Haiyu Mao

Homepage: https://hybol1993.github.io/
Phone: +1 (352)-226-2663
Email: maohaiyu1993@gmail.com

OVERVIEW I am a 4th-year Ph.D. candidate at Tsinghua University. My research interests mainly include the emerging Non-Volatile Memories (e.g., Racetrack Memory, Phase Change Memory and Resistive Random-Access Memory) and the Processing In Memory (PIM). I focus on solving problems when the NVMs are integrated into the traditional memory hierarchy, leveraging characteristics of NVMs to enable high-performance PIM and also exploring security hazard of NVMs when they are applied in both the memory system and PIM.

EDUCATIONAL BACKGROUND

Tsinghua University Beijing, China
Ph.D. of Computer Science and Technology; Aug. 2015 – July. 2020 (expected)
Advisor: Jiwu Shu; GPA: 3.7/4.0

University of Florida Florida, United States
Visiting Scholar of Electronic and Computer Engineering; Oct. 2017 – May. 2019
Advisor: Tao Li

Peking University Beijing, China
Research Intern of Computer Science and Technology; Feb. 2015 – June. 2015
Advisor: Guangyu Sun

Northeastern University Shenyang, China
Bachelor of Software Engineering; Aug. 2011 – July. 2015
GPA: 90/100 (**Rank: 1/201**)

PUBLICATIONS

Haiyu Mao, Jiechen Zhao, Jiaqi Zhang, Tao Li, Fan Yang, Jiwu Shu. "For the double-blind principle, the title is hidden." *Under Submission to **ISCA 2019**.*

Fan Yang, Youyou Lu, Youmin Chen, **Haiyu Mao**, Jiwu Shu. "No Compromises: Secure NVM with Crash Consistency, Write-Efficiency and High-Performance." *To appear.* (**DAC 2019**).

Haiyu Mao, Mingcong Song, Tao Li, Yuting Dai, and Jiwu Shu. "LerGAN: A Zero-Free, Low Data Movement and PIM-Based GAN Architecture." *In 2018 51st Annual IEEE/ACM International Symposium on Microarchitecture* (**MICRO 2018**), pp. 669-681. IEEE.

Haiyu Mao, Xian Zhang, Guangyu Sun, and Jiwu Shu. "Protect non-volatile memory from wear-out attack based on timing difference of row buffer hit/miss." *In Proceedings of the Conference on Design, Automation & Test in Europe* (**DATE 2017**), pp. 1627-1630. European Design and Automation Association.

Haiyu Mao, Chao Zhang, Guangyu Sun, and Jiwu Shu. "Exploring data placement in racetrack memory based scratchpad memory." *In 2015 IEEE Non-Volatile Memory System and Applications Symposium* (**NVMSA 2015**), pp. 1-5. IEEE.

Guoqi Liu, **Haiyu Mao**, Baoming Pu, Yongfeng Zhu, Jin Huang. "Fault Diagnosis of Wind Turbines Based on Wavelet Neural Network." *Journal of Chinese Computer Systems, 2015.* (First student author) In Chinese.

PROJECTS

Enhance The Endurance of NVM-based PIM Devices 04/2018-Present

Key idea: Utilize the characteristics of both NVM cells and neural networks to prolong the lifetime of PIM device.

- Analyse the write behavior of updating the weight matrix in the NVM-based PIM array when PIM is employed in neural network training.
- Analyse the accuracy behavior of neural networks when changing the particular percent of weight values into given values during different training epochs (by modifying Caffe framework).
- Design a micro-architecture on both software and hardware levels to lengthen the lifetime of PIM device. Leverage the characteristic of NVM cells that stuck-at-fault cells can still be used in analog computing and one resistance can represent different numerical values with different sense amplifier references. Combine the inherent fault-tolerant ability of neural networks and their particular weight updating behaviors.
- Evaluate the proposed micro-architecture, including the lifetime simulation, performance, energy consumption, and software/hardware overhead breakdown.

