"Efficacy of open source tools for recovery of unconventionally deleted data for forensic consideration"

Vinay Singh*, Lav Kesharwani, Vaibhav Saran, A K Gupta, E.P. Lal, Amrita Verma

Sam Higginbottom Institute of Agriculture, Technology & Sciences

Deemed to be University, Allahabad

786.vinaysingh@gmail.com; lav2222@yahoo.com; dr.vaibhavsaran@gmail.com

***Corresponding Author: Vinay Singh**

ABSTRACT

Overwriting of digital data can be defined as obliteration of evidence in digital world. Most computer users today are familiar with the risk of their files being recovered even after deletion and hence many often use concept of overwriting their files instead of deleting them in order to prevent the unwanted recovery of those files. Although there are specialized tools available for the successful retrieval of unconventionally deleted data but they are highly costly and require experienced technical efficiency. In order to deal with the variations of digital evidences in eccentric situations it was necessary to have a variety of tools and not to rely on any single technology or methodology. In the present study it was found that unconventional methods of deleting the digital data were an efficient method of data destruction. The present study was done to evaluate the effectiveness of open source tools for the recovery of deleted data through the unconventional methods.

Keywords: Open Source Tools, Data Destruction, Unconventional Method, Recovery

INTRODUCTION

Computer crime can be defined as any intentional act associated in any way with computers where a victim has suffered or could have suffered a loss and a perpetrator made or could have made a gain.

Casey, (2004), [2] defined digital evidence as any data stored or conveyed using a digital device that sustain or contradict a theory of how an offense occurred or that deal with critical element of the offense such as intent or alibi.

Nelson *et al.*, (2004) [9]**,** explained that digital forensic involves scientifically investigating and scrutinizing data from digital device storage media so that the data can be used as an evidence in court.

Goel, (1985) [6], has clearly stated that software reliability can be assured if software faults do not cause a failure during a specified exposure period in a specified environment. Understandably, unpredictable digital forensic software will lead to untrustworthy results which may put at risk the whole forensics investigation.

Carrier, (2002) [1] and **Dan *et al.*, (2007),** [3] in their respective works have raised an argument whether digital forensics investigation using open source tools would be better.

Farrell (2009) [4] has clearly mentioned that open source tools that perform specific functions were persistently being developed and distributed in the academic as well as in other field of interest and most significantly these new functions were ultimately integrated into larger analysis suites. These open source tools can be Graphical user interface based programs or command line based programme that allow an analyst to explore and search the data on a hard drive.

Panchal, (2013), [10] stated that a computer operating system makes use of a directory that consists of the name and placement of each file on the drive. When a file was deleted, several events took place on the computer. While the user can no longer see the file listed in any directory, in reality nothing has been done to the file itself. Technically speaking the file location was just removed from index of the file directory or also better known as File Allocation Table (FAT) and this newly available space is called free or unallocated space and until the free space is overwritten by another file, the forensic specialist can retrieve the file partially or even completely, in some cases. Overwriting might be caused by a variety of user activities, such as adding a new program or creating new documents that happen to be written to the space where the "deleted" files exist.

WRITING ON A HARD DISK

Data is written to the hard disk drive in clusters (the default size is 512 MB) by the drive's read/write head(s) that float on a head support of air above the platters. A read/write head can never randomly move from a cluster in one track to a cluster in a different track directly beside it without rotating the platter one full turn, thus clusters are written to the hard drive in a checkerboard fashion. Often the clusters are not completely filled up with data, creating a blank space which is commonly known as "slack space." When files are stored, the operating system physically writes the files to clusters on the platters, as well as logically writing a path in the operating system. This action occurs so that the computer will know which cluster is housing a specific file. The maintenance of such a logical or registry based entry of any data is the sole purpose of file allocation table, which remembers the logical path of specific cluster or specific sector in which the file is stored. When we just delete a file or any data through conventional methods, simply the registry entry of location of the specific file is deleted but the data itself still remains present in the hard disk.

