Joakim Kävrestad

Fundamentals of Digital Forensics

Theory, Methods, and Real-Life Applications

Springer

Joakim Kävrestad

Fundamentals of Digital Forensics

Theory, Methods, and Real-Life Applications

Springer

Joakim Kävrestad
School of Informatics
University of Skövde
Skövde, Sweden

ISBN 978-3-319-96318-1 ISBN 978-3-319-96319-8 (eBook)
https://doi.org/10.1007/978-3-319-96319-8

Library of Congress Control Number: 2018948608

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Introduction

This is a book written for the sole reason that when I wanted to hold a course on digital forensics, I could not find a textbook that seemed to fulfill my requirements. What I needed a book to cover was:

- Sound forensic thinking and methodology
- A discussion on what computer forensics can assist with
- Hands-on examples

My answer to my own needs was, well, to write my own book. It has become obvious to me that writing a book that fulfills those demands is not a very easy task. The main problem lies within making proper hands-on examples. For that reason, I decided to put emphasis on what digital forensics is at its very core, and to make this piece of literature relevant worldwide, I have tried to omit everything that only seems relevant in a certain legislation. That being said, this is the book for you if you want to get an introduction to what computer forensics is, what it can do, and of course what it cannot do. It did feel good to use some sort of well-known forensic software for the examples in this book. Since forensic software can be quite expensive, I decided to use two options interchangeably. The first collection of tools are the proprietary AccessData Forensic Toolkit that was chosen for the sole reason that AccessData provides the ability to get certified, free of charge, at the time of writing. Using the predecessor of this book in teaching shows that this book can in fact be used to prepare for the AccessData certification test. Further, this book uses a collection of various open source or otherwise free tools that can accomplish the same as the proprietary AccessData tools.

This book begins with setting the stage for forensics examinations by discussing the theoretical foundation that the author regards as relevant and important for the area. This section will introduce the reader to the area of computer forensics and introduce forensic methodology as well as a discussion on how to find and interpret certain artifacts in a Windows environment. The book will then take a more practical turn and discuss how's and why's about some key forensic concepts. Finally, the book will provide a section with information on how to find and interpret several artifacts. It should at this point be noticed that the book does not, by far, cover every single case, question, or artifact. The practical examples are rather here to serve as demonstrations of how to implement a forensically sound

way of examining digital evidence and use forensic tools. Throughout the book, you will find real-world examples where I provide examples on when something was used or important in a real-world setting.

Since most computers targeted for a forensic examination are running some version of Windows, the examples and demonstrations in this book are presented in a Windows environment. Being the most recent flavor of Windows, Windows 10 was used. However, the information should to a very large extent be applicable for the previous version of Windows.

Also, most chapters in the book come with a "Questions and tasks" section. Some are questions with a right or wrong answer, and some are of more exploratory nature. Whatever the case, answers or discussions are found in Appendix A—Solutions. Complementing the book, there are video lectures covering most of the book content available for viewing at YouTube: https://www.youtube.com/playlist?list=PLEjQDf4Fr75pBnu8WArpeZTKC9-LrYDTl.

Happy reading!

Part I
Theory

Now that the book has kept you interested this far it is time to discuss what digital forensics actually is. This will be done in a very theoretical manner but I have tried to keep it short. This part begins with an overview of what digital forensics and cybercrime is, before discussing some computer theory that is necessary for a forensic examiner to be familiar with and highlight how to find and interpret forensic artifacts that the author deems to be common and important. The final chapters will discuss how to collect digital evidence in a structured manner, how to analyze digital information and write forensic reports.

What Is Digital Forensics?

1

So then, what is digital forensics? Well, the most simple explanation could be that it is the examination of digital storage and digital environments in order to determine what has happened. "What has happened" in this context could be whether or not a crime was committed, whether or not someone remote controlled a certain computer, when a picture was taken or if a computer was subject to intrusion. That being said, it can be basically anything.

However, looking at the target of some actual forensic investigations it is evident that saying "What has happened" is not covering the entire field of computer forensics since forensic examiners also look into what is currently happening. There have, for instance, been several cases in Sweden and globally were forensic examiners monitored network traffic in order to capture data that was later used to identify sexual predators. There are also situations when forensic examiners, during house searches, record what is currently happening on a computer. The case of using digital forensics to monitor activity in real time may be even more apparent in the corporate world where it is common to examine intrusion attempts and malware behaviors as it is happening.

Looking to the scientific community, Reith et al. (2002) described digital forensics in the following way:

> Digital forensics is a relatively new science. Derived as a synonym for computer forensics, its definition has expanded to include the forensics of all digital technology. Whereas computer forensics is defined as "the collection of techniques and tools used to find evidence in a computer

Today, this definition seems a bit old, but it does hold a few key aspects. To begin with, it describe that computer forensics is a collection of techniques and tools. While those are defiantly two important aspects, this definition does not fit my personal beliefs as it kind of omits the methodology and mind-set that, for me, is the foundation of digital forensics. However, it does capture that digital forensics extends to all digital technology and that is an important aspect as today,

J. Kävrestad, *Fundamentals of Digital Forensics*,
https://doi.org/10.1007/978-3-319-96319-8_1

important evidence may be found in everything from thumb drives to computers or the cloud.

A more recent description is found on www.forensiccontrol.com (2017):

> Computer forensics is the practice of collecting, analyzing and reporting on digital data in a way that is legally admissible. It can be used in the detection and prevention of crime and in any dispute where evidence is stored digitally. Computer forensics follows a similar process to other forensic disciplines, and faces similar issues.

What is noticeable in this description is that it determines the tasks involved during a forensic investigation: collecting, analyzing and reporting. It also describes that computer forensics is comparable to other forensic disciplines and that does suggest that methods used and conclusions drawn during a computer forensic investigation should face the same scrutiny as an analysis of a fingerprint or DNA test. The rest of section will discuss each of these, beginning with establishing a model that could be used to describe a digital forensic examination.

1.1 A Forensic Examination

As we just established, the foundation of digital forensics is that it is the practice of collecting, analyzing and reporting on digital data. It does, for sure, also impose that there is some data that we target for examination and a reason for the examination. It does also impose that, unless we do the examination for the fun of it, there is someone that we should report back to. I have collected those aspects and formed the very abstract model as shown in Fig. 1.1 that does try to summarize the named aspects in a graphical way.

Figure 1.1 reflects the discussed processes and the inputs or outputs that should be present in each process. From top to bottom, *Collect* should be the process of collecting digital evidence. I would also say that in this process you do target a person or a data source that would commonly be a device.

