

Brief introduction to counter-forensics

How investigators can find themselves looking for what isn't there
By Noemi Kuncik and Andy Harbison
This article was originally featured in Digital Forensics Magazine

With the ever-increasing growth of computer performance and storage abilities, together with the expansion of networks and falling costs, digital crime is rapidly becoming an everyday worry not only for personal users but also for corporations of all sizes.

Thanks to the sophistication of today's computers and networks, and the spread of knowledge about their workings, digital criminals can not only carry out their plans and mostly remain anonymous but also cover their tracks. By doing so, they make it extremely difficult and time-consuming for a digital forensic investigator to put the pieces together and solve the puzzle.

In the year 2000, the main focus of digital forensic practitioners was probably cyber-crime investigation. Large-scale electronic discovery did not really come into widespread use until 2003 or 2004,¹ and the large majority of precedent-setting cases date from 2003 or later. The highly influential decisions in *Zubulake v Warburg* were made in 2003 - 2005, and the U.S. Federal Rules of Civil Procedure were modified to take into account the discovery of electronically stored information in December 2006.

Digital forensic investigation, especially in criminal cases, requires a considerable knowledge of computer science. Nowadays it can be argued that most people working in the digital forensic field are electronic discovery practitioners, many of who have insufficient knowledge about computers beyond what they see in front of them. Many, perhaps most of them are trained on a narrow range of commercial tools. Platform-based certifications such as the EnCase Certified Examiner (EnCE) are widespread.

Such training focuses on the correct operation of the software in question and the interpretation of the results it displays. It typically does not consider in detail how computers operate behind the scenes – how they save data and where, what happens when they are turned on or off, and so on. This can be a handicap when up against advanced computer users who deploy sophisticated methods to hide their identities, and cover their tracks by means of counter-forensics.

¹ Nelson, S.D.; Olson, B.A. and Simek J.W. The Electronic Evidence and Discovery Handbook American Bar Association 2006.

In most cases misuse, abuse and suspicious activities that take place inside a computer system can be traced back to their perpetrators by an adept digital forensic investigator. However, before such an investigation can even be started the investigator needs to be aware that some form of evidence manipulation has occurred. Often this is more a matter of knowing what evidence should be present but is not, rather than detecting the overt use of counter forensic tools. Consequently, investigative experience and a thorough grounding in computer science is essential in detecting the use of counter-forensics.

We have seen a considerable upsurge in the use of counter forensics in the last two years, but this has not been reflected in the literature. It is our concern that with the proliferation of electronic discovery, where less attention is typically paid to forensic and ephemeral data, counter-forensics may be underused. We have begun to see sophisticated counter-forensic measures being deployed in many of our cases. It has now become routine procedure for us to check for prior evidence tampering in every disk we analyse.

Large-scale electronic discovery did not really come into widespread use until 2003 or 2004.

There are many possible causes of this upsurge. Public awareness of digital forensics has greatly increased with its growing use in civil and criminal cases. Most developed countries have now seen a number of high-profile legal cases where digital forensics techniques have been employed with great success. It is equally possible that the proliferation of police procedural TV shows in recent years, such as “CSI”, “Forensic Detectives”, and many others (as well as books), have caused this increase in counter-forensic sophistication. These shows regularly include examples of digital forensics techniques in use and have made people aware of the resources at the disposal of investigators, and thus of their own risk of exposure. Another cause may be simply the growing level of IT understanding among the general public in most developed countries.

In this article we will give a brief outline of counter-forensics, and then discuss the different locations where data can be found in a computer system, and the different types of data present. We will focus mainly on the Microsoft Windows family of operating systems which are installed on most of the world’s computers, especially PCs. Consequently most evidence elimination and counter-forensic tools are oriented towards removing data from Windows and its NTFS file system. The large majority of documents retrieved in electronic discovery procedures will also be acquired from computers running some version of Windows.

Counter-forensics

Counter-forensics (or “anti-forensics” as it is often termed in the USA) is the collective term for techniques intended to complicate, inhibit, subvert, or delay forensic techniques for finding evidence. It can be seen as encompassing a broad range of techniques from subtle and highly sophisticated data altering techniques to methods as crude as smashing evidential hard drives with a hammer. The purpose of counter-forensics is to make sure that evidence is not discovered and subsequently disclosed to a court, arbitrator, or some other forum. Additionally, in most cases at least some attempt is made to disguise the fact that evidence is being altered or withheld. In the vast majority of cases such tampering with evidence will damage the interests of those using these counter-forensic techniques.

The forensic material on computer hard drives can be broken up into a number of broad categories.