DATA DELETION

The foremost conventional effort to delete a data is to remove all of the files by simply moving the files to the recycle bin and then emptying the bin such that all the programs and files are permanently erased that is supposed to contain valuable and sensitive information. But by just deleting the personnel files and emptying the recycle bin the traces of those sensitive data cannot be completely removed. The delete function only removes file names from a directory list (also known as File Allocation Table) and frees the sectors for the next file to occupy space on the hard drive. Meanwhile, these files actually continue to exist. Apart from these the method of deletion of data, the other method which can be applied to delete the data from a hard disk is "Degaussing". In this process the hard disk is exposed to a strong and variable magnetic field, which disturbs the intrinsic (i.e. magnetic) nature of a hard disk. But it has also been seen in many cases where after successful degaussing the drive becomes unusable **Rosencrance, (2007)** [11].

DATA OVERWRITING

Wiping a computer clean is not as easy as it may appear. Recently, it has been clearly observed during research that a large number of computers easily found in secondary markets contained information such as consumer's names, credit card numbers, and social security numbers **Jones, (2005)** [7].

There are many specifications by different government agencies for the successful overwriting of data rendering it difficult to retrieve later on by the forensic experts, which are mentioned below:

- A single pass overwriting with either zeros or ones or random data.
- The 1995 DoD standard of 3 passes, the first with either ones or zeros, the second with the opposite of the first, and the third pass should write random data.
- The current DoD standard of 7 passes.
- The Guttmann standard of 35 passes using different algorithms.

MATERIALS AND METHODOLOGY

The aim of research was to evaluate the efficiency of open source softwares for the recovery of unconventionally deleted data. During this research the open source software were tested in condition of deletion of file in ways which were not generally adopted by a normal user such as wiping of the storage media, deletion of data using overwriting method and use of specialized software to delete the data.

The methods conventionally used for the deletion of files or any digital data of valuable information consist of either pressing 'DELETE' button or using 'SHIFT+DELETE' in combination.

MATERIALS USED

❖ **Workstation**

- Processor : AMD E1-2100 APU

- RAM : 4 GB
- System Type : 64-bit Operating System
- Graphics Card:1GHz Radeon™ HD
- Operating System: Windows 7 Professional
- Service Pack : Service Pack 1

❖ **Used Product**

- Kingston DataTraveler 4 Gb Pen drive
- Toshiba 4 Gb Micro SD card

❖ **Brand New Product**

- San Disk 16 Gb Pen Drive
- Transcend 1 Gb Micro SD card

❖ **Recovery Software**

- Win-Hex :Version 16.7
- Pro-Discover Basic :Version 7.0.0.3
- Recuva :Version 1.52.1086
- Stellar Windows Recovery : Version 6.0

❖ **Erasing Software**

- Eraser :Version 6.2.0.2969
- Stellar File Wipe :Version 4.1

TEST DATA USED

For the purpose of this research, a known data set was prepared so that it would facilitate the recovery process and also allow us to classify the open source data recovery products for the recovery of data deleted through unconventional method. The known data set comprises of common file types (viz . . . pptx, .xlsx, .docx, .zip, .rar, .txt, .pdf, bmp, .pub, .jpg, .psd, .mp3, .mp4) as illustrated in the table below:

