12-3-2012

Evidence Examination Tools for Social Networks

Brian Cusack
Edith Cowan University

Jung Son
AUT University

DOI: 10.4225/75/57b3afc1fb861
10th Australian Digital Forensics Conference, Novotel Langley Hotel, Perth, Western Australia, 3rd-5th December, 2012

EVIDENCE EXAMINATION TOOLS FOR SOCIAL NETWORKS

Brian Cusack[1] and Jung Son
AUT University, Auckland, New Zealand
[1] SRI - Security Research Institute, Edith Cowan University, Perth, Western Australia
brian.cusack@aut.ac.nz; jung921@gmail.com

Abstract
Social networking (SNS) involves computer networks and billions of users who interact for a multiplicity of purposes. The web based services allow people to communicate using many media sources and to build relationship networks that have personalized meanings. Businesses and Governments also exploit the opportunity for economical consumer interaction. With the valued use of SNS services also comes the potential for misuse and legal liability. In this paper three software tools are tested in the laboratory to assess the capability of the tools to extract files from the four most popular web browsers while browsers are being used to surf the three most popular SNS sites, Facebook, Twitter, and LinkedIn. The results showed that the capability for evidence extraction differed markedly between tools indicating that the use of a particular tool has a material impact if the files are being extracted for evidential purpose.

Keywords
Social Networking, Networks, Evidence Extraction, Web Browser Investigation, Internet, Digital Forensic Tool Testing

INTRODUCTION

Digital forensics is a professional practice and study that concerns the extraction evidence of a digital nature from systems and devices (Oh, Lee & Lee, 2011). Compliance with technical procedures and legal requirements is foremost in the preparation of investigative reports and central to research studies (Ravi, Kumar & Jain, 2007). In this research tools are tested for the capability to extract evidence from social networking websites (SNS). SNS are online forums in which users come together at their convenience to share information in the form of digital text, graphics, links, or sometimes just to chat. SNS are defined as a web based services that users to create a public or semi-public profile allowing sharing and connections with other users (Boyd & Ellison, 2007). Just like emails and instant messaging technology, SSN Sites are an important communication medium providing organizations, businesses and governments with a means to interact with the public. The most popular social networking websites used at the time of writing include Facebook, Twitter and LinkedIn (Alexa, 2012).

The popularity of social network sites demonstrates the power of user-created content. However, since networked computers allow social networks to expand and grow in ways that were previously unanticipated, more criminals utilise SNSs to achieve their goals (Coyle & Vaughn, 2008). With the popularity of social media, many people willingly publicise where they live, their religion, their medical status, their friends, personal email addresses, phone numbers, photos of themselves and status updates, which informs others where they are and what they are doing. Criminals are able to use the communal information as an aid to commit crime. The access to and the preservation of digital evidence is a challenging technical and inter-jurisdictional problem. Currently, there exists no all-in-one tool kit, investigation methodologies or standardized procedures that forensic investigators can follow for SNS. Each attempt at digital evidence collection requires tool testing and audit against existing standards and professional guidelines. Social network forensics will therefore be a major focus for research and development in the future (Haggerty, 2010).

This paper is structured to first define the problem of sns digital forensic investigation. a research methodology is then specified to answer the research question: how is effective sns forensic investigation undertaken? A complex experiment is then executed to demonstrate how evidence may be extracted from the three most popular sns (Facebook, Twitter, and Linkedin). The three software tools are tested to assess the capability of the tools to automatically extract files from the four most popular web browsers while surfing ssns. The results show that there is large variation in tool capability and reason to insist standardized procedures are followed when doing ssn digital evidence extraction.

NETWORK FORENSICS IN SSNS

SSNs form a new subset of network forensics and provide new challengers for investigators. The challenges are less from the nature of social activity and more relate to the global and specialist use of networks and the related software. A user's social network activity is distributed over networks, tools and devices that are multiple, distributed and can be volatile and discontinuous. Consequently investigation targets the full scope of potential evidence retention locations but may only be able to access some. Internet history analysis is a primary technique and involves examining and analyzing a suspect's Internet activity. This is usually achieved by investigating the Web browser used by a suspect to access and interact with the World Wide Web (WWW). All of the well known Web browsers such as Mozilla Firefox, Google Chrome, Apple Safari, and Microsoft Internet Explorer (IE) save detailed information of activities in a cache, in the internet history list and in cookies in order to improve the user experience and save browsing time (Daniel & Daniel, 2012). Table 1 displays various Internet artifacts specifically related to SNS evidence which may be discovered on a suspect's operating system used to interact with a SNS. Possible artifact locations that include files or disk areas are also presented.