PIM-based High-performance/Low-power GAN Training 10/2016-04/2018

Key idea: Remove the structured zero insertion and shorten the interconnections when training a GAN in PIM.

- Analyse and then find that the zero-inserting operation incurs serious redundant storage and computation, which can bot be solved by traditional compression since the data in PIM are all structured for both storage and computation.
- Observe that the interconnection in PIM is a bottleneck when training a GAN in it, since the long routings inside the GAN training hinder the performance of PIM.
- Propose a zero-free data reshaping scheme alongwith a structured data mapping scheme to save both storage and bandwidth in PIM.
- Propose a 3D reconfigurable interconnection structure in PIM to radically shorten the routings when training a GAN.
- Evaluate the proposed zero-free data reshaping scheme and the 3D-connected PIM, focusing on the performance and energy consumption, as well as the overhead.

Wear-out-based NVM Security Protection 09/2015-10/2016

Key idea: Reveal particular information through the difference between row buffer hit and miss.

- Find a vulnerability of NVM that we can reveal the information about particular data position through side channel, basing on the different read latency of row buffer hit and miss.
- Conduct a wear-out attack on this particular data successfully, even through the NVM is under the production of the state-of-the-art wear-leveling scheme.
- Build up a protection scheme to defend this attack. This protection scheme further prolongs the lifetime of NVM (compared with the state-of-the-art wear-leveling scheme), and only introduces a little hardware overhead.
- Evaluate the lifetime of NVM under the protection of proposed scheme, as well as the performance/energy consumption of NVM on the SPEC 2006.

Employ Racetrack Memory as The On-chip Scratchpad Memory 02/2015-08/2015

Key idea: Explore the data placement to enable the read/write port of Racetrack Memory to move the least.

- Observe that the read/write ports of Racetrack Memory are the most time/energy consuming part when it is used as the on-chip cache.
- Propose a Scratchpad Memory based data placement scheme to reduce the movement of read/write ports in Racetrack Memory, employing the genetic algorithm.
- Further optimize the data placement scheme by providing the initial genes through the help of *Fisrt Come First Store scheme*, *Most Access in The Middle scheme* and *Most Access in The Front scheme*.
- Evaluate the proposed Scratchpad Memory based Racetrack Memory on the SPEC 2006.

SELECTED AWARDS

MICRO-51 Student Travel Award — ACM SIGMICRO, 2018
Comprehensive Secondary Scholarship — Tsinghua University, 2017
Guanghua Scholarship — Tsinghua University, 2016
Scholarship Provided by The Mayor of Shenyang **(6 in the whole university)** — Mayor of Shenyang, 2015
Top 10 Excellent Undergraduates **(10 in the whole university)** — Northeastern University, 2014
Outstanding Undergraduate in Shenyang **(0.26%)** — Shenyang, 2014
Outstanding Student Pacesetter **(0.5%, three times)** — Northeastern University, 2012/2013/2014
National Scholarship in China **(1%, three times)** — Ministry of National Education, 2012/2013/2014

TECHNICAL SKILLS

Programming Languages Proficient: C, C++, Java, Python
Used: Javascript, Go, MPI, OpenMP, CUDA, Matlab
Frameworks: Caffe, TensorFlow, PyTorch, Hadoop, Spark
Simulator: Gem5, DRAMSim, NVSim, NVmain, CACTI

ENTREPRENEURIAL EXPERIENCE

One of five founders in the startup company called *Feisong* 07/2013-10/2014

- Design and Code the Andriod-based application called *Feisong*, used as an online convenience store for a community.
- Publish the app *Feisong* on the Andriod application market, hire deliveryman, find sponsors and start the test run in our campus with the other four founders.
- Popularize *Feisong* to other communities and sign license contract with Student Unions of Shenyang Pharmaceutical University, Shenyang University and East China Jiaotong University to use *Feisong*.