2014

Up-dating investigation models for smart phone procedures

Brian Cusack
_Auckland University of Technology, brian.cusack@aut.ac.nz_

Raymond Lutui
_Auckland University of Technology, raymond.lutui@aut.ac.nz_

DOI: 10.4225/75/57b3e1a3b87c

Originally published in the Proceedings of the 12th Australian Digital Forensics Conference. Held on the 1-3 December, 2014 at Edith Cowan University, Joondalup Campus, Perth, Western Australia.

This Conference Proceeding is posted at Research Online.

UP-DATING INVESTIGATION MODELS FOR SMART PHONE PROCEDURES

Brian Cusack, Raymont Lutui
Auckland University of Technology, Auckland, New Zealand
brian.cusack@aut.ac.nz, raymond.lutui@aut.ac.nz

Abstract
The convergence of services in Smart Technologies such as iPhones, Androids and multiple tablet work surfaces challenges the scope of any forensic investigation to include cloud environments, devices and service media. The analysis of current investigation guidelines suggests that each element in an investigation requires an independent procedure to assure the preservation of evidence. However we dispute this view and review the possibility of consolidating current investigation guidelines into a unified best practice guideline. This exploratory research proposes to fill a gap in digital forensic investigation knowledge for smart technologies used in business environments and to propose a better way to approach smart technology investigations.

Keywords
Investigation, Models, Smart Devices, Evidence, Preservation

INTRODUCTION
At present digital forensic investigators are faced with many different digital forensics investigation process models advocating best practices for extracting and preserving evidence. Smart technologies have created a problem where an investigator must apply many previously used models to collect and preserve digital evidence. The proliferation of investigation process models has arisen from the rapid and continuous innovation of devices, systems and applications for business use. Different proprietary designs, software, and access controls have influenced the adoption of digital forensic investigation models and the continuing revision of best practice. Individuals and businesses are very much dependent on computers, corporate networks, mobile devices and the Internet to conduct their daily tasks. The new generation of digital mobile devices are known as smart devices because of their processing power, memory and storage spaces are very similar to that of a desktop computer. These smart devices are capable of storing, transmitting and processing large amounts of private and confidential data (Owen & Thomas, 2011, p.25). Over the past decade, these smart devices have become a target for criminal and civil evidence gathering. As a result, it is very important that digital forensic investigators can complete their investigation effectively and efficiently within the constantly changing technological environment. In order for the investigators to achieve best practice goals the forensic investigation process models require constant updating and adapting to the new challenges. According to Tanner & Dampier (2009), digital forensic models are divided into three categories as Investigative models, Hypothesis models and Domain models (p.291).

Digital forensics comprises of various areas that relate to different technologies. There are four main areas and these are Computer forensics, Network forensics, Mobile forensics and Cloud forensics (Lin et al., 2011, p.387). Computer forensics is defined as the use of specialised techniques for recovery, authentication and analysis of electronic data when a case involves issues relating to reconstruction or computer usage, examination of residual data, authentication of data by technical analysis or explanation analysis of technical features of data and computer usage (Hankins et al., 2009, p.233). Network forensics on the other hand is defined as the use of scientifically proven techniques to collect, fuse, identify, examine, correlate, analyse and document digital evidence from multiple, actively processing and transmitting digital sources for the purpose of uncovering facts related to the planned intent, or measured success of unauthorised activities meant to disrupt, corrupt, and or compromise system components as well as providing information to assist in response to or recovery from these activities (Palmer & Corporation, 2001, p.27). Mobile phone forensics is defined by the National Institutes of Standards and Technology as, the science of recovering digital evidence from a mobile phone under forensically sound conditions using accepted methods (Jansen & Ayers, 2007, p.6). Final type of digital forensic is known as Cloud forensics. Cloud forensic is defined as a mixture of traditional computer forensics, small-scale digital device forensics and network forensics. Therefore, Cloud forensic is the application of digital forensic science in the cloud computing environments (Ruan et al., 2013, p.38).

