Online Social Networks As Supporting Evidence: A Digital Forensic Investigation Model and Its Application Design

Norulzahrah Mohd Zainudin, Madjid Merabti, David Llewellyn-Jones
School of Computing and Mathematical Sciences
Liverpool John Moores University
Liverpool, United Kingdom
n.mohd-zainudin@2008.ljmu.ac.uk, {M.Merabti, D.Llewellyn-Jones}@ljmu.ac.uk

Abstract—The growth of online social networks has encouraged new ways of communicating and sharing information and is used regularly by millions of people; it has also resulted in an increase in its use for significant criminal activities and perpetrators are becoming increasingly sophisticated in their attempt to use technology in order to evade detection and perform criminal acts. Hence a systematic model for forensic investigation of online social networks is required in order to obtain optimum results from the networks’ investigation. We have reviewed the existing literature of digital forensic investigation models and frameworks, most have quite similar approaches, and some of the models are generic which do not focus on the purpose of the investigation. In addition, there is no standard and consistent model, only sets of procedures and tools, thus many digital crime investigations are performed without proper guidelines. Moreover, there is no model built specifically for online social networks but in contrast digital crimes related to them are growing rapidly. To address these challenges, we have developed a standard model of digital forensic investigation for online social networks in this research. This model incorporates the existing traditional frameworks, allowing us to compile a comprehensive digital forensic investigation model specifically for the networks. In order to implement the model in online social networks investigation, we will design its application that will be implemented to investigate, document and report important information in the networks without having to search manually.

Keywords—computer security; digital forensic; online social networks

I. INTRODUCTION

Recent years have seen a massive increase in the number of online social networks such as MySpace, Facebook, Twitter and Friendster which facilitate a high degree of user personalisation and user intercommunication. Online social networks are defined as web-based services that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system [1]. Criminals are becoming increasingly sophisticated, attempting to use technology in order to evade detection and perform criminal acts. This happens in virtual worlds using computers as a communication medium allowing online transactions and giving fraudsters new methods for attacking systems [2]. For this reason, it emerges of value to present a systematic approach that can be used by digital forensics investigators attempting to resolve such types of network-based cyber crime, to help ensure that any digital evidence recovered can be used in a court of law.

Computer forensics is used as an investigative tool in order to allow the investigator to determine what has occurred, when it occurred, where it occurred, why it might have occurred and hopefully who is responsible [3]. All of this information is required to ensure that there is sufficient evidence to prosecute criminals. The main challenges among the process of computer forensics include performing the analysis and reporting the results to ensure that the evidence is consistent and reliable for prosecution of criminals in a court of law.

In this paper, we present a comprehensive digital forensic investigation model specifically for online social networks. The model is separated into two environments which are physical and digital environments that reflect the scene of investigation. The novelty of this work is therefore the development of this model, along with development of a tool that will support online social networks investigation and analysis process.

The organisation of this paper is as follows. Section II explains previous digital forensic investigation model. In Section III we present our proposed model and Section IV
II. PREVIOUS DIGITAL FORENSIC INVESTIGATION MODEL

There are various number of authors that have proposed digital forensic investigation models and frameworks for example Kruse and Heiser [4], The U.S. Department of Justice [5], Carrier and Spafford [6], Lee [7], National Institute of Justice [8], Reith [9], Ciardhuáin [10], Leong [11] and Yong-Dal [12] and Cohen [13]. In recent years, as online social networks have developed and used widely, a number of applications have been developed [14], [15], mainly is for visualising the networks and all of them can be applied only on single network.

From the existing models mentioned above, a few issues can be raised. Firstly, the models proposed are basically developments of earlier models and most have quite similar approaches. Secondly, some of the models are generic and do not focus on the purpose of the investigation [16]. Obviously there is no standard and consistent model, only sets of procedures and tools, thus many digital crime investigations are performed without proper guidelines. Moreover, there is no model built specifically for online social networks but in contrast digital crimes related to these networks are growing rapidly [17-20]. This aspect will be discussed thoroughly in the next section.

III. THE PROPOSED MODEL

In Section II we have reviewed the existing literature of digital forensic investigation models and frameworks. We have found that although there are a number of digital forensic investigation models already developed, most have quite similar approaches, there is no standard and consistent model, only sets of procedures and tools, thus many digital criminal investigations are performed without proper guidelines. Moreover, there is no model built specifically for online social networks. In contrast digital crimes related to these networks are rapidly growing in number.

