Cyber Crimes Analysis Based-On Open Source Digital Forensics Tools

Article · January 2013

CITATIONS
6

READS
659

5 authors, including:

Shafi'i Muhammad Abdulhamid
Federal University of Technology Minna
110 PUBLICATIONS **1,502** CITATIONS
SEE PROFILE

Victor Onomza Waziri
Federal University of Technology Minna
36 PUBLICATIONS **81** CITATIONS
SEE PROFILE

Audu Isah
Federal University of Technology Minna
23 PUBLICATIONS **68** CITATIONS
SEE PROFILE

Olawale Surajudeen Adebayo
Federal University of Technology Minna
16 PUBLICATIONS **124** CITATIONS
SEE PROFILE

CYBER CRIMES ANALYSIS BASED-ON OPEN SOURCE DIGITAL FORESICS TOOLS

[1]Victor O. Waziri PhD,
Department of Cyber Security Science;
School of Information and Communication Technology; Federal University of Technology, Minna-Nigeria

Okongwu N. O
Economic and Financial Crimes Commission, Nigeria

Audu Isah PhD,
Department of Mathematics/ Statistics
School of Federal University of Technology, Minna-Nigeria

Olawale S. Adebayo
Department of Cyber Security Science;
School of Information and Communication Technology;
Federal University of Technology, Minna-Nigeria

Shafi'i Mohammed Abdulhamid
Department of Cyber Security Science;
School of Information and Communication Technology;
Federal University of Technology,
Minna-Nigeria

Abstract:
In this paper, we are present the digital forensic open source tools: Fiwalk, Bulk_Extractor, Foremost, Sleuth Kit, and Autopsy which are all Linux based forensic tools to extract evidences that could be presented in the court of law. Fiwalk reads a disk image and outputs a block of XML containing all the disk image of resident and deleted files. Foremost recovers files by using their headers, footers and data structures. The Sleuth Kit and Autopsy perform various aspects of file system analysis. The Autopsy Forensic Browser is a graphical web interface that presents the results generated by Sleuth Kit. This research project demonstrates the usefulness of the above-mentioned forensic tools for analysis and recovery of obliterated data from hard drives. This paper found that Sleuth Kit, Autopsy Forensic Browser, Fiwalk, Bulk_Extractor, and Foremost all provide effective file system analysis and recovery tool sets. The increasing complexity of storage devices requires that the investigator employs different forensic tool set to complement his arsenal of tools. No single digital forensic tool would be sufficient for an entire digital forensic investigation case. With this consideration, this paper employs various forensic tools. The demonstration of the effectiveness of these digital forensic tools utilized in this paper could serve as an alternative for investigators looking to expand their digital forensic tool set functionality in the court of law. Details of the experiments are fully given at the expense of bulkiness since this works is aim at enhancing the utilities of open source forensics tools applications.

Keywords: *Digital Forensics, Fiwalk, Foremost, Sleut Kits Bulk_Extractor, Autopsy, Linux, Ontologies*

1.

2. Introduction

To start writing on well known discipline such as digital forensics is always ideologically complex to do so. Forensics Computing may be construed as a methodical series of techniques and procedures for accumulating evidence. As in (Anthony et al, 2007) and (Garfinkel, 2009 A), Cybercrimes is on the rise and could be detected using Digital Forensics tools for which the crimes are extracted from various storage devices and digital media. Such systematic analyzed extractions could be presented in the court of Law in a sequential and meaningful format as evidences. Two types of test

questions (Brian Carrier, 2002) should be applied by investigators for both computer forensics and traditional forensics to survive in a court of law. These are:

Authenticity: Where does the evidence come from?

Reliability: Is the evidence reliable and free of flaws?

With these basic test questions, computer crime investigations should be predetermined through policy and what is acceptable risk" to every organization. Cybercrime includes the followings as outline in (Marcella et al, 2008), (Anthony et al, 2007) and (David, 2008):

1. **Theft of Intellectual Property.** This pertains to any act that allows access to potent, trade secrets, customer data, sales trends, and any confidential information
2. **Damage of Company Service networks:** This could occur if someone plants a Trojan horse, conducts a denial of service attack, installs an unauthorized modem, or installs a backdoor to allow others to gain access to the network or system
3. **Financial Fraud:** This denotes to anything that may use fraudulent solicitation to prospective victims to conduct fraudulent transactions
4. **Hacker System Penetration:** These occur via the use of sniffers, rootkits and other tools that take advantage of vulnerabilities of the systems or software.
5. **Distribution of Execution Virus and Worms:** These are Some of the most common forms of Cyber crime

Cyber Crime maybe spelt out into three comprises known as the "3Ts":

i. Tools to commit crime;
ii. Targets of the crime (Victim); and
iii. Material that is tangential to the crime

Divergent methods of cyber crimes abound and include amongst others, the following without further explanations due to space constraints: The Computer Facilitated Crimes that involves both insiders and external attacks. All these could be categorized to into various examinational patterns for analyses.

2.2 Rules of Computer Forensics

A good forensics investigator is expected to follow these suggestive rules as outlined in (Anthony et al., ibid):

i. Examine original evidence as little as possible. Instead, should examine the duplicate of the original evidence known as the image
ii. Should follow the rules of evidence and do not temper with the evidence
iii. Always prepare a chain of custody, and handle evidence with care
iv. Never exceeds the knowledge based on the investigative laid down rule by the court
v. Should document any changes of evidence

The rest of this paper runs as follows; Section 2 reviews the related works based on digital forensics devices of open source tools; section 3 outlines the methodologies of the research; in section 4, we perform some experiments based on the open sources tools for the computer forensics. Section 5 discusses the general computing outputs while section 6 gives further hints for future research investigations

3. Related Works

It is admissible that computers have become part of our lives worldwide (Brian,

2003). With the exploitation of web Technologies, so also vast exploits of the technologies have been established by criminals to commit crimes. Due to large and complex involvement of computers in web businesses and other intricate utilities in computing, computer is emerging as legal evidence in both civil and criminal cases.

