Erej *et al.LGU (IJECI)* 2022

LGU (IJECI)
ISSN: 2522-3429 (Print)
ISSN: 2616-6003 (Online)

LGU International Journal for
Electronic Crime Investigation

Research Article

Vol. 6 issue 2 Year 2022

Data Carving - The Art of Retrieving Deleted Data as Evidence

Fatima Fatima and Erej Azeem
Government College University, Lahore
erejazeem00@gmail.com,

Abstract:
This paper proposes an approach to extract the hidden information as sensitive data can be hidden by the criminal in free space or slack space. But their might be cases when there exists no file system meta data information for file recovery. For this purpose Data Carving techniques are used by forensic examiners. If a file is carved in a forensically sound manner, it is then acceptable in the court of law. Many automated tools exist to carve data out of a hard drive. In this paper we looked on how to carve data in an old fashioned way followed by carving data using tools.

Keywords: Digital Forensics, Hard disk forensics, Data Carving, File system, Autopsy, Forensic Explorer, Unallocated space, file signature, manual extraction

1. Introduction

Today we regard ourselves as in a computerized world, where most data is made, caught, communicated, put away, and handled in computerized structure. From house to the industry, computers have become a everyday thing and there has been a massive increase in the crimes associated with it. This is where Digital Forensics come in. The focus of the forensic science discipline known as "digital forensics" is the use of digital information produced, saved, and conveyed by computers as a source of evidence in investigations and legal procedures. Digital forensic analysis makes it possible to recognize the type of crime committed and the culprit behind the crime. The main source of evidence against such crimes is the computer hard disk [1]. Investigator can perform forensic investigation of the file system on the hard disk to gather evidence against the criminal. Suspect might hide some sensitive information in the free space or the slack space of the file system therefore there is a need of forensic investigation of these spaces to retrieve the sensitive information [2]. Data storage on hard disk drive are organized by the file system, it is responsible for allocating free space to the files and to keep track of those files. To extract data

from unallocated space examiners use the technique of Data carving. Even though it is advantageous to comprehend the procedure and have the ability to carry it out manually if necessary, there are a lot of utilities that can carry out this activities for us.

There are different situations in which we carry out the process of Data carving. Such as, if you open a file, it can appear to be one thing when in reality it's two things since the contents may have been stuffed into another file. You might have data that is simply stored outside the reach of several widely used operating system functions in situations where you have forked data, such as an alternate data stream or a resource fork [3]. We therefore carry out data carving. However, in order to extract data from a hard drive, we must first understand what we are looking for and how far to search. Since we are dependent on byte patterns on the drive, adopting this data carving strategy has several drawbacks. It cannot be guaranteed that byte patterns like "ff d9" will only appear in the files we are looking for. The most frequent instance of carving in an actual investigation is an attempt to recover deleted data for which the associated metadata is either missing or no longer linked.

2. Data Carving

A wise forensic examiner once said "when all else fails, we carve". Data carving, also known as File carving is a forensic procedure that is used for reconstructing files in unallocated space. It is also an effective skill for the recovery of deleted data from unallocated memory space by locating the file signatures or the magic numbers.

There are several issues that an examiner may face during the process of Data Carving. Such as, the majority of tools would be able to locate the header, but it is possible that the footer would not be in the same or subsequent cluster. Thus, the carved file may be incomplete and may not be viewable. Data carving process can also result in many false hits, therefore, Footer analysis may reduce this problem. For several reasons file system information may not be available at some point, digital media could have been formatted to destroy the file system or a specific file might have been removed or deleted such that the file system indexes stop referring to the file content or maybe a file might be hidden in slack space or unallocated clusters [4]. A file recovery process or the process of recovering deleted data by locating the file signatures is basically known as File carving or Data carving. Raw bytes of the disk are scanned for the file carving process and then reassembled by examining the file signature.

