The Next Generation of Digital Forensic Capability

N.L. Clarke[1,2], F. Li[1], M. Al-Fahdi[1], D. Joy[1], S. Alqahtany[1], G. Alotibi[1] & H. Mohammed[1]

[1]Centre for Security, Communications and Network Research, Plymouth University, United Kingdom
[2]Security Research Institute, Edith Cowan University, Western Australia

Abstract

Digital forensics has become increasingly important as both cyber and computer-assist crime increases. Whilst existing tools provide a range of functionality to enable investigators to apply forensic analyses (e.g. hashing, data carving, application-level parsing), a range of challenges exist which severely limit their application in the future. These challenges include: the volume of data, where investigators simply do not have the resource to inspect all data; cases that no longer involve one computer but a multitude (e.g. desktop, tablet, mobile, removable storage, online storage, games console, smart TV), where a need exists to inspect holistically rather than on an individual basis; the use of anti-forensics, that brings the reliability of evidence found in computer forensics into question and thus a need for more effective suspect-independent sources of evidence and the tools to analyse them (e.g. network forensics); and the introduction of cloud computing provides a paradigm shift where current forensic processes and procedures simply do not work.

This paper presents an overview and discussion of three research efforts being at the Digital Forensic Laboratory (DFL) within the Centre of Security, Communications and Network Research (CSCAN) at Plymouth University. Together they are focussed upon the creation of the next generation of digital forensic tools. The first seeks to reduce the investigator time and cognitive effort in analysing large (in terms of data and sources) and complex cases using artificial intelligence to automatically sort, identify and prioritise evidence from benign files. Preliminary analysis (of a single source) using unsupervised clustering shows that up to 93% of evidence can be correctly classified. The second project looks to develop a novel approach to network forensics through the identification of user actions based solely on network metadata (i.e. independent of end-to-end encryption). This enables an investigator, in a completely encrypted environment, to be able to appreciate not only the services that a user is using but also what the user is doing within the service (i.e. posting on Facebook, watching a video or using instant messenger). The approach is also taken to the next step where this information can be used to create behavioural profiles. These profiles are used to identify users. The ability to identify a user from network traffic results in a Network Forensic-Analysis Tool (NFAT) being able to filter based upon a user rather than an IP address. In most incidents, it is the user the investigator is interested in; however, in practice the IP is used to assume it links to an individual. In mobile computing this assumption does not hold true, so being able to filter and view a user's traffic is crucial. Experimental studies involving 27 users over a 2 month collection period provide a True Identification Rate of the best performing application of 87.4%. Compared with existing NFATs such as Wireshark and Xplico, this provides for a fast and efficient approach to identify user traffic and examine their actions. The third body work being undertaken is in the development of a cloud-based forensic framework. Focussing upon Infrastructure as a Service (IaaS) cloud model, the approach provides a robust framework for the acquisition and creation of forensic images of virtual appliances that comply to the well-accepted computer forensic procedures. Moreover, the approach achieves this will a huge reduction in data over existing research, with an increase in the granularity of the resulting image and in a cloud-service provider (CSP) independent fashion. Enabling the owner of the data to remain in complete control.

Drawing these research efforts, the paper continues to discuss the existing challenges and proposes several suggestions as to where future research efforts should be placed.