Graduate School of Information Science, University of Hyogo

16th International Research Seminar

PARADIGM SHIFT IN AI SECURITY

Wed. 20 Aug. 2025 (13:00 ~ 14:00) JST

IN-PERSON/ONLINE SEMINAR

AI security is a multidisciplinary research area between AI and information security, and many kinds of attacks and their countermeasures have been found until now. In this talk, we discuss several major attacks and their countermeasures on AI in the past decades, including the author's results. Especially, we focus on backdoor attacks whereby an adversary embeds vulnerabilities in machine learning models through mislabeled data and show that this attacks accelerates other attacks against AI. We also discuss security applications of AI in recent years, including cyber-physical systems, and show several applications as the author's results. Furthermore, we focus on that large language models (LMMs) have been utilized in various research fields and then discuss future challenges in security orchestration for LLMs.

Guest Speaker

Naoto Yanai

Principal Engineer, Technology Division

Panasonic Holdings Corporation

Panasonic

Yanai got the PhD degree in theoretical cryptography in 2014. He then joined Osaka University as an assistant professor in 2014 and became an associate professor in 2021. He has also researched in various topics in information security and has also published several top citations papers in blockchains and AI security. He has also organized an international workshop for AI security, SECAI from 2023. From 2024, he moved to Panasonic Holdings Corporation as a principal engineer and has researched in both academic and industrial topics. In today's topic, he present paradigm shifts in multidisciplinary area between cybersecurity and AI, including his recent results.

Register here (free)

https://shorturl.at/H9zHM

Contact: rashed@gsis.u-hyogo.ac.jp

Kobe Campus for Information Science, Computational Science Center Building, Large Lecture Hall (720), 7th Floor

https://www.u-hyogo.ac.jp/about/access/

For more details:

https://researchmap.jp/naotoyanai