C32 Automated Standards-Based Normalization and Correlation of Mobile Device Evidence

Eoghan Casey, PhD, University of Lausanne, Lausanne, Vaud 1015, SWITZERLAND; Martina Reif, MS*, University of Lausanne, Lausanne, SWITZERLAND; Quentin Rossy, PhD, School of Criminal Justice, University of Lausanne, Lausanne 1015, SWITZERLAND*

Learning Overview: This presentation will provide attendees with a more efficient, reliable, standards-based approach to automatically normalize, combine, and correlate digital and multimedia evidence. This presentation provides a fit-for-purpose solution using the evolving Cyber-investigation Analysis Standard Expression (CASE) standard to represent information extracted using mobile device forensic tools and to combine this information into correlation analysis tools. This work concentrates on treating the following categories of information: (1) communications (Short Message Service [SMS], chat messages, email); (2) identifiers (telephone numbers, email addresses, social media accounts, Internet Provider [IP] addresses); (3) temporal indicators (timestamps, call durations); and (4) spatial indicators (Global Positioning System [GPS] coordinates, addresses).

Impact on the Forensic Science Community: This presentation will impact the forensic science community by: (1) saving time correlating and analyzing digital/multimedia evidence from mobile devices; (2) reducing risk of errors and omissions combining and correlating digital/multimedia evidence from mobile devices; (3) increasing completeness of forensic analysis of digital/multimedia evidence; and (4) tracking chain of evidence throughout the export and combination process to allow forensic analysts to track findings back to their origin.

Increasingly, mobile devices contain large amounts of digital evidence relevant to criminal investigations.[1] This digital evidence can be analyzed to make inferences about identities, locations, chronologies, and relationships between relevant entities to address critical questions in criminal investigations.[2] However, there are severe limitations in current capabilities to combine and correlate all available digital evidence. First, no single mobile device forensic tool can extract all types of digital evidence. Second, tools export data in different formats, without consideration for inter-operability. Third, it is necessary to combine the results from multiple tools to obtain comprehensive visibility across all digital evidence. Fourth, mobile device forensic tools do not maintain chain of evidence throughout the export and combination process, making it difficult to track forensic findings back to their source. Dealing with these problems, forensic practitioners reformat and combine information from different sources by hand, which is a laborious and time-consuming process that can result in errors and omissions.[3] For example, using spreadsheet software or database applications to import and format data from various sources can result in items such as date-time stamps being altered, entries not being imported, and other problems that negatively impact forensic analysis. There is a need for automated combination and correlation between datasets processed by mobile device forensic tools.[4]

The specific educational objectives of this presentation are to: (1) raise attendee awareness of current limitations in mobile forensic tools and how to address these limitations; (2) teach the attendees how digital/multimedia evidence extracted using mobile forensic tools is automatically translated into the CASE standard; (3) inform attendees about open source resources for implementing CASE in existing tools and systems; (4) demonstrate the value of importing normalized digital/multimedia evidence from mobile forensic tools into a platform for correlation and analysis; and (5) provide a roadmap of work to strengthen and expand CASE adoption across the digital forensic community.

This work addresses challenges in the way different tools represent extracted data, including missing information, only displaying one party in communications, and variations in the format of common information (e.g., timestamps, phone numbers).

By implementing the CASE standard, this solution does not require tool developers to alter their data model. Instead, it is necessary to translate their data model into a community-developed ontology for representing cyber-investigation information. This approach allows visualization tools, such as Cellebrite® and Analyst's Notebook®, to automatically (rapidly) import data from all sources into a cohesive and comprehensive picture to support selection, correlation, and analysis.

Reference(s):

1. McMillan, Jack E.R., Glisson, William B., Bromby, Michael. 2013. Investigating the Increase in Mobile Phone Evidence in Criminal Activities. In: *2013 46th Hawaii International Conference on System Sciences*, Wailea, HI, USA, January 2013, pp. 4900–4909. IEEE. DOI: 10.1109/HICSS.2013.366.
2. Casey, Eoghan; Ribaux, Olivier; Roux, Claude. 2018. The Kodak Syndrome: Risks and Opportunities Created by Decentralization of Forensic Capabilities. *Journal of Forensic Sciences*, Volume 64, Issue 1.
3. Casey, Eoghan; Barnum, Sean; Griffith, Ryan; Snyder, Jon; van Beek, Harm; Nelson, Alex. 2017. Advancing coordinated cyber-investigations and tool interoperability using a community developed specification language. *Digital Investigation*, Issue 22.
4. Reif, Martina. 2019. *L'implémentation de CASE aux extractions téléphoniques.* Master's Project, University of Lausanne.

Digital/Multimedia Evidence, Standards-Based Correlation, Chain of Evidence

*Presenting Author