Artefact	**Evidence Description**	**Primary Data Location**
Internet History	List of websites URLs visited	Browser Database (eg. index.dat)
Session	Cookies and other session data created by SNS interaction	Browser profile files Browser cache
Web Pages	Web site data and files. such as html	Browser cache pagefile.sys, hiberfil.sys and unallocated space
Images	Pictures and other images, such as jpeg images	Browser cache pagefile.sys, hiberfil.sys and unallocated space
Video	Video files, such as flash video	Browser cache pagefile.sys, hiberfil.sys and unallocated space
Emails	Electronic mail provided by SNSs	Various email client data
Downloads	Material downloaded from SNS	Browser cache Temporary Files Unallocated Space

Table 1: Social Network Artifacts and Location of Potential Evidence

EVIDENCE EXTRACTION

Retrieving evidence from SNS interaction is a relatively complex task as the potential data and information stored in a suspect's device is dependent on a number of variables. Additionally, suspects may actively remove potential evidence (using anti-forensic techniques such as private browsing) to obstruct the investigation process. Thus, there is the prospect that a digital forensic investigator may overlook evidence, or fail to determine traces of evidence, and that insufficient proof will be procurable to support a case. The following 3 SNSs were chosen, namely Facebook, Twitter and LinkedIn, to gauge whether existing tools can adequately examine data and information created during SNS interaction. The most likely tool a user has for interaction is a web browser (Oh, et al., 2011) and highest probability of use is Microsoft's Internet Explorer (IE) (43%), Mozilla Firefox (28%), Google Chrome (21%) and Apple Safari (5%) (StatCounter, 2012).

Three different digital forensic tools were chosen to perform and establish SNS evidence extraction capabilities. These are off the shelf tools that are readily available in a Lab and none of the three tools selected are specifically designed for extracting evidence from SNSs; rather they have the ability to recover, examine or analyse data from other technologies which are also present in SNS Forensics Table 2 displays the 3 chosen tools: CacheBack, Internet Evidence Finder (IEF) and EnCase Forensic.

Name	Description
CacheBack (version 3.7.5)	Internet analysis tool including browser cache, history and chat discovery. Supported Browsers: IE, Firefox, Google Chrome, Opera and Safari. SNS Support: Supports Facebook Chat (using RMC).
Internet Evidence Finder (version 4.3)	Computer forensics tool to recover Internet related evidence. Supported Browsers: IE, Firefox, Google Chrome, Opera and Safari. SNS Support: Social Network artefacts from Facebook, bebo, MySpace, Twitter and Google Plus.
EnCase Forensic (version 7.03)	Commercial digital forensic framework produced by Guidance Software. Supported Browsers: IE, Firefox, Opera, Safari and Chrome. SNS Support: Browser history analysis and various EnScripts available to parse data.

Table 2: Digital Forensic Tools for SNS Evidence Extraction

Test cases were constructed to perform function based testing. A total of 4 test cases were identified to provide the information needed to determine actions and events relating to social networking. Table 3 displays the identified test cases, each with an accompanying test case reference number and description of the evidence data set required for tool testing.

Test Case #	**Test Case Name**	**Tested Tool Functionality**
TC01	SNS History Analysis	Provide detailed list of SNS URLs accessed.
TC02	Browser Cache Analysis	Automatically examine and decode Web browser cache for SNS information, data and files.
TC03	SNS Session Analysis	Locate Internet session artefacts created by SNS interaction.
TC04	Facebook Chat Analysis	Automatically examine evidence for Facebook chat messages.
TC05	Repeatability & Reproducibility	Tool achieves same results consistently.

