Proceedings of the Seventh International
Workshop on Digital Forensics & Incident Analysis (WDFIA 2012)

Forensic Analysis of User Interaction with Social Media: A Methodology

J. Haggerty[1], M.C. Casson[2], S. Haggerty[3] and M.J. Taylor[4]

[1]School of Computing, Science & Engineering, University of Salford, Greater Manchester, M5 4WT
[2]Henley Business School, University of Reading, Reading, RG6 6UD
[3]School of Humanities, University of Nottingham, Nottingham, NG7 2RD
[4]School of Computing & Mathematical Sciences, Liverpool John Moores University, Liverpool, L3 3AF
e-mail : J.Haggerty@salford.ac.uk; m.c.casson@reading.ac.uk; sheryllynne.haggerty@nottingham.ac.uk; M.J.Taylor@ljmu.ac.uk

Abstract

The increasing use of social media, whereby users interact online, ensures that it will provide a useful source of evidence for the forensics examiner. Due to the dynamic nature of this environment, current approaches for its analysis are not without their limitations. This paper posits a novel inter-disciplinary methodology for the forensic analysis of user interaction with social media. In particular, it presents an approach for the quantitative analysis of user engagement to identify relational and temporal dimensions of evidence that will be relevant to an investigation. In this way, it may be used 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. In order to demonstrate the applicability of this methodology, this paper applies it to a case study of users posting to a social media Web site.

Keywords

Digital forensics, social media, social network analysis, regression analysis

1 Introduction

Social media, whereby users interact online, plays an increasingly important role in our lives due to accessibility from heterogeneous computing devices. Sites that are widely used range from those where users can post comments in near-real time to social networking services. For example, Twitter, the micro-blogging site, has 100 million users with over 230 million 'tweets' sent each day (Business Insider, 2011). Facebook has 800 million users, of which more than 350 million access the service through mobile devices (Facebook, 2012). Interaction with social media will therefore provide a useful source of digital evidence during an 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 (see for example BBC, 2011). Current

forensics tools find this environment problematic as they focus on the extraction of evidence from storage media.

This paper posits a novel inter-disciplinary approach for the forensic analysis of social media. In particular, it presents a methodology for the quantitative examination of social media users pertinent to a forensics investigation through temporal social network and regression analysis. In this way, it may be used 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 *with* the social media service itself to identify those actors of significance over time. Therefore, the proposed methodology enhances a forensics investigation by analysing the relational *and* temporal dimensions of actors' interactions with social media. Moreover, it identifies those actors that have a statistically significant relationship with the social media under investigation to triage evidence.

This paper is organised as follows. Section 2 discusses related work. Section 3 posits the methodology and describes two models for the analysis of social media. Section 4 presents the results of applying the methodology to a case study. Finally, we make our conclusions in section 5 and discuss further work.

2 Related work

Social network analysis has been proposed as an aid for digital investigations, especially in those that involve the interaction of actors online. For example, Haggerty *et al* (2011) use the Enron email corpus as a case study to propose a method for the triage and analysis of email data. Dellutri *et al* (2009) focus on the identification of social networks to reconstruct a user's profile by combining a smartphone's data with social relationships found on the Internet. However, these approaches have in common that they focus on the details of extracting and identifying evidence.

Current tools for social network analysis that may be used for digital forensics investigations involving social media often provide static visualisations. Applications such as *Pajek* (Vlado, 2012) visualise network information via the connection of vertices through arcs and edges. In addition, weighting may be applied to network edges to represent strength of ties (Perer and Schneiderman, 2009). The common issue with these static social network visualisations is that they do not represent temporal changes in the network, including an individual actor's engagement over time. Thus, Falkowski *et al* (2006) suggest an approach for analysis of subgroup evolution in social networks. This approach uses a number of views to facilitate analysis and displays the network in a graph, such as those used in static approaches, but laid out along a temporal plain. Hu and Gong (2010) present a visualisation of individuals' spatial-temporal social networks through three-dimensional graphs. Belingerio *et al* (2010) use three-dimensional graphs combined with hierarchical

trees to identify eras in social networks. These approaches have in common that they centre on relationships between actors, i.e. they assume direct communications between them. However, in many social media environments data is posted on the site for all to see, rather than direct communication. Therefore, the relationship is often one of actor-to-Web-site rather than actor-to-actor. Thus, other approaches must be used for forensic analysis of social media.

