2016 Symposium on Colossal Data Analysis and Networking (CDAN)

Recovering Deleted Browsing Artifacts from Web Browser Log Files in Linux Environment

Anuradha P.
Department of Computer Science and Engineering,
College of Engineering,Kallooppara
Pathanamthitta, Kerala, India
Email: anuradhap8@gmail.com

Raj Kumar T.
Department of Computer Science and Engineering,
College of Engineering, Kallooppara
Pathanamthitta, Kerala, India
Email: rajcek@gmail.com

Sobhana N. V.
Department of Computer Science and Engineering,
Rajiv Gandhi Institute of Technology
Velloor P O,Pampady, Kottayam
Email: sobhana.nv@gmail.com

***Abstract*— In the present day scenario when it comes to Information interchange no one can even think about it without the use of Internet. Today there are uncountable numbers of websites in existence with loads of information present in them. To acquire this information one has to use a web browser in which these websites can be browsed according to our need. These web browsers store the users browsing history in specified log files which can be used in digital forensic investigation to acquire lots of information regarding the suspects browsing history. But these log files can be manipulated and cleaned by the users as and when required. In such a situation when suspects involved in any cyber crime deliberately cleans the traces of their browsing activity from any computer system further in depth examination of the crime becomes challenging and typical for the investigator concerned. This research is mainly focused on recovering such deleted browsing artifacts from the installed browser (Google Chrome in Linux environment) by an in depth analysis of the image of the hard disk used by the suspects involved in the crime.**

Keywords—web browser; browsing history; log file; digital forensic investigation; deleted browsing artifacts

I. INTRODUCTION

Nowadays, users used the web browser not only for browsing web pages but also to perform on line tasks. As a result of rapid adoption of the Internet globally, computer crimes such as extortion, child pornography, money laundering, fraud, software pirating, corporate espionage are increasing in a rapid rate. The main fact is that the people who do such things try to leave no forensic evidence. Sometimes the criminals make use of the Internet which makes tracking the users of a web page complex and difficult[1].

A web browser is a program that allows users to access web applications and web pages on the Internet. Each web browser stores in files or in databases which contain a large number of potentially useful information for investigators. Regarding the web browser Mozilla Firefox for example, it is possible to obtain information about the Web Pages visited, the content entered into form fields, the bookmarks and downloads performed by a given user. Footprints extracted from web browsers allows to know the user's interests (based on query to search engines, bookmark and visits, etc.) to identify potential accomplices (malicious file downloaded from a remote server) [2].

978-1-5090-0669-4/16/$31.00 © 2016 IEEE

Computer Forensics deals with the collection and analysis of data from computer systems, networks, communication streams (wired and wireless) and storage media in a manner admissible in a court of law [3]. With the rapid growth and use of Internet, Web browser forensics has become an integral part of computer forensics. Forensic artifacts left by web browsers are therefore relevant to the investigation for all sorts of on line activity. Much of the research into the forensic reconstruction of web browser history has been focused on the identification and extraction of forensic artifacts of individual browsers. Examining a suspects web browsing history could provide critical clues to solving a case since criminal, corporate or investigations involving illegal or improper web usage usually requires expert analysis of the information stored by a web browser as a result of a suspects Internet activity[4].

A Browser might contain direct or indirect evidence of the specific crime. Nowadays there are more utilities existing that allow users to wipe some of the artifacts of their activities. So the main question arising is "How to recover deleted browsing artifacts".But fact is that knowledgeable analysts may still be able to find indications of useful artifacts. Based on the above fact, my research focus on partially recovering deleted browsing artifacts by performing a simple web browsing session on Google Chrome and manually delete it and later perform a forensic analysis of hard disk to recover browsing artifacts.

The remaining part of this paper is formulated as follows.Section II defines Web browser forensics it also presents the different types of evidences in Web browsers. Section III discusses major web browsers in Linux environment.Section IV details the implementation and experiments of the system. Result analysis in section V and its discussions are described in Section VI. Conclusions are in Section VII,with suggestions for future work.

II. WEB BROWSER FORENSICS

Web browser forensics is a section of a larger field of study, known as computer forensics. The goal of computer forensics is to identify digital evidence,collect and preserve it. Next step is to analyze evidential data in a manner so that the integrity of the evidence collected must be preserved. Web browser forensics is concerned with analyzing and extracting evidence related to a user's Internet browsing activities. Web browser forensics of a suspects computer has become an essential artifact of many forensic investigations.