Computer evidences are admitted in courts of law and these evidences could be anyfile or fragment recovered from the storage devices such as email, browsing history, graphics, photographs, or application documents. These files could be extracted from the hardisks and imaged to recover undeleted and deleted files. Deleted file recovery would require special techniques to retrieve them and this is what we are set out to achieve. These are professionally retrieved in a non-destructive technique. Evidence may be recovered from storage medium installed in digital equipment such as computers, cameras, PDAs, or cell phones [Gialanella et, 2008]. All forensics work should be strategically documented in a clear a system extraction; a principle known as chain of custody; in other to for the evidence to be admissive in the court of law

Computer devices that can establish evidence in the court of law are part of our lives. There have been a lot of some works on digital forensics community to create a common file formats, schemas and ontologies [13]. Despite all these efforts for a common need of affiliation, there has been little concrete standardization. As stated [13], DFRWS started the common Digital Evidence Storage Format (CDESF) Working group in 2006; in which the group created a survey of disk image storage formats in September, 2006. Due to lack of resources, the group disbanded in August, 2007. Hoss and Carver discussed ontologies to support digital forensics (carver and Hoss, 2009), but did not propose any concrete ontologies that can be used. Garfindel introduced an XML, representation for the system metadata (Garfindel et al, 2009), but it has not been universally adopted.

In another development, Richard and Roussev reviewed requirements for the "Next Generation Digital Forensics". Their works emphasized on system requirements with the argument that that inefficient system design, wasted CPU cycles, and the failure to deploy distributing unified techniques could introduce significant and unnecessary delays that directly translate into uncecessary delays (Richard and Roussev, 2006).

(Politt,2007) reviewed 14 different models for digital forensics investigation but did not attempt to evaluate or catalog them given time constraints. Bradford et al, 2004) argue that it is unwise to depend on upon "audit trails and internal logs" and further postulate that the digital forensics will only be possible on future systems if those systems make proactive efforts at data collection and preservation. Hey proposed a mathematical model for deciding the content and frequency of proactive forensic event recorders. Politt et al., further discussed how virtualization software and techniques could be productively applied to both digital forensics research and education (Polit et al., 2008). They argued that any discussion of virtualization with respect to digital forensics would face an unwelcomed tautology. In effect, the impact of virtualization on forensic examination could virtually be ignored-except when it could not. This is due to the fact that virtualization, and sometimes the subject of virtualization is the subject of the forensic examination, and sometimes the

virtualization as a tool is used by forensic examiner.

The literature like other disciplines goes on with different opinions and approaches. For instance, Turnbull et all performed a detailed analysis on the specific-digital media formats being collected by the South Australian Police Elctronics Crime section, theirs appears to be more the first quantitative analysis of its kind (Turnbull et al., 2009)

This paper is concerned with the application of some open sources for digital forensics as evidence in the court of law. We are applying Fiwalk, Bulk_Extractor, Foremost, Sleuth Kit, and Autopsy which are all Linux based forensic tools that could downloaded as open software.

3 Methodology

This section describes the practical experimental methods carried out in this research paper using digital forensic open source tools.

3.1.1 Experimental Analysis

The following tools listed below were used in the experimentations.

1. **Foremost version;**
2. **Fiwalk**;
3. **Bulk_Extractor** were compiled in Ubuntu 10.10; and
4. **Sleuth Kit and Autopsy** was compiled on both Ubuntu 10.10 and Windows 7 (Cygwin).

3.1.2 Steps to recover files from pen drive

1. Using the mount command, the pen drive has been assigned the mount point /dev/sdc, on /media/HHH and file type is fat; as illustrated from Fig 3.0.

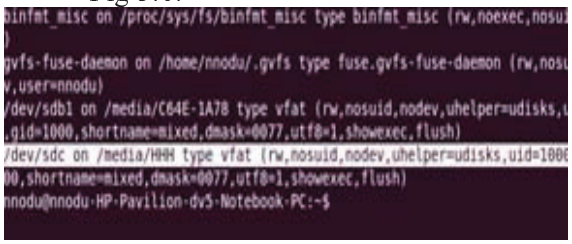


Fig 3.0: use of the mount command to display current mount points

We then use the 'cd' command to navigate the pen drive and display contents of the drive. The full command is given as follows:

1. Command to navigate to pen-drive: **cd /media/HHH**

   Command to list contents of a device/folder and show space it occupies on disk: **ls –s**

2 The command for launching the FIWALK:

   fiwalk -X <file> < diskimage> -v -X<file> = XML output to a <file> (full DTD)

3 The command for launching the bulk_extractor:

bulk_extractor -o output_dir [options] image;

4. Command to launch autopsy: **/autopsy**

Web link for autopsy forensic browser: **http://localhost:9999/autopsy**

Location of the evidence locker: **/cygdrive/j/evidence.**

5. **Command to create an image hard drive**
   The command for launching the creation of image hard drive:

dd if=/dev/sdb ibs=512 of=/media/Passport/flashimage.img.dd

At the pen-drive's directory, we used the command above to display file contents and block size in bytes. The files are listed below:

0321680561.pdf, and 0321680561.rar, bluehills.jpg, waterlillies.jpg, sunset.jpg, winter.jpg

Foremost should not be run from the folder/device you wish to recover data from. So

navigate to a folder we created on the desktop called recovery. A folder called 'output' will be the result of our recovery.