Table 1: Data set consisting of Document files and Media files

Documents Files				Media Files			
File No.	**File Name**	**File Extension**	**File Size**	**File No.**	**File Name**	**File Extension**	**File Size**
1.1	Department	.pptx	1.44 Kb	**1.1**	AB-1	.bmp	1.47 Mb
1.2	Kotwali	.pptx	4.70 Mb	**1.2**	AB-2	.bmp	1.80 Mb
1.3	Rubber Industry	.ppt	7 Mb	**1.3**	AB-3	.bmp	1.11 Mb
1.4	FP Analysis	.ppt	6.45 Mb	**1.4**	AB-4	.bmp	1.17 Mb
2.1	Exam Report	.xlsx	12 Kb	**2.1**	Cake	.pub	273 Kb
2.2	List of items purchased	.xlsx	12 Kb	**2.2**	Bday card	.pub	941 Kb
2.3	Bank Statement	.xlsx	12 Kb	**2.3**	Daniel	.pub	562 Kb
2.4	Electricity Bill	.xlsx	9.32 Kb	**2.4**	Card Front	.pub	251 Kb
3.1	Reema	.docx	16 Kb	**3.1**	DSC_0247	.jpg	12.5 Mb
3.2	Seema	.docx	16 Kb	**3.2**	DSC_0247	.jpg	12.2 Mb
3.3	Insane	.docx	12.7 Kb	**3.3**	DSC_0146	.jpg	12.2 Mb
3.4	Assignment	.docx	27.4 Kb	**3.4**	DSC_0143	.jpg	12.7 Mb
4.1	Woodbury_lo	.zip	6.30 Mb	**4.1**	Ram	.psd	874 Kb
4.2	Sternberg	.zip	3.52 Mb	**4.2**	Kids	.psd	102 Kb
4.3	Chawla	.rar	10 Mb	**4.3**	Shayam	.psd	75.8 Kb
4.4	McCullo	.zip	13.5 Mb	**4.4**	Eyes	.psd	386 Kb
5.1	Traffic Jam	.txt	190 Bytes	**5.1**	Ring-A	.mp3	1.06 Mb
5.2	Sleeping & Walking	.txt	4 Kb	**5.2**	Ring-B	.mp3	554 Kb
5.3	Changing Mood	.txt	706 Bytes	**5.3**	Ring-C	.mp3	258 Kb
5.4	Way of living	.txt	563 Bytes	**5.4**	Ring-D	.mp3	454 Kb
6.1	Bath Bombs	.pdf	228 Kb	**6.1**	VID-1	.mp4	2.97 Mb

6.2	Blast Shelter	.pdf	457 Kb	**6.2**	VID-2	.mp4	1.58 Mb
6.3	Explosive	.pdf	176 Kb	**6.3**	VID-3	.mp4	277 Kb
6.4	Smoke Bomb	.pdf	214 Kb	**6.4**	VID-4	.mp4	4.57 Mb

METHODOLOGY

The methodology used for this study was divided into four sections. These were the Sanitization Phase, Erasure Phase, Recovery Phase and finally Software Phase respectively. These different Phases are explained in further details:

- ❖ **<u>Sanitization Phase</u>:** In this phase the drives selected for the performation of this experiment would be sanitized properly in accordance to the method of Department of Defence (DoD). In this way we can ensure the data that existed previously, (if any), was totally wiped off and there was no interference in the successful carrying out of our examination.
- ❖ **<u>Erasing Phase</u>:** This was the phase in which the predefined set of data was stored in the external storage device and then erased using the overwriting method as suggested by US Department of Defence and Guttmann.
- ❖ **<u>Recovery Phase</u>:** This was the most crucial phase in our experiment where the deleted predefined set of experimental data will be recovered using the open source softwares, downloaded from their respective sites.
- ❖ **<u>Software Phase</u>:** During this phase, the experimental data will be deleted using the specialized software that claim to successfully delete all the data and then the deleted data will be recovered using open source tools.

RESULT

During this experiment 48 test files of various extensions comprising of both document files and media files were used. The following result was obtained when deleted data was being recovered through various Recovery Softwares:

Table 2: Files Recovered Through Different Softwares in Brand New Products

Mode of deletion	**Brand New Products**							
	Stellar Windows Recovery		**Recuva**		**Pro Discover**		**Win- Hex**	
	16 GB	**1 GB**	**16 GB**	**1 GB**	**16 GB**	**1 GB**	**16 GB**	**1 GB**
Normal Delete	48/48	48/48	48/48	48/48	48/48	48/48	20/48	20/48
Shift Delete	48/48	48/48	45/48	45/48	48/48	48/48	20/48	20/48
1 Overwrite pass	0/48	0/48	0/48	0/48	0/48	0/48	0/48	0/48
Eraser	0/48	0/48	0/48	0/48	0/48	0/48	0/48	0/48
Stellar	0/48	0/48	0/48	0/48	0/48	0/48	15/48	15/48

Table 2 revealed that the recovery through different software like Stellar Windows Recovery, Recuva and Pro Discover, recovery was possible and it was seen that it successfully retrieved the data deleted through the process of Normal and Shift Deletion which was in compliance to results obtained by **Sansurooah *et al.*, (2013)** [12]. But it was observed that files and its content were recovered along with few modifications like changes in the name of the files (especially media files),the existence and date of creation whereas these softwares fails to recover deleted data through the process of overwriting and by the process of specialized wiping software Eraser and Stellar. Win-Hex was unable to retrieve all the test data files due

to the limitation of the software to recover files of size greater than 200Kb. It was observed that file and its content were recovered along with the changes in the name of the files (especially media files). Win-Hex software fails to recover deleted data through the process of overwriting and by the process of specialized wiping software Eraser. But it showed the existence and date of creation and modification when the data was deleted through Stellar File Wipe, which was in resemblance to results obtained by **Kuepper, (2002)[8]** and **Geiger, (2005)** [5]. Although it showed the file as recoverable but when it was recovered we were unable to open the file, which means either it was damaged or corrupted.

