

Criminal Defense Challenges in Computer Forensics

Rebecca Mercuri

Notable Software, Inc.
P.O. Box 1166, Philadelphia, PA 19105
notable@notablesoftware.com

Abstract. Computer forensic techniques may be unfairly applied in order to tip the scales of justice in the direction of prosecution. Particular areas that are known to be problematic for defense experts include: erroneous allegations of knowledgeable possession; misuse of time stamps and metadata; control and observation of the discovery process; authentication issues; deficiencies and the lack of verification for proprietary software tools; deliberate omission or obfuscation of exculpatory evidence; and inadvertent risks resulting from the use of legitimate services. Examples in the author's caseload are used to illustrate these inequities in an effort to encourage reform.

Keywords: Computer Forensics, Forensic Investigation, Criminal Defense, Expert Witness, Digital Evidence.

1 Background

In criminal investigations where computer or digital evidence is involved, the defense usually begins from a position of disadvantage. Their fight may be an uphill battle that hopes to be tipped in favor of the accused by the facts, combined with the knowledge and experience of the legal and forensic team. Prosecution maintains a considerable upper hand, especially in terms of time and resources, so any impediment or obfuscation they may introduce into the discovery process can be costly or even disastrous to defense. This unbalance is anathema to a fair trial and such inequities promise to be further exacerbated as the complexity of handling and evaluating computer-based forensic materials continues to grow.

Since forensic efforts are likely to expose information that may deny or corroborate various claims of both the prosecution and the defense, it is the responsibility of investigators and expert witnesses to analyze and report on evidence data in an unbiased fashion. Techniques must be rigorous and repeatable, using accepted scientific methods.

One would think that both sides would necessarily draw the same conclusions from the same data. But in reality, since it is usually not possible to thoroughly examine every element of a large set of digital materials, a directed search is necessary to ferret out what is relevant to the case at hand. These decisions (what to look for, how and where to look, what to disregard, etc.) can spell the difference between exoneration and conviction, so they should be made with great care. In the end, the evidence tells its own story, and what it reveals should ultimately lead to the truth.

2 Challenges

In the course of working as a defense computer expert witness over the last decade, I have compiled a list of areas that demand attention in terms of litigation equity in cases involving digital evidence. These items are illustrated here with actual examples, omitting names and details for reasons of confidentiality and privacy.

2.1 Possession Is 9/10 of the Law

Many computer forensic matters involve issues of possession of digital contraband, such as child pornography (CP), unlicensed software or multimedia. Defendants are often shocked to learn that even though they never knew about, looked at, used, or shared the controversial materials, they can still receive a felony possession charge, with counts multiplied by the number of items found (including copies of the same item, virtual links to copies, or even deleted copies). For example, in order to justify increased penalties, a single video can even be assessed with multiple counts, based on the discrete images within.

In a Federal case in which I testified, the possession elements consisted only of two small blurry thumbnail images found via law enforcement's data carving of a thumbs.db system file, with the original pictures not present anywhere on the drive. My courtroom demonstration showed how a downloaded or copied folder containing a thumbs.db file could inadvertently conceal artifacts from files that had previously been deleted, prior to transmission from someone other than the accused.

Here's how this could have happened. Say a file folder on Jane's hard drive contained 4 pictures, 2 of naked children and 2 of other content. Jane deletes the

2 naked children pictures and then emails the folder to Joe. Joe receives the folder containing 2 legal pictures and copies it to his drive, but is unaware that a thumbs.db file has also been transmitted and copied (since the system files are, by default, not shown in Windows directories). As it happens, the thumbs.db file transferred to Joe's drive will contain small, lower resolution versions of all 4 of the photos originally in the directory on Jane's drive, but he will only see the 2 thumbnails that correlate with the 2 picture files that were copied. If Joe's drive is later impounded and data carved by a forensic laboratory to extract its contents, this thumbs.db file will produce 2 additional thumbnails that Joe never knew existed, these being the thumbnails of Jane's 2 naked children pictures.