Having a person as a target would be the normal state in a criminal investigation where you have someone that is suspected of a crime. You would then, after getting a search warrant, start searching for devices that belong to the suspect. In a corporate setting, it could be more common to target a device rather a person, and it would all depend on the reason for doing the investigation.

In this process, it is important to mention that in order to collect the correct data you need a proper order. The order in this case would include the target person or devices to collect data from, but it should also include the reason for the investigation, at least on an abstract level. This is because you would look for vastly different data sources if you are investigating a suspected malware attack of a child abuse case. It is also important to know if you should prepare to collect information from volatile data sources such as memory circuits or only need to care about static media such as hard drives. Another technical consideration is if you should expect encrypted data or not. While there will be a more detailed technical discussion on

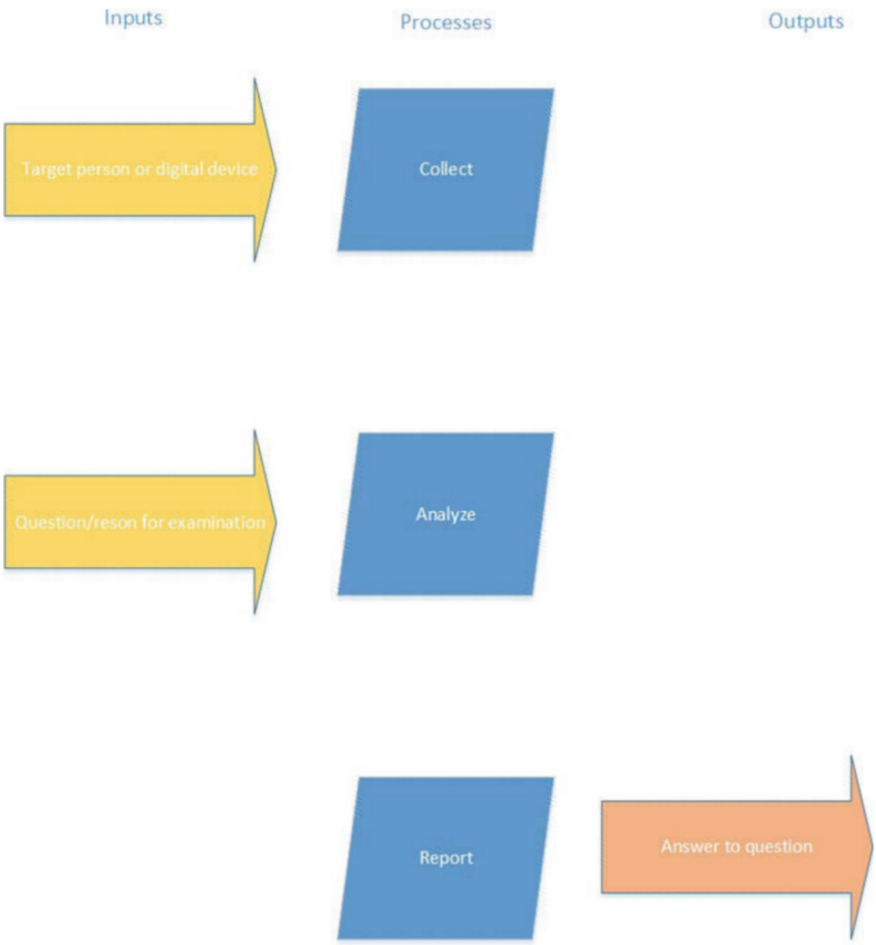

Fig. 1.1 Overview of forensic processes

data collection later in this book it is important to mention that you need to come prepared. The preparation steps should help you determine what to expect and should at least include figuring out the reason for the forensic examination and a background check on the person from whom you are collecting data.

The process of *Analyzing* data is more concerned with finding out what has happened in a digital environment or what was done using a digital device. In a corporate environment, a forensic expert would normally be quite free to conduct whatever examinations she wants. However, with a precise question the examination will without doubt be more efficient. It is worth mentioning that the input to this phase is commonly found in a discussion between the person ordering the examination and the forensic expert. Also, it is common that new questions and follow-up questions will arise during the investigation. As one example, during an

investigation of a computer during a drug case the initial request was to find out if the computer had been involved in any activities related to selling drugs online. The investigation clearly showed that it had been, but a large portion of the evidence was found in folders shared among several computers. In this case, a follow-up question was to determine who, more than the computer owner, had access to the folders in question. As a final note, it is important to mention that, in a criminal investigation, depending on the local legislation the questions that are taken as input to this process may be more or less important in setting the rules of what the forensic expert is allowed to do.

In the final process, *reporting*, the findings from the analysis are reported. The purpose of this step is mainly to report well-grounded answers to the questions given to the examiner in the previous step. In this step, it is very common that new questions will arise in light of the provided answers, and for that reason the last two steps are commonly iterative. It is also worth mentioning that it is of great importance that the conclusions drawn in this step are actually conclusions that are backed up by the findings during the examination. Each of the phases and considerations relating to each phase will be discussed in greater detail in Chaps. 4 and 5.

1.2 How Forensics Has Been Used

To deepen the introduction on the concept of digital forensics, this final section of the first chapter is dedicated to describing two criminal cases and one corporate case that the author has been involved with. The intention is to continue the introduction to the area with some examples of how digital forensics and forensic methods have been used in reality.

The first case was a case where a person (A) got suspected of computer intrusion by having tricked the victim (B) into giving up the credentials to his Web site. A had then used the credentials to modify B's website in malicious ways and later destroy it completely. This case started with a report to the police, and since B has very good indications of the actual identity of A, a house warrant was issued, A got arrested, and his computers got seized. In this case, the forensic examination was done by the author of this book. What is interesting about this case is that the police investigators did not know anything about computer crimes and the forensic examiner had to assume the role of co-investigator. In the first interrogations with A, is became evident that A had been in contact with B using chat clients. As such, the first forensic task was to map the communication between A and B by searching for usernames related to B. The result of this process was that it became evident that A had contacted B and said that he was a Web designer who offered to aid B with his Web site. After some communication, A managed to trick B into giving away the credentials to said Web site.

The second step in the forensic examination was to actually find evidence of A being involved in the malicious modifications to B's Web site. In this case, searching for URLs and HTML code related to B's Web site revealed that there were modified versions of B's Web site located on A's computer. Moreover, one of the modifications to B's site involved including pictures with sexual content on the Web site. By using forensic tools to search for identical pictures, it was revealed that the pictures did not only reside on A's computer but was also taken with A's iPhone, resulting in A being convicted of computer intrusion.