Table 3: Files Recovered Through Different Softwares in Used Products

Mode of deletion	**Used Products**							
	Stellar Windows Recovery		**Recuva**		**Pro Discover**		**Win- Hex**	
	8 GB	**4 GB**	**8 GB**	**4 GB**	**8 GB**	**4 GB**	**8 GB**	**4 GB**
Normal Delete	48/48	48/48	48/48	48/48	48/48	48/48	20/48	20/48
Shift Delete	48/48	48/48	45/48	45/48	48/48	48/48	20/48	20/48
1 Overwrite pass	0/48	0/48	0/48	0/48	0/48	0/48	0/48	0/48
Eraser	0/48	0/48	0/48	0/48	0/48	0/48	0/48	0/48
Stellar	0/48	0/48	0/48	0/48	0/48	0/48	15/48	15/48

Table 3 reveals that the recovery through different software like Stellar Windows Recovery, Recuva and Pro Discover, recovery was possible and it was seen that it successfully retrieved the data deleted through the process of Normal and Shift Deletion which was in compliance to results obtained by **Sansurooah *et al.*, (2013)** [12]. But it was observed that files and its content were recovered along with few modifications like changes in the name of the files (especially media files), the existence and date of creation whereas these softwares fails to recover deleted data through the process of overwriting and by the process of specialized wiping software Eraser and Stellar. Win-Hex was unable to retrieve all the test data files due to the limitation of the software to recover files of size greater than 200Kb. It was observed that file and its content were recovered along with the changes in the name of the files (especially media files). Win-Hex software fails to recover deleted data through the process of overwriting and by the process of specialized wiping software Eraser. But it shows the existence and date of creation and modification when the data was deleted through Stellar File Wipe, which was in resemblance to results obtained by **Kuepper, (2002)** [8] and **Geiger, (2005)** [5]. Although it shows the file as recoverable but when it was recovered we were unable to open the file, which means either it was damaged or corrupted.

Table 4: Comparison of features of open source recovery softwares

TOOLS / **FEATURES**	**Stellar Windows Recovery**	**Recuva**	**Pro-Discover**	**Win-hex**
Recovery after Normal Delete	✓	✓	✓	✓
Recovery after Shift Delete	✓	✓	✓	✓
Recovery after Deletion through 'Overwriting'	✗	✗	✗	✗
Recovery after Deletion through 'Erasure'	✗	✗	✗	✗
Recovery after Deletion through 'Stellar File Wipe'	✗	✗	✗	✗

Preview file prior to recovery	✓	×	✓	✓
Filename remains unaltered	×	×	×	×
File's Existence after deletion	×	×	✓	✓
File Information (Size, Creation & Modification Date)	✓	×	✓	✓
Operational Ease	✓	✓	✓	×
Special Functions (Imaging, Hash Calculation)	✓	×	✓	✓
File Size Recovery Liberation	✓	✓	✓	×

DISCUSSION

Few study suggested that various governmental agencies can recover data that has been overwritten any number of times, but most data recovery companies claim they cannot recover data that has been overwritten even once.

In the lights of the result obtained through this research work, it can be said in compliance with the data recovery companies that even one overwrite pass can delete the data and render it unrecoverable through the Open source tools.

The result obtained in this research work are also in accordance with **United States Defense Security Service, (2006)** [13], which claims that deleting the data through one pass overwriting with either one or zero or any random number renders the data insufficient to be retrieved back.

The results obtained during this research work are also in compliance with **Sansurooah *et al.*, (2013)** [12], who had claimed in his earlier works that simply deleting the data in question was not enough to ensure that the data cannot be recovered.