This argument, though entirely plausible and consistent with the facts in the Federal case, was not convincing to either the jury or the Judge. While this case is on appeal, the client is currently serving a three-year prison sentence, along with a lifetime Megan's Law (pedophile registry) conviction requirement.

2.2 Lack of Knowledge Is No Excuse

While most laws with respect to contraband are typically phrased in terms of "knowing" possession, in practice, the accused may be deemed to have had "knowledge" purely on the basis of involuntary activity performed by the computer's operating system, as in the thumbs.db situation above.

An example in another recent case involved accusation of multiple counts of possession, where the circumstantial evidence used to demonstrate “knowledge” consisted of undated Windows Media Player logs. This case is being prosecuted despite the fact that the law enforcement forensic report indicated that the computer had been infested with numerous Trojan Downloader viruses, including one known to adversely affect the Windows Media Player.

2.3 Confusing Time Stamps

Time stamp metadata is not required to be provided in a uniform fashion. Law enforcement forensic reports may freely mix Universal (or Greenwich Mean) Time, Standard Time and Daylight Savings Time, sometimes even without annotation. If this information is given only in the form of derivative evidence, there may be no easy way for the defense attorney to correlate the various time conventions with the actual data in order to establish proper timelines.

Rarely is there any indication in the report as to whether the system clock was properly functioning, or what its offset may have been to real-time. If the system was live when confiscated, its clock would likely be viewable on the display, and this information should be recorded, but the collection of such data is often curiously absent when times are critical to an effective defense. In one case, hundreds of file time stamps extended for a half-day after the time when a live computer was impounded, but the police investigator failed to even mention or account for this disparity in his report.

Although Microsoft generally discredits the reliability of the “last accessed” timestamp, since it is easily altered by system operations that are not directly user-initiated, either or both prosecution and defense may choose to use this metadata if it is helpful to their construction. Best practice should be to always disallow it for any use.

2.4 Prosecution May Impede or Observe the Defense Discovery Process

Laboratory reports issued on behalf of prosecution can take advantage of limitations imposed on the forensic process, which essentially prohibits access by defense to the original digital evidence. District Attorneys work closely with the police investigators and their offices are often co-located in the same facility as the impounded materials.

In CP cases, current practice treats the defense legal and forensic team as if they were criminals with intent to possess or distribute contraband. Whereas law enforcement staff, working in behalf of prosecution, has their choice of equipment and software tools and may set up computer data queries to run on evidence drives for days or weeks, defense is usually prohibited from using their own laboratory and materials, and time constraints are often severely imposed (sometimes because of funding). Defense can be limited to viewing only derivative evidence at a law enforcement impounding site, typically one of the Regional Computer Forensic Laboratories (RCFLs) run by the FBI in conjunction with State Police. Defense may not even be allowed to retain copies of text-based materials (such as file directory trees) generated from the impounded data, for use as exhibits,

without prior prosecution approval. Prosecution can further insist that all onsite defense forensic activities be monitored by law enforcement, thus potentially revealing key strategy, while defense has no such complementary window for perusal of prosecution's investigation or discovery process.

The defense investigation team may even be subjected to subtle or overt harassment. In one instance, our defense team was provided with a small, unoccupied, examination room at a RCFL in order to peruse copies of evidence drives in a CP matter. Each evening, when we left this room, the officer who monitored our investigation locked the door in front of us, while verbally indicating that integrity would be maintained. Each morning, when our team arrived, the door was already wide open. On one of the days, all of the chairs in the room were missing. Although there was a pile of chairs in the nearby hallway, we were informed that "these were needed for a class." A telephone call to the public defender on the case resulted in some chairs being returned to the room after a few hours.

2.5 Defense Is Unable to Authenticate Materials and Copies

Whereas law enforcement usually has the opportunity to gather, impound and examine the original evidence shortly after (although occasionally even before or during) the time when an incident has occurred, the defense team may have access only to partial derivative materials, sometimes months or years after the event, with no ability to restore the situation as it was during the alleged crime. Protests about the lopsided nature of this process generally are regarded as insufficient reason to dismiss charges or suppress evidence.