Another criminal case where the forensic involvement was much smaller but played an important role was in a murder case where the suspect had shot a person. There were some pieces of evidence pointing to the suspect, but he was given alibi by his girlfriend who said that he was at home at the time of the crime. Home in this case was about 90 min away from the murder site, by car. In this case, the suspect's telephone was examined and the IMEI number of the phone was identified. It was then possible to get records displaying what IMEI numbers had been connected to the mobile towers in close proximity to the murder site at the time of the murder. Turning out, the suspect's phone was connected to a mobile tower very close to the murder site, at the time when he said that he was at home. This was a key piece of evidence that led to the suspect being sentenced to lifetime in prison for murder.

A final example from the corporate sector was a case when an employee of a company was suspected of placing a Trojan horse in the company network. The employment had been terminated, and the suspicion was that the employee had placed a Trojan horse to get back at the company for sacking him. The Trojan horse was detected and analyzed by the company's IT department, and it was evident that it was configured to send information to an IP address located close to where the former employee lived. Since search warrants and tracing IP addresses are off limits for companies, other actions had to be taken. After careful examination of how the Trojan horse got inserted into the network, it seems as if it had been copied from a USB stick. It was also possible to determine the unique identifier for the USB stick.

A USB device that was issued by the company and used by the employee was examined, and the unique identifier was the same as for the USB stick that was used to distribute the malware. When the employee was confronted with the evidence he admitted to having injected the Trojan horse, and a civil lawsuit was filed.

1.3 Questions and Tasks

Here are the questions for the first chapter, and for these questions you may benefit from answering them in a group discussion!

1. Consider in what types of criminal investigations that computer forensic experts may be involved and in what way.
2. Consider when a computer forensic expert may be needed in a corporate environment.

3. Brainstorm on what types of devices may be interesting to a computer forensic expert.
4. To whom are the findings of a computer forensic examination of interest?

References

Forensic Control. (2017). Beginners guide to computer forensics. Available online: https://forensiccontrol.com/resources/beginners-guide-computer-forensics/ (Fetched: 2017–07-06).

Reith, M., Carr, C., & Gunsch, G. (2002). An examination of digital forensic models. *International Journal of Digital Evidence, 1*(3), 1–12.

Cybercrime, Cyber Aided Crime and Digital Evidence

2

Before dwelling deeper into forensics, it seems reasonable to have a discussion on what signifies cybercrime. Or, maybe more importantly, how and when digital evidence comes into play during criminal investigations. I choose to include this discussion due to the fact that during my work as a forensic examiner, I was often faced with the misconception that my daily work was with cybercrime in the sense of computer hacking and that sort of things. In reality, digital evidence is present in crimes of almost every kind. Even so, it can be of importance to understand the different roles of digital evidence in different types of crimes. On the topic of cyber and cyber-related crimes, one could divide the types of crimes in the following way:

- Cybercrime
- Cyber aided crime
- Crimes with digital evidence.

To further understand the difference of these different categories, a short discussion on what a crime actually is would be useful. To begin that discussion, it is interesting to look at what Rogers wrote back in 2000. He uses the traditional approach of means, motive and opportunity to discuss cyber criminals. In this discussion, motive is the reason for why someone is committing a crime. Take defrauding for example, the common motive for defrauding someone is to earn money. Means would be the tools used to commit the crime, and opportunity could be described as the possibility to commit the crime. One could argue that a crime begins in motive and that the means and opportunity are mere results of the easiest way to achieve what is wanted as motive.

So, for a crime to happen there has to be means, motive and opportunity. Further, there has to be someone who is committing the crime, a criminal. And someone who is targeted by the crime, a victim. Finally, there is some kind of relationship between the criminal and the victim, something is happening between them. It should be noted that this is a simplified view that is not always 100%

J. Kävrestad, *Fundamentals of Digital Forensics*,
https://doi.org/10.1007/978-3-319-96319-8_2

accurate, but it is good to have this basic notion of what a crime is in order to dwell deeper into what a cyber or cyber aided crime would be.

2.1 Cybercrime

Looking at what a true cybercrime could be, Interpol (2018) provides the following definition;

> Advanced cybercrime (or high-tech crime)—sophisticated attacks against computer hardware and software

Going from that definition, one could say that cybercrimes are crimes where computers are used to do crimes at other computers. This would include, for instance, computer intrusions or denial of service attacks. This gives that cybercrimes are crimes that can only be committed by someone who has a fair knowledge on how computers work. Looking to the discussion on means, motive and opportunity for a crime and applying those to cybercrime, it is reasonable to say that for a crime to be a true cybercrime, the means and opportunity would involve computerized tools and knowledge for a crime to be cyber. This discussing becomes important when an investigation is trying to find suspects, as the knowledge needed to commit a crime would be part of the suspect's profile. That, however, is not a topic to be discussed in a book about computer forensics.

2.2 Cyber Aided Crime

A much broader category of crimes would be cyber aided crimes. As discussed by Interpol (2018), those are traditional crimes that make use of the Internet in some way. This is exemplified by Kävrestad (2014) who studied the difference between online and offline fraud. In that study, the process of a fraud was modeled as shown in Fig. 2.1.

In brief, Fig. 2.1 depicts that a fraud is when a fraudster deceives a victim into giving up something of monetary value. For this to be possible, there has to be a delivery method for the actual fraud. The delivery method can be e-mail, telephone or real-life contact. What decides if a fraud is online or offline was found to be how the delivery is carried out, an online fraud will be carried out using digital means of communication and an offline fraud would use offline means of communication. This distinction is important because it helps the forensic expert or the investigator to understand how a crime was committed and thus, how to best investigate it. That is, where to look for evidence.

Looking at how crimes are committed today, most crimes have been around for a long time and are committed by criminals that does not necessarily hold any high grade of computer skills. However, the use of computers created a new arena where

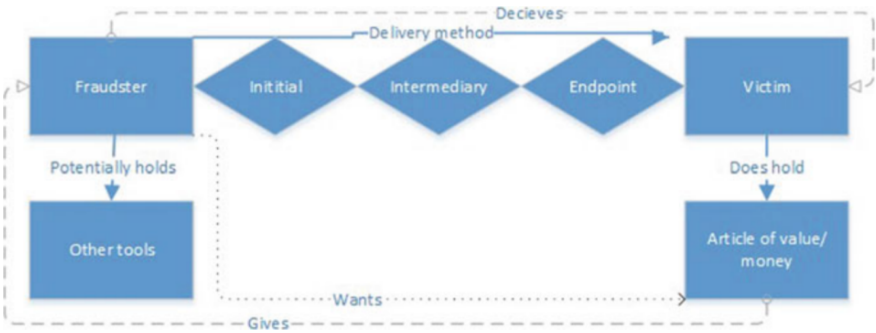

Fig. 2.1 Model of a fraud (Kävrestad 2014)

it appears convenient to embark on criminal activities such as frauds, drug trades, child abuse or whatever. As such, the criminals are not typically computer experts, rather, the online and offline fraudsters and drug dealers are the same type of criminals.