Nowadays cyber crimes reported on the Internet are in a booming rate. Web forensics relates to these cyber crimes. Web forensic analysis brings out some details like when a suspect browse a web page, in what sequence did somebody access a web page ,time of particular access and so on. Web history can also provide crucial evidence in crimes unrelated to computers, such as in the case of Neil Entwhistle who was convicted of murdering his wife and baby daughter after forensic investigators found a Google search for ”how to kill with a knife” in his computer’s web history[5].

Web browser forensics is not new in digital forensics research. Due to the fast development in web technologies, web browsers have been adaptive with continuous version releases. This poses a great challenge to the digital forensics community because they have to continuously experiment with new web browsers to learn how to forensically analyze their artifacts[6].

2.1 Types of Evidence in Web Browsers

During an investigation,following are the different kinds of evidences that an investigator would be looking for

- Surfing history: Surfing history of a user would mainly contain typed URLs, redirects and also the number of visits to a particular site.
- Bookmarks: This would mainly contain shortcuts or bookmarks created to specific websites by the user.
- Downloads: An investigator would mainly need to check for downloaded file in the default locations, also in the user defined locations or sometimes files are downloaded to default locations and then are moved or copied to user defined locations.
- Cookies: These are files created by web sites that are stored on users computer hard drive when he or she visits that particular site. It contain a wealth of information about the user. It would contain information like user names, passwords and web session information.
- Cache: It is a temporary area on the disk which is used to store most recently visited web sites. An investigator should check the temporary files because criminals may forget to delete the information the computer stores.
- Favourites folder: This would contain URL’s of sites that user wants to remember.

III. MAJOR WEB BROWSERS IN LINUX

There are many web browsers available, each having different methods of storing a user’s web usage history. The browsers prevailing today in the Internet are Internet Explorer, Firefox, Google Chrome, Safari and Opera. The history files produced by the browsers are not in a format readable by the humans and parsing them requires external tools. For example, Internet Explorer stores usage history in several index.dat files, located in various different places depending on the version of operating system[7]. Based on web browser statistics report published on December 2014 ,most frequently used web browsers are Mozilla Firefox and Google chrome[8]. From version 3 onwards database used in Firefox to store web browser history is SQLite[9]. Google Chrome uses SQLite databases similar to Firefox but with notable differences explained below. All databases are stored in the directory Default, with history and web data being the most important databases stored here. Chrome does not add extensions to its files. Chrome is not consistent with which datetime format it uses, some tables use standard Unix Epoch time (microseconds since 1st January 1970) while in others it uses its own variation of Windows Filetime (100 nanosecond intervals since 1st January,1601 UTC divided by 10). These two files are hidden files. So, in order to examine them, the browser should be setup to show both hidden files and system files.In this paper, I focus only on one popular Web browser Google Chrome on Linux environment.

Profile location of Google Chrome on Linux is:

/home/<$user>/.config/google-chrome

IV. IMPLEMENTATIONS AND EXPERIMENTS

This section provides a brief overview of web browsing sessions performed on Google Chrome.

4.1 Tools and setup

The following are the tools used for the assessments, acquisitions, examinations and analysis:

Hardware

- 1-Desktop (PC – 2GB RAM)
- 1-Desktop (PC – Forensic Workstation -2GB RAM)
- 150 GB SATA hard drive
- 1- Bootable USB drive(8 GB)
- 1- External USB drive(8 GB)
- 1- USB Write Blocker (IDE/SATA)

Software

- Ubuntu 14.04 LTS(64 bit)
- Google Chrome
- AccessData Forensic Tool Kit (FTK) 3.0 - used to analyze forensic images.

The objective of this experiment was to determine whether we can partially recover the user’s deleted browsing history from the hard disk. Therefore, all the experiments were handled in a forensically sound manner, as if we were handling real evidence. All procedures were properly documented, and evidence was safely preserved.

The experiment began by taking the hard drive and forensically wiped with zeroes using DD tool to ensure that no previous artifacts remained in disk . After the disk was successfully wiped, then installed it with Ubuntu 14.04 LTS(64 bit) from a bootable pen drive. Next, the disk was installed with only one specific browser from external USBdrive. The web browsers installed was Google Chrome and simple web browsing sessions were carried out.

The scenario consisted of:

- Watching videos on Youtube.
- Searching for images on Google Image Search.
- Browsing items on Amazon
- Accessing Gmail

Below table describes the Web Browsing session performed on Chrome.