In the Appendix we list the assessed digital forensic investigation models that span an evolution from computer forensics in the 1990s until smart phone investigation models of 2012. These models form the basis of the
analysis and we observe the different variations. The remainder of this paper is structured to review previous literature on digital forensic investigation models, potential problem areas, verification of the problem, an improved model and a discussion of its application.

PROBLEM AREAS

The convergence of computing and communications on mobile devices along with the application services and commercial services delivers an information rich environment for the user. The consequence is that much evidence is available but it is often stored in many locations and is in large quantities. Mobile smart devices have advanced functionalities with the ability to combine many functions onto one device such as, camera, video, Internet access, calendar, address book just to name a few. The devices are running on operating systems similar to a PC which allows users to install third party applications. Security and privacy protection became a major concern when business and private users’ realised the amount of private information and data that these smart devices hold (Lin, et al., 2006, p.386). The smart devices have the ability to establish wireless connectivity and most of them also have the ability to utilise the cellular network. These devices also support multimedia applications and messaging services with GPS, gyroscopes, and accelerometers sensors built in. Smart devices advancement and growth in usages and popularity gives rise to very large data sets (Wang, et al., 2012, p.52; Leavitt, 2011, p.11). Also, the pervasiveness and ubiquitous nature of these smart devices increase the complexity of the situation for the forensic investigators (Bednar, et al., 2008, p.3). While business and private users embrace the mobility and advancements of these technologies, criminals also find other ways to utilise these devices to conduct illegal activities (Lin, et al., 2011, p.386; Dezfouli, et al., 2012, p.186). As a result, to deal with this emerging and growing new phenomenon, previous digital forensic investigation guidelines require revisiting and reviewing (Hankins, et al., 2009, p.230).

Most smart technology devices access cloud environments for the information services. Cloud computing has been defined in various ways for instance, Furht (2010, p.3) defined Cloud computing as, “a new style of computing in which dynamically scalable and often virtualized resources are provided as a services over the Internet”. According to Mollah, et al. (2012, p1), Cloud computing is a, “TCP/IP based development and integration of computer technologies such as fast microprocessor, huge memory, high-speed network and reliable system architecture.” The National Institute of Standards and Technology released their Special Publication 800-145 and defined Cloud computing as “a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction”. The cloud is said to be a network of data centres working together to provide powerful applications, platforms and services that can be accessed by its users over the Internet (Abhishek & Mahasweta, 2011, p.3). For investigation purposes a number of problems arise (Mell & Grance, 2011, p.2). Cloud computing has four various deployment models.

<table>
<thead>
<tr>
<th>Private Cloud</th>
<th>Community Cloud</th>
<th>Public Cloud</th>
<th>Hybrid Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>This model refers to a cloud infrastructure that may be owned and operated by an organisation for private use only.</td>
<td>This model refers to a cloud infrastructure that is owned, managed and used exclusively by a community with similar concerns such as security requirements, policies or mission.</td>
<td>This model refers to an infrastructure that is open to the general public. This infrastructure may be owned and operated by an academic institution, government organisation or a business.</td>
<td>This model refers to an infrastructure which is a combination of two or more of the other three models. This particular model allows the infrastructure to remain exclusive while they are bound by standards or branded technologies.</td>
</tr>
</tbody>
</table>

Table 1. Cloud computing various deployment models (Mell & Grance, 2011, p.3).

Each structure has to be evaluated prior to investigation and the relevant evidence preservation assurances taken. In some instances the cloud presents insurmountable problems for evidence acquisition on account of the structures.