In accordance to the claim made by **Kuepper, (2002)** [8] **and Geiger, (2005)** [5] which states that wiping software may fail to erase the file associated with slack space, in this study it was able to retrieve the header of the deleted file, which shows the existence of the file at some point of time in the system.

CONCLUSION

On the basis of present study it can be concluded that open source tools can be effectively used as a preliminary examination tool at the lower level of investigation as it successfully reveals the existence of file through their header file along with date of their creation and their modification. The best advantage of implementing such tools in digital forensic for the purpose of investigation was that it will decrease the stress and pendency's of cases at higher level. Hence it can be stated that open source tools although unable to recover data in all the eccentric situations still can be used efficiently due to the following reasons:

- ❖ No cost of purchasing the tools.
- ❖ Can be downloaded from any open source platform.
- ❖ Less specialized training required to operate the softwares.

Open source recovery softwares reviewed throughout this research work, (viz Stellar Windows Recovery, Recuva, Pro-Discover and Win-Hex) showed the potential to recover the deleted data.

On the basis of the parameters described in **Table 4**, the following inference can be drawn in relation to the effectiveness of the data recovery softwares evaluated:

Pro discover was best suitable option available for the recovery of data deleted through various means and method. The advantage of using this tool over others can be stated as below:

- ❖ Retrieved the deleted files in maximum cases.
- ❖ Accurate file information of the test data set used was evident.
- ❖ Allows the investigator to open and check the file even before recovering it.
- ❖ In cases, where file could not be recovered, it showed its existence in the system.
- ❖ Information about clusters of the physical

drive and algorithm used to delete the data could be seen easily.

Win Hex can be regarded as an effective recovery tool as it displayed the potential to recover most of the files in eccentric situations, but was unable to do the same due to size limitation of file to be recovered. Win Hex being a hex editor tool also gives privilege to the investigator to look into the binary values of file but it requires expertise training.

REFERENCE

1. **Carrier, B., (2002). "**Open Source Digital Forensic Tools: The Legal Argument." http://www.digital-evidence.org/papers/opensrc_legal.pdf
2. **Casey, E., (2004)** "Digital evidence and computer crime forensic science, computers and the internet" 2nd Edition P 101 *London: Academic Press*.
3. **Dan, M., Anna, C., Steve, R., Alain, G., Matthew, K., and Jeremy, T., (2007)** "Is the Open Way a Better Way? Digital Forensics Using Open Source Tools." Paper presented at the System Sciences, *40th Annual Hawaii International Conference*.
4. **Farrell, P., (2009)** "A framework for Automated Digital Forensic Reporting" Available at: http://cisr.nps.edu/downloads/theses/09thesis_farrell.pdf
5. **Geiger, M., (2005).** "Evaluating Commercial Counter-Forensic Tools." Digital Forensic Research Workshop. http://www.dfrws.org/2005/proceedings/geiger couterforensics.pdf
6. **Goel, A. L., (1985)** "Software Reliability Models: Assumptions, Limitations, and Applicability", *IEEE Transactions on Software Engineering*, SE-11(12), 14111424.
7. **Jones, A., (2005)**. "How Much Information Do Organizations Throw Away?" *Computer Fraud and Security*. **5**, (3), 4-9.
8. **Kuepper, B., (2002).** "What You Don't See On Your Hard Drive". *SANS Security essentials GSEC Practical Assignment.*
9. **Nelson, Bill, Phillips. A, Enfinger. F and Steuart. C, (2004)** Guide to Computer Forensics and Investigations.
10. **Panchal, E. P., (2013),** "Extraction of Persistence and Volatile Forensics Evidences from Computer System," *International Journal of Computer Trends and Technology (IJCTT),* available at: http://www.ijcttjournal.org
11. **Rosencrance, L., (2007)**. "EBay Auction Yields Drive Holding Political Data". Computer World. accessed from http://www.computerworld.com
12. **Sansurooah. K, Hope. H, Almutairi. H, Alnazawi. F and Jiang. Y, (2013**), "An investigation into the efficiency of forensic data erasure tools for removable usb flash memory storage devices", originally published in the Proceedings of the *11th Australian Digital Forensics Conference*. available at http://ro.ecu.edu.au/adf/127
13. **United States Defense Security Service, (2006)** National Industrial Security Program Operating Manual (NIPSOM). Washington: GPO.