As well, the Rules of Evidence are fairly stringent with regard to the admissibility of duplicates of evidence being constrained to only those copies that can be authenticated. For digital materials, many law enforcement labs have standardized on the use of MD5 and SHA1 hashes for this purpose, despite the fact that both have been considered weak for some while. Nevertheless, even these hash computations are not allowed to be observed when the forensic duplicates are made for defense use. Defense may (on request) receive a printout, supposedly generated at the time of the copying, that shows the hash value computed from a copy of the original compared with the copy made of that copy. Rarely is there a report confirming that the original was hashed prior to any law enforcement examination or manipulation of the materials.

Essentially there is a dark hole of time between impounding and hashing, where contamination could occur. In matters where stings are involved, it is conceivable that data unfavorable to prosecution could be deleted or overwritten, in order to conceal possible entrapment activities. These authentication steps should be considerably improved, but this is unlikely to happen until a defense challenge on these authentication issues is successful in court, which has yet to occur.

2.6 Proprietary Software Tools and Problematic

A decade or more ago, most data recovery laboratories were able to create full forensic image (or raw) copies of drive media using the UNIX/Linux `dd` utility. While

dd-based tools are still considered by many computer scientists to be the gold standard for such activities, law enforcement laboratories have shifted to using the commercially available FTK and EnCase software for evidence archiving (and examination). This is in part because of the increased utility, support and training provided by the manufacturers, but also likely because these products have graphical user interfaces instead of being command-line driven tools. Evidence drive images are increasingly provided in FTK or EnCase formats, and there are many problematic aspects to this, not the least of which is that defense's investigation should not be constrained to specific methodologies.

That these products are also proprietary creates an additional problem wherein it is difficult to perform any independent validation other than black-box testing (such as hashing a data set for which the hash value has also been computed using an open-source tool). As is well described in computer security literature, assurances of correctness based on black-box testing are of dubious value, since comprehensivity also must include white-box (code) examination.

Vendors of these proprietary products have been rather coy in reporting their validation assessments. AccessData (the manufacturer of FTK) issued a white paper [1] that touts the use of its products "by more than 30,000 investigators" along with citations from court testimony in order to establish its validity as a forensic tool under Daubert. [2] In answering the question "Has FTK Technology Been Reliably Tested?" the AccessData white paper cites results of NIST's testing of their disk imaging module, and with regard to error rates, their response is that these "are not generally a relevant category within the field of computer forensics, because there are essentially no errors in data acquisitions."

In fact, the referenced NIST/NIJ report [3] tells a slightly different story. The results note four "anomalies" (which could be interpreted as types of errors), as follows:

1. "If a logical acquisition is made of an NTFS partition, the last eight sectors of the physical partition are not acquired."
2. "The sectors hidden by a host protected area are not acquired."
3. "The sectors hidden by a device configuration overlay are not acquired."
4. "The location of corrupted data in an image file is not reported."

Results from the same NIST/NIJ study for Guidance Software's EnCase data acquisition product indicated similar problems with slightly different results, such as only the last sector of NTFS partitions not being acquired. This essentially means that the same image acquired by FTK and by EnCase would not match, thus the reported hash values would not be equivalent. Also disconcerting is that the EnCase 4.22a test results report [4] stated that "for some partition types (FAT32 and NTFS) that have been imaged as a logical (partition) acquisition, if a logical restore is performed there may be a small number of differences in file system metadata between the image file and the restored partition."

In contrast, the NIST/NIJ study for the dd utility provided with FreeBSD 4.4 indicated that "no anomalies were detected" and that "for all 32 test cases that were run, the dd utility produced an accurate bit-stream duplicate or an image on disks or partitions of all disk sectors copied." [5] It is not necessarily the case that open source products are more accurate (nor more secure) than closed source, but at least with those released within certain research communities there

is often an extensively documented code review history, which is not available for the closed source counterparts, where outside examination may be prohibited by trade secrecy.