Table 1:Web browser session performed for experiments

User Scenario Performed		
Website	***Activities***	***Keywords***
www.images.google.com	-Searching for "Parker pens" -Searching for "Mini Cooper" -Viewing several search results	"Mini Cooper" "Parker pens"
www.gmail.com	-Access the existing gmail account -Send e-mails with attachment	"Gmail ids"
www.amazon.in	-Searching for "Samsung galaxy " -Searching for "Idea Net Setter" - Browsing through the list of items	"Samsung galaxy" "Idea Net Setter"
www.youtube.com	Searching for "Mookambika Temple" - Watching several search videos	"Mookambika"

4.2 Forensic acquisition and analysis

After completing the web browsing session using Google Chrome, the web browser history was deleted manually .Then computer was shut down and disconnected from the power and the disk was carefully removed. This disk was individually connected to the Desktop using a hardware-based write blocker so that alteration of any data can be prevented.

The Desktop PC hard disk was developed with Ubuntu 14.04 LTS and FTK 3.2 to make it a dedicated computer forensic workstation. An image of the evidence drive was created using DD command line tool and was verified by several different hashes. The converted images were in raw dd format which are suitable for analysis phase. The analysis was performed using FTK. Each image file (in raw format) was added to an FTK case. A keyword search was then performed using FTK's "Live Search" and "Indexed search" function. The results of this search are outlined in the Result Analysis section.

V. RESULT ANALYSIS

This section decsribes the analysis of the disk images of Chrome using Forensic Tool kit. Lot of positive hits were obtained after conducting search on the disk images based on the keywords specified in Table 1.

5.1 **Analysis of Google Chrome Disk Image**

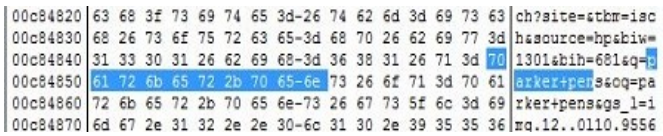

Figure 1: Hex view of the hard disk image for the the key word search 'parker+pen'

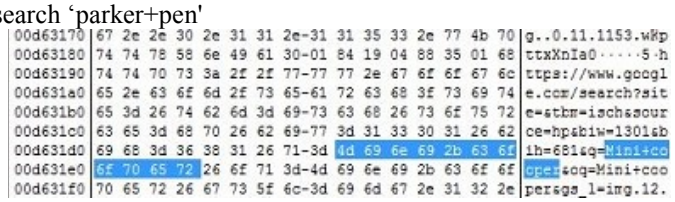

Figure 2: Hex view of the hard disk image for the the key word search 'mini+cooper".

Conclusion obtained from Figure 1 and 2

By analyzing the above two hex images in Figure 1 and 2, It was clear that user used Google as the search engine and he/she also view images of the parker pen and mini cooper.

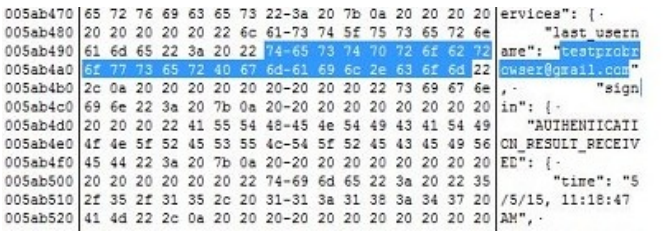

Figure 3: Hex view of the hard disk image for the the key word search '@gmail.com".

Conclusion obtained from Figure 3

An email id "testprobrowser@gmail.com" was found. It was clear that someone was sign in with this id and an authentication was received from Gmail server at 05/05/15 11:18:47 am.

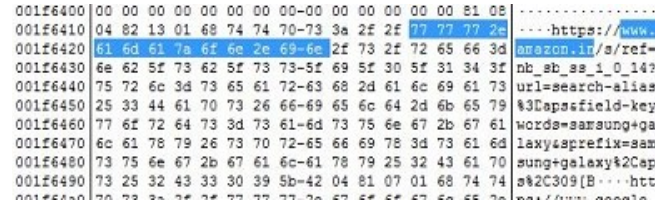

Figure 4: Hex view of the hard disk image for the the key word search 'samsung+galaxy'.

Conclusion obtained from Figure 4

So the conclusion obtained was that Samsung galaxy mobile details were browse through Amazon web site may be for a purchase.

