Contribution ID: **26** Type: **not specified**

Towards Interpretable Android Malware Detection with Transformer-Based Models

Tuesday 10 June 2025 16:15 (45 minutes)

As Android continues to dominate the global mobile market, cybercriminals increasingly target its vast user base with sophisticated malware. In this presentation, we propose an interpretable framework for Android malware detection that leverages language model to analyze a range of features—including app manifests, API calls, and opcode sequences. By integrating feature analysis techniques, our approach not only achieves high detection accuracy but also provides critical insights into which features drive classification decisions. We will share empirical results demonstrating the method's effectiveness on real-world datasets, discuss the benefits of interpretability for security practitioners, and explore how these findings can inform the next generation of mobile threat defense systems.

Length

Optional: Speaker / convener biography

Hantang Zhang is a doctoral student in Software Engineering and Security at Umeå university. His research interests include natural language processing (NLP) and software security. Currently, he focuses on leveraging transformer-based language models and feature analysis frameworks to enhance the accuracy and transparency of Android malware detection.

Primary author: ZHANG, HANTANG (Umeå University)

Presenter: ZHANG, HANTANG (Umeå University)

Session Classification: Talks

Track Classification: Talks and presentations