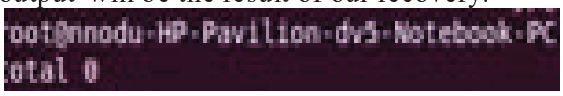


Fig3.1: using the command formost to display the pen drive contents, this is currently empty

3.2 Using Fiwalk to Process a 80-Gigabyte Disk Image in Order to Produce a Digital Forensics XML

We use Fiwalk to produce a digital Forensics XML in this sequential order:
The disk image called diskimage.img is stored on an external hard drive. The external hard drive's name is Passport. The syntax to invoke Fiwalk is already given in the last subsection.

3.3 Using Bulk Extractor to Analyze Disk Image for Domain Names, Wordlist, Log-file, and Emails Accessed from Drive

The disk image of 80-gigabyte hard drive is processed below. The Syntax for using bulk extractor is stated in the subsection above.

We navigate to the hard disk drive containing the 80 gigabyte hard drive image called diskimage.img, and invoke bulk extractor to start processing. The output directory is in /media/local/

3.4 Using autopsy and Sleuth Kit to perform Volume and File System Analysis on a 80 gigabyte hard disk

Procedure:

Invoke the autopsy by the use of the command see (Appendix B):

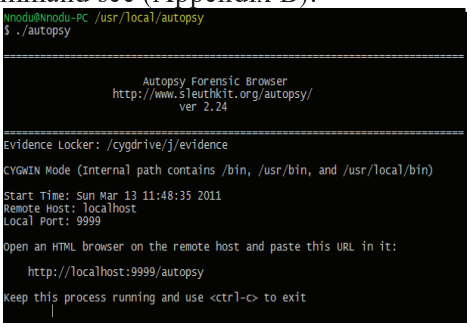


Fig3.2: Autopsy Forensic Browser showing that the invocation was successful

1. From Fig3.2, we will open a HTML browser and paste the address as depicted in the command above

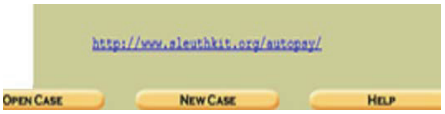


Fig3.3: The autopsy forensic browser interface opens on a web browser

2. We already have a previous case as shown from Fig 3.4

Fig3.4: Interface showing a previous case file which was called analysis_80g_hdd and was described as having crashed (damaged)

3. A new image is added which must be in the evidence locker and autopsy, accounts for the image file by creating a symlink (symbolic link)

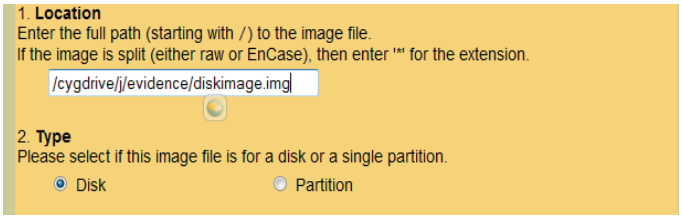


Fig.3.5: The image file path for the investigation is added and the type '.img' is also stated so autopsy can know what kind of image file is being investigated

From Fig3.5, the image file path for the investigation is added which has an extension of .img From Fig3.6 Analysis of the file system of the 80 gigabyte Hard disk shows two partitions:
Partition 1 of mount point C: is of type NTFS, sector range from 2048 to 149837823. Partition 2 is of type RAW,

sector range from 149837824 to 156299263 With mount point at /2/

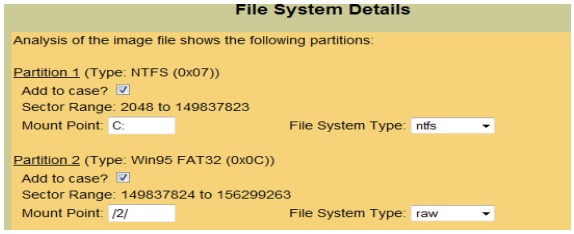


Fig3.6: After importing the diskimage: Autopsy details a summary of the File System and Image File details

All the units are in 512 byte sectors
The add image button was clicked and the image was successfully added linked to the evidence locker and linked as shown in Fig3.8. The ok button is then clicked again to continue.

Fig3.7: Autopsy shows the mount points (partitions) and names with file system type

From Fig 3.7, The 80 gigabyte hard disk is displayed on three mount points

3.4 Investigation and Analysis

In this section investigation and analysis of the experiments performed are conducted

3.4.1 Analysis of mount disk of name "DISKIMAGE.IMG-DISK"

a. **PROCEDURES**

1. The tab labeled Analyze when clicked to start the analysis reveals another window opens with three modes of analysis namely:
   Keyword Search, image details and Data Unit

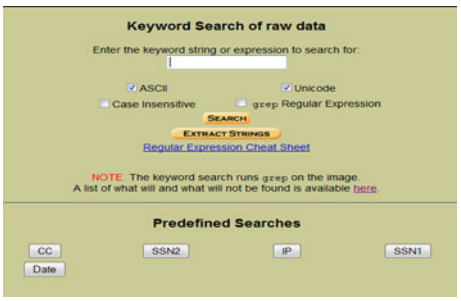


Fig3.7.1: Autopsy shows the keyword option for searches available, predefined searches are also listed to help the investigator

2. A search is performed with the keyword 'bank' but first to make the search faster, the entire strings in the disk image is extracted. This is done by the use of the 'EXTRACT STRINGS' tab.

b. **Observation**

The extraction of string was taking a long time, maybe because of the size of the 80 gigabyte hard drive. Even when the extraction was carried on a quad core windows 7 64 bit system with ram of 8gigabyte, it was still slow. So I decided to try a smaller storage device.
The storage device to be used is the same drive in which I performed a search using Foremost. The state of the pen drive though has changed. An open office document of size 244 megabytes was added to the pen-drive.

3. The **dd** command was utilized in Ubuntu to make an image of the pen-drive the full syntax is shown Appendix B

4. A new case file is created for the 244mb Pen drive as shown it **Fig3.92.** Autopsy recognized the file system type to be fat 16.