Similarly lot of positive hits were obtained for the key words specified in Table1.Each among them were not explained in this paper. In addition to that, some images were also carved from the hard disk using FTK toolkit.

Figure 5

Figure 6

Above two images are carved from the Google chrome disk image.

Summary of Conclusion of Google Chrome disk image

Based on the indexed search and live search of keywords displayed on Table 1, lot of positive hits were obtained from the unallocated drive space of the disk. Hex representation of some among the disk images containing positive hits are listed above. Images can also be able to carve from the unallocated spaces. By analysing the disk images and the carved images obtained, it was clear that the user accessed Google web site, perform searches with key words Mini cooper, Parker pens, watch some videos on YouTube regarding Mookambika temple, Access Amazon site with login id, Access gmail account, time stamp regarding gmail access was also obtained.Carved images of Mini Cooper and Samsung obtained also help to claim that he/she performed image searches on Google. Overall result analysis help to trace complete web accessing history of the user. So from the above conclusions it should be clear that even though, an user deleted his/her destroys his browsing activity, the forensic analysis of suspects hard disk helps to recover almost all deleted browsing artifacts.

VI. RESULTS AND DISCUSSIONS

The disk images of Google Chrome on Linux environment were tested and analyzed in order to extract digital evidences. As we can see from the result analysis section , lot of web browsing footprints were obtained from installed version of Chrome . Some browsers left enough information to establish an affirmative link and some did not. From the disk images of Chrome, almost all URL history,email accounts, timestamps and images etc was recovered. Passwords,playable videos are not able to recovered. Most of the data was recovered from free space/slack space areas. Results obtained from the browser help to trace complete browsing activity of a user.

VII. CONCLUSION AND FUTURE SCOPE

The Internet is used by almost everyone, including suspects under investigation. A suspect may use a Web browser for different purposes such as to collect information, to hide his/her crime, or to search for a new crime method. Acquiring evidence from web browser is an important process for digital forensic investigation. After analyzing a trace of web browser usage by a suspect,it is possible to determine the objective, methods and criminal activities of a suspect. The existing tools and research related to web browser forensics are reviewed and uncovered their problems. In response a new methodology has been proposed to remove some of the limitations that exist in this field.

In case of removing web browser log files it was found that from the disk images of Google Chrome, it was possible to recover almost all URL history, email accounts, timestamps and images etc. So from the conclusions obtained after analyzing disk images, it should be clear that even though an user deleted his/her destroys his browsing activity, the forensic analysis of suspects hard disk helps to recover almost all deleted browsing artifacts. In the case of history deletion ,experts opinionated that if forensic acquisition of the web browsing traits is performed soon after the web browser session being investigated, then more evidence can be retrieved.

Although this project confined only on a single web browser running in Linux environment,future research will focusing on developing a tool that perform integrated analysis on all other Web browsers running in Linux environment and also to develop a recovery software in the case of history deletion.

REFERENCES

[1] Sandeep Kumar Khanikekar, " Web Forensics", unpublished

[2] Cruz-Cunha,Maria Manuela ,"Handbook of Research on Digital Crime," Cyberspace Security and Information Assurance, pp.233, 2014

[3] Natarajan Meghanathan, Sumanth Reddy Allam and Loretta A. Moore, "Tools and Techniques for Network Forensics", Department of Computer Science, State University international Journal of Network Security & Its Applications (IJNSA), Vol .1, No.1, April 2009

[4] Samuel Pyne. T , Dr. Hongmei Chi, "Internet explorer forensics: Reconstructing Internet Activity Using Pasco and Galletaerm ,Project CIS 5390, *Digital Forensics , Fall 2007*

[5] Junghoon Oh, Seungbong Lee, Sanjin Lee, "Advanced evidence collection and analysis of web browser activity", Elsevier Ltd 2011

[6] Sarah Lowman,Ian Ferguson,"Web History Visualisation for Forensic Investigations", Forensic Focus, 2011

[7] Bunting S M, "Understanding index.dat files",Computer Forensics Resources, 2010

[8] Browser Statistics, Retrieved on December 2015 from http://www.w3schools.com/browsers/browsers_stats.asp

[9] Pereira, M. T, "Forensic analysis of the Firefox3 Internet history and recovery of deleted SQLite records",Digital Investigation. Vol 5. Issue 3-4, pp. 93-10, 2009