File Signatures are also termed as **Magic numbers** which is a constant used for identifying the file format and distinguishing between the file formats, this means two different file types cannot have same File signatures. For example a JPEG file begins with "**0xFFD8**" and ends with "**0xFFD9**". These constants are called File signatures or Magic numbers. File signatures can be altered which can result in fake file type identification.

Depending upon the situation, various tools for Data carving can be used by the examiner. However, it is required that the examiner clearly understands the features of the tool and has a clear concept of carving a file. Currently, many Data carving commercial tools exists.

Namely, Encase, Win-hex, Access-data FTK, foremost, scalpel and many more.

If we talk about Windows Operating System, space is allocated on the hard disk drive as adjoining sectors in the form of group which are also known as clusters or allocation units. Whenever a new file is created, available space is found by the system and that space is allocated to that file. Unallocated space can be defined as the space that is not allocated within the file system to the active files [5]. This space is also sometimes called free space and on a hard drive it is a logical space on which Windows operating system can write to. It is basically opposing of allocated space, where files are already written by the operating system.

Data might by hidden by the criminal in slack space. The remaining storage on a computer's hard disk drive is the Slack space and it when an operating system allocates a space to the file and all of that space is not needed by the file. In other words, The portion after the end of a file but before the end of a cluster or block is usually referred to as slack space. In computer forensics, the analysis of Slack space is a very important aspect. In forensic investigations, it can be an important form of evidence. For example, if an entire hard disk cluster is filled by a file and the user deletes that file, and then a new file is saved by the user that does not entirely fill the hard disk cluster then the leftover space would not be necessarily be empty, it may contain data from the deleted file. Forensic examiners can extract that information using computer forensics tools.

In the following figure 1, one cluster is shown that is of 4096 bytes containing 8 sectors of 512 bytes each represented as S1, S2, S3 .. S8 respectively.

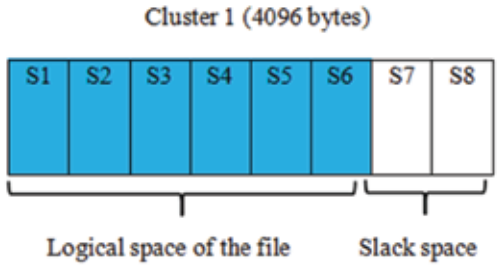

Figure 1: Slack space explanation

To destroy the evidence, file deletion is one of the effortless way. Whether by using "DELETE" or "SHIFT + DELETE" button. Whenever the file is being deleted, the contents of the file are not wiped. Windows use the concept of pointers to track where the files exist on the hard disk drive. Every file and folder has a pointer. When a file is deleted, the pointer is removed and the sectors that contain the data of the file are marked as "available". The file content is recoverable until and unless those sectors are overwritten. If a person does not wants his/her data to be recovered, they can use tools that wipes hard disk drives free space.

2.1 Difference between File Recovery and File Carving

One might be confused between the term file recovery and data carving. By using some forensics tool we can recover the deleted file until and unless it is not overwritten by any other file [6]. File system information is used for the purpose of file recovery and many files can be recovered by using the file system information. Whereas, file carving or data carving works on raw data. File system information is

not used during the process of file carving. In Data carving, a file is recovered on the basis of content and the structure of the file without the involvement of any matching file system meta data. If there is a case like corrupt directory entries or the missing directory entries, file carving technique is effectively used. All we are looking at when it comes to data carving is a collection of bytes on a disc. It basically comes down to finding the drive's header bytes and then just extracting data from succeeding bytes.

2.2 Importance of Data Carving

Data carving is a crucial aspect in Digital forensics because it is a considerate technique in detecting a deleted file. A file can be hidden or concealed anywhere in areas like lost clusters, slack space, unallocated clusters of the hard disk drive. Forensic investigators may be sometimes in a situation where they are required to recover data. But why Data Carving? When the data is there, but can't be correctly interpreted due to absent or damaged meta-data. For examples: File system corruption, Device formatting, Unknown proprietary formats, Files removed or deleted (unintentionally or intentionally).