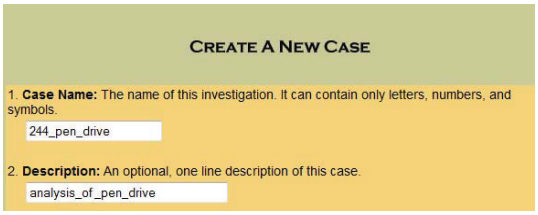


Fig3.7.2: Opening a new Case in Autopsy

3.5 Materials used in Experimental Analysis

a. **Programs Used**

Foremost Version 1.1, Fiwalk, Bulk_Extractor and Autopsy

b. **Operating Systems Used**

Windows 7 Version 6.1.7600 Build 7600, System Type: X64 based PC.

Installed Physical Memory = 3.00 Gigabyte; Ubuntu 10.10 (Linux) and Cygwin

3.5.1 Choice of Materials

Foremost was used because the program size is not large and easy to compile in Linux based system. **Fiwalk and Bulk_Extractor:** was used because the programs were easy for use and system resources. The recovery completed without any problems. **Ubuntu, Cygwin:** These operating systems that are Linux based. Most of these Forensic Tools are originally ported from Linux and therefore are easy to compile on Linux based environments.

3.5.2 Foremost

Foremost was used in this research paper to recover deleted data from a 512 megabyte pen drive.
This software recovers files using their headers, footers, and data structures.
The syntax for foremost usage see section 3.1

3.5.3 Fiwalk

This was used to analyse a 80-gigabyte Hard drive and an XML report of the entire structure was generated. Bulk_Extractor was used to recover email address, web domain addresses and histogram reports accessed from the Hard drive. Fiwalk makes it easy for non-experts to do significant forensic research and write powerful forensic tools (DEEP.). Based on Sleuth Kit, XML, and the Python programming language, this approach makes it easy for programmers to create tools that perform forensic processing without the need to master domain-specific knowledge (Garfinkel, 2009).

4. The Experimental Results

This section presents the results and findings from the materials and method of section 3. The results are from the experimental analysis performed in the methods sections are hereunder given in this pattern.

4.2 Results of recovery operation on Pen drive using Foremost

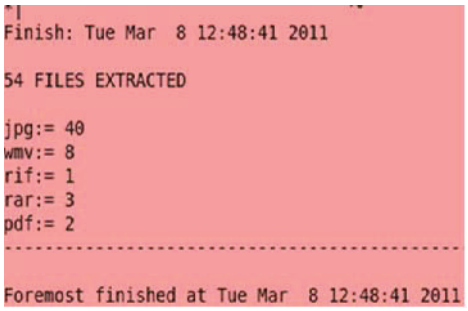


Fig4.0: **This screen shot shows the result of the recovery**

54 files were recovered.

Jpg = 40; wmv= 8; rif = 1;
rar = 3; pdf = 2

A summary of the audit.txt contains a report of what formost has done using the command in section 3. We will view some contents of the audit.txt file which is displayed below:

Foremost started at Tue Mar 8 12:48:34 2011
Invocation: foremost -T -v -t all -i /dev/sdc
Output directory: /home/nnodu/Desktop/recovery/output_Tue_Mar__8_12_48_34_2011
Configuration file: /etc/foremost.conf
File: /dev/sdc
Start: Tue Mar 8 12:48:34 2011
Length: 244 MB (256900608 bytes)

……..
……..
Finish: Tue Mar 8 12:48:41 2011

54 Files extracted which are not given here for want of space:
jpg:= 40; wmv:= 8; rif:= 1; rar:= 3; pdf:= 2
Foremost finished at Tue Mar 8 12:48:41 2011
We will then navigate to the output file containing the recovered files with this command as stated in section 3:
The output file contains six files of size 24 bytes: namely audit**.txt, pdf, rar wmv, avi, jpg**

4.2.1 Examination of JPG Files recovered from Pen drive using Foremost

Procedures:

a. **Step1**.

To examine the .jpg folder, (Appendix C3)' , I navigate to the jpg folder (Appendix C 4)

b. **Step2.**

So I then copied the 40 jpg files (Appendix C) from the folder to another folder on my desktop called jpgfolder.

I did this so I would be able to view the jpg files. So from Fig4. I now display the jpg files recovered. The files recovered were tested and found to be in good condition.

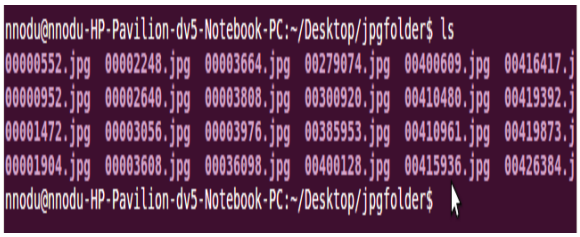


Fig 4.1 showing 40 jpg files recovered

4.3 **Discussion of Results**

The results obtained from the preceding experiments are discussed

a. **Results of Recovery operation performed on Pen Drive**

We discovered that the intial jpg's on the pen drive before the deletion, namely:

Blue hills, water lilies, sunset.jpg and winter jpg were recovered and renamed according to 0003608.jpg, 00003664.jpg, 00003808.jpg and 0003976.jpg by foremost. But because of copyright issues we will not display them.

b. **Examination of Pdf Files recovered**

Just like in the examination of the jpg files I will navigate to the jpg folder by using the 'cd' commnd and repeat the two step process. Using ls I will list the contents of the file .