Table 3: Test Cases for Forensic Tool Testing

The test environment comprised of the target computer and the investigation computer. Initially, the target computer's Hard Disk Drive (HDD) was forensically sanitized using EnCase Forensic to write zeroes (00) to the entire storage area to ensure no data remanence. The target computer, a Dell Laptop, was then installed with Microsoft Windows Professional. The four different Web browsers were then installed, all of which were the most recent versions available at the time of testing. A test data set was authored based on the requirements of the test cases as presented in Table 3. The target machine was used to actively interact with the prescribed social networks, carrying out the necessary tasks and interaction to generate the required data for the various test cases. After data generation was completed for each social network using all 4 Web browsers, the target computer was powered off. Following correct digital forensic procedures the target laptop was documented, and the HDD extracted to perform a forensic image. A forensic image of the target HDD was created using AccessData FTK Imager (version 3.0.1) and a Tableau T35e Forensic Bridge to ensure write protection of the original target device. The forensic image was acquired in EnCase E01 format, stored on an external HDD and forensic verification conducted by performing an acquisition and verification hash (MD5 & SHA1) value comparison. Figure 1 shows an overview of the test environment.

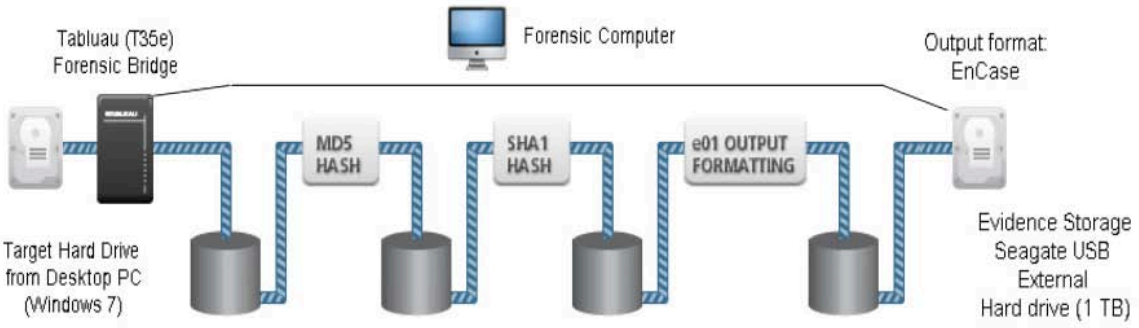

Figure 1: Testing Environment Overview

The analysis process was based on the assumption of the importance of required evidence in a digital forensic investigation. Additionally, data analysis was also based on working backwards by asking what information is required, by selecting search queries and defining specific locations to analyze the link between the collected data and the characteristics of a case. Figure 2 displays the overall procedure used to perform analysis on the collected data in order to establish the capabilities of forensic tools to automatically extract digital evidence.

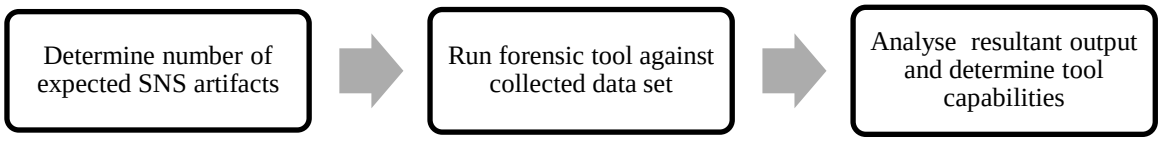

Figure 2: Data Analysis Procedure

The tool ranking method was adopted from literature (NIST, 2001). A mixed methodology consisting of both qualitative as well as quantitative elements was used to conduct the analytical comparison of the digital forensic tools to be tested. The quantitative elements consisted of how many instances of a specified artifact was found (by the tools function) for each test case compared to the total number of artifacts contained within the created data set. The weighted numerical total of each test case was collected for each product to provide robust feedback for analytical comparison indicating tool quality, functionality and reliability. Table 4 summarizes the test rating scale of 0 to 3, with the associated description, standard and percentile grouping.