Traditional data recovery techniques are based on the file system information and the metadata information is used to recover deleted files. However their might be some cases where there is no metadata information available and for such cases advanced forensic techniques are used such as Data carving. Files can be recovered through data carving as long as they are not overwritten.

A file system can store information in a variety of methods, and that information and data may persist long after the user thinks it has gone. This is crucial information for forensics experts because a suspect can try to conceal evidence, so you'll need to know not just where it might still be but also how to get to it. Data carving skills are crucial for this reason.

3. Acceptable Evidence

Any document, physical item or testimony that can be used for proving a fact under the rules of evidence in a court of law. The admissibility of the evidence depends on various components. Those evidence that does not fall under the law of evidence are entitled as "inadmissible". During file carving there is a complication that has to do with fragmented files. By fragmented files we mean a file that is stored on the disk at two or more different physical locations. And some techniques cannot reconstruct these types of files. Carving is done on the basis of file signatures and unfortunately not all file types have a standard footer signature and therefore locating the end of file is difficult. But if a file is carved in a forensically sound manner, it is then acceptable in the court of law [7]. Forensic soundness provides assurance that during the investigation the evidence was not destroyed or corrupted.

4. Manual Data Carving

Many automated tools exist to carve data out of a hard drive. In this section we will see how an investigator or examiner can perform manual Data carving by locating the file signatures. Basically, We will look on how to carve data in an old fashioned way. To extract data we will use UNIX utilities [8]. First we created an image of a 4GB USB rather than using a raw

partition using the command;

sudo dd if=/dev/sdb1 of=image.dd

Figure 2: Creating Forensic image

Now we will find the headers of PNG file using the command;

grep -oba IHDR image.dd

And after running this command we get number of headers in the image file. “IHDR” is the PNG file header.

Figure 3: Checking Headers of PNG file

And similarly we can find the footers of PNG file using command;

grep -oba IEND image.dd

The ”IEND” chunk must appear LAST. It marks the end of the PNG data stream.

Figure 4 shows how we find out our starting sector, ending sector and block size.

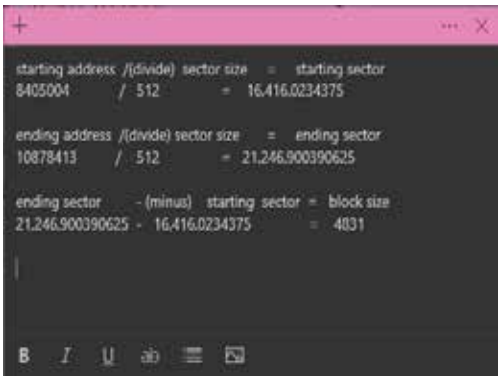

Figure 4: Starting/ending sector, block size

Now we have all the information about our PNG file. let’s carve our PNG file from image.dd file using command;

dd if=image.dd of=img.png bs=512 skip=16416 count=4831 command.

Figure 5: Carving file

We successfully extracted the PNG file from image.dd File named img.png.

Figure 6: File extracted

5. Data Carving using HEX Editor

File carving is a technique for getting erased or reassembled computer files. It entails looking for a file within a data stream. This step is important in digital forensics because the forensics expert must examine all of the file system files and check for any deleted or formatted files that need to be further investigated. We can use any HEX editor such as WinHex and HxD to perform the process of Data Carving. Download the image file from here;

http://sceweb.sce.uhcl.edu/abeyseker-a/ITEC4381/images/Mantooth.E01

Steps to perform Data Carving using HEX Editor;

i. Our objective in manual data carving is to locate a JPG file. Now in the HEX we will find header of a JPG file by simply right clicking or "CTRL + F" and the header for a JPG file is "FF D8".

Figure 7: Location JPG header

ii. We got the starting point and on the left side we can see the digital offset "90640896" in figure 3.