There are currently two recovered PDFs namely 00030904.pdf and 00273880.pdf in the pdf/ folder. We then repeated steps in examination of jpg files, by using the 'cp' command, we will copy the recovered pdf files from their current folder to a folder on the desktop I created called pdffolder. The command is shown in Appendix C 6:

Fig 4.2 shows the recovered .pdf folders

Fig4.2: A screen shot of recovered pdf files

c. **Findings**

1. 00030904.pdf and 00027880.pdf are ' actually one and the same file, which is the same as 0321680561.pdf that was deleted
2. The two files recovered are the same pdf, probably saved at different times on the pen drive. And on examination the pdfs were found to be in good condition.

d. **Examination of WMV Files recovered**

All the procedures above in recovery for the pdf and jpg's are repeated.

Fig4.3: showing eight recovered WMV files

From Fig4.3 eight WMV files were recovered. I then copy them to a folder called wmvfolder created on Desktop. This is for easy examination of the files.

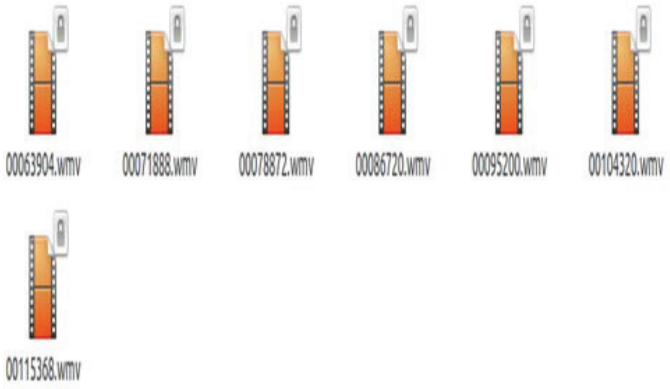


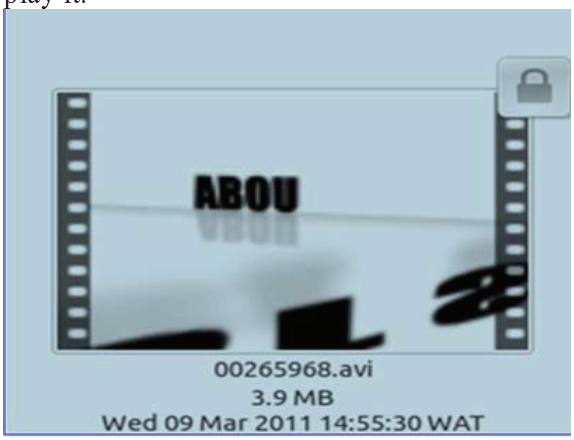
Fig4.4: showing eight recovered wmv files

From Fig4.4, the wmv files recovered were in good condition

e. **Examination of AVI Files recovered**

To examine the avi files recovered, I would simply repeat the two step followed

in the recovery of jpeg. An avi file of size 3.9mb was recovered and it is displayed in the screen shot of Fig4.5. The avi file was in good condition when I used VLC to play it.



f **Examination of RAR Files Recovered**

Repeating the procedures of recovering jpg files, 3 rar files were recovered namely, 00004184.rar, 00122704.rar, and 00139776.rar . This is displayed in Fig4.6

Fig 4.7 is a screen shot of the product of the extraction of the rar files, namely folder 00122704, 00139776 and 0321680561.pdf

Fig4.6: Showing three rar files

recovered with their sizes also displayed

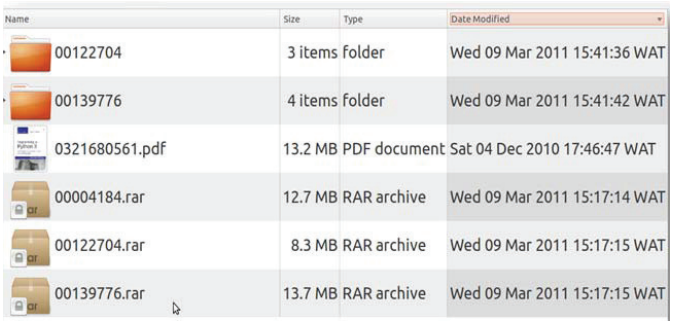


Fig4.7 shows the products of the

extraction of the rar files

4.3.1 Findings from Fiwalk Recovery Process

1. Archive file 00004184.rar is the same as archive 0321680561.rar I deleted at the start of the experiment.
2. All the archive rar files were tested and found to be in good condition.

The XML report is saved in Passport as diskimage.xml

The result of an extract from the file using fiwalk is displayed below briefly:

fiwalk xmloutputversion="0.3">
<metadata>
<dc:type>Disk Image</dc:type>
</metadata> <creator>
….
…..
<command_line>fiwalk -X0 /media/Passport/diskimage.img -v</command_line>
<uid>0</uid>
<username>root</username>
<start_time>Thu Mar 10 10:29:27 2011</start_time>

The result of the bulk_extractor process is briefly shown below:

All Threads Finished!
Phase 2. Creating Histograms
ccn ccn_track2 domain email kml rfc822 telephone url zip 0:
make_histogram(://([^/]+),services) -> /media/local//url_services.txt
…….
……..
inputs: 154966024 outputs: 209336
total time: 17288230 msec

elapsed time: 9182.6 seco

Fig4.8: Screen shot shows bulk_extractor's generated result

From Fig4.8, a summary of the results indicate the following files recovered:

1. **Report.xml** = showing an XML report of the extraction process.
2. **Zip.txt** = shows zip files files described by length, compression method and version.
3. **Url.txt**= a histogram of all URL's by domain
4. **Url_searches.txt** = a histogram of all search items, including Google, Yahoo, Bing
5. **Url_histogram.txt** = shows frequency distribution of the url sites accessed from the drive
6. **Telephone_histogram** = shows frequency distribution of telephone addresses used for a transaction on the system
7. **Telephone.txt** = shows the telephone addresses used for a transaction on the harddisk drive
8. **rfc822.txt**= shows the documents saved on the harddrive, such as letters, memo's etc.
9. **email_histogram.txt** = shows frequency distribution of email addresses accessed from the hard disk
10. **email.txt** =shows frequency distribution of email addresses accessed from the hard disk
11. **domain_histogram** = shows frequency distribution of domain addresses accessed from the hard disk
12. **domain.txt** = shows domain addresses accessed from the hard disk
13. **ccn.txt** = this file reports Federal Express Account numbers

a. **Observation**

In the domain_histogram.txt file, domain frequencies of occurrences were listed from the most occurring to the least occurring.