Counter-forensics is sometimes seen as being mostly about evidence destruction or erasure, but this is not the whole story. In many instances, particularly in respect to evidence in civil cases, it may not be necessary for counter-forensic techniques to destroy or erase data on evidential media. It is enough if they make it more difficult for an investigator or analyst to recover the data. Commercial digital forensics specialists usually operate within time limits and charge an hourly or daily rate for their services. Slowing an investigator's rate of progress, by disrupting the evidence or converting it to a format that is difficult to search, increases the costs of an investigation for the clients or legal professionals who pay the bills. This can deter them from pushing enquiries as far as they otherwise might.

As we have already pointed out, many modern digital forensics practitioners are trained to use specialised software applications such as EnCase and FTK, but do not understand how the computers they examine actually work at a deep level. Some counter-forensic tool developers have even designed their applications specifically to defeat common forensic analysis applications. They realise that many unsophisticated users will unquestioningly believe outputs given to them by their tools, and so can be easily deceived. Detecting the use of counter-forensic methods is often a matter of knowing what should be on a hard drive, but is missing from the drive being investigated. If an investigator does not understand exactly what should be there in the first place, he is not going to know if it has been removed.

For example, some counter-forensic applications will remove the Windows file table entries associated with deleted files, making it considerably less straightforward to identify. The forensic material on computer hard drives can be broken up into a number of broad categories tampered or erased files. Some such programs will simply remove the first few characters of each file table entry in the knowledge that Guidance Software's widely used EnCase application will not then resolve the entry. However, the mere fact that such file table entries are missing should warn the investigator that something is amiss.

Another major problem is that very little research has been carried out on identifying the telltales left by general purpose counter-forensic tools and determining what data specific tools actually leave behind. The counter-forensic tools used by computer hackers are often highly specific, targeting particular types of forensic evidence. A skilled hacker can surgically remove practically every trace of his presence on a computer. The general purpose evidence removal tools used by non-hackers are different, and are often either too general – removing data that does not need be removed – or too specific – failing to remove data that does. It should be possible to work out which tool has been deployed to destroy data by examining what has been removed and what has been missed. Unfortunately most research appears to be done in the field of hacking counter-forensics, and the more general tools used by less sophisticated IT users has not really attracted much attention.²

A related issue is that most research on counter-forensics is done in the context of computer security and counter-hacking rather than the more mundane arena of civil law, despite the fact that counter-forensics can potentially do far more damage when employed in civil cases than it does in hacks. (The prosecution rate for hackers was extremely low even before counter-forensic techniques became common, so their introduction has not greatly altered the overall picture).³

² Andy Clarke, Inforenz, How effective are Evidence Eliminators? Presentation to COSAC Conference, Kildare, Ireland 2002.

³ <http://www.usdoj.gov/criminal/cybercrime/cccases.html> data is stored in clusters and what happens after data is deleted.

In hacking, it may prevent an investigator pursuing a hacker, but deployed in legal cases, it directly acts to subvert the fair resolution of cases. Its purpose is to directly “load” the scales of justice, by altering the “database” of information available to the court and to the parties to any dispute, necessarily influencing the likely eventual outcome.

Forensic evidence on computers: What is there to destroy?

We all know that computers typically store large numbers of word processing files, spreadsheets, databases, cached web pages, emails and other “working” documents as part of their normal operation. These documents are the material upon which most legal activity is based. It is also becoming well known that most operating systems in current use do not delete files very efficiently, and that deletion does not by any means guarantee complete erasure of all data from the disk. The “empty” portions of a hard drive can, in fact, be full of fragments (more or less complete) of files deleted earlier – sometimes, much earlier.

Computers are designed to retain and retrieve information very efficiently and to protect the integrity of files and documents against internal failures and external errors. In consequence they tend to retain surprising numbers of copies of the files stored on their disks. These are usually deleted when the software has finished with them, but they persist in the “empty spaces” of disks long after they are supposedly gone.

Anti forensics and counter-forensics

“Anti-forensics” is a less satisfactory term than “counter-forensics”. The term “counter-forensics” implies that measures are taken to complicate, inhibit or subvert forensic investigation. “Anti-forensics” suggests that these measures actually prevent forensics being performed, which rarely if ever happens. Even if an evidential hard drive is actually replaced, it is still possible to determine this fact, providing useful evidence to the investigation. Anti-forensics is a term originally coined by the computer hacking community who tend to see forensics in a negative way.

Moreover the term “anti-forensics” appears to deny Locard’s Exchange Principle, one of the fundamental tenets of all forensic science. In essence, the principle simply states that “Every contact leaves a trace.” In digital forensics this means that any action on a computer device can change the data on that device

Active data

Active data is the working data on a computer: the operating system, programs, and working files stored on hard drives. The documents, emails, spreadsheets and other data people use day-to-day is active data, and consequently it is the material that is most often used in administrative, civil and criminal litigation.