Recent work has identified regression analysis as an approach to investigate user interaction with social media. For example, Shwu-Min Horng (2010) uses multiple regressions for the analysis of connections between users' behaviour on social network sites and data collected from Google Analytics. Kumar and Saha (2009) use regression analysis with other techniques for the data mining of Web blog entries to detect user sentiments. Beck (2011) uses logistic regression analysis to detect spam and phishing attacks over the Twitter network. However, these approaches have in common that they are not focused on the requirements of a forensics investigation. This paper therefore posits a methodology to incorporate regression analysis into a forensics investigation of social media.

3 Forensic analysis of user interaction with social media

Social networks formed through social media are dynamic, evolve over time and react to endogenous and exogenous events suggesting relational *and* temporal dimensions of this type of media. They therefore not only provide evidence of user engagement but also an indication of how the network has developed over time. Of interest to the forensics examiner investigating social media would be the following:

- Who are the key actors identified through the social media?
- What is an actor's relationship with the social media under investigation?
- What statements can be asserted about an actor's social media usage?
- Are there key periods of activity and relationships providing evidence?
- Are there specific actors that drive the narrative on the social media?
- Is there a pattern of interaction that may be identified?
- Network reactions to endogenous and/or exogenous events?

It should be noted that actors who post on social media have varying levels of relationship with the other actors in the network(s) observed. In some cases, the actors will form personal relationships, whereas in others, they do not. What they do have in common is that they have a relationship *with* the social media itself. Therefore, current tools and techniques used in the analysis of social networks, for example in email communications, do not fully meet the requirements for the analysis of social media. Moreover, data is made publicly available by the actors actively engaging with the social media. This is contrast to social networks formed through other communications media, such as email, whereby forensic analysts must adhere to relevant laws to protect privacy.

As discussed in the previous section, the common issue with static social network visualisations is that obviously they do not represent temporal changes in the

network. They are therefore limited in answering the questions above if used in isolation. Temporal social network analysis aims to answer the needs of forensics examiners by analysing change over time by highlighting 'real' relationships; actors in contact *at* a particular point in time and shown *over* a period of time. It aims to provide an interactive visualisation of time-varying (social) network data for examiners to interact with. This interaction can provide new questions for the examiner. Importantly, it identifies macro trends in networks to meet the shortcomings of static social network analysis tools.

Two models are posited which examine different aspects of network behaviour and identify statistically significant actors. Model 1 identifies the actors who most regularly interact with the social media and the periods that experience the most user interaction. This has the advantage that the results provide information that can be used to test whether the variations between periods of time and between actors are statistically significant or not, i.e. to determine whether it is likely that they were produced purely by random fluctuations or whether systematic factors, such as endogenous and exogenous events relevant to the investigation, are at work. This is achieved through a panel regression. Panel data analysis is an increasingly popular form of longitudinal data analysis among social science researchers. A panel is a cross-section or group of individuals who are surveyed periodically over a given time span.

Model 1 assumes that there are N individuals indexed $i = 1,\ldots, N$ and T times indexed $t = 1,\ldots, T$. A matrix Y (known as the interaction matrix) is created with N columns and T rows, and therefore with NT cells. The entries in the cells are binary: one if a particular individual interacts in a particular period and zero otherwise. All actors who belong to the network at any time should normally be included. A panel regression model is specified in which the probability that individual i interacts in period t is

$$\text{Prob}(y_{it} = 1) = c + a_t + b_i + u_{it} \quad (1)$$

where y_{it} is the binary element in the ith column and tth row, c is a constant, measuring the relative frequency with which an actor on average interacts with the social media, a_t is a period-specific factor reflecting the above-average popularity of the site at time t, and b_i is an actor-specific factor reflecting the disposition of the ith actor to interact more frequently than average. The variable u_{it} represents an unobservable random disturbance which can be either positive or negative; it is assumed to have zero mean and a constant variance, independent of i and t.