It should be noted that the NIST/NIJ study is primarily functional (hence also black-box) but even its limited scope has shed insight on the incompatibilities and problems of data acquisition. One must assume that similar problems exist with many other forensic software tools, such as those that recover and report metadata, perform data carving, restore system information, and so on. NIST has only recently issued its first drafts of the file identification and deleted file recovery tool specifications, and no testing in this area has yet begun. Certainly the performance of verification testing on the full arsenal of software tools (and their numerous releases) used by forensic examiners, against the evolving set of storage media types and formats (especially RAID arrays), clearly will easily exceed the resources of any testing laboratory, even NIST's. But efforts toward the establishment of minimal requirements for same are commendable. Nevertheless, the Court's general acceptance (as "authentic") of derivative evidence produced by such software tools must be questioned.

2.7 Exculpatory Evidence May Be Uncollected, Withheld or Destroyed

The pseudo-documentary television series "To Catch A Predator" depicts the vigilante "Perverted Justice" group luring alleged pedophiles into situations where a young teen is left alone. In some of these episodes, an announcer may read a particularly lurid portion of the chatroom discussion to the suspect, in an effort to cajole them into confessing the reason for their appearance at the home.

Some of these matters and other similar scenarios played out by FBI, state or local police, have wound up as part of my caseload. By the time defense discovery is allowed, the computers used to conduct the sting have all conveniently been decommissioned, or are deemed unavailable (on the grounds that perusal of other retained data could compromise unrelated cases). Although defendants may be charged with destruction of evidence for merely reformatting a drive, the Court turns a blind eye to the disappearance of digital media that may have contained exculpatory evidence or could be used to demonstrate entrapment. Prosecution's reports and derivative data materials are also carefully filtered to exclude or obfuscate evidence that could corroborate the defendant's claims.

2.8 Access to Legitimate Services Can Carry a High Degree of Risk

File sharing services are particularly problematic, since the method whereby data is transferred also allows snooping by law enforcement agents for file names or hashes that match known contraband. The U.S. General Accounting Office's 2003 study [6] demonstrated that "in searches on innocuous keywords likely to be used by juveniles, we obtained images that included a high proportion of pornography" including CP and child erotica. Regardless, inadvertent instances of contraband within a file share may be deemed possession with intent to distribute.

Researcher Michael Caloyannides [7] has delineated numerous benign legitimate activities (email, web surfing, freeware, encryption, disk wiping, etc.) that

can be used to demonstrate “an obvious pattern of whatever the accuser wants a court to believe.”

3 Conclusions

These aforementioned forensic challenges have lent a subtle shift toward presumed guilt, rather than an assumption of innocence in today’s court proceedings, which demands correction in order to reinstitute fairness and balance. Some of the issues described herein may be addressed on an individual basis in an effort to improve the body of case law for defense, but others might require more broad attention via explicit tort reform. Recognition that such challenges exist is the first step toward combating these problems.

References

1. AccessData Corporation, The Rules of Digital Evidence and AccessData Technology, http://www.accessdata.com/downloads/media/Rules_of_Digital_Evidence_and_AccessData_Technology.pdf
2. Daubert v. Merrill Dow Pharmaceuticals, 509 U.S. 579 (1993)
3. National Institute of Justice, Test Results for Digital Data Acquisition Tool: FTK Imager 2.5.3.14, NCJ 222982 (2008)
4. National Institute of Justice, Test Results for Digital Data Acquisition Tool: EnCase 4.22a, NCJ 221168 (2008)
5. National Institute of Justice, Test Results for Disk Imaging Tools: dd Provided with FreeBSD 4.4, NCJ 203095 (2004)
6. United States General Accounting Office, Child Pornography Is Readily Accessible over Peer-to-Peer Networks, GAO-03-537T (2003)
7. Caloyannides, M.A.: Forensics Is So “Yesterday”. IEEE Security & Privacy 7(2) (2009)