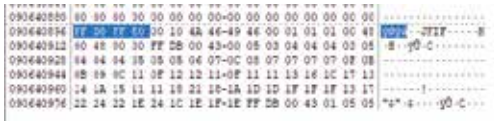

Figure 8: Header of a JPG file

iii. Now the next step is to find the footer for the JPG file. The footer for the JPG file is "FF D9" and we will follow the same procedure we followed for locating the file header that is by simply right clicking or "CTRL + F".

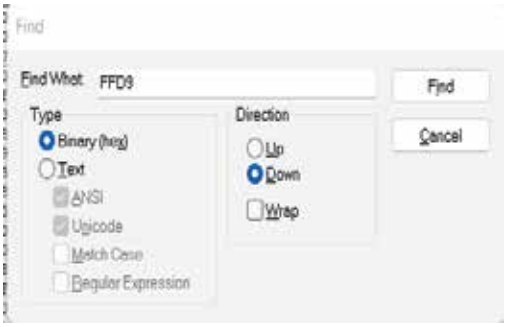

Figure 9: Locating the footer of JPG file

iv. We have successfully located the footer of the JPG file, that is the Ending of the file. As it can be seen in figure 5, "FF D9" has been highlighted with its digital offset 90659280". In other words we have got the starting and the ending point of the JPG file.

Figure 10: Footer of the JPG file

v. If we notice figure 3 and figure 5, we have digital offset 90640896 in figure 3 and digital offset 90659280 in figure 5. We will subtract 90640896 from 90659280 and we get 18400 which is the selection size of the JPG file.

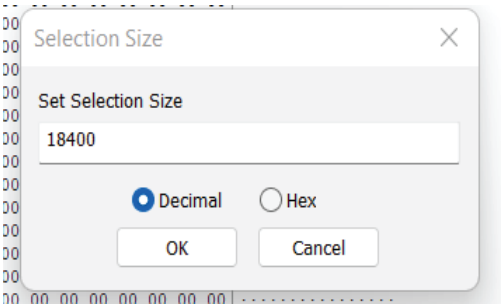

Figure 11: Defining the selection size of the JPG file

vi. We then save the file with JPG extension. And file is therefore recovered. Viewing the JPG file, we can see this is one of the JPG file that was also carved using Forensic Explorer and Autopsy software.

Figure 12: JPG File recovered by manual Data carving process.

5.1 Data Carving using other Tools

Tools like Forensic Explorer and Autopsy can be used for the purpose of Data Carving. Forensic Explorer is a tool that can be used by both novice and experienced investigators. This tool provides easy to use graphical user interface (GUI) followed by keyword search, data recovery, script technology, sort and filter. Large volume of data is quickly processed, complex investigative tasks are efficiently automated, detailed reports are produced and productivity is thereby increased. On the other hand, A digital forensic platform, Autopsy is used by corporate examiners and investigators, law enforcement agencies and military for digital forensic investigations for data carving.

Carving support offered by Forensic Explorer is for mare that 300 file types. It supports Cluster based file carving, Sector based file carving and also the Byte based file carving. In FAT or NTFS, which are the cluster based file system, a new file must start in a new cluster. The file signature then appears near the file boundary so the file signatures are therefore searched near the file boundaries and the carving speed is then achieved. Their might be some situations in which performing lower level search for sector aligned file signatures may be advantageous, additional files can be recovered. Time needed to complete search is increased when carving in sector mode. And in some situations, carving data on byte by byte level may be of great importance [9]. A byte based data carve is used when we are searching for a file that exists within a file. For example, within a backup file. A **Robust Pascal Scripting** engine is offered by Forensic Explorer where writing of data carving scripts is possible.

Whereas, Autopsy comes with a module named **Photo Rec Carver** that is used for Data carving and carves files from unallocated space [10]. The module works on the same principle of locating the File signatures i.e. headers and footers. For using this module the examiner just needs to select the checkbox in the ingest module settings and then the Photo

Rec Carver is enabled. Under the tree of Data Sources, results of carving are shown with the heading "$CarvedFiles".

