

# SHILEI TIAN

✉ tianshilei1992@gmail.com · ☎ +1 (631) 590-0814 · in shilei-tian

## EDUCATION

---

**Stony Brook University (SBU)**, New York, United States 2019/08 – Present

*Ph.D.* in Computer and Information Sciences

**Shanghai Jiao Tong University (SJTU)**, Shanghai, China 2015/09 – 2018/03

*M.S.* in Computer Technology

**Xi'an University of Posts and Telecommunications (XUPT)**, Shaanxi, China 2010/09 – 2014/07

*B.S.* in Computer Science and Technology

## PUBLICATIONS

---

- **Shilei Tian**, Johannes Doerfert, Barbara Chapman, Asynchronous OpenMP Offloading on NVIDIA GPUs, LLVM Performance Workshop, San Diego, USA, February 23, 2020.
- **Shilei Tian**, Haotian Wang, Sha Li, Fan Wu, Guihai Chen, Trajectory-Based Multi-Hop Relay Deployment in Wireless Networks, COCOA 2017, Shanghai, China, December 16–18, 2017.
- Haotian Wang, **Shilei Tian**, Xiaofeng Gao, Lidong Wu, Guihai Chen, Approximation Designs for Cooperative Relay Deployment in Wireless Networks, ICDCS 2017, Atlanta, GA, USA, June 5–8, 2017.

## WORKING EXPERIENCE

---

**Intel Asia Pacific Research and Development Co.**, Shanghai, China 2018/04 – 2019/08

Software Engineer, Intel Compilers and Languages (ICL)

**Intel Asia Pacific Research and Development Co.**, Shanghai, China 2017/03 – 2018/04

Software Engineer Intern, Machine Learning and Translation (MLT)

## SELECTED PROJECTS

---

**LLVM and OpenMP** 2019/08 - Present

*Contributor*

- Implemented multiple streams in OpenMP which enables asynchronous OpenMP offloading on NVIDIA GPUs.
- Designed and implemented unshackled tasks, and implemented the clause `nowait` for directive with unshackled tasks. Each `nowait` target will be implicitly transformed to an unshackled tasks.

**Data Parallel C++ (DPC++) Compiler** 2018/05 – 2019/08

*Software Engineer, Validation Owner*

- Led the validation team of DPC++ compiler responsible for quality assurance, test development, etc.
- Implemented 10+ new features for DPC++ runtime library, such as `cl::sycl::stream`, host side of `cl::sycl::half`, asynchronous handler, kernel and device description for sub-group, and fixed 20+ product issues.
- Enabled **all** regular tests and analyzed thousands of product issues. Made Khronos SYCL conformance test suite work with Intel SYCL and developed 10 new test cases for Intel specific extension sub-group.

## TensorFlow (MKL-DNN Backend)

2017/12 – 2018/05

*Software Engineer (Intern)*

- Fixed a critical issue in the MKL-DNN backend, which seriously affected the convergence.
- Designed and implemented a bug pinpoint tool, CoSim, which is an operator-wise comparison tool that can locate the operator that first introduces large computing error between the target and reference. The corresponding poster has been published in the Intel Validation Summit 2018.
- With the help of CoSim, found and fixed two issues in the operator AddN that potentially affected the convergence. Besides, fixed a number of failures in the unit test.
- Made the TensorFlow (MKL-DNN backend) work on macOS, which was not supported before.
- Designed and implemented a test framework for the MKL-DNN backend, and correspondingly added hundreds of unit test cases. After that, no integration issue found any more.

## Intel Chainer

2017/03 – 2017/12

*Software Engineer Intern*

- Optimized the layers of dropout and softmax using Intel intrinsics and trickily designed parallel computing, yielding 20x and 8x performance improvements respectively.
- Designed and implemented a performance tool, Layer Performance Comparison (LPC). This tool is to measure the performance of different layers, like conv or relu, between different frameworks. We found that there existed a large performance gap in the layer of sum and relu with the help of LPC. The later corresponding optimization gained 10% performance improvement in GoogLeNet.
- Designed and implemented a general test infrastructure for deep learning frameworks based on Jenkins, Node.JS, MongoDB, Mongoose, covering the features of task trigger, distribution, scheduling, execution; data collection, process, storage; and report presentation.