Temporary or replicate data

This is the mass of copied data stored on hard drives. It is produced in large quantities by most popular applications. For example, Microsoft Word automatically makes copies on the hard disk of whatever documents are currently being written or edited. It does this so that, if the program or computer crashes, the working document will not be lost. As soon as the completed document is saved, Word will delete its temporary copies and the user will often never be aware that they ever existed. But on modern computers, deletion is not the same as erasure and the data from Word’s temporary copies of a document can hang around for a long time in the empty spaces of the disk.

Another common source of replicate data is Internet browser software like Internet Explorer, Firefox, or Safari. These programs usually store the component parts of the web pages they download from the Internet in an area called the “browser cache”. They do this so they can reuse the data if the user revisits the web page at a later time. This sometimes saves the computer having to download the full web page a second time. This was a big time saver in the past when Internet connections were typically a lot slower than today. Of course, it also means that a user’s web browsing can often be literally reconstructed from the browser cache, often to his considerable embarrassment. (Recently, browser suppliers have started providing optional features for deleting much of this potential evidence).

Residual data

Residual data is the data left behind in the “empty spaces” of the drive. The two principal repositories of residual data on any computer are the “unallocated” and “slack” spaces.

To understand these we must briefly review how a hard drive operates. In simple terms, hard drives work the same way as old fashioned libraries. The files are arranged across the hard drive much as books are organised in the shelves of a library, and like a library the disk maintains an index system, usually called the file table. When a user wants to access a file, the computer does not search through the disk looking for it – that would take far too long. Instead it goes to the file table (the “card index”) looks up exactly where the file is, and then goes straight to it.

When a file is deleted most file systems do not overwrite the space on the disk where the file was stored, in effect “removing the book from the shelves”. Instead a small note is made on the file table entry (the “index card”) that the file is now “deleted” and the card and space on the disk is available for reuse.

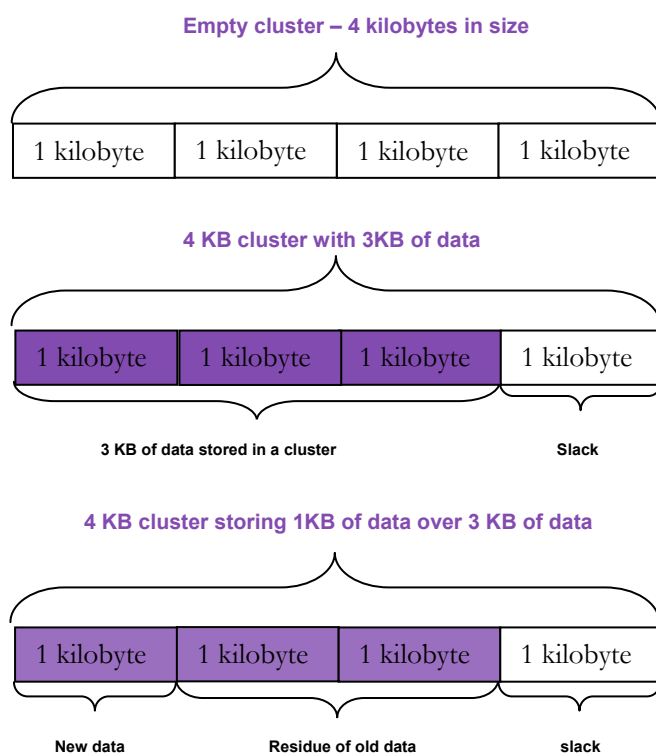
The computer takes this short cut to save time. Hard drives operate very slowly indeed compared to the computer’s processor and memory. They are therefore a potential bottleneck, throttling system performance. Overwriting deleted files would take a lot of time, and in any case they will, in theory, be overwritten with the passage of time, so the file system does not do it. In the meantime, however, a digital forensics specialist may still be able to retrieve the abandoned fragments of those deleted files.

Sometimes the space on the hard drive used by a deleted file is reused before the corresponding file table entry has been reused. In this case a digital forensics specialist will still be able to determine the name, creation data, size, and other characteristics of the deleted file, even if she is unable to recover the deleted file itself. On other occasions the file table entry is reused, leaving the data intact on the hard disk. The file is then said to be in “unallocated” space. The great majority of the empty space on a hard drive is made up of this “unallocated space”. It is normally a junkyard of different file fragments from documents, system files, and other ephemera. Slack space is more persistent. It, too, is a by-product of the way hard disks are organised. In order to further speed up the process of finding the location of a file on the hard drive, the computer divides the drive’s address space into a large number of units called “clusters”. On Windows, clusters are usually 4 kilobytes (4,096 8-bit bytes) in length. The start of any document can only lie at the start of one of these clusters. Of course this is rather inefficient in terms of storage capacity. It means a 1 kilobyte file will still take up 4 kilobytes on the disk. 3 kilobytes of data will be wasted. This “wasted” space is called the cluster slack (or sometime cluster tip).

