“Efficient GAN-based Adversarial Example Generation Against ML-Based Network Intrusion Detection Systems”

By **Darren Hartono**

Abstract:

In the realm of network security, Network Intrusion Detection Systems (NIDS) are essential for identifying and mitigating malicious activities targeting networked devices. Traditionally, these systems have relied on signature-based and anomaly-based detection techniques. However, the increasing complexity and adaptability of cyber threats have driven the adoption of Machine Learning (ML) approaches in modern NIDS, significantly improving their ability to detect a wider range of attack vectors. Despite these advancements, ML-based NIDS remain vulnerable to adversarial examples—deliberately crafted inputs designed to mislead models and trigger incorrect classifications. Originally identified in the field of computer vision, adversarial examples now pose a critical threat to the reliability and robustness of ML-based cybersecurity systems. This thesis explores the use of Generative Adversarial Networks (GANs) to generate adversarial examples that closely resemble legitimate network traffic while effectively evading detection. By leveraging the generative power of GANs, the research aims to produce realistic and functional adversarial samples efficiently. The primary objective is to adapt, enhance, and evaluate a proposed GAN-based adversarial example generation algorithm to improve its effectiveness in testing and validating the resilience of ML-based NIDS. Experimental tests demonstrate that the proposed algorithm outperforms existing methods, achieving better results with reduced computational overhead.

Date: Monday, June 2nd, 2025

Time: 11:00 AM – 1:00 PM

Location: 14-238B

Committee: Dr. Fang, Dr. Debruhl, Dr. Anderson, and Dr. Xu