Chapter 14
A Framework for the Forensic Analysis of User Interaction with Social Media

John Haggerty
School of Computing, Science and Engineering, University of Salford, Manchester, UK

Sheryllynne Haggerty
School of Humanities, University of Nottingham, Nottingham, UK

Mark C. Casson
Henley Business School, University of Reading, Reading, UK

Mark J. Taylor
School of Computing and Mathematical Sciences, Liverpool John Moores University, Liverpool, UK

ABSTRACT

The increasing use of social media, applications or platforms that allow users to interact online, ensures that this environment will provide a useful source of evidence for the forensics examiner. Current tools for the examination of digital evidence find this data problematic as they are not designed for the collection and analysis of online data. Therefore, this paper presents a framework for the forensic analysis of user interaction with social media. In particular, it presents an inter-disciplinary approach for the quantitative analysis of user engagement to identify relational and temporal dimensions of evidence relevant to an investigation. This framework enables the analysis of large data sets from which a (much smaller) group of individuals of interest can be identified. In this way, it may be used to support the identification of individuals who might be 'instigators' of a criminal event orchestrated via social media, or a means of potentially identifying those who might be involved in the 'peaks' of activity. In order to demonstrate the applicability of the framework, this paper applies it to a case study of actors posting to a social media Web site.

DOI: 10.4018/978-1-4666-4006-1.ch014

INTRODUCTION

Social media, applications or platforms that allow users to interact online, has risen in prominence in recent years. Sites that are widely accessed range from those whereby users can post comments in near-real time to social networking services. For example, Twitter, the micro-blogging site, has a reputed 500 million users (Dugan, 2012) with on average more than 2,200 'tweets' (messages) sent every second (Twitter Engineering Blog, 2011). The social networking service, Facebook, has over 900 million active monthly users of which over 500 million are active daily (to 31 March 2012). Moreover, more than 300 million photos are uploaded to the service every day (Facebook, 2012). Interaction with social media will therefore provide a useful source of evidence during a digital forensic investigation. For example, a number of people used social media to encourage rioting, criminal damage and theft during civil unrest in the UK during August 2011 and have since been sentenced to significant terms in prison under section 44 of the Serious Crime Act 2007 ("Southampton Facebook riot messages," 2012).

Current forensics tools find this environment problematic for two reasons. First, they focus on the extraction and analysis of evidence from storage media rather user interaction with online environments. Second, the forensic examiner might not be able to gain access to the storage media itself, such as the Web server hosting the content, as it may be located beyond their jurisdictional powers. However, an advantage of social media analysis is that data is made publicly available as the actors engage with this environment in full view of active and passive users through their online posts. This is in contrast to user interaction formed through other digital communications media, such as email or Internet telephony, where analysts must adhere to relevant laws to protect privacy.

This paper posits a novel inter-disciplinary framework for the quantitative analysis of active user interaction with social media in order to identify actors of potential interest to an investigation. In particular, it uses temporal social network and regression analysis to support the identification of individuals who might be 'instigators' in a criminal event orchestrated via social media, or a means of potentially identifying those who might be involved in the 'peaks' of activity. Unlike previous approaches to social network analysis in digital forensics whereby the relationships *between* actors are identified and analysed, this approach focuses on the relationship of actors *with* the social media service itself to identify those actors of significance over time (whether minutes, hours, days or weeks, etc.). Moreover, it identifies those actors that have a statistically significant relationship with the social media under investigation to triage evidence. In this way, the approach can enable analysis of large data sets from which a (much smaller) group of individuals of interest can be identified.

This paper is organised as follows. In the next section, we discuss approaches to social network and social media analysis relevant to a digital forensics investigation. Following this, we posit the framework for the analysis of social media interaction using temporal social network and regression analysis. We then demonstrate the applicability of the framework through a case study and discussion of results. Finally, we make our conclusions and discuss further work.

SOCIAL NETWORK AND SOCIAL MEDIA ANALYSIS

Computer forensics tools, such as EnCase (Guidance, 2012) and the Forensic Toolkit (Access Data, 2012), are used by examiners to recreate files and data from a suspect's computer. An