Rating	Description	Rating Standard	Percentage Found
0	Miss	Unable to find any evidence	0%
1	Below	Sometimes able to find evidence but not accurate	1 - 29%
2	Meet	Able to meet the search requirement	30-59%
3	Above	Able to meet the requirement and provide excellent results	60 - 100%

Table 4: Digital Forensic Tool Testing Ratings for SNS Evidence Extraction

The accuracy of a forensic tool was assigned a quantitative value between 0 and 3 for each of the 4 test cases. If a tool failed to recover any data in a particular area, it was rated a 0 for that category, while a rating of 3 indicates that the tool exceeded the expected result including recovering deleted data and/or more information than other tools were able to recover.

The results

In order to determine the capabilities of digital forensic tools for performing extraction of SNS artifacts the prescribed testing methodology was implemented. The collected data was analyzed to determine the expected number of SNS artifacts, and then each tool was run against the data set to determine the number of extracted SNS artifacts for each test case. The findings illustrate the capabilities of the tools based on the ability to recover SNS artifacts from the different Web browsers used to access different SNSs. Test Case 01 (TC01) was designed to test the functionality of forensic tools to automatically examine and parse data from a suspect's computer and provide a list of SNSs accessed. Visited websites are unique Uniform Resource Locator (URL) addresses of an Internet resource, for example, www.twitter.com. Test Case 02 (TC02) was planned to test the ability of forensic

tools to examine and extract SNS artefacts from the Web browsers temporary cache storage. TC02 also tests the ability for tools to automatically examine a browser cache by identifying and decoding the data to provide images, html files and other files in temporary storage. Test Case 03 (TC03) was designed to test the ability of tools to automate the discovery of session related SNS artefacts such as cookie files. Test Case 04 (TC04) is a specific test scenario for Facebook related information which is designed to test the functionality of automated Facebook chat data extraction. Finally, Test Case 05 (TC05) was prescribed to determine the reliability of the chosen tools for testing by repeating the same test to ensure consistent results were obtained. Repeatability is defined as the closeness of agreement between independent test results under repeated conditions that are as constant as possible. Test results must be repeatable and reproducible as it is not possible to estimate experimental errors without them (NIST, 2001).

Figure 3 shows the results in visual comparison form and Table 5 the numerical results.

Tools tested in this research must always acquire the same results, thus making a case result reproducible and reliable. To ensure this, TC05 was conducted to determine the reliability of the SNS digital evidence found by the various forensic tools. All 3 test case scenarios were performed 3 times with each selected tool in order to observe the accuracy and consistency of the results. All 3 forensic tools obtained exactly the same results for each test case which was repeated. Additionally, the same results were produced on 3 different investigation computers: Windows 7 on Mac OSX (parallel), Windows 7 on Dell laptop, and Windows 7 on a Lenovo Desktop PC.