Then, a summary of the XML report detailing the Bulk_Extractor processes just executed has been summarized in the Fig 4.9. This shows the **url** values values extracted have the highest frequency of 334847 while the lowest are the **zip** files . While the telephone records extracted have the lowest value of 1915 records.

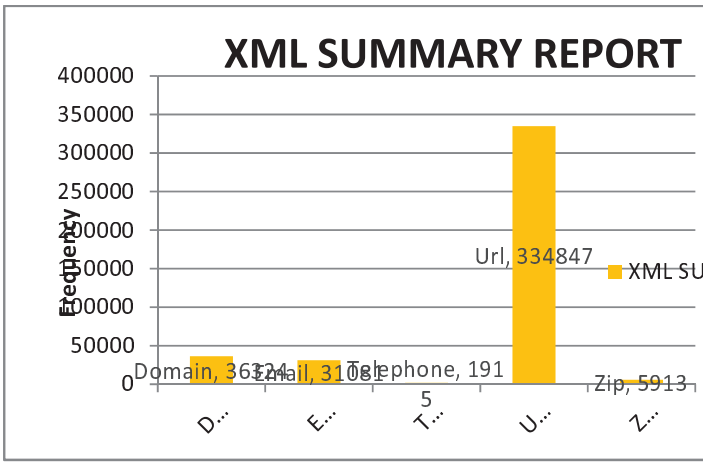


Fig 4.9 The summary of the XML report detailing the Bulk_Extractor processes just executed, the X axis contains the report recovered.

4.3.3 Results of Keyword Search Performed with Sleuth-Kit

1. An attempt is made to extract the strings for the keyword search to be faster
2. The extraction of strings was successful as seen from Fig4.91

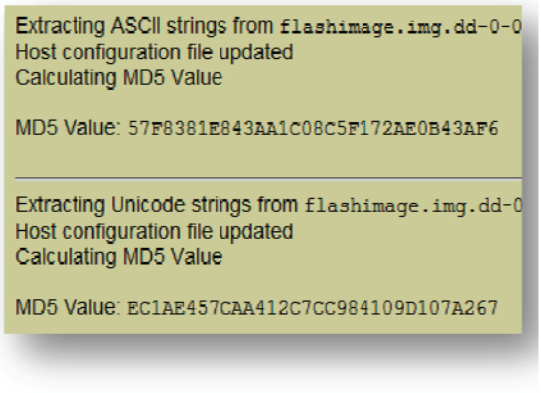


Fig4.9.1: The successful operation of string extraction is detailed with the MD5 values stated for ASCII string and Unicode string extraction

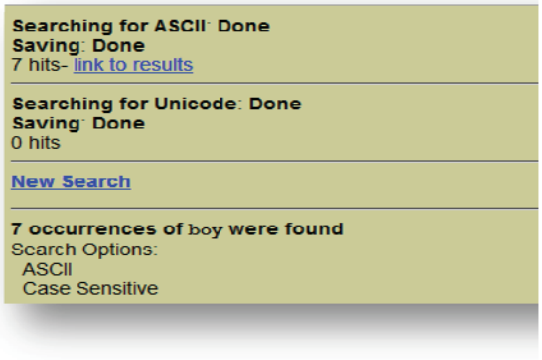


Fig4.9.2: Showing some hits of search for 'boy' under ASCII, but none for the Unicode search

3. A search for boy was executed and the seven hits were obtained from the ASCII section as shown from Fig4.92, but none from the Unicode search. 7 occurrences of boy were found from the search. Sectors 12572, 199482,256193, 305054, 380781, 432376, 452504,

The contents of sector 12572 are of type ASCII. The result can be exported/ saved; this is useful for recovery of documents. In addition, notes can be added to the recovered sector by the investigator to note points of interest. The notes tab is close to the EXPORT CONTENTS TAB.

