LITHUANIAN COMPUTER SOCIETY
VILNIUS UNIVERSITY, INSTITUTE OF DATA SCIENCE AND DIGITAL TECHNOLOGIES
LITHUANIAN ACADEMY OF SCIENCES

16th Conference on

DATA ANALYSIS METHODS
for Software Systems

November 27–29, 2025

Druskininkai, Lithuania, Hotel "Europa Royale"
https://www.mii.lt/DAMSS

VILNIUS UNIVERSITY PRESS
Vilnius, 2025

Contacts:

Dr. Jolita Bernatavičienė
jolita.bernataviciene@mif.vu.lt
Prof. Olga Kurasova
olga.kurasova@mif.vu.lt
Tel. (+370 5) 2109 315

https://doi.org/10.15388/DAMSS.16.2025
ISBN 978-609-07-1200-9 (digital PDF)

Forensics of Corpus Messages in Popular Android Device Applications

Matas Niewulis, Virgilijus Krinickij, Linas Bukauskas

Cybersecurity Laboratory
Institute of Computer Science
Faculty of Mathematics and Informatics
Vilnius University

matas.niewulis@mif.stud.vu.lt

With the rapid evolution of mobile technologies, smartphones have become an essential part of modern communication and personal data storage. Increasing functionality, accessibility, and integration of mobile devices with cloud-based ecosystems have significantly expanded the scope and sensitivity of the data they handle. The Android operating system dominates the global market, serving as the foundation for a wide variety of devices with differing architectures, firmware, and storage structures. As devices used by a user accumulate data containing personal communications, browsing history, application interactions, and metadata, they have become a critical focus for digital forensics and data-driven research. Due to frequent system updates, encryption mechanisms, and diverse manufacturer implementations, retrieving and analyzing Android data remains a complex and often inconsistent process. Existing tools are typically tailored for specific device types or Android versions, resulting in compatibility limitations and only partial data extraction. Such limitation creates a strong need for a universal and automated solution that can reliably extract and process data across different Android environments.

To address these challenges, this research aims to model and develop a Linux-based solution for more automated data acquisition and analysis from Android devices. The provided solution integrates various technological components to establish communication with the target device through Android Debug Bridge, enabling full access to its file system and corpus. The proposed solution would capture a complete physical bit image of the device's data partition and organize

extracted information into structured directories for further processing. Key data sources include a corpus which in turn undergoes analytical processing using a Naive Bayes classifier to identify correlations in an aim to visualize relational structures within the corpus. Unlike manual examination methods, an automated workflow would reduce human error and significantly improve the efficiency and repeatability of mobile data analysis. The classifier's output and analytical summaries would be presented to the analyst in a human readable format, allowing clear interpretation and standardized reporting. The research outcome would result in a more stable and automated data extraction and evaluation for forensics analysts.