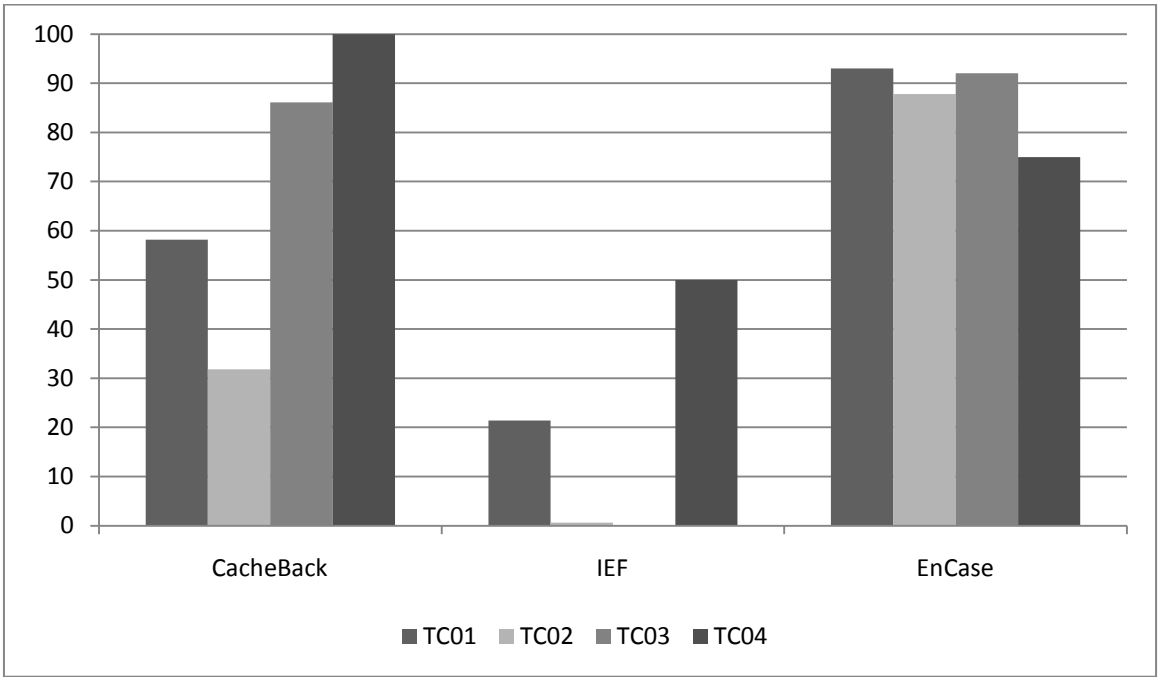

Figure 3: Comparison of Automated Evidence Extraction Tools

Table 5 displays the overall capabilities of forensic tools to automate the process of evidence examination of SNS artifacts. The results of each test case scenario are displayed and aggregated to determine the overall capability of the tested forensic tools.

Scenario	CacheBack		IEF		EnCase	
	Score	**Rating**	**Score**	**Rating**	**Score**	**Rating**
TC01	**58.2%**	**2**	**21.4%**	**1**	**93.2%**	**3**
TC02	**31.8%**	**1**	**0.6%**	**1**	**87.8%**	**3**
TC03	**86.1%**	**3**	**0%**	**0**	**92%**	**3**
TC04	**100%**	**3**	**50%**	**2**	**75%**	**3**
TC05	**100%**	**3**	**100%**	**3**	**100%**	**3**
Total Weighted	**75.2%**	**12**	**34.4%**	**7**	**89.6%**	**15**
Ranking	**2nd**		**3rd**		**1st**	

Table 5: Summary of Findings: Automated Examination Capabilities of Tools

DISCUSSION

The findings of this research have identified the capabilities of digital forensic tools to perform examination and extraction of social networking artifacts based on various test cases designed to test tool functionality. Although each of the three tools is not specifically designed for extracting evidence from SNSs, a large proportion of digital evidence of SNS interaction was able to be examined.

The overall top performing tool was EnCase Forensic which was able to automatically examine and extract an average of 89.6% of all SNS artefacts. It also had a perfect rating of 3 for all test cases. The comprehensive Internet history search was an exceptionally useful technique to isolate SNS related artifacts. However, the lack of support for certain Web browsers (specifically the version of Safari tested) may provide limited results. It is important to note that the EnCase version used (7.03) was the first to support Google Chrome Web browser. It should also be noted that EnScripts, a scripting language provided with EnCase, were used during 2 test case scenarios. Firstly, during TC02 an EnScript was used to automatically examine the Safari Web browser cache contents, and secondly for TC04 to extract Facebook chat information. The ability to configure software and perform advanced forensic investigation techniques is an advantageous addition to a forensic tool. However, it requires experienced investigator knowledge, as well as thorough testing to ensure reliable results and viable evidence is produced.

CacheBack also performed well at automatically examining SNS artifacts, being the second highest with an average of 75.2% of all artefacts extracted. Additionally, CacheBack proved to be the most automated examination tool, requiring minimal manual analysis by the investigator. The interaction with the CachBack tool was uncomplicated and straightforward and provided simple yet powerful automated analysis techniques such as filtering by host. However, there were also some disadvantages noted about the CacheBack tool. A major drawback was the reliance on additional third party forensic tools. For example, Cacheback does not have the ability to directly extract data from a forensic image; instead the image has to be mounted and then CacheGrab is used to extract Internet artefacts. Additionally, the RMC tool used for Facebook chat examination is not included with CacheBack.

IEF performed poorly overall, discovering an average of 34.4% of all SNS artifacts. However, the low rating is mainly due to lack of functionality for TC02 and TC03 and the testing methodology used. On the plus side, IEF was exceptionally easy to use and provided rich reporting capabilities. The major flaw of the tool is that it is designed to be run on a live system, and cannot process forensic image files as used in this testing. Therefore, a third party tool is always needed to mount the forensic image for evidence extraction.