A third problem is the current divisions that are made regarding digital forensic areas of expertise (see figure 1). Each area of investigative expertise has evolved in keeping with technological developments and the systems developments. When any of the components of the illustrated digital forensic areas are involved in an investigation, the investigator needs to follow proper investigation procedure. These are scientifically proven techniques and methods to obtain and analyse digital evidence is such a way that publically accepted standards
are complied. These standards are often written as professional guidelines and declared in the forensic report to substantiate admissibility. The adoption of a scientifically proven method to preserve, acquire, analyse, document evidences obtained from digital sources will help with the admissibility of the evidence in the court of law (Ademu, Imafidon & Preston, 2011, p.175). The identified problems for investigators in relation to smart technologies are the volumes of data involved, the distribution of that data in different systems, formats and jurisdictions and the constraints provided by different best practice guidelines.

**Figure 1. Digital forensic expertise divisions**

**PILOT STUDY VERIFICATION OF PROBLEMS**

To prove that the problem areas located in the literature analysis exist in practice a pilot study was set up with an iPad and a case scenario. The pilot study was set up to confirm (or otherwise) the issues, problems and gap identified in the literature analysis for a smart device digital forensic investigation. The test bed in Figure 2 shows three different wireless accesses to three different types of wireless network environments. The test bed was set up to reflect a crime scene in which the criminal accesses a private company’s information system via a business mobile smart device and exercises the following actions. Access the company’s web server and defaces the company’s website. Accessed the company’s mail server and sent out fake e-mails to the company’s suppliers and downloads the company’s sensitive documents and uploads them in to a cloud account via wireless access from a cafeteria. The actions were executed using the iPad 4 with Wi-Fi and 3G capabilities.

**Figure 2. Test bed for the pilot study.**

The machine that was employed during the forensic investigation process was running on a 64bit Microsoft Windows 7 Professional with service pack one. The computer also ran on Intel core i7-2600 processor with
eight gigabytes of memory. To examine the iPad, a logical acquisition approach was the technique employed in this test case scenario. Logical acquisition is defined as a bit-by-bit copy of objects stored logically such as directories and files stored on a logical store of the device such as the file system partition (Jansen & Ayers, 2007, p.13). Three different tools were installed on the forensic computer. These were iTunes 11.0.5.5 which is a free backup utility provided by Apple that comes with the iPad 4. Since our pilot study utilises an Apple mobile device and also only required to take a logical acquisition of the device, we take advantage of the free logical backup utility (iTunes) provided by Apple (Bader & Baggili, 2010, p.1; Said, Yousif & Humaid, 2011, p.122). This was followed by an in-depth examination and analysis of the acquired backup copy. To analyse the acquired data, SQLite database browser was also employed to read the databases and the plistEditor Pro v2.1 was employed to read the .plist files (Bader & Baggili, 2010, p.1). Prior to acquiring data from the iPad, the automatic synchronisation feature of iTunes was disabled. The iPad was then connected to the computer through the USB cable. The data acquisition process was then initiated manually, once completed, the iPad was disconnected to avoid further unwanted processes from taking place. Data acquired from the iPad goes to the iTunes default backup location which is C:\Users\Admin\AppData\Roaming\Apple Computer\MobileSync\Backup\. The name of the folder containing the data extracted from the iPad is very long which is a combination of forty hexadecimal characters “5a062e5a92472a3efc14a31d4a01752a8a3a4157” representing the unique identifier of the iPad. The names of the acquired files also adopted the same naming convention which signifies the unique identifier for each data source obtained from the iPad (Bader & Baggili, 2010, p.7).

The extracted data showed in Figure 3 came in three different file formats; the plist file, mddata files and the mdinfo files. The plist files are Apple’s property list file format which stores data in plaintext and can be read using plist editor software. The mddata files stores data in raw binary format while the mdinfo file contains encoded metadata for the corresponding binary mddata files. In general, the iPad operating system (iOS) stores data in binary list and database files. Other information such as the device’s status, application settings and user’s configuration preferences are stored in XML plist files. These includes time zone, pairing records with devices and computer, email accounts, network identification, browser history, cookies and bookmarks. Information such as text messages, email messages, contacts list, call logs, notes, calendar are stored in SQLite database files. However, to read the binary files, a parsing tool called “iPhone Backup Extractor” is used (Bader & Baggili, 2010, p.7). Various tools and techniques are applied and the iPhone backup extractor was the analysis tool that is employed to read the extracted binary files into a readable format as it shows in Figure 4.
During the test, an application called 2xClient was used to access the private network from the iPad. The record was located in a folder named "com.2X.2XClient" as it showed in Figure 5.