This view is further discussed by Rogers (2001), who described different types of computer criminals. On the topic of online fraudsters, he argued that online fraudsters are simply fraudsters that commit their crimes online. The same can be said about criminals that sell drugs online and that are involved in child exploitation crimes and a wide range of other criminals. They are committing traditional crimes and have traditional motives, but they see the opportunity to commit the crimes from the comfort of their own house, using the internet. Also, as of today, the means to commit the crimes become owning a computer and most people have a computer already.

2.3 Crimes with Digital Evidence

It goes without saying that a forensic expert can look for, and expect to find, digital evidence in cybercrimes and cyber aided crimes. Actually, it is interesting to notice how a lot of information that was extremely hard or even impossible to come by in offline crime is quite often captured when the crimes are committed in an online environment. Consider, for example, a drug trade. A traditional drug trade would involve two people meeting in the streets to exchange drugs and money. Often times, there would not be a single trace of that transaction ever taking place unless it was monitored. However, doing the very same trade online would involve e-mail or chat between the buyer and seller as well as a digital transaction of money. This digital communication and money transaction will leave digital traces that can be uncovered and used as evidence.

As an end to this brief cybercrime discussion, we should not forget how digital evidence can play a big role even in crimes that are totally offline. Thing is, in modern society it is very hard to do anything without leaving digital traces. Even if you are doing something totally offline, in the heat of the moment or whatnot, there is a great chance that there can be digital evidence to support what happened. This can involve communication logs that can show what the criminal did after the crime was committed. Maybe he looked up punishments for the crime he committed or even talked to some friend about what he did? I have even seen an example where a cell phone was used to tie a suspect to a crime scene, when the cell phone was not even used, it was just present!

2.4 Questions and Tasks

The task for this chapter is to get hold of two verdicts, then read them and consider how digital evidence what used in the cases. Try to get one verdict about a traditional cybercrime such as hacking or copyright infringement and one about something unrelated to the digital world, such as theft. In Sweden, you can call a local court and have them send you verdicts over e-mail and you are often able to find verdicts online, just make sure you do not break any local laws!

References

Interpol. (2018). Cybercrime. Available online: https://www.interpol.int/Crime-areas/Cybercrime/Cybercrime. Fetched April 9, 2018.

Kävrestad, J. (2014). Defining, categorizing and defending against online fraud.

Rogers, L. (2000). *Cybersleuthing: Means, motive, and opportunity*. Available online: http://www.sei.cmu.edu/library/abstracts/news-at-sei/securitysum00.cfm. Fetched May 1, 2017.

Rogers, M. K. (2001). *A social learning theory and moral disengagement analysis of criminal computer behavior: An exploratory study*. Doctoral dissertation, University of Manitoba.

Computer Theory

3

Up until this point, we discussed what computer forensics is and pretty much concluded that is it about examining and deducing what happened on a computer or in a computer system. That is all well and good but to move on further you do need a bit of background knowledge. The intent of this book is not to provide you with a summary of computer science. Rather, I expect you to have a fair "know-how" on computer stuff. But there are a few areas that I found that IT people commonly do not know that much about, but that are important to a computer forensic expert. Those areas are covered, in brief, here. Note that each sub-section is an overview. For a complete understanding—follow the references!

3.1 Secondary Storage Media

Secondary storage media refers to media where data is stored for long-term preservation. This is in contrast to primary memory, which includes random-access memory and cache memories, which is used for short-term storage. Secondary storage includes hard drives, CD/DVD, USB flash drives and memory cards. This discussion refers mainly to hard drives but is also (commonly) applicable for USB flash drives and memory cards.

The first thing that is important to know is the physical size of the storage media. This is because it is important to know that you can account for all the storage area on a computer. Say that you find a computer that appears to have a "C:\" partition of 200 GB but a physical examination of the hard drive reveals that it is supposed to be able to house 250 GB of data. This could mean that there is another hidden partition present on the hard drive or that the hard drive was reformatted. Either way, the remaining 50 GB may contain valuable evidence.

This is also a good place to comment on how hard drive formatting is commonly handled by the operating system. It is easy to assume that if you repartition your hard drive, the existing data is overwritten. That is, most times, not the case. Rather,

J. Kävrestad, *Fundamentals of Digital Forensics*,
https://doi.org/10.1007/978-3-319-96319-8_3

the hard drive is made up from sectors and clusters that can be allocated to a file or a partition. When you partition a hard drive you create a master boot record or GUID partition table (other versions exist as well, but seems rare) that contains a partition table. The partition table houses information about partitions on the hard drive including starting and ending sector for each partition. If you resize your partitions, the only thing that will happen is that the partition table gets updated. The actual data on the hard drive is often unaffected. While this makes the data that was on the hard drive inaccessible by the operating system, it is still possible to recover it using forensic tools.

It should also be mentioned that it is quite common that a hard drive that may appear empty is just reformatted. When a hard drive is reformatted, it happens every so often that only the partitioning table is removed. The partitions, that we will discuss next, still remain on the disk. The reformatting only made it possible for the computer to put new data in the sectors that made up the partitions. But until that happens, the old data is still fully readable and the partitions can be recovered using forensic tools. Further information on hard drives and partition tables is beyond this book, but a good source of information is available at www.ntfs.com/ntfs (NTFS 2017).

3.2 The NTFS File Systems

As we just discussed hard drives and partitions, the next logical step becomes discussing file systems. A file system is essentially a structure used to control how data is stored and retrieved on a storage device and is the common content of a partition. So to make things clear: a hard drive contains partitions, a partition commonly contains a file system and a file system is used to structure data. I am saying that a partition *commonly* contains a file system because that is not always the case. For instance, a partition may contain some semi-organized data such as swap space, that is the case for swap partitions in the world of Linux—but that is another story.

As for the file system, there are several different file systems out there such as ext4 (common on Linux), NFS (common for network storage) and FAT32 (common on surveillance video and thumb drives). However, we will dig into the NTFS file system that is used on moderns Windows-based computers for the sole reason that NTFS is the most common encounter for a forensic examiner and this book is aimed at examinations of Windows-based computers.

