by providing missing primitives: fast switching and memory sharing.
  - Improves GPU utilization for hyper-parameter tuning by 2.38×, and for DL inference applications by 42× over not sharing the GPU.
  - Code open sourced at <https://github.com/SymbioticLab/Salus>
- Deep Tree: SQL Injection Detection by the Power of Deep Learning *From Sep., 2016 to Mar., 2017*
  - Tree-based CNN for SQL statements classification with 94.7% accuracy for injection detection.
  - Compiled new SQL statements dataset of 4161 samples.
  - Code open sourced at <https://github.com/Aetf/tensorflow-tbcnn>
- System Design of Streaming Video Analysis Application in Storm *From Apr., 2016 to July, 2016*
  - Storm topology with several video classification, captioning and object tracking workloads.
  - Analysis of latency and throughput on 3 GPU servers, to understand the relationship between parallelism hint and performance.
  - Summer research in Clarity Lab.
- Evaluation of Graphical Keyboard User Interface *From Sep., 2015 to Dec., 2015*
  - Evaluated of two GKUI applications completing different tasks. Operating systems were also included as a variable in the experiment.
  - The result indicates significant improvements using GKUI in both tasks.
  - Course research project for Introduction to HCI Research at University of Michigan.
- Neural Network Classifier with Generalized Correntropy Loss *From Jan., 2015 to May, 2015*
  - Bachelor's thesis.
  - Implemented a neural network classifier using generalized correntropy loss function.
  - Analyzed the classifier's behavior under varied order parameters in generalized correntropy loss function.
- Application data separation using SEAndroid in NSKeyLab *From Nov., 2013 to Nov., 2014*
  - Focused on the implementation of multi-domain data separation.
  - Ported the SELinux policy compile tool chain to Android.
  - Implemented domain management and storage system service in both native user space and Android framework.
- Summer Practice in NSKeyLab (the Ministry of Education Key Lab for Intelligent Networks and Network Security on Network Traffic Capture and Analysis) *From July, 2012 to Sep, 2013*
  - Mainly engaged in the application development for local host network traffic capture and analysis, as well as the network traffic reconstruction algorithm.
  - Mastered related techniques such as WinPcap, WPF, TCP stream reassembly and HTTP reconstruction.
  - Analyzed 5 high speed download traffic samples and tens of HTTP traffic samples.
- Graphical Data Quality Management System based on IP-MAP *From Nov., 2012 to May, 2013*

- Professional graphical software that offers specialized platform for data quality management based on IP-MAP.
- Funded by the national innovation fund project of College Students of Xi'an Jiaotong University.
- Awarded the third prize by Xian Jiaotong University in the "Tengfei Cup" undergraduate curriculum academic science and technology competition.
- Participated as team leader and programmer.

## Project Experience

- LLDB Support for KDevelop (Google Summer of Code) *From May, 2016 to Aug., 2016*
  - Extended KDevelop C/C++ debugger architecture to allow multiple debugger backends.
  - Added LLDB backend for KDevelop, enabling C/C++ debugging through LLDB.
  - Successful Google Summer of Code project. Link: <https://summerofcode.withgoogle.com/archive/2016/projects/6014826014834688/>
  - Keep contributing to the KDevelop after the GSoC period.
- Contribution to open source Mozilla project *From June, 2014 to 2015*
  - Fixed several bugs (features) in JavaScript JIT Engine.
  - Implemented recover instruction for multiple MIR instructions.
  - Got Mozillian membership.
- Activities on GitHub *From 2012 to present*
  - Side projects including a bencode library, network traffic capture and analyze, torrent list migration between uTorrent and qBittorrent.
  - Contributed to several projects: KeepassXC, Mono, kmscon, CuteMarkEd, hid-apple-patched.
- Software Low Power Mode Based on Protean Code *From Sep., 2015 to Dec., 2015*
  - A runtime power saving optimization platform based on Protean Code triggered by OS power events.
  - Implemented devectorization pass which disables SLP instructions when batteries are running out.
  - Course project for Advanced Compiler in University of Michigan, worked in a group of 4.
- Pixel Cube Game for LeapMotion *From Nov., 2013 to June, 2014*
  - 3D pixel painting using hand gestures, powered by LeapMotion
  - Course project for Software Project Management in Xi'an Jiaotong University, worked in a group of 8.
- Mouse Control with Kinect *From May, 2012 to July, 2012*
  - Developed algorithm to smooth mouse movement by projecting hand movement onto a cylindrical surface.
  - The algorithm can adapt to different body parameters.

## Extracurricular Experience

- President of Research and Development Department *From Sep, 2012 to Feb, 2013*
  - Microsoft Student Technology Club of Xi'an Jiaotong University
- Commissary in charge of studies *From Sep, 2012 to June, 2015*

- My class (Software 14) in Xi'an Jiaotong University
- Social investigation on the effect of volunteer work around Xi'an *From Feb, 2012 to Jun. 2012*
  - First prize (group) for "My College Life" Students Social Investigation