Figure 5. 2xClient SQLite file.

Figure 6 shows that the connection record, stores the username used (Administrator), the connection ID, the access port number and IP address of the server that the iPad accessed.

Figure 6. plist file record on com.comcsoft.iTransferPro.

An online Unix Time Conversion tool was used to convert "397282107" to "Wed, 04 Aug 2013 04:08:27 GMT". The Pilot study shows the scope of digital forensic investigation on a smart device (with scenario tests). It indicates that the three problems identified from literature are present. The problem of large data quantities can only be resolve through automation and the problem of cloud connectivity can be managed by setting limits to investigation. However the problem of many independent approaches remains outstanding where there is apparent redundancies between approaches and a requirement to be updated to the new technologies, services and related service integrations.

EVALUATING MODELS

(Note this section reviews and discusses briefly the 13 Digital Forensic investigation models shown in the Appendix that had to be removed for the file size requirement. It is available from the authors.) Each model has been evaluated for its principles, phases and the other elements included to locate necessities, redundancies and
gaps between the models. Figure 7 provides example of the analytic framework and the types of resolution that may be achieved using reference phases and investigation interaction criteria as the units of analysis. The first model was the Computer Forensic Investigation Process model (1995). This model focuses exclusively in the investigation process beginning with data acquisition. The model does not define how the investigator can approach a crime scene. The model also puts emphasis on the evaluation stage by providing three extra investigation steps within the evaluation stage. Digital evidence must be analysed without bias or modification (Reith et al., 2002, p.3). As an improvement, the six phase Investigative model from the DFRWS was developed for computer and network forensics (Palmer & Corporation, 2001, p.17). The DFRWS investigative model addresses the short comings of the computer forensic investigation processes in the Computer Forensic Investigation model (1995). The DFRWS model was also developed to cover not only forensic investigation on computers but networks as well. The progressive development and comparative analysis of each model can be traced until the recent models developed for cloud environments. Martini & Choo proposed a digital forensic investigation framework for cloud computing in the year 2012. This cloud investigation framework consists of four phases which are the evidence source identification and preservation phase, collection phase, examination and analysis phase. Due to the fact that virtualization is the key element in implementing cloud computing, this provides forensic investigators with more challenges. The decentralised nature of how data is processed in the cloud creates new disruptive challenges for investigators. As a result, the traditional ways of acquiring data are no longer practical (Birk & Wegener, 2011, p.1).
Valjarevic and Venter (2012) concluded that there are significant disparities between the existing digital forensic investigation process models. These inequalities relate to the following; the number of phases of the model, the scope of the model, the similarities of the phases name, the hierarchy levels and the principle behind the construction of the investigation process model. The proposal for a new working model is presented in the next section.

THE PROPOSED WORKING MODEL

The investigation model evaluation conducted by Valjarevic & Venter in 2012 shown in Figure 7 was organised around the attributes of reference phases and investigation interaction. Our evaluation of the same models was conducted by focusing on the relevancy of the model to an investigation involving a smart device. As a result it was evident that existing models have different concepts regarding the purpose of the use application. Most of the existing models have different names for the phases but a similar purpose. We believe that in order to integrate all of the digital forensic expertise divisions without compromising the efficiency and the effectiveness of the investigation, a revised digital forensic investigation process model is required. This new model needs to be able to provide the investigator with a clear definition of the investigation path, and also to clearly define the external links to other repositories of evidence. However, most importantly, these processes must be achieved without sacrificing the integrity and the credibility of the evidences. Figure 8 shows the result of our analysis.