As previously discussed, the partitions are stated in the partition table found in the master boot record. Next, a partition formatted with the NTFS file system begins with a metadata file called the partition boot sector. What we need to know about this file is that it contains the Master File Table (MFT) that is basically a dictionary of all files and folders on the NTFS partition. The most important content, for a forensic examiner, in the MFT is the file records. All files and folders on the partition have one! For each file or folder on the partition, the MFT record contains

information about the name and the actual file data. However, a MFT record cannot be bigger than 1024 bytes so files that are bigger than about 600 bytes (about 400 bytes are reserved for file name and such) cannot reside in the record. In these cases, the MFT record describes what clusters on the hard drive that house the file (Guidance Software 2016). Files contained in the MFT are called resident, and files not contained in the MFT are called non-resident. Before we move on you should also know that there is a backup MFT, commonly located at the end of the partition (TechNet 2017).

So how are files created and deleted? Well, when you create a file or folder it will get a MFT record. If the file is small enough, it will be stored in the MFT and if it is too big the computer will allocate clusters and store the file in the clusters. When you delete the file, it is actually the MFT record that gets deleted and the data in the allocated clusters remains there until they are overwritten. This allows a forensic examiner to recover deleted files using forensic tools. Do note that there is a technology known as trim that overwrites clusters that are unallocated by the MFT, this is quite commonly used for SSD hard drives.

3.3 File Structure

To be able to recover and understand files, you need to know a little bit about how files are commonly structured. You should know that a file does not need to follow a certain structure so what you read here is not always the case. Well then, the common structure of a file is that it begins with a header containing metadata and then comes the actual data and finally a trailer. The metadata commonly contains what is called a file signature that tells the computer what kind of file the file is, such as a JPEG or PDF. By knowing this, you can search a hard drive for headers and trailers to find files even if they are deleted from the MFT. You can do this by searching for the hexadecimal or alphanumerical file signature depending on your software.

An example of a file signature is given in Fig. 3.1, which shows the file signature for a JPEG file. The left-hand side shows the file offset in hexadecimal (not relevant at the moment), the middle column shows the file data in hexadecimal and the right-hand side shows the file data in alphanumerical format. As you can see, the file begins with FF D8 FF E0 and this is what you would search for if you wanted to look for deleted JPEG files. You could also search for JFIF which is part of the alphanumerical file signature. Another example is given in Fig. 3.2, which shows a part of the header for a PDF file. In this case, you would search for 25 50

```
0000 FF D8 FF E0 00 10 4A 46-49 46 00 01 01 01 00 78 ÿØÿà··JFIF·····x
0010 00 78 00 00 FF DB 00 43-00 02 01 01 02 01 01 02 ·x··ÿÛ·C·····
0020 02 02 02 02 02 02 02 03-05 03 03 03 03 03 06 04 ·····
```

Fig. 3.1 Header of a JPEG file

```
000000 25 50 44 46 2D 31 2E 34-0D 25 E2 E3 CF D3 0D 0A %PDF-1.4·%âãÏÓ··
000016 34 20 30 20 6F 62 6A 0D-3C 3C 2F 4C 69 6E 65 61 4 0 obj·<</Linea
000032 72 69 7A 65 64 20 31 2F-4C 20 35 30 31 35 32 37 rized 1/L 501527
000048 2F 4F 20 36 2F 45 20 34-39 38 31 31 31 2F 4E 20 /O 6/E 498111/N
000064 31 2F 54 20 35 30 31 33-32 38 2F 48 20 5B 20 34 1/T 501328/H [ 4
```

Fig. 3.2 Header of a PDF file

44 46 2D or %PDF-1.4 since that is the file signature for a PDF file in hexadecimal and alphanumerical.

It is also worthwhile to mention that there are different approaches on how to store files. Most file formats, including plain text files and many picture formats, store files as plain files. However, some files including Microsoft Office files and compressed file formats such as ZIP are stored as compound files. Compound files are files that maintain some structured storage approach of their own (Microsoft 2017-1). That means that there is a local file structure within the compound file. This is the common case for compressed files. What is special about compound files is that they cannot be fully examined when they are in their "packed" state. Instead, they must be unpacked to be fully analyzed. The reason is that the data in the compressed state is represented in a different way than in the original, unpacked state.

3.4 Data Representation

This section contains a very brief discussion on how data is stored and represented in a computer system. This is simply to make you understand that the data may have different meaning depending on how you interpret it.

To begin, the data stored on any storage media is stored in binary, with zeroes and ones. You may group the bits into groups of eight called bytes and a byte may also be represented with two hexadecimal signs. To make life complicated, different applications may store data in different order. To begin, when we are looking at a single byte, containing 8 bits, the order is always the same. You interpret the bits with the leftmost bit having the highest significance and the rightmost having the lowest, as depicted in Fig. 3.3.

That is all well, but when you have a data set consisting of more than one byte we get in trouble. There are two ways to store consequent bytes. The first is called big-endian, storing bytes with the biggest end first making the first byte the most significant. In contrast, we have little-endian storing data with the smallest end first, reading from left to right (Cohen 1980). To give an example—consider the word "troll" in little- and big-endian in Fig. 3.4.

That is that on binary and hexadecimal representation. You should just know that depending on what kind of data you are looking at you may want to look at it in binary or hexadecimal.

Bits:	1	1	1	1	1	1	1	1	1
Value:	256	128	64	32	16	8	4	2	1

Fig. 3.3 Bits and values

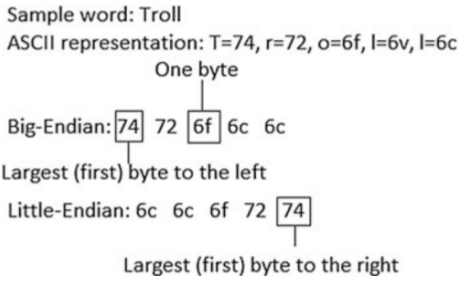

Fig. 3.4 Example of little- and big-endian

The final part on data representation is that you should know that computers have different ways of representing characters, called different ways of encoding data. While I have no intention of discussing different ways of encoding text or data you should know that different ways of encoding data exists, such as ASCII, UTF-8 and UTF-16. What the encoding decides is basically how a sign is represented in binary or hexadecimal code. For instance, the letter "A" is represented as "feff0041" in UTF-16 and as "41" in ASCII, using hexadecimal code. The reason for why this is important is that if you open a data set that is encoded in ASCII with a program that expects something else, the result will we screwed up.

3.5 Windows Registry

The Windows registry is a hierarchical database that stores information about users, installed application and the windows system itself (Microsoft 2017-2). That makes it a very important place for forensic examiners to look and something for this book to provide an overview of.

