Title: **Adversarial Malware Generation using AI: From Evasion to Detection Enhancement**

Reethika Ambatipudi, RISE

Abstract:
Traditional malware detection systems—whether signature-based or AI-driven—struggle to keep up with the rapid evolution of modern threats. In this work, we explore how machine learning can be used not just for detecting malware, but for generating it—specifically **adversarial malware** designed to evade state-of-the-art detectors. We compare two state-of-the-art approaches: **MalGANs**, which use Generative Adversarial Networks to craft feature-level malware variants that fool classifiers, and **Reinforcement Learning**, where an agent iteratively learns evasion strategies through feedback from the detection system. While MalGANs are efficient and automated, RL-based methods offer flexibility to modify malware behaviour, though at higher computational cost. Our roadmap includes a systematic empirical evaluation of these approaches, the development of a unified adversarial malware generation framework, and leveraging these insights to harden detection systems. By using offensive AI as a tool for defense, we aim to expose and patch the blind spots in modern cybersecurity infrastructure.