# On the Development of Automated Forensic Analysis Methods for Mobile Devices

Panagiotis Andriotis[1], Theo Tryfonas[1], George Oikonomou[1], Shancang Li[1], Zacharias Tzermias[2], Konstantinos Xynos[3], Huw Read[3], and Vassilis Prevelakis[4]

[1] University of Bristol, MVB, Woodland Road, Clifton, Bristol BS8 1UB, UK,
[2] FORTH-Institute of Computer Science, N. Plastira 100, 70013, Heraklion, Greece,
[3] University of South Wales, Pontypridd, Wales, UK, CF37 1DL, UK,
[4] Technical University, Hans-Sommer-Street 66, 38106, Braunschweig, Germany

**Abstract.** We live in a connected world where mobile devices are used by humans as valuable tools. The use of mobile devices leaves traces that can be treasured assets for a forensic analyst. Our aim is to investigate methods and exercise techniques that will merge all these valuable information in a way that will be efficient for a forensic analyst, producing graphical representations of the underlying data structures. We are using a framework able to collect and merge data from various sources and employ algorithms from a wide range of interdisciplinary areas to automate post-incident forensic analysis on mobile devices.

**Keywords:** Steganalysis, Sentiment Analysis, SMS, Social Media

## 1 Introduction

The basic types of data we can retrieve during a forensic analysis on devices are text and images, the metadata for which are usually stored internally in SQLite databases [1]. A forensic investigation deals with the problem of merging all useful information, in order to provide evidence at a court of justice. In this project we aim to automate this process and decrease the analysis time using data mining methods to extract sentiment polarity from short messages. Also, we highlight connections and interactions between entities that exist in the Smartphone Ecosystem and in various social media communities, providing graphical representations that demonstrate the proximity of their relationships [3]. Finally, we propose a lightweight classification mechanism that distinguishes suspicious JPEG images that might exist in the device's internal memory [2].

## 2 Methodology

The data aggregation mechanism, called DEViSE, provides a platform where the data from various sources can be stored in a homogeneous format using XML files. All these information can be stored in a central database and therefore, used

upon the request of the visualization tools. For the social media module of our platform, we developed a crawler that can be enriched by data derived by mobile devices. Furthermore, we extended the functionality of a graph representation of interactions between entities by highlighting the 'closest friends' of the person under investigation. The short text messages can be further analysed to produce the Sentiment Timeline View and depict the emotional polarity between entities for a given timeframe. Our approach to this problem is the use of a bag-of-words schema that utilizes special features like the existence of emoticons and the lexicon's word valence evaluation. Finally, the automated system we propose is able to perform steganalysis on the JPEG images that exist in the internal memory of the mobile device, using our model for colour images derived by the empirical Benford's Law.

## 3 Results and Conclusion

Regarding the results derived from the JPEG images classification, our approach reaches hit rates of 70% - 100%, depending on the algorithm used to create the stego-carrier. The short text Sentiment Analysis module can correctly identify the emotional polarity of around 69% of the messages with a false positive rate reaching approximately an average of 25%. Finally the graphical representation of the entity linking, results to informative graphs which can be further enhanced by clustering algorithms that various visualization tools provide by evaluating measures like centrality. To conclude, we have developed an analysis automation platform able to perform specific tasks based on data collected from various sources. This system could be a helpful asset to the community of forensic analysts, but of course it cannot substitute their expert judgement to a court. It can instantly produce informative constructions derived from a wide data pool associated with the person under investigation, but it cannot act as a judgement tool in itself.

## Acknowledgement

This work has been supported by the European Union's Prevention of and Fight against Crime Programme "Illegal Use of Internet" ISEC 2010 Action Grants, grant ref. HOME/2010/ISEC/AG/INT-002 and the Systems Centre of the University of Bristol.

## References

1. Andriotis, P., Oikonomou, G., and Tryfonas, T.: Forensic Analysis of Wireless Networking Evidence of Android Smartphones. In: WIFS, pp. 109-114 (2012)
2. Andriotis, P., Oikonomou, G., and Tryfonas, T.: JPEG Steganography Detection with Benford's Law. Digital Investigation, 9(3), 246–257 (2013)
3. Andriotis, P., Tzermias, Z., Mparmpaki, A., Ioannidis, S., and Oikonomou, G.: Multilevel Visualization Using Enhanced Social Network Analysis with Smartphone Data. IJDCF, 5(4), 34–54 (2013)