Model 2 identifies particular time patterns in interaction. It tests for the existence of a deterministic linear time trend by which interaction either increases or decreases over time at a constant absolute rate. It also tests for persistence in interaction patterns, whereby interaction at the immediately previous period increases the probability of interaction in the current period. It also tests for delayed persistence, whereby interaction in the last-but-one period increases the probability of interaction in the current period. Testing for persistence also provides a test for alternation, whereby

interaction in the previous period (or the period before that) discourages interaction in the current period and non-interaction in the previous period (or the period before that) increases it.

A disadvantage of model 1 is that although it contains a large number of time dummies, it does not directly address the question of whether there are systematic time patterns in interaction. This can be remedied by replacing the time dummies with more meaningful variables. These include a linear time trend, t, whose values range from 0 (at the start of the period) to $T - 1$ at the end, as well as lagged values of the dependent variable.

$$\text{Prob}(y_{it} = 1) = c + a_0 t + a_1 y_{it-1} + a_2 y_{it-2} + b_i + u_{it} \quad (2)$$

A time trend variable, t, is formed by stacking N sequences on top of each other. A variable y_{-1}, representing a single lag in the dependent variable is generated by taking the values of the dependent variable for each individual for periods from 0 to $T - 1$, adding an empty cell at the beginning, and then stacking them as before. A double lag variable, y_{-2}, is captured by taking the values of the dependent variable for each individual for periods 0 to $T - 2$, adding two empty cells at the beginning, and then stacking them in the same way. The estimated regression is

$$y = \text{c} + \Sigma_t a_t w_t + a_0 t + a_1 y_{-1} + a_2 y_{-2} + \Sigma_i b_i x_i \quad (3.1)$$

where for some $i = j$,

$$b_j = 0; \quad (3.2)$$

The parameters a_0, a_1, a_2, b_i are estimators of the parameters of the model represented by equation (2).

The social media interpretation of the results obtained from employing these models to actors engaging *with*, rather than passively *following*, social media is as follows. Positive results identify actors that engage with the network during key periods and when the network is more active. This could occur for a number of reasons, including: actors interact with the social media because something is affecting the network requiring the whole community to react; actors joining in an upswell of popularity of the social media; are online community leaders whose virtual presence is required. Positive results suggest that these actors are independent of the network but will interact when it is in their interest to do so. Negative results identify actors who engage with the network when interaction is less popular or re-engage even when interaction is low. This could be for a number of reasons, including: actors new to this particular social media and wish to engage with others in the online community; interact with the network in a decline from a popular period (actors who heard the social media was active but interact as engagement decreases); actors who require the status that the social media interaction provides; leading actors of the network itself. Negative results suggest that these actors are more dependent on the network and so make every effort to interact.

4 Case study and results

To demonstrate the applicability of the proposed methodology, this section presents a case study of users posting comments on a social media Web site. Depending on the investigation, the time periods will be hours for fast-moving events, such as riots, or days for longer-term interactions, such as suspects involved in the dissemination of indecent images of children. For ease of reference, we assume hours in this case study. As the aim of this paper is to posit the quantitative methodology, qualitative analysis is not discussed. However, it is recognised that qualitative analysis, such as reading the posts made by users, would be an important element of the overall investigation.

Figure 1 applies the network data to a temporal social network analysis tool, *Matrixify* (Haggerty and Haggerty, 2011). This tool visualises the network data as a two-dimensional matrix to show individual user interaction with the network over time. It also provides a number of menus to provide alternative views and analysis of the network, such as graph layouts, network statistics and qualitative analysis tools. The aim of using temporal analysis is to visually explore the data in order to provide a more nuanced and sophisticated overview of the network, assess actor interaction and identify change over time. It does not aim to *answer* questions; but to *raise* questions around the data not evident in other forms. These questions can then, for example, guide the forensic examiner to relevant sources of further analysis or challenge existing hypotheses and theories.