The usability of a forensic tool is another exceptionally important aspect to consider. CacheBack and IEF have intuitive and easy-to-use user interface while EnCase is more complex. However, the usability also depends on the forensic investigator's experience and knowledge of procedures and forensic tools. In addition, forensic tools need continual upgrading so that additional functionality and improvement of evidence examination techniques

can be provided. For example, the Web browsers used in the research are constantly evolving. So that the data created by such applications can be automatically parsed and examined the forensic tools also need to evolve. It is difficult to conclusively state the single best forensic tool for SNS investigation of a suspect's computer. Each tool has strengths and weaknesses for performing specific functions. However, it can be recommended that EnCase has the most functionality due to the additional built-in tools which can be implemented. In summary, the correct tool choice for automated SNS investigation should be based on the scenario and functionality required.

CONCLUSION

In conclusion, there are tools currently available that have the ability to automatically examine, extract and provide digital evidence of SNS activity using known forensic investigation procedures. The tools tested and findings reported, illustrate that the functionality provided by the selected tools was able to be used to suit SNS investigations. Additionally, a high proportion of SNS artefacts were able to be automatically examined, thus providing valuable digital evidence.

However, there are still a number of issues surrounding SNS investigations. More recently, SNSs have become accessible not only via a web browser, but also by mobile devices such as smartphones. This raises a whole new realm of issues and using conventional digital forensic tools for collecting and analysing evidence is not appropriate.

In summary, the process of automated examination and extraction functionality provided by tools is exceptionally important, especially with the increasing amount of data present in forensic investigations. Social networking is part of the modern age. With it comes the ease and prospect of criminal or malicious usage. Having the tools to process the data and produce viable digital evidence is crucial.

REFERENCES

Alexa. (2012). Alexa Top 500 Global Sites. Retrieved 19 February, 2012, from http://www.alexa.com/topsites

Berghel, H. (2003). The discipline of Internet forensics. Communications of the ACM. 46(8). 15-20.

Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication. 13(1). 210-230.

Coyle, C. L., & Vaughn, H. (2008). Social Networking: Communication revolution or evolution? Bell Labs Technical Journal. 13(2). 13-17.

Daniel, L. E., & Daniel, L. E. (2012). Chapter 31 - Internet History (Web and Browser Caching). Digital Forensics for Legal Professionals. 213-218. Boston: Syngress.

Guo, Y., & Slay, J. (2010). Data Recovery Function Testing for Digital Forensic Tools. Advances in Digital Forensics VI. 297-311. Springer Boston.

Haggerty, J. (2010). Digital Forensics Investigations of Social Networks: Learning from Other Disciplines. Paper presented at the Centre for Security, Communications and Network Research. University of Plymouth.

Lyle, J.R., White, D.R. & Ayers, R.P. (2008). Digital Forensics at the National Institute of Standards and Technology. Gaithersburg, Maryland.

NIJ. (2008). Electronic Crime Scene Investigation: A guide for first responders. Washington, DC. U.S. Department of Justice.

NIST. (2001). General Test Methodology for Computer Forensic Tools. Retrieved 21 Oct 2011, from http://www.cftt.nist.gov/Test Methodology 7.doc

Pollitt, M. (2008). Applying Traditional Forensic Taxonomy to Digital Forensics. In I. Ray & S. Shenoi (Eds.), Advances in Digital Forensics IV (Vol. 285, pp. 17-26): Springer Boston.

Ravi Kumar Jain, B. (2007). Web Browser as a Forensic Computing Tool. ICFAI Journal of Information Technology. 47-57.

Oh, J., Lee, S. & Lee, S. (2011). Advanced Evidence Collection and Analysis of Web Browser Activity. Digital Investigation. 8(S). 62-70.

StatCounter. (2012). Top 5 Browsers on 2011 | StatCounter Global Stats. Retrieved 19 February, from http://gs.statcounter.com/#browser-ww-yearly-2011-2011-bar

Venter, H., Labuschagne, L., & Eloff, M. (2007). New Approaches for Security, Privacy and Trust in Complex Environments. In Proceedings of the 22nd International Information Security Conference. Sandton, South Africa