C19 Android™ App Forensic Evidence Database

Chao-Chun C. Cheng, Iowa State University, Ames, IA 50011; Chen Shi, MS, Iowa State University, Ames, IA 50011; Brody Concannon, Iowa State University, Ames, IA 50011; Zhenqiang Gong, PHD, Iowa State University, Ames, IA 50011; Yong Guan, PhD, Iowa State University, Ames, IA 50011*

Learning Overview: After attending this presentation, attendees will understand how to use this new Android™ Center for Statistics and Applications in Forensic Evidence-App Evidence Database (CSAFE-AED) in their casework investigation. This presentation will introduce the basics, challenges, and limitations of the current mobile device forensics, and demonstrate how to take advantage of the CSAFE-AED database and search/recover the possible evidence from the possible locations on and outside the mobile devices being investigated.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by demonstrating how this new Android™ CSAFE-AED can simplify and speed up the investigation procedures.

This project seeks to develop a set of automated Android™ app analysis tools (EviHunter) to discover all the possible evidence an Android™ app generates in the forms of files in the local storage, SQLite databases, or data sent to remote third party server(s). Authors from the CSAFE worked with researchers from the National Institute of Standards and Technology (NIST) to establish a dictionary-like Android™ app-evidence database (CSAFE-AED) that includes apps themselves (Android™ Application Package (APK) code, version, and meta data, (e.g., the number of downloads at the time the APK was collected)) and all possible evidential data (type, location(s), and evidence format/syntax) that each app can likely generate and store on the device or remote servers.

This presentation will introduce the basics, challenges, and limitations of the current mobile device forensics practice. A detailed explanation about the usefulness and availability of information about the potential evidence types and their locations on Android™ mobile device or remote servers (e.g., /data/data/com.app.foo/shared_prefs/goo.xml → Visited URLs, Time) will also be provided. A demonstration will be given to explain how the usage of CSAFE-AED improves the overall mobile forensic analysis process via a set of carefully designed case works. In addition, this presentation will elaborate on the methodology and large-scale experimental evaluation of the approaches used to build CSAFE-AED. Overall, digital forensic investigators will learn how to take advantage of the CSAFE-AED database and search/recover potential evidence from the possible locations on and outside the mobile devices being investigated.

Based on the investigation of various app stores available globally, the number of various real-world apps has exceeded seven million so far. Commercial mobile device forensic tools, such as Cellebrite's Universal Forensic Extraction Device (UFED), support the profiles of approximately 6,000 apps and may not sufficiently support real-world mobile forensic case investigations. When digital forensic practitioners analyze mobile devices that had apps installed that are not currently supported by these tools, manual investigation has to be conducted over every single file extracted from the storage medium of the mobile device. Such manual casework may be error-prone and time-consuming. For example, a five-year-old Nexus® 7 tablet with 90 apps installed (both user and system space) can easily have approximately 20,000 files extracted from the device image. Performing manual forensic analysis on such a mobile device is arduous work, oftentimes simply infeasible to be accomplished within certain required time periods, which in turn may lead to more serious completeness and quality problems.

To tackle the challenges addressed above, a completely different approach was proposed and evaluated to provide much better coverage and precision guarantees. Both static and dynamic program analysis approaches were applied on analyzing/resolving evidentiary data, such as the file path and its corresponding evidence types. Currently, some existing work on privacy leakage problems are similar, but are not applicable for mobile app forensic analysis. To preserve the advantages (e.g., better time efficiency) provided by existing tools, a set of refined/improved automatic program analysis algorithms to analyze Android™ apps were implemented to create the CSAFE-AED. The CSAFE-AED also has the capability of handling new, updated versions of Android™ applications that are updated and published from time to time. The advantages of leveraging AED are: (1) mitigating false negatives caused by manual investigation; (2) improving the chances of identifying evidentiary data stored in specific format (e.g., custom encoding work); and (3) fast-tracking the evidentiary date of real-world apps.

The CSAFE-AED app forensic evidence database will be the first of its kind in terms of size (number of apps in the database) and accuracy/completeness of app analytical results. The CSAFE-AED database is expected to create fundamental and effective changes to current digital forensic practice, in particular mobile (Android™) device forensics.

Mobile Forensics, Android™ Apps, Digital Evidence