4.3.4 File Type Sorting By Category

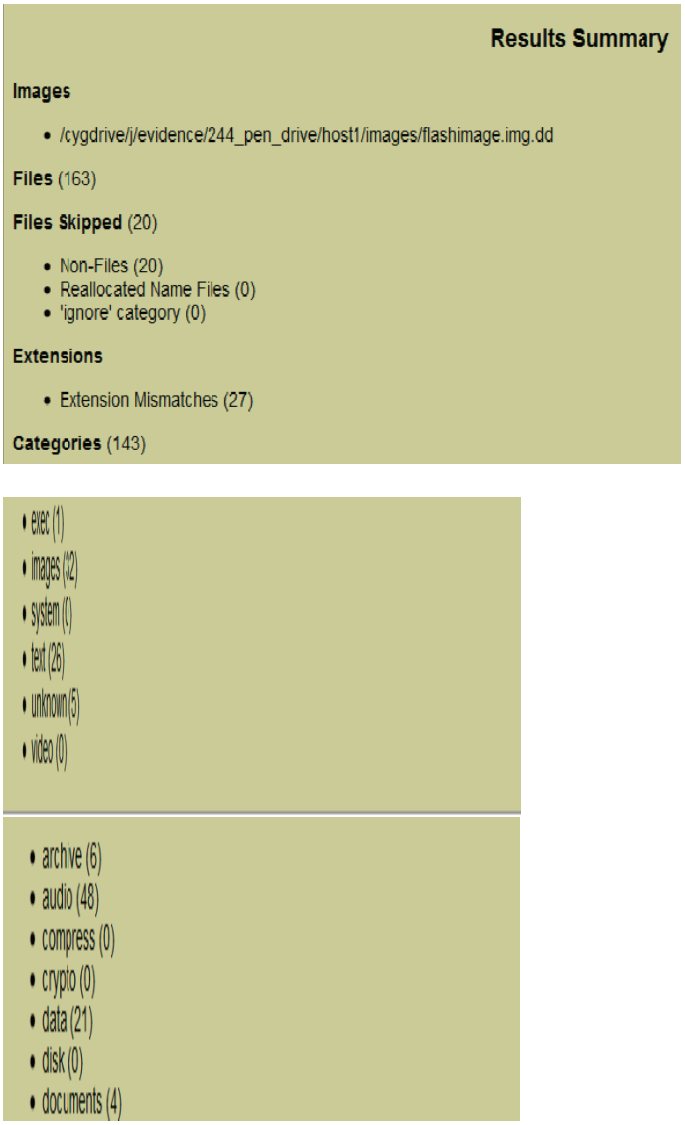


Fig4.9.3: A result of the file sorting of the pen-drive

In Fig4.9.3, For the Categories, this explains the types of files identified from the pen drive, 163 files were discovered from the file-sorting process.

Archive files found = 6, audio file = 48, compressed files = 0, crypto = 0 Data files = 21, Disk images = 0, Documents = 4

Executable files =1 Images recovered = 32

System = 0 Text = 26 Unknown = 5

Video = 0

S/No	Detail	Number(s)
1	Archive	6
2	Audio	48
3	Compress	0
4	Crypto	0
5	Data	21
6	Disk	0
7	Documents	4
8	Exec	1
9	Images	32
10	System	0
11	Text	26
12	Unknown	5
13	Video	0
	TOTAL	**143**

Table 4.0: Tabular form of report in Fig4.36 summarizing the categories of files found after the File sorting process

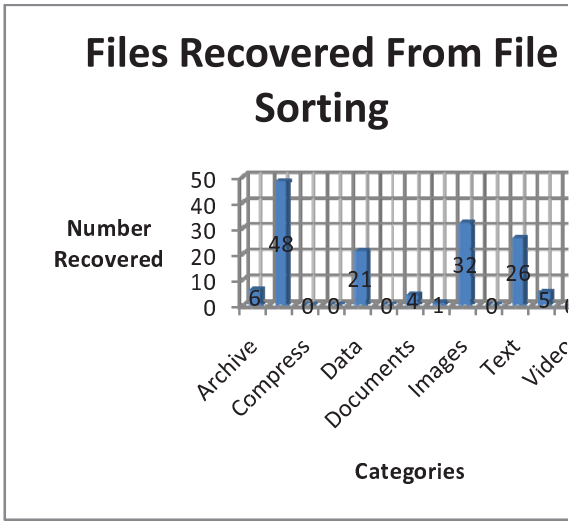


Fig4.94: Bar Chart form of report summarizing the categories of files found after the File Sorting process

4.3.5 **Interpretation of Bar Chart of Recovered Files from Pen Drive**

From Fig 4.9.4 the bar chart shows that the audio files have the highest occurrence rate of 50 files among the recovered items. The cypto, compressed, Disk, executable files, Executable files, and video files were the lowest with zero files recovered.

4.3.6 **Comparing Foremost and Sleuth Kit/Autopsy Recovered Items From Pen-Drive**

the similar search attributes of Foremost and Sleuth Kit named below and detailed in

Table 4.1:

1. Compressed items
2. Images Recovered
3. Video Recovered

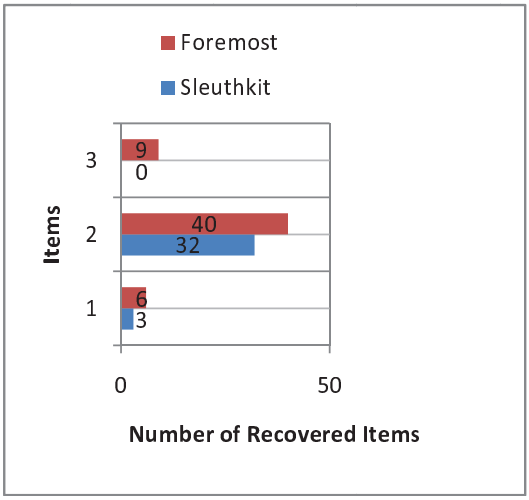
Program	Number of Compressed Items Recovered	Number of Images Recovered	Number of Video Recovered
Sleuthkit	3	32	0
Foremost	6	40	9

Table 4.1: comparison of recovered items by Foremost and Sleuthkit

a. **Analysis of Fig 4.95 Bar Chart**

The Bar Chart shows the frequency distribution of three common recovered items between Sleuth Kit and Foremost. The largest frequency of recovered items is for images, with Foremost having a higher number of 40 items as compared to 32 items of Sleuth Kit. Foremost recovered twice as much compressed items from SleuthKit.

Fig 4.95 A bar graph showing the comparative analysis of the common recovered items between Sleuthkit and Foremost



Legend of Barchart

1 = Video

2= Images

3=Compressed Files

4.4 Summary of Findings

Fifty-four files were extracted by Foremost from the pen drive. Bulk Extractor took 153 minutes to complete its operation, which finished with detailed 18 reports of its activity. Performing searches with Sleuth Kit with large devices seems slow. String extraction is encouraged to be a preliminary activity before performing a search, this make the subsequent searches faster. Foremost recovered more items than Sleuth Kit

5 Experimental Discussion

The case study presented distinct challenges, with different aspects of The Sleuth Kit and Fiwalk toolset. They were utilized to effectively perform a file system analysis. The focus of this research project was based on a case study that is employed to help demonstrate the usefulness of The Sleuth Kit, Autopsy Forensic Browser, and Fiwalk as file system Analysis toolsets.