5.1.1 Comparison between Carved files using Forensic Explorer and Autopsy

Following table shows a comparison of carved files between Forensic Explorer and Autopsy using the same Forensic image "mantooth".

	Forensic Explorer	Autopsy
GIF	0	1
REG	0	2
JPG	5	5
DOC	6	6
TXT	368	0
DAT	2	0
NTFS	1	0
PNG	0	1

It can be seen that 1 GIF, 2 REG, 5 JPG, 6 DOC and 1 PNG file is carved by Autopsy. Whereas, 5 JPG, 6 DOC, 365 TXT, 2 DAT and 1 NTFS files are carved by Forensic Explorer. A plus point for Forensic Explorer is that it provides a Disk view and a category graph which provides convenience for the examiner to categorize the file types.

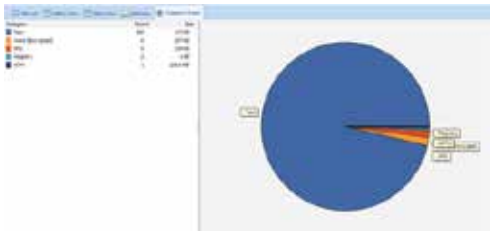

Figure 8: Forensic Explorer Category graph

6. Conclusion

Digital forensics, field of file recovery is still growing and has made progress in the recovery of transient data. In conclusion, this paper covered the basics of Data carving in which we described data carving along with other important concepts of slack space, unallocated space and Magic numbers. We performed file carving using the tools and HEX editor and on the other hand we also touched the area of manual data carving using UNIX utility. Furthermore advancement of tools for data carving process will have a greater impact. However, data carving remains a beneficial technique for the recovery of files and potential evidence without using any file system meta data information during digital investigations. There is a lot of research yet to be done in the area of data recovery. File carving technique is greatly used by the forensic experts and the examiners to squeeze every bit of data out of the media. However, carving is impossible if a new file has already been overwritten in the unallocated area where the old file was placed.

7. References

[1] Povar, Digambar, and V. K. Bhadran. "Forensic data carving." International Conference on Digital Forensics and Cyber Crime. Springer, Berlin, Heidelberg, 2010.

[2] Meshram, Bandu B., and Dinesh N. Patil. "Digital Forensic Analysis of Hard Disk for Evidence Collection." International Journal of Cyber-Security and Digital Forensics, vol. 7, no. 2, Apr. 2018, pp. 100+. Gale Academic OneFile,

link.gale.com/apps/-doc/A568570241/AON-E?u=anon~9c67264d&sid=-googleScholar&xid=28bb2ba3. Accessed 23 Aug. 2022.

[3] Alherbawi, Nadeem, Zarina Shukur, and Rossilawati Sulaiman. "Systematic literature review on data carving in digital forensic." Procedia technology 11 (2013): 86-92.

[4] Povar, Digambar, and V. K. Bhadran. "Forensic data carving." International Conference on Digital Forensics and Cyber Crime. Springer, Berlin, Heidelberg, 2010.

[5] Povar, Digambar, and V. K. Bhadran. "Forensic data carving." International Conference on Digital Forensics and Cyber Crime. Springer, Berlin, Heidelberg, 2010.

[6] A. Pal and N. Memon, The Evolution of File Carving, IEEE Signal Processing Magazine, no.March, pp. 59-71, 2009.

[7] Meyers, Matthew, and Marc Rogers. "Computer forensics: The need for standardization and certification." International Journal of Digital Evidence 3.2 (2004): 1-11.

[8] Cantrell, Gary D., and Joan Runs Through. "Teaching Data Carving Using The Real World Problem of Text Message Extraction From Unstructured Mobile Device Data Dumps." Journal of Digital Forensics, Security and Law 14.4 (2020): 4.

[9] Beek, Christiaan. 2011. Introduction to File carving. McAfee

[10] Pahade, Raj Kumar, Bhupendra Singh, and Upasna Singh. "A survey on multimedia file carving." International Journal of Computer Science & Engineering Survey 6.6 (2015).