To begin, the Windows registry is a tree structure where each node in the tree is called a key and every key may have a value or sub-keys. A registry tree can be as deep as 512 keys (Microsoft 2017-2). The values that a key can contain are just arbitrary data, and it is up to the application that stored the value to decide the format and how it is to be interpreted. The registry is made up of several files, so-called hives (Guidance Software 2016). Each hive contains a set of data, the hives that are most commonly of interest to a forensic examiner are called SAM, SECURITY, SYSTEM and SOFTWARE. There is also another file associated with each user called NTUSER.dat. There is one NTUSER.dat for each user on the system, and this file is located in the user directory (…\Users\<username>). The other registry hives are located in the …\System32\config\folder. You may extract

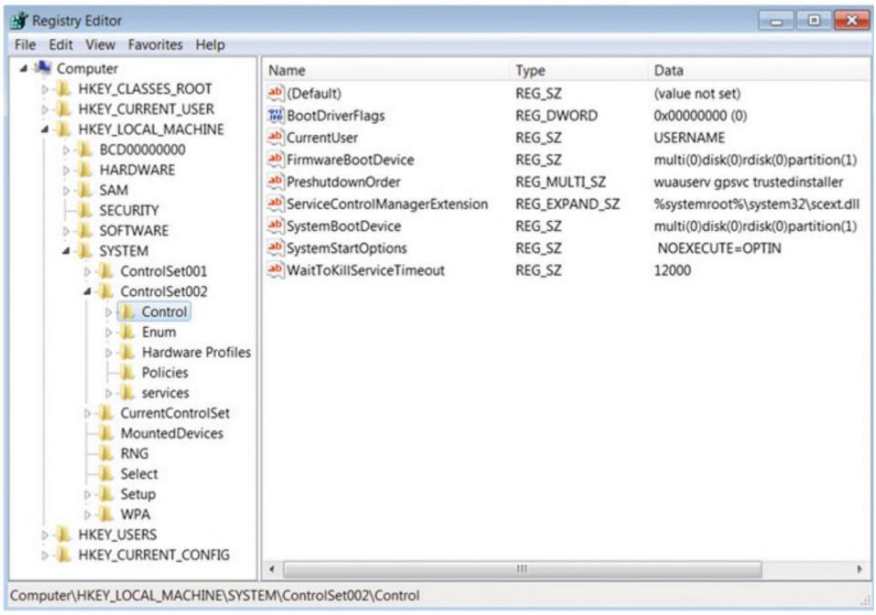

Fig. 3.5 Regedit overview

the hives and analyze them with a forensic tool, such as AccessData Registry Viewer or Registry Explorer as done in later sections of this book. You may also examine the registry of a running Windows system through the built-in utility regedit. In Fig. 3.5, presenting an example of regedit, you can see that it presents the registry hives in a format that is a bit different than you may think. This is because regedit shows the registry as seen by the running computer. HKEY_-CURRENT_USER contains the data stored in NTUSER.dat for the current user and data from the other hives is present in HKEY_LOCAL_MACHINE. In the picture, you can see that there are several keys in the tree at the left and some values in the pane at the right. In this case, the values are located under the key "Control" that in turn is a sub-key to the key ControlSet002 that is in the SYSTEM hive.

As you may understand, the registry can be a huge database and many programs store data in the Windows registry. It is strongly suggested to work with the registry to learn what kind of information that can be found in it. The rest of this section will cover each registry hive and the information found, in brief.

NTUSER.dat is a hive that stores information about a specific user account. This hive can, for instance, contain information such as the user's browser settings and history and data related to user applications.

SOFTWARE is the go-to hive for information related to applications. This includes data stored by Windows and data stored by other applications. A common piece of information to fetch here is the Windows version and install date, located in the sub-key \Microsoft\WindowsNT\CurrentVersion. This key will also tell you

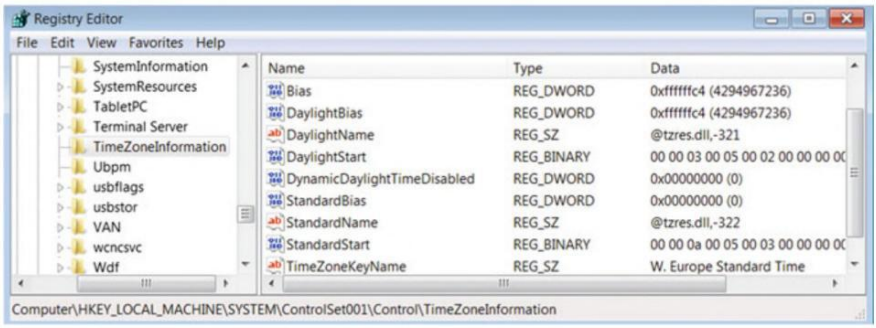

Fig. 3.6 Time zone information in the registry

the registered owner of the computer and it is surprisingly common that a real name is set here. Note that dates are commonly not stored in human-readable format. For instance, the install date is stored as a UNIX time stamp—seconds that have passed since midnight on the first of January in 1970—this needs to be converted.

SYSTEM will contain information about the system including USB devices that have been connected to the system, time zone settings and information about networks that the computer has been connected to. An example is given in Fig. 3.6 that shows you time zone information stored in the key \SYSTEM\ControlSet001 \Control\TimeZoneInformation\TimeZoneKeyName.

The *SAM* and *SECURITY* hives are protected by the Windows system and cannot be browsed using regedit on a running computer. However, extracting them from a forensic image and browsing them using a forensic tool is no problem. The *SAM* hive basically stores information about users. Examining this hive you can, for instance, find the users on the local machine, information about when they last logged on, when each account was created and password hashes. Finally, we have the *SECURITY* hive that stores some information about the system, perhaps mainly the system audit policy and the Syskey that you will need in addition to the *SAM* hive if you need to crack user passwords.

3.6 Encryption and Hashing

There are tons of good books for you to read if you want to get down and dirty with encryption and hashing, but for the purpose of this book I will just discuss the terms very briefly at an abstract level. Encryption and hashing are cryptographic techniques used to hide data. Understanding how this works is crucial for a forensic expert because, well, criminals usually do not want their data to be found and analyzed. Also, in modern computers, encryption and hashing are usually built-in, fundamental parts of the normal computer behavior making encrypted data a normal part of the forensic examiners daily work.

3.8 Questions and Tasks

Here are the questions for this chapter.

1. Brainstorm on what secondary storage media devices there are that can be of interest during a forensic investigation.
2. What happens when you delete a file from a NTFS file system and how can you recover deleted files?
3. What is meant with resident and non-resident files?
4. Why do you need to know the difference between little- and big-endian?
5. Use regedit to find out what time zone your computer is set to use.
6. What is hashing and what signifies a secure hash algorithm?