Our aim is to develop a specific model for investigation in online social networks and then we will develop a prototype that reflects the forensic investigation process in online social networks based on the model that has been developed before. Fig. 1 shows the model of digital forensic investigation for online social networks that we have proposed.

The proposed model comprises the whole process of online social network investigation. Hence we have divided the whole process of investigation into two environments.

The physical environment consists of activities that are carried out before the investigation. These are preliminary activities including notification from the enforcement body, planning of how to conduct the investigation and also surveying of any physical crime scene and evidence present. After these activities have been completed, investigators will proceed to the digital environment where they will carry out investigation and analysis of online social networks using an application prototype that will be developed. The next activity will shift back to the physical environment where all the evaluation process takes place.

The following list describes all activities in detail. The explanations of each process are as follows.

A. Preliminary

This is the first process in the model which consists of:

1) Acknowledgment: This is the first step of an online social network forensic investigation where a case or an audit is requested from an external organisation such as the police, customs, or a company. The process consists of establishing details of an event from the organisation and what they expect from the investigation. There is no technical component in this process.

2) Construction: After all the details are collected from the organisation, a thorough plan must be constructed. Planning will include operations, infrastructure, and authorization from relevant people/organizations. The difference between our construction process and previous models is in the infrastructure component, where the investigator needs to determine the type of online social network to be investigated, determine if there is any data including the suspect or victim’s profile or any other information gathered from the network that can be used in the next
process of investigation. Besides, this activity will carry out any relevant issues such as checking whether the investigation needs authorisation or any special equipment required in the investigation process. This process also includes authorisation if required from the online social network’s authority which will require a process of application for permission to investigate confidential information in that particular online social network.

3) Notification: This activity will depend on the investigation. For digital crime investigations this activity will be waived. For audit purposes, on the other hand, the audited person will be notified.

4) Survey: There is only one model that was developed by Carrier and Spafford [6] involving a survey of the physical crime scene and the identification of pieces of physical evidence. In our proposed model, besides carrying out an investigation of physical issues, the survey also involves study of the machine used and also the social network itself.

B. Investigation

1) Identification: This activity will be carried out by implementing the prototype to be developed. First we will identify any evidence or supporting information that might be available in an online social network. For example the name of a suspect and a victim will be given to enable us to conduct a thorough investigation of a case. For the case of auditing, we need to identify the person that will be audited.

2) Searching: Based on the relevant data gathered from the investigation process, we will run a thorough search that enables us to discover relevant data automatically. There are a large number of different types of data that can be collected and used as evidence or supporting information that might be extracted from an online social network, as follows:

3) Filtering: The filtering activity will scale down and focus the investigation on relevant information and discard any irrelevant information.

4) Capturing: Information collected through filtering will be captured in the best way to ensure the integrity of the data is sustained. The data itself will be analysed in the next process.

C. Analysis

A thorough analysis will be carried out based on the information collected from the previous activities. This activity will be supported by a module in the prototype to be developed.

1) Hypothesis: This activity consists of developing a hypothesis for the case to support any discovered evidence.

2) Reporting: The reporting activity will involve documenting the analyzed data and evidence gathered from the previous process, as well generating a detailed report of a suspect (or audited person) and others related to the case.

D. Evaluation

1) Presentation: The report that has been prepared in the previous activity will be presented to the relevant people. For example, if a police case is executed, the report will be presented to the jury. If it is a company investigation, the report might be presented to the company management.

2) Justification: In this activity, investigators will have to rationalize the validity of the evidence and will need to defend it against any doubt or challenge.

3) Review: If the evidence presented has any reasonable doubt, the investigation will be reviewed. The investigators will decide whether they should revert to one of the previous activities in order to discover more evidence. Otherwise, this activity can be waived.

IV. REQUIREMENTS FOR IMPLEMENTATION OF THE MODEL

In Section III we have discussed thoroughly about the model that we have designed. The model is divided into two environments: the physical and digital environments. We have also considered how the prototype can be developed to execute the activities in a digital environment. While in the previous section we have discussed about online social networks related applications that we have examined and discussed about the features that can be applied in developing our tool prototype. Based on those discussions, we will develop the prototype that reflects the process in our model and the prototype must be able to fulfill our requirements in order to produce an efficient tool for online social networks forensic investigation.