Now, say a 4 kilobyte file is deleted, and the cluster is reused by the computer later to store a 1 kilobyte file. Obviously the first kilobyte of the cluster will contain the data of the new file, but the remaining 3 kilobytes will contain the remaining three kilobytes of the old file. This old data

will be preserved on the disk until the new file is itself deleted. Hence, even if a user thinks she has deleted a document, parts of it can persist in slack space for months or years afterwards. See Figure 1 for a visual explanation of how data is stored in clusters and what happens after data is deleted.

Figure 1 – Storing data in clusters



Systems data

Modern operating systems accumulate a lot of tracing data within themselves, usually in an attempt to make the computer more user-friendly and to help users work more productively. A lot of this data can be of immense value to a Digital forensics investigator. For example it can tell them:

- the files and documents most recently used;
- what folders were opened, and when;
- the creation, last access, and last modification dates of the files stored on the system;
- what users logged onto the computer and when;
- what devices, such as USB pens drives have been connected to the computer, and when;
- what web-addresses have been typed into web-browsers; and
- which programs have been used on the computer, by whom, and how often (if the system offers per-user authentication). Often this data replicates and corroborates other data stored in the active spaces, logs and historical data, making it a useful resource for the forensic investigator.

Logs and historical data

Most computers regularly log the activity and performance of both the operating system and the applications running on it. This is done to help administrators diagnose problems on the system, to help users remember what they did in the past – or for security reasons. Obviously this data

can be enormously helpful to a forensic investigator trying to assemble a timeline of events. For example the key logs found on a Windows computer are:

Event Logs, where applications and the operating system record events such as hardware and software errors, system shutdowns and restarts, and many others.

History Logs, which record every web address, component and cookie file accessed by the computer together with the time and date of access. The history logs also record a certain amount of file access data.

API Logs, which record the connection of devices (such as USB drives) to the computer,

Application Logs, which contain details of events logged by applications such as media programs. The events to be written are determined by the developers of each program, not the operating system.

Conclusion

It is dangerous for digital forensics specialists to ignore the possibility that evidence on the computers they analyse has been tampered with or deleted. Our own experience is that a significant proportion of the computers we analyse have undergone some form of evidence modification before coming into our possession.

The safest approach for an investigator is not automatically to assume that his forensic applications are telling him the whole story. Even the most apparently complete tools have their weaknesses, and investigators should not be afraid to look at the raw data to double-check whatever the software is telling them. Of course, this requires that the investigator understand where the tools he is using are getting the data they display on the screen.

Above all, a complete digital forensics skill set does not begin and end with a platform-specific certification. A good investigator should take the opportunity to expand her knowledge, and should not be afraid to look more deeply at the underlying operating principles of the systems she investigates.

A useful piece of advice, even for commercial digital forensics specialists - for whom every hour must be accountable – is not to be scared to follow your nose. If a forensic tool gives findings that are difficult to explain, or look inconsistent with the other evidence you are seeing, do not be afraid to look “under the bonnet”.

Author bios

Noemi Kuncik is an IT Forensics Specialist with a BA (Honours) degree in Computer Science and Masters in Computer Science and Informatics from University College Dublin. Noemi worked with Mick Moran of Interpol to create a training program countering child grooming and is researching the use of data mining in conjunction with Digital Forensic Investigations.

Andy Harbison is a Director and IT Forensic Lead holding a BSc in Electronic Engineering and MSc's in Business Administration and Information Technology. Andy lectures at the University College Dublin, Law Society of Ireland and Dublin City University and has written articles on computer fraud, electronic litigation and data privacy. He is a regular speaker at conferences.

Contact

Andy Harbison

Director, Forensic & Investigation Services

D +353 (0)1 6805 766

T +353 (0)1 6805 805

E andrew.harbison@grantthornton.ie

24-26 City Quay, Dublin 2

Noemi Kuncik

IT Forensics Specialist

D +353 (0)1 6805 662

T +353 (0)1 6805 805

E noemi.kuncik@grantthornton.ie

Offices also in Limerick and Kildare

This briefing is provided for general information purposes only and is not a comprehensive or complete statement of the issues to which it relates. It should not be used as a substitute for advice on individual cases. Before acting or refraining from acting in particular circumstances, specialist advice should be obtained. No liability can be accepted by Grant Thornton for any loss occasioned to any person acting or refraining from acting as a result of any material in this briefing. Grant Thornton, Irish member of Grant Thornton International, is authorised by the Institute of Chartered Accountants in Ireland to carry on investment business. www.grantthornton.ie