Time periods are placed on the X axis and actors that engaged with the network during this time are represented on the Y axis by name. Their involvement and role in the network is indicated by coloured dots. Thus, individual engagement with the network is represented horizontally and user engagement on a period-by-period basis is viewed vertically. Ten-hour markers are included to aid reference. In addition, the analyst can specify colours within the tool. In this case, normal users are coloured blue, whilst those users with a higher level of access, for example, moderators or administrators are coloured red. This aids analysis of actor engagement by role.

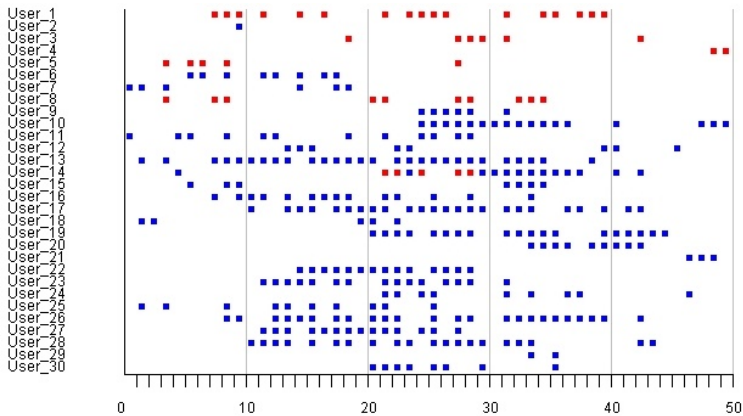

Figure 1: Temporal view of user interaction over time

Figure 1 highlights the trends in interaction whereby engagement is low initially, it then sees a period of higher interaction between 15-35 hours, before engagement decreases. In addition, it raises a number of questions about network engagement that may be useful during an investigation. First, why are some actors active for only a short time and others active for longer? Do some not find the social media useful, and if so, why? Do they interact elsewhere? Conversely, do the long-term actors dominate the network, and do they constitute an online clique? Second, why is the network denser in particular periods (in this case, the 10s, 20s and 30s)? Why does this change significantly in the mid-30s? Are exogenous or endogenous events driving this? Third, what are the causes for the relative lack of actor involvement in certain periods? In the case study this occurs in the 0s and 40s. Why are these periods particularly lacking in long-term actor involvement? Are actors interacting with other social media? Fourth, how does shifting actor engagement impact on access to information or the reaction to exogenous events, particularly those relevant to the forensics investigation? Fifth, is there a long-term survival of the network(s) or has it a natural life-cycle?

Of interest in the results for statistically significant actors and for network analysis is whether an actor has a positive or negative relationship with the social media. Therefore, models 1 and 2 are applied to the social media engagement data. As discussed in the previous section, a positive result indicates statistically significant actors who interact with the social media during key periods and when it is increasingly or decreasingly active. A negative result indicates those actors who engage with the social media even when user engagement is low. This overview by 10-hour periods enables the investigator to assert statistical trends in user interaction and changes in the social media usage. Figure 2 illustrates the total number of actors within a subset and the numbers of positive and negative statistically significant actors.