The main feature of our working model is the relevant pathways that may be taken in relation to a particular investigation. An investigator is to start with the device at the incident detection process. The relevant pathways allow a traditional forensics method and/or to branch out at the external links process to a cloud environment or to a computer and networking environment or to both. In this way the redundancies and exceptions created in the analysis of the appended models are removed so that an investigator has clear direction for investigation processes and decision making. Once the acquired data from a smart device is analysed and criminal activities on private network and cloud environment are found, the working model clearly defines the required investigatory steps. In the verification study the results and IP address of servers that were accessed in a private network were clearly identifiable and also the cloud account that was used to transfer the scenario stolen documents. In such discovery, there is no need for the investigator to pick up another model but just follow processes defined in this model. For instance, there is no need to go through processes such as deciding strategies on how to approach a new thread of investigation. The investigator only needs to acquire the required access permission and go straight to the IP address that was found on the mobile device. This minimises the volume of data that the investigator has to acquire and analyse. Effectiveness and efficiency is an important element of an investigation because, it minimises the chance of making a mistakes. The iterative features built into the model assure duplication of processes is eliminated and regardless of the number of pathways required to complete the investigation one report will be delivered.

The following definitions assist the interpretation of the proposed working model (see Figure 8):

- **Incident Detection Process:** This phase initiated an investigation which is usually triggered by a phone call reporting a crime or a mobile device found in a crime scene.

- **First Response Process:** Another part of the initialisation process in this model deals with the first awareness of an incident, acknowledging the incident and starting the process by involving the stakeholders. This can be achieved by a system or an individual and involves further reporting to the system administrator or the stakeholders or investigator.

- **Approach strategy:**
This phase is concerned with the development of a method or strategy on how to approach the investigation to maximise the collection of potential untainted evidence while minimising the impact to the victim.

- **Evidence Source Preservation:**
  This phase is concerned with identifying sources of potential evidence in a digital forensics investigation. This involves the identification of potential evidence collection methods whether by traditional or digital forensics. Irrespective of the identified source of potential evidence, forensic investigator needs to ensure the proper preservation of the evidence. This will involve isolating, securing and preserving the state of both potential physical and digital evidences.

- **Potential Evidence Collection:**
  This phase involves the physical crime scene and the acquisition of potential digital evidences by employing standardised and accepted methods.

- **Examination:**
  This phase involves an in-depth systematic search of potential evidences that relate to the alleged crime.

- **Trace Potential Evidences External Links:**
  After an in-depth systematic examination of potential evidences on a smart mobile device, alleged crime is traced down based on evidences acquired from previous phases such as:
  
  i. IP addresses which linked to a private local area network.
  
  ii. Name of cloud provider or username that identified a link to cloud service provider that was used in the alleged crime.

- **Analysis:**
  This phase is concerned with reconstructing the fragments of data, drawing conclusions from the evidence found and determining their significance. An iteration feature is implemented in this phase to allow the investigator to go back to the preservation to reconfirm or further investigate potential evidences found in the data.

- **Presentation/Report:**
  This phase involves summarising and explaining the conclusions of the investigation.

- **Return Evidence/Storage/Chain of Custody**
  This phase ensures that the chain of custody, storage and the return of physical and digital evident follows the proper procedure for handling evidence.

The preparation phase for Network and Cloud forensics is as follows:

- **Preparation phase:**
  In addition to preparing for an investigation in a private local area network or the Cloud environment, this phase also involves obtaining the required authorisation from local legal bodies for further investigation and access to more information from this environment.
CONCLUSION

The literature analysis of digital forensic investigation frameworks showed three main problems when one or more selections of the frameworks were applied to smart technology investigations. The pilot study verification of the problems confirmed that big data, cloud environments and the division of focus into many models weaken the capability of the investigator to apply an efficient and effective approach to smart technology investigations. Our contribution shows that a relevant model can be constructed (figure 8) from the necessities, redundancies and gaps in the other established models. This is exploratory research and further work is proceeding to test the model in practice.

REFERENCES


