Data Science Transdisciplinary Area of Excellence

Data Salon 2025 — 2026

Resolving the tension between model efficiency and adversarial robustness for AI algorithms

Adnan Siraj Rakin, School of Computing

Friday, Nov. 7, 2025, 12:15-1:15 PM
AD-148, with lunch served
Zoom: https://binghamton.zoom.us/j/93495116544

Abstract

AdderNet emerges with superior hardware efficiencies compared to Convolutional Neural Networks (CNNs) by replacing the multiplicative operations with energy-efficient additive alternatives. Nevertheless, AdderNet's security exploration remains underexplored, especially against emerging adversarial weight perturbation attacks, such as the Bit-Flip Attack (BFA), which injects memory faults into weight bits. In our recent paper, we investigate AdderNet's vulnerabilities to BFA for the first time and discover that it is more susceptible to BFA than its CNN counterpart. To utilize the hardware efficiency of AdderNet and at the same time improve resiliency against BFA, we propose a novel Non-Negative AdderNet model. Proposed method consists of two key components: i) it incorporates a non-positive weight encoding technique, and ii) a novel accelerator design with an integrated non-positive decoder (NPD) to achieve hardware-level lightweight and runtime detection and correction of BFA. To further test the resilience of our proposed defense, we perform both a baseline BFA and an advanced BFA (BFA+), where the attacker has sufficient knowledge about the proposed defense. Our proposed NNAN makes BFA obsolete and requires many more attack cycles to be compromised against BFA+, successfully enabling the design of both hardware-efficient and secure AdderNet against BFA.

About the speaker: Adnan Siraj Rakin is an Assistant Professor in the School of Computing at Binghamton University (SUNY). Previously, he completed his Ph.D. (2022) and Masters (2021) in Computer Engineering from Arizona State University (ASU). His research interests include deep learning, computer vision and security. He has been the author/co-author of over many publications in IEEE/ACM top-tier journals and conferences (e.g., CVPR, ICCV, T-PAMI, USENIX Security) on this broad topic of Machine Learning Security. He has received the 2022 dean's dissertation award from the dean of Arizona State University (ASU) in recognition of his contribution to Machine Learning Security.

About the Data Salon: Data Salon is a dynamic venue designed to foster the exchange of ideas and the formation of new collaborations. Each gathering includes a brief talk to inspire discussion, but the emphasis lies on the social dimension — creating an open and welcoming space where scholars, researchers, and practitioners can engage in dialogue, discover shared interests, and explore opportunities for collaboration. More than a lecture series, Data Salon is a catalyst for community-building and cross-disciplinary connection.