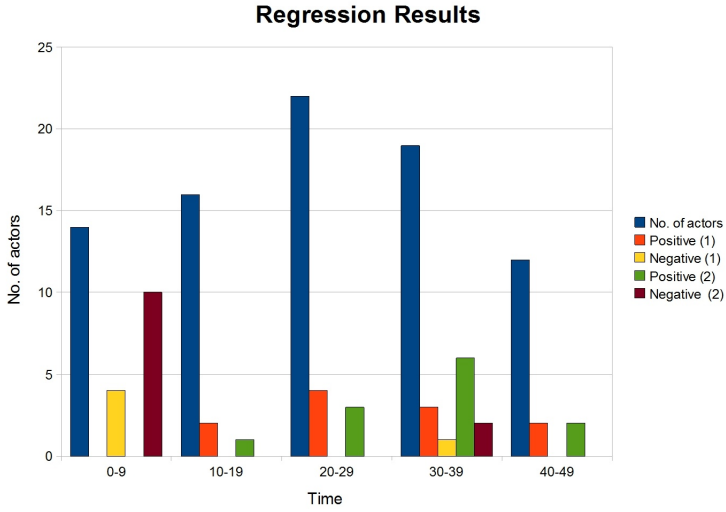

Figure 2: Results of the regression analysis

As figure 2 demonstrates, relationships with the social media change over time, as indeed, do the number of actors involved in engaging with the social media. During 0-9 hours, statistically significant actors have a negative relationship with the social media suggesting a core of actors who interact with the social media even at less popular times. This is reversed in the next two periods, reflecting an upswell of user engagement. This would direct the forensic examiner in their analysis as it indicates a change in user interaction and may provide an indication that endogenous or exogenous network events relevant to the investigation are occurring. Therefore, they will be interested in the qualitative analysis of the posts being made at this time. The period 30-39 indicates another change in relationships with the social media where interaction occurs at both popular and less popular times. Finally, the network declines with users sporadically interacting with the network.

In order to direct (and triage) the qualitative examination, individual actors highlighted by the regression analysis are clustered by period to identify those that are actively reacting to endogenous and exogenous events. Figure 3 illustrates the relational view by 10-hour periods. These actors are clustered by their relationship with the social media, where non-statistically significant actors are on the left, positive actors are top right and negative actors are bottom right. This provides a static network representation and is produced in the *Matrixify* tool. If an actor posts in the same time period as another, a relationship (represented by a line) is formed. This can also be used by the forensics examiner to identify actors that form relationships through the social media, for example, reacting to posts by other actors or to network events (although relational analysis is not the focus of this paper).

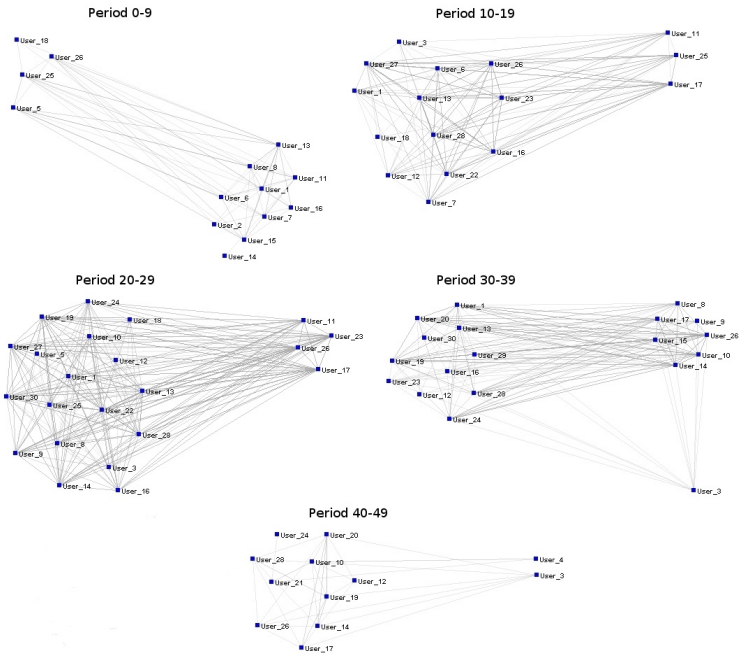

Figure 3: Network cluster analysis identifying key actors

As illustrated in figure 3, the majority of actors who post in the period 0-9 have a negative relationship with the social media. These may be actors who have a long-term relationship with the social media. This changes in the next two periods where actors of note are positive and fewer in numbers. Of interest to the forensic examiner is User_11 who is identified in all three periods. In addition, User-17 is highlighted in periods 10-29. User_26 is highlighted in periods 20-39. Finally, in period 40-49, User_3 and User_4 are highlighted as interaction with the social media decreases. These actors would be of note as they are contributing to the evidentiary evolution of the network, fuelling or reacting to endogenous and exogenous events. This would direct the forensics examiner to potential posts requiring qualitative analysis.

This section has presented a case study to demonstrate the applicability of the proposed methodology for the analysis of relational and temporal dimensions of evidence in social media. The temporal social network analysis highlights macro trends in user interaction and raises further questions regarding user engagement. The regression analysis identifies the nature of actor relationships with the social media over time. Finally, the cluster analysis highlights those individual actors that have a statistically significant relationship with the social media. In this way, the methodology posited in this paper supports 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. Moreover, the methodology could be utilised to triage potentially large data sets that the forensics examiner may encounter prior to a qualitative analysis.

