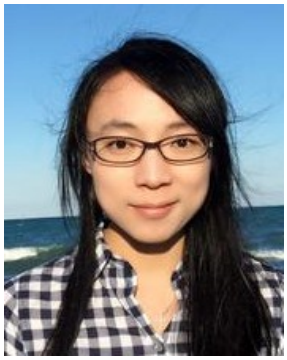
UNIVERSITY of WISCONSIN
UWMILWAUKEE

Special Interest Group
on Big Data

Invited Seminar

4pm - 5pm, Thursday, March 28, 2019
UWM EMS 715

Secure Learning in Adversarial Environments



Bo Li, Ph.D.
Assistant Professor
Department of Computer Science
University of Illinois at Urbana-Champaign
Home page: http://boli.cs.illinois.edu/

Abstract:

Advances in machine learning have led to rapid and widespread deployment of software-based inference and decision making, resulting in various applications such as data analytics, autonomous systems, and security diagnostics. Current machine learning systems, however, assume that training and test data follow the same, or similar, distributions, and do not consider active adversaries manipulating either distribution. Recent work has demonstrated that motivated adversaries can circumvent anomaly detection or classification models at test time through evasion attacks, or can inject well-crafted malicious instances into training data to induce errors in classification through poisoning attacks. In this talk, I will describe my recent research about evasion attacks, poisoning attacks, and privacy problems in machine learning systems. In particular, I will introduce an example of physical attacks in autonomous driving recognition system, and discuss several potential defensive approaches as well as robust learning models.

Short Bio:

Dr. Bo Li is an assistant professor in the department of Computer Science at University of Illinois at Urbana–Champaign, and is a recipient of the Symantec Research Labs Fellowship. Previously she was a postdoctoral researcher in UC Berkeley. Her research focuses on both theoretical and practical aspects of security, machine learning, privacy, game theory, and adversarial machine learning. She has designed several robust learning algorithms, a scalable framework for achieving robustness for a range of learning methods, and a privacy preserving data publishing system. Her recent research focuses on adversarial deep learning and generative models, as well as designing scalable robust machine learning models against adversarial attacks