PURDUE UNIVERSITY | Elmore Family School of Electrical and Computer Engineering

Faculty Candidate Seminar -- Purdue Computes: Systems-Software

Teodora Baluta

Ph.D. Candidate, Computer Science
National University of Singapore

Monday, March 4, 2024
10:30 A.M. – 11:30 A.M.
LWSN 3102

New Algorithmic Tools for Rigorous Machine Learning Security Analysis

Abstract

Machine learning security is an emerging area with many open questions lacking systematic analysis. In this talk, I will present three new algorithmic tools to address this gap: (1) algebraic proofs; (2) causal reasoning; and (3) sound statistical verification. Algebraic proofs provide the first conceptual mechanism to resolve intellectual property disputes over training data. I show that stochastic gradient descent, the de-facto training procedure for modern neural networks, is a collision-resistant computation under precise definitions. These results open up connections to lattices, which are mathematical tools used for cryptography presently. I will also briefly mention my efforts to analyze causes of empirical privacy attacks and defenses using causal models, and to devise statistical verification procedures with 'probably approximately correct' (PAC)-style soundness guarantees.

Bio

Teodora Baluta is a Ph.D. candidate in Computer Science at the National University of Singapore. She enjoys working on security problems that are both algorithmic in nature and practically relevant. She is one of the EECS Rising Stars 2023, a Google PhD Fellow, a Dean's Graduate Research Excellence Award recipient and a President's Graduate Fellowship recipient at NUS. She interned at Google Brain working in the Learning for Code team. Her works are published in security (CCS, NDSS), programming languages/verification conferences (OOPSLA, SAT), and software engineering conferences (ICSE, ESEC/FSE). More details are available on her webpage: https://teobaluta.github.io/.

CS Host: Ninghui Li ~ ninghui@purdue.edu
ECE Host: Charlie Hu ~ ychu@purdue.edu