The prototype that will be developed will have several functions. In order to accomplish the crucial process of a forensic investigation in online social networks, the application prototype should have these functionalities:

1) Auto generating: The application prototype should be able to generate data based on queries given by examiners, thereafter it will do the rest of the process in searching, analysis and reporting of a particular examination. Hence, there will be very minimal human involvement in application of the prototype.
2) **Ability to search and filter data**: A technique to search data automatically will be developed according to specific conditions demanded by an examiner. Subsequently the searched data will be filtered in order to discover relevant data from the searching process.

3) **Ability to report comprehensively**: The prototype that will be developed should be able to create a report based on the previous process and will provide significant information from the investigation.

4) **Ability to provide a time-efficient prototype**: We will concentrate on techniques able to fulfil steps within the digital forensic investigation in online social networks that have suitable complexity. The aim is to ensure they can be used over realistic timescales.

5) **Ability to run and perform analysis of multiple searches of an individual’s online social network accounts**: The prototype will be able to search and analyse different networks of an individual to increase the quantity of any supporting information that can be gathered.

V. DESIGN OF THE APPLICATION PROTOTYPE

In this section we will discuss on designing the application prototype that will be a tool to represent the processes in digital environment as described in our model. Basically we will divide our prototype into two modules that are investigation and analysis modules. Each module will follow the processes in the model.

A. **Investigation Module**

This module will focus on investigation in online social networks. Section III has described the processes involved in this module which are identification, searching, filtering, and capturing. Fig. 2 illustrates the flowchart that can clarify the whole process in this module. As we can see the flowchart corresponds to the investigation activity in our proposed model as well as fulfils the requirements as explained in Section IV. The application will begin with options for the online social network(s) to be investigated. The user can opt for more than one networks to be investigated hence more clues can be gathered. Then the user needs to answer a list of queries, which can be answered based on the information given by the enforcement team that requested for this investigation or the physical investigation that was carried out in the previous activity. Fig. 3 shows in detail of the query process in the application.

The querying process starts with the name(s) to be investigated if there is any. The user will enter the name(s) and another query about any particular date or any range of date that needs to be searched will appear. Next the user will be asked whether there is any more data that he/she wants to enter that can aid the searching process. Should the user has more data, a list of options will be displayed and this process can be repeated again based on the options provided. After this process is completed, the auto-searching process will be carried out. The application will filter all relevant information based on the data entered previously and this information will be captured and stored in a data capture device.

![Flowchart of investigation module](image)

**Figure 2. Flowchart of investigation module**

B. **Analysis Module**

This module will focus on analysing and generating a comprehensive report of our prototype. Fig. 4 shows the flowchart of the module. Even though the flowchart looks simple, the process behind this module is challenging because first we need visualise and map the collected information that will be able to clarify connections among the people that are being investigated. And then we will produce a hypothesis based on the information we have found, for instance, if the mapping shows that two people that are being investigated have been exchanging messages, they have some similar images tagged to them, the application will generate a hypothesis that shows they are connected. Finally the application will produce a comprehensive report based on all information that have
been collected which can be used as supporting evidence in that particular investigation.

VI. CONCLUSIONS AND FUTURE WORK

In this paper we have reviewed the existing literature of digital forensic investigation models and frameworks. We described existing digital forensics investigation models and frameworks and found these to generally involve the process of identifying, preserving, analysing and presenting digital evidence. For the purposes of general investigation (e.g. analysis of a hard disc), there are various tools available because they are produced according to general investigatory requirements. However, to conduct investigations in online social networks, these tools are not suitable because they do not provide specific functions and options as discussed in the previous section. To deal with these shortcomings, there is a need to establish a standardized forensic investigation process for these networks, thus we have developed a comprehensive online social network digital forensic investigation model and we will develop an application prototype to fulfill the essential requirements of online social network digital forensic investigations. And then we have designed the application prototype by developing the algorithms to ensure that the objective of systematic investigation and analysis process as described in our model is accomplished.

Since this is an ongoing project, we intend to work further in a number of directions. We will carry out a number of case studies to validate the design of the application prototype. Subsequently, we will develop the design and we will carry out evaluations of the prototype to make sure the purpose of developing this model is fulfilled and that the functionalities meet the essential requirements of the online social networks digital forensic investigation model.

REFERENCES