References

Cohen, D. (1980). ON holy wars and a plea for peace. IETF. Available online: https://www.ietf.org/rfc/ien/ien137.txt. Fetched July 6, 2017.

Guidance Software. (2016). *EnCase Computer Forensics II*. Guidance Software.

Microsoft. (2017-1). Compound files. Available online: https://msdn.microsoft.com/en-us/library/windows/desktop/aa378938(v=vs.85).aspx. Fetched July 6, 2017.

Microsoft. (2017-2). Structure of the registry. Available online: https://msdn.microsoft.com/en-us/library/windows/desktop/ms724946(v=vs.85).aspx. Fetched July 6, 2017.

NTFS. (2017). NTFS—New technology file system designed for Windows 10, 8, 7, Vista, XP, 2008, 2003, 2000, NT. Available online: http://www.ntfs.com/ntfs.htm. Fetched July 6, 2017.

TechNet. (2017). File systems. Available online: https://technet.microsoft.com/en-us/library/cc938949.aspx. Fetched July 6, 2017.

Notable Artifacts

4

Following the discussion on computer theory, it is important to have a discussion on some of the more notable forensic artifacts that can be of great importance during a forensic examination. A forensic artifact is basically a piece of information that holds forensic value. Quite often the forensic artifacts are pictures, word documents, text messages or some other information where the importance is quite evident. A picture showing drugs will always be a picture showing drugs. However, in a Windows operating systems there are several artifacts that track the usage of the computer in a way that can be of great interest for a forensic examiner. What is interesting, and often problematic, about those artifacts is the fact that Microsoft provides little or no documentation about how those pieces of information actually work. Thus, the function of the artifacts described in the remaining of this chapter has been examined, tested and understood with experience. The chapter is written based in Windows 10 version 1709, and there may be some differences in how the artifacts work in earlier and later versions of Windows. It is important that you, as a forensic expert, ensures that you understand the artifacts that you use to draw conclusions on your own, and research is necessary if you are uncertain.

In combination to this chapter, Chap. 8 presents even more artifacts in a more hands on manner, including guides on how to find them using forensic tools. The rest of this chapter is devoted to descriptions and explanations of common important forensic artifacts.

4.1 Metadata

One of the single-handed, most important sources of forensic information is metadata. Metadata is basically information about information and most objects, such as files and folders on a computer system will also have metadata. On a computer running Windows and the NTFS file system, the file system will record metadata for every file created on the computer. The metadata will include

J. Kävrestad, *Fundamentals of Digital Forensics*,
https://doi.org/10.1007/978-3-319-96319-8_4

information such as when the file was created and last modified and who created it. Several file types will store additional metadata. For instance, Microsoft Office files will store information about the author name, title of the document, how many times it has been modified and more. The author name is the name that was registered as the owner of the office application that was used to first create the document in question. To view file metadata in Windows, you can open the properties menu and select the "details" tab; however, this view will only provide a subset of the metadata that is actually stored. However, most forensic software's will parse and present all available metadata. A more practical discussion on what the metadata can tell you is provided in Chap. 11.

4.2 EXIF Data

EXIF data is metadata stored in pictures and deserves a section of its own because of its great importance for computer forensic experts. It is very common for a forensic examination to include looking for certain pictures, and when finding interesting pictures, a given follow-up question is to determine where, when and with what device the pictures were taken. This is done by examining the pictures EXIF data.

EXIF data was originally developed to help photographers record when they took a certain picture, what camera they used and what settings the used (Mansurov 2018). However, the data stored as EXIF data is also very valuable to a forensic examiner. First off, it is up to the camera manufacturer to decide what information to store as EXIF data and it is often possible for a user to turn off the storage of this information. Also, websites commonly exclude EXIF data when pictures are published online. However, when you find interesting pictures on a computer, it is often possible to find important information about it by examining the EXIF data. You may access this data, as you would any other metadata put there are also special built parsers and forensic tools available for EXIF data analysis. Among the more important pieces of information that can be recorded as metadata is

- Camera make and model
- Device name
- Time when the picture was taken
- GPS coordinates describing where the picture was taken
- Serial number of the device that took the picture
- Name of the person who took the image.

Note that the information stored in the EXIF data is, of course, the information that was available to the device. If the user configure the device name to be "Jacksons Iphone," then that will appear in the device name field. Figure 4.1 shows sample metadata including EXIF data viewed through Windows.

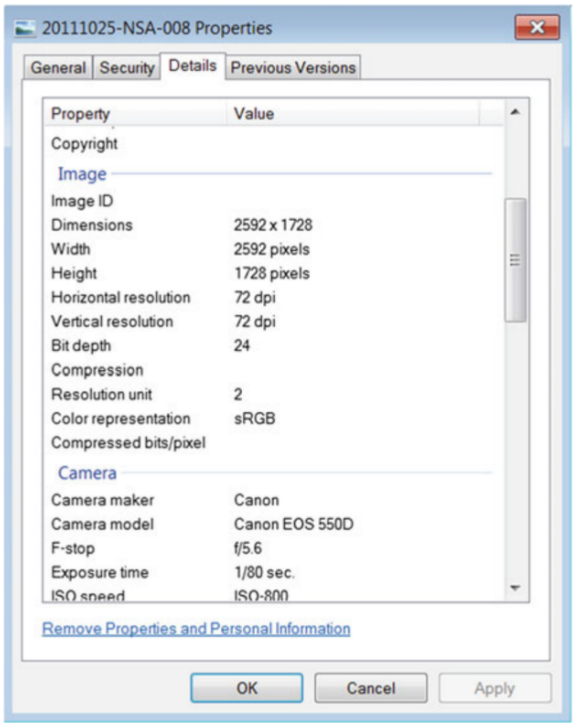

Fig. 4.1 Picture metadata including EXIF data

4.3 Prefetch

Prefetching, in Windows terminology is the process of bringing data and code pages into memory before it is needed. The idea is to track normal application usage and load the data that an application usually needs during runtime when the application is loaded. This process was implemented to increase performance of applications that used is a similar manner every time it is used (Nair 2012).

Prefetch data is stored in prefetch files located in the "Prefetch" folder under the system root (commonly c:\Windows). The most significant function of the prefetch files, from a forensic perspective, is that they contain information about how many times an executable was run, and when it was last run. The file name of a prefetch file begins with the name of the executable followed by a hash of the location where the executable is stored. For instance, a prefetch file for FTK imager could be named "FTK IMAGER.EXE-1B23CEFA.pf". If there is a second instance of FTK imager installed somewhere else there would be a second prefetch file with the same executable name but another hash value. There will be a "modified" time stamp for the prefetch file and that time stamp reflects the last runtime of the application, as the prefetch file is updated when the application is executed. The data in the prefetch file contains information about how many times the application

was used, what hard drive it resides on, and what files and directories it referenced. The data format is somewhat cumbersome to read but there are several good and free to use parser available including one by Erik Zimmerman that is presented in Chap. 14.