The scope of case study employed provided a good test of the functionality of the Sleuth Kit, Fiwalk, Bulk_Extractor and Foremost toolkits. Rarely is a single forensic tool ideal for an investigation or recovery process. Therefore, a combination of tools should be applied for flexibility and faster digital forensic investigation process. One major problem affecting all forensic tools is the increasing size of storage media. This often increases the time required for a complete analysis.

The seven objectives mentioned in the chapter titled “introduction” were all achieved and some observations noted briefly discussed below.

The research found that The Sleuth Kit Autopsy Forensic Browser and Fiwalk all provide effective file system analysis toolsets. The flexibility of the tools contained within The Sleuth Kit often lead to complex command line strings, the complexity of which is overcome by the automation provided by the Autopsy Forensic Browser. Not only do The Sleuth Kit and Autopsy Forensic browser provide an effective toolset, they also offer an affordable alternative to expensive commercial or proprietary based toolsets.

Foremost was found to be a program that though compact and small, was a good software for data recovery. Foremost was found not to be complex as compared to Sleuth Kit. Foremost was useful in recovering executable files, video files.

While Sleuth Kit was useful for analyzing a storage device, its recovery feature seems more suited for text files and documents. Observed strengths lies in the “what is contained” and “where is it located” , rather than “how can I extract it?”. Also string extraction in preparation for Keyword search with Sleuth Kit and Autopsy was found to be time consuming for large storage media. The web interface is for Autopsy greatly simplified the investigative process.

The demonstration of the effectiveness of The Sleuth Kit and Autopsy Forensic Browser may be used by individuals or Law Enforcement as part of an
Evaluation, when looking to extend their current Digital Forensics toolset, either as an alternative or complement to their current tools
Bulk_Extractor was found to be very efficient in extracting emails, domain addresses, etc. It also seemed fast not minding the size of the storage device. The histogram text files recovered served as detailed and informative statistical tools.
Fiwalk was good for XML generation of a disk image. It was also fast and efficient.

6 Suggestion for Further Research Works

Due to the complexity of new and improving storage devices, an investigator would be well equipped, to be knowledgeable in Hardware and software architecture of various operating system and media devices. Also a working

knowledge of programming languages would further his area of expertise in digital forensic science. Since most of the forensic tool-kits are Linux based, a firm grasp of the physiology of multiple platforms would be an added advantage. Further research should be employed in making these tools used in this thesis to be part of a software suite. This means integrating all the tools into one program. In addition, the tools should have a way of communicating with each other, i.e. a symbiotic association between the tools.

References

Anothony Reyes, Jack Wiles (2007): Cybercrime and Digital Forensics Syngress Publishing Inc, Elsevier Inc, 30 Corporate Drive, Burlington, MA 01803; ISBN 13:978-1-59749-228-7; pgs 734

Bradford Phillip G, Brown Marcus, Perdue Josh, Self Bonnie. Towards proactive computer-system forensics. In Proceedings of the international conference on information technology: coding and computing (ITCCa^{TM}04); 2004.

Brian Carrier (2002). Open Source Digital Forensics Tools; the Legal Argument

Brian Carrier (2003).Defining Digital Forensics Examination and Analysis Tools Using AbstractionLayers. International Journal of Digital Evidence. Vol1, Issue 4, Windar, 2003. Available at: http://www.ijde.org/docs/02_winter_art2.pdf

Carv er DL Hoss AM. Weaving ontologies to support digital forensic analysis. 2009

Cohen MI. PyFlag: an advanced network forensic framework. In: Proceedings of the 2008 Digital Forensics Research Workshop. DFRWS, http://www.pyflag.net; August 2008 [accessed 06.03.09].

Corey Vicka, Peterman Charles, Shearin Sybil, Greenberg Michael S, Bokkelen James Van.Network forensics analysis. IEEE Internet Comput 2002;6 (6). ISSN: 1089-7801:60

Garfinkel,S.(2009A)"Automating Disk Forensic Processing with Sleuthkit, XML and Python". (SADFE 2009).

Garfinkel Simson L, Farrell Paul, Roussev Vassil, Dinolt George (2009). Brining Science to Digital Forensics with standardized Forensic corporal. In: Proceedings of the 9th Annual Digital Forensic Re search Workshop (DFRWS); August 2009. Grenier Christophe. Data carving log, http://www.cgsecurity.org/wiki/Data_Carving_Log n;

Gialanella, David, (2008); New Tech, Old Problem. ABA Journal, 94(8), 35

Marcella, Albert J, Menedez, Doug (2008). Cyber Forensics. Boca Raton, FL; Auerbach Publications Taylor & Francis Group

Nance Kara, H a y Brian, Bishop Matt. Digital forensics: defining a research agenda. In: Proceedings of the 42nd H awaii international conference on system sciences; 2009.

Pollitt Mark, Nance Kara, Hay Brian, Dodge Ronald C, p Craiger Phili, Burke Paul, Marberry Chris, Brubaker Bryan Virtualizati on and digital forensics: a research and education agenda. J Digit Forensic Pract 2008;2 (2). ISSN: 1556-7281:62-73.

Pollitt Mark M. An ad hoc review of digita l forensic models. In: Proceedings of the second inter national workshop on systematic approaches to digital forensic engineering (SADFE'07); 2007

Simon L. Garfinkel (2010): Digital Forensics Research: The Next 10 Years. Journal homepage: www.elsevier.com/located/diin

Turnbull Benjamin , Taylor Robe rt , Blundell Barry. T he anatomy of electronic evidence a quantitative analysis of police e-crime data. In International conference on ava i l ability, reliability and security, (ARES '09); March 16-19-2009. p.143-9. Fukuoka.