5 Conclusions and further work

The increasing use of social media, whereby actors interact with Web sites, ensures that it will provide a useful source of evidence for the forensics examiner. Currently, an examiner will conduct a time-consuming qualitative analysis of social media relevant to their investigation. Current tools, such as those used for social network analysis, do not fully meet the requirements of this type of investigation.

This paper presents a novel inter-disciplinary approach for the forensic analysis of social media. In particular, it posits a methodology for the quantitative analysis of social media users to identify the relational and temporal dimensions of evidence that will be relevant to an investigation. Given the potentially large data sets in this environment, the proposed methodology may also be used to triage data. In order to demonstrate the proposed methodology, this paper has applied it to a case study of users interacting through social media. In this way, the forensics examiner is able to gain a more sophisticated and nuanced view of user interaction and is able to assert those actors that have a statistically significant relationship with the social media under investigation. Future work will further develop the techniques presented in this paper. In particular, it aims to incorporate tools for automated qualitative analysis.

6 References

BBC (2011), http://www.bbc.co.uk/news/uk-england-hereford-worcester-16185152. (Accessed 7 February, 2012)

Beck, T. (2011), "Analyzing Tweets to Identify Malicious Messages", *Proceedings of the IEEE International Conference on Electro/Information Technology*, Minnesota, USA, 2011, pp. 1-5.

Berlingerio, M., Coscia, M., Giannotti, F., Monreale, A. and Pedreschi, D. (2010), "Towards Discovery of Eras in Social Networks", *Proceedings of the M3SN 2010 Workshop, in conjunction with ICDE2010*, California, USA, 2010, pp. 278-281.

Business Insider (2011), http://articles.businessinsider.com/2011-09-13/tech/30148448_1_dick-costolo-twitter-s-ceo-tweets. (Accessed 7 February, 2012)

Dellutri, F., Laura, L., Ottaviani, V. and Italiano, G.F. (2009), "Extracting Social Networks from Seized Smartphones and Web Data", *Proceedings of the 1st International Workshop on Information Forensics and Security*, London, UK, 2009, pp. 101-105.

Facebook Statistics (2012), http://www.facebook.com/press/info.php?statistics. (Accessed 9 January, 2012)

Falkowski, T., Bartelheimer, J. and Spiliopoulou, M. (2006), "Mining and Visualizing the Evolution of Subgroups in Social Networks", *Proceedings of the International Conference on Web Intelligence*, Hong Kong, 2006, pp. 52-58.

Haggerty, J. and Haggerty, S. (2011), "Temporal Social Network Analysis for Historians: A Case Study", *Proceedings of the International Conference on Visualization Theory and Applications (IVAPP 2011)*, Algarve, Portugal, 2011, pp. 207 - 217.

Haggerty, J., Karran, A.J., Lamb, D.J. and Taylor, M.J. (2011), "A Framework for the Forensic Investigation of Unstructured Email Relationship Data", *International Journal of Digital Crime and Forensics*, Volume 3 Number 3, September 2011, pp. 1-18.

Hu, B. and Gong, J. (2010), "Modeling Individual-Based Social Network with Spatial-Temporal Information", *Proceedings of the International Conference on Management and Service Science*, Wuhan, China, 2010, pp. 1-4.

Kumar Pal, J. and Saha, A. (2010), "Identifying Themes in Social Media and Detecting Sentiments", *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, Odense, Denmark, 2010, pp. 452-457.

Perer, A. and Schneiderman, B. (2009), "Integrating Statistics and Visualization for Exploratory Power: From Long-Term Case Studies to Design Guidelines", *IEEE Computer Graphics and Applications,* May/June, 2009, pp. 39-51.

Shwu-Min Horng (2010), "Analysis of Users Behavior on Web 2.0 Social Network Sites: An Empirical Study", *Proceedings of the 7th International Conference on Information Technology: New Generations*, Nevada, USA, 2010, pp. 454-459.

Vlado, A. (2012), http://vlado.fmf.uni-lj.si/pub/networks/pajek/. (Accessed 7 February, 2012)