4.4 Shellbags

Next topic to handle is shellbags. Shellbags are used to store information about GUI settings for explorer, that is used to browse files and folders on a Windows-based computer. That means that they store information about what preferences a user sets for viewing certain directories. This can, for instance, be how to list files in the directory. To further explain the use of shellbags, if you browse to a folder and set viewing options to "detailed list," then close the folder and browse back to it you will notice that your setting are still there. This is shellbags working for you.

The forensic significance of these artifacts comes from the fact that a shellbag for a certain folder is created when a user is actually viewing that folder. Thus, the existence of a shellbag for a certain folder is a very good indication that the user in question has visited that particular folder. Also, the shellbags are stored in NTUser.dat and another user-specific file called UsrClass.dat, located in …/AppData/Local/Microsoft/Windows/UsrClass.dat. That makes the shellbag data user specific. On a third notice, it seems as if shellbags are not deleted and can therefore serve as evidence of deleted folders and since they collect information about network shares, mounted encrypted volumes and removable media, they can provide information about that as well. Further, experiments done by the author indicates that for Windows 10 version 1709, UsrClass.dat is the best source of shellbag data. However, both locations should be examined.

As said, shellbags are stored in registry in the following keys

- USRCLASS.DAT\Local Settings\Software\Microsoft\Windows\Shell\BagMRU
- USRCLASS.DAT\Local Settings\Software\Microsoft\Windows\Shell\Bags
- NTUSER.DAT\Software\Microsoft\Windows\Shell\BagMRU
- NTUSER.DAT\Software\Microsoft\Windows\Shell\Bags

However, the names of the shellbag keys are numbers and the values of the keys are in binary format making manual interpretation hard. It is more feasible to use a tool designed to parse shellbags, and there is one made my Erik Zimmerman that will parse out shellbag information from a registry hive. The tool is called "Shellbags Explorer" and is presented in Chap. 14. Sample output from this tool is presented in Fig. 4.2.

As shown in Fig. 4.2, there are traces of the computer being used to browse several folders on several drives. Also, in the right pane there are time stamps that provide additional information. "Created on" would reveal the first time that a folder was visited and "Modified on" should be updated if the user makes changes

Value Name	Value Type	Data
a	RegSz	\\VBOXSVR\VMDShare
MRUList	RegSz	ba
b	RegSz	\\VBOXSVR\My_Pictures

Fig. 4.4 Map Network Drive MRU again

(Default)	REG_SZ	(value not set)
a	REG_SZ	\\VBOXSVR\VMDShare
b	REG_SZ	\\VBOXSVR\My_Pictures
MRUList	REG_SZ	ab

Fig. 4.5 Map Network Drive MRU yet again

Then, if \\VBOXSVR\VMDShare was to be mounted again, "MRUlist" would be updated to reflect that. This is demonstrated in Fig. 4.5, using Windows built-in regedit.

To complicate matters a bit more, the MRU keys containing the entry "MRUListEx" works in a similar manner but looks quite different. Instead of naming the values with letters, numbers are used and the data format is in hexadecimal. For this reason, using a Registry Explorer that is capable of interpreting this data is very handy. Examples of such tools will be given in Sect. 4.2. However, even if the format of these MRU keys is different, they work in the same way. Whenever an event occurs, an entry is created with a number. The order of the events is recorded in numbers stored in DWORD (four bytes) format in "MRUListEx". The order in "MRUListEx" can tell in what order the events recorded in the listing appeared. As an example, consider Fig. 4.6 that shows the MRU key for opened.txt files.

Drag a column header here to group by that column

Value Name	Valu...	Data
0	Reg...	54-00-6F-00-20-00-62-00-65-00-20-00-64-00-65-00-6C-00-65-00-74-00-65-00-64-00-2E-00-74-00-78-00-74-00-00-00-80-00-32-00...
MRUListEx	Reg...	03-00-00-00-05-00-00-00-04-00-00-00-02-00-00-00-01-00-00-00-00-00-00-00-FF-FF-FF-FF
1	Reg...	72-00-65-00-6D-00-6F-00-74-00-65-00-74-00-65-00-73-00-74-00-2E-00-74-00-78-00-74-00-00-00-78-00-32-00-00-00-00-00-00-00...
2	Reg...	6C-00-61-00-73-00-74-00-2E-00-74-00-78-00-74-00-00-00-66-00-32-00-00-00-00-00-00-00-00-00-00-00-6C-61-73-74-2E-74-78-74...
4	Reg...	66-00-69-00-72-00-73-00-74-00-20-00-63-00-72-00-65-00-61-00-74-00-65-00-64-00-20-00-6C-00-61-00-73-00-74-00-20-00-6F-00...
5	Reg...	69-00-6E-00-20-00-74-00-68-00-65-00-20-00-6D-00-69-00-64-00-64-00-6C-00-65-00-2E-00-74-00-78-00-74-00-00-00-80-00-32-00...
3	Reg...	6C-00-61-00-73-00-74-00-20-00-73-00-61-00-76-00-65-00-64-00-2E-00-74-00-78-00-74-00-00-00-78-00-32-00-00-00-00-00-00-00...

Type viewer | Slack viewer

```
            00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F 10 11 12 13 14
00000000    6C 00 61 00 73 00 74 00 20 00 73 00 61 00 76 00 65 00 64 00 2E   l.a.s.t. .s.a.v.e.d..
00000015    00 74 00 78 00 74 00 00 00 78 00 32 00 00 00 00 00 00 00 00 00   .t.x.t...x.2.........
0000002A    00 00 6C 61 73 74 20 73 61 76 65 64 2E 74 78 74 2E 6C 6E 6B 00   ..last saved.txt.lnk.
0000003F    00 56 00 09 00 04 00 EF BE 00 00 00 00 00 00 00 00 2E 00 00 00   .V.....ï¾.............
00000054    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   .....................
00000069    00 00 00 00 00 6C 00 61 00 73 00 74 00 20 00 73 00 61 00 76 00   .....l.a.s.t. .s.a.v.
0000007E    65 00 64 00 2E 00 74 00 78 00 74 00 2E 00 6C 00 6E 00 6B 00 00   e.d...t.x.t...l.n.k..
00000093    00 22 00 00 00                                                   ."...
```

Fig. 4.6 RecentDocs\.txt in Registry Explorer