Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/NA10.20250109
Published: June 17, 2025

Evaluating the Capabilities of LLMs in Security Analysis for Mobile Applications

Joyce Champie
MS Graduate Student
Department of Computer Science
Florida Polytechnic University
Lakeland, Florida, USA
cjoycemaptuetagne00@floridapoly.edu

Karim Elish
Associate Professor
Department of Computer Science
Florida Polytechnic University Lakeland
Florida, USA
kelish@floridapoly.edu

Abstract

Large Language Models (LLMs) have demonstrated significant potential across various domains, including digital security. Recent studies have explored their ability to assist in mitigating code vulnerabilities; however, their effectiveness in mobile application security analysis remains an open question. This work aims to systematically evaluate the capabilities of LLMs in two key areas: mobile code vulnerability detection and malware family classification. First, we assess the performance of multiple LLMs in identifying vulnerabilities in Android code using an open dataset of over 100 vulnerable code samples from the Open Worldwide Application Security Project (OWASP). The evaluation focuses on the models' ability to detect security flaws and verify whether the permissions declared in an application's manifest file align with its actual behavior. This analysis will provide insights into the accuracy, strengths, and limitations of different LLMs in static code security assessment. Second, we investigate the potential of LLMs for malware family classification. Using a compiled dataset of malware samples labeled by family, we evaluate whether LLMs can accurately categorize malicious applications into their respective families. The study examines each model's classification accuracy, consistency, and effectiveness in identifying distinguishing characteristics of malware families. The findings of this research will provide a comprehensive understanding of LLMs' effectiveness performance in mobile security analysis, highlighting their strengths, limitations, and areas for improvement. Also, the findings will contribute to the growing body of knowledge on AI-driven security solutions and help to advance and automate mobile threat detection.

Keywords

Large Language Models, Vulnerability Analysis, Malware Classification, Mobile Application Security, AI in Cybersecurity.

Biographies

Joyce Champie is a graduate student at Florida Polytechnic University, where she is currently pursuing her master's degree in computer science and working as a graduate assistant. She holds a bachelor's degree in software engineering from Coventry University in 2022 and was recognized for Outstanding Research Achievement among a cohort of 74

students. From September 2023 to June 2024, she worked as a research assistant in the Computing Department at Coventry University. She also worked as an Online Marketing Officer for Snippet Limited Ghana. Her current research focuses on security analysis for mobile applications using AI and ML. Joyce's technical interests include business analytics and data analytics. She has also worked with various companies, including United Bank of Africa and Sopadit Transit, as a Database Intern. Joyce's projects include designing a website for Omega Community Church and creating a database for menstrual cycle tracking using MySQL.

Dr. Karim Elish is an associate professor of Computer Science at Florida Polytechnic University. He obtained his Ph.D. and MS in Computer Science from Virginia Tech. Before joining Florida Poly in 2016, he was an assistant professor at the Department of Computer Science at Purdue University-Fort Wayne. He received the prestigious Florida Poly ABLAZE Award for Excellence in Teaching in 2018. Dr. Elish's research portfolio is centered at the intersection of cybersecurity and software engineering. His expertise encompasses a wide array of topics, including cybersecurity analytics, software security, mobile security, and applied AI and ML. His work in these areas has garnered substantial recognition, with over 1,630+ citations on Google Scholar. He also has showcased his proficiency in acquiring funding, securing a total of around $716,000 in grants as both PI and Co-PI. Dr. Elish is serving as a reviewer for many high impact ISI journals and conferences such as IEEE Transactions on Dependable and Secure Computing (TDSC), IEEE Transactions on Information Forensics & Security, IEEE Transactions on Mobile computing (TMC), and others. He holds senior membership in both IEEE and ACM.