Title: Nano-Device and Architecture Interaction in Machine/Deep Learning

Abstract: This presentation discusses the challenges and prospects of using resistive switching devices in machine/deep learning accelerators. First, we introduce the crossbar array architecture to efficiently implement the weighted sum and weight update operations that are commonly used in the neuro-inspired machine/deep learning algorithms, and described the desired device characteristics of the synaptic weights and neuron nodes. Then, we present “NeuroSim”, a device-circuit-algorithm co-design framework to evaluate the impact of non-ideal device effects on the system-level performance (i.e. learning accuracy) and trade-offs in the circuit-level performance (i.e. area, latency, energy). Last, we propose to binarize the neural network algorithm to allow very low precision weights and neurons, thereby making it compatible to “digital” synapses (i.e. binary NVM) and showed our efforts for large-scale prototype demonstration.

Bio: Shimeng Yu received the B.S. degree in microelectronics from Peking University, Beijing, China in 2009, and the M.S. degree and Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA in 2011, and in 2013, respectively. He is currently an assistant professor of electrical engineering and computer engineering at Arizona State University, Tempe, AZ, USA.

His research interests are emerging nano-devices and circuits with a focus on the resistive memories for different applications including machine/deep learning, neuromorphic computing, monolithic 3D integration, hardware security, radiation-hard electronics, etc. He has published >60 journal papers and >100 conference papers with citations >4900 and H-index 31.

Among his honors, he is a recipient of the DOD-DTRA Young Investigator Award in 2015, the NSF Faculty Early CAREER Award in 2016, and the ASU Fulton Outstanding Assistant Professor in 2017.

He served the Technical Program Committee for IEEE International Symposium on Circuits and Systems (ISCAS) 2015-2017, ACM/IEEE Design Automation Conference (DAC) 2017, and IEEE International Electron Devices